

**UNIVERSIDAD DE COSTA RICA  
SISTEMA DE ESTUDIOS DE POSGRADO**

**ANÁLISIS NO DIRIGIDO DE LA EVOLUCIÓN DE METABOLITOS DE CACAO  
(*THEOBROMA CACAO* L.) DURANTE EL TIEMPO DE FERMENTACIÓN, APLICANDO  
CROMATOGRAFÍA LÍQUIDA DE ULTRA ALTA RESOLUCIÓN ACOPLADA A  
ESPECTROMETRÍA DE MASAS (CLUR-EM)**

**Tesis sometida a la consideración de la Comisión del Programa de Estudios de  
Posgrado en Ciencia de Alimentos para optar al grado y título de Maestría  
Académica en Ciencia de Alimentos**

**ANA LUCÍA MAYORGA GROSS**

**Ciudad Universitaria Rodrigo Facio, Costa Rica**

**2018**

## 1. DEDICATORIA

*A Fabrice Vaillant, mi mentor profesional:  
quien transformó mi visión de la ciencia y del quehacer científico;  
quien me enseñó que nada es imposible cuando se trabaja con estrategia y ahínco.*

## 2. AGRADECIMIENTOS

*A mis pilares, Dios y a la Virgen*

*A mis papás, a Abue, a Ale, a David, a tía Martha, tía Norma y mi prima Laura M.*

*A mis amistades, en particular a Xioma, Cate, Elías y Andrés Jara+*

**AI CITA:** por todas las oportunidades de crecimiento que me han dado, en particular a Carmela Velázquez, Ana Mercedes Pérez, Fabrice Vaillant e Ileana Morales. A mis colegas Marvin Soto, Carolina Rojas y a Wilfredo Flores quienes me apoyaron en distintas etapas de este proceso.

**A CIPRONA:** por la apertura y espíritu de servicio y colaboración, en especial a Rosaura Romero, Lorena Hernández y Juan Carlos Brenes.

**A la Escuela de Química:** por motivarme en la continua búsqueda de la excelencia científica, en especial a Giselle Tamayo y Luis Manuel Quirós, quienes fueron piezas clave en esta investigación.

**A la Facultad de Farmacia:** en particular a José Manuel Fallas por transferirme sus aprendizajes y conocimientos en metabolómica.

**Al Programa de Maestría en Ciencia de Alimentos, y al Sistema de Estudios de Posgrado (SEP):** por permitirme participar en distintos eventos académicos a nivel internacional que contribuyeron a mi formación profesional y a la proyección de la Universidad de Costa Rica en temas vanguardistas, como lo es la metabolómica.

“Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Ciencia de Alimentos de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Ciencia de Alimentos.”



---

Ph.D. Rosaura Romero Chacón

**Sistema de Estudios de Posgrado**



---

Ph.D. Giselle Tamayo Castillo

**Directora de Tesis**



---

Ph.D. Fabrice Vaillant Barka

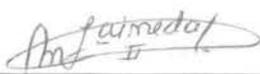
**Asesor**



---

Ph.D. José Manuel Fallas Ramírez

**Asesor**

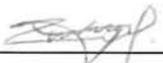


---

M.Sc. María Lourdes Pineda Castro

**Directora**

**Programa de Posgrado en Ciencia de Alimentos**



---

Ana Lucía Mayorga Gross

**Candidata**

### 3. TABLA DE CONTENIDOS

	<i>Página</i>
1. DEDICATORIA .....	ii
2. AGRADECIMIENTOS .....	iii
3. TABLA DE CONTENIDOS .....	v
RESUMEN .....	vii
ABSTRACT .....	viii
4. LISTA DE CUADROS .....	ix
5. LISTA DE FIGURAS .....	xi
6. LISTA DE ABREVIATURAS .....	xv
7. JUSTIFICACIÓN .....	1
8. OBJETIVOS .....	6
8.1. Objetivo general .....	6
8.2. Objetivos específicos .....	6
9. MARCO TEÓRICO .....	7
9.1. Cacao ( <i>Theobroma cacao</i> L.) .....	7
9.1.1. Perspectiva histórica del cacao .....	7
9.1.2. Cultivo .....	8
9.1.3. Clasificaciones .....	9
9.1.4. Morfología y composición del fruto de cacao .....	10
9.1.5. Procesamiento de cacao .....	15
9.2. Metabolómica .....	26
10.8.1. Generalidades metodológicas del análisis metabolómico .....	27
10.8.2. Aplicación de la CL-MS en análisis de metabolitos .....	35
10.8.2.1. Cromatografía líquida .....	35
10.8.2.2. Espectrometría de masas .....	39
10.8.2.3. Espectrometría de masas tándem (MS/MS) y mecanismos de fragmentación .....	45
10. MATERIALES Y MÉTODOS .....	52
10.1. Localización .....	52
10.2. Descripción de la materia prima .....	52
10.3. Metodología .....	54
10.3.1. Procesamiento de cacao .....	54

10.3.2.	Muestreo .....	57
10.3.3.	Extracción de muestras .....	57
10.3.4.	Preparación de controles de calidad .....	59
10.3.5.	Determinación instrumental .....	60
10.3.6.	Procesamiento de la megadata .....	62
10.3.7.	Análisis estadístico.....	63
10.3.8.	Clasificación de la información .....	63
10.3.9.	Identificación tentativa de los iones.....	64
11.	Resultados y discusión .....	65
11.1.	Fermentaciones de cacao .....	65
11.2.	Evaluación de las corridas cromatográficas .....	67
11.3.	Evaluación del modelo estadístico.....	68
11.4.	Análisis de los iones discriminantes .....	73
12.	CONCLUSIONES.....	94
13.	RECOMENDACIONES .....	95
14.	BIBLIOGRAFÍA .....	96
15.	ANEXOS .....	106

## RESUMEN

La fermentación es una etapa crítica en el procesamiento de cacao, principalmente por su impacto en la calidad. En esta etapa se desarrollan los precursores de aromas y sabores, mediante reacciones bioquímicas dentro de las semillas. A pesar de que se han identificado diversos compuestos, todavía se requiere mayor investigación en el tema. Es por esto que en esta investigación se empleó una estrategia de metabolómica, con un abordaje no dirigido, para identificar los metabolitos que presentaron los mayores cambios en tres etapas: a) de 0-2 días, b) de 3 a 4 días y c) de 5 a 6 días de fermentación.

Se estudiaron tres tipos de cacao, los cuales se fermentaron bajo el sistema de cajón. El clon IMC-67 se fermentó en tres ensayos independientes, en una caja de poliestireno de 23 x 22 x 17 cm y con una carga promedio de 40 kg. Cacao de Trinidad y Tobago, con genética desconocida, se fermentó en un ensayo en un cajón de madera con una carga de 30 kg. Finalmente, cacao de Upala (Costa Rica), con genética desconocida, se fermentó en un cajón de madera de 1,0 x 1,0 x 0,5 m, con una carga cercana a 100 kg.

Los muestreos se realizaron diariamente, y las muestras se congelaron hasta finalizar las fermentaciones. Posteriormente se secaron hasta un 7% de humedad. Las muestras se descascarillaron, molieron y desgrasaron. Finalmente, se extrajeron con una solución de acetona:agua:ácido acético (70:29,5:0,5%). Los extractos filtrados y evaporados, se diluyeron en agua y almacenaron en congelación hasta su medición.

Estos extractos se separaron y midieron en un UPLC-ESI-Q-ToF-MS, con una columna C18, y con ionización tanto positiva como negativa. Los datos obtenidos se preprocesaron con el programa MZMine, obteniéndose 2979 iones. Estos iones se emplearon como variables para realizar modelos estadísticos multivariados; las cuales se normalizaron con el algoritmo de Pareto y se transformaron logarítmicamente.

El análisis de componentes principales (PCA) evidenció que, en etapas avanzadas de la fermentación, las muestras tendían a agruparse, independientemente de su procedencia. Adicionalmente se vio una agrupación según los días de fermentación: 0-2 días, 3 a 4 días y 5 a 6 días; lo cual fue coherente con las etapas biológicas: inicios-fermentación anaerobia, fermentación aerobia, final de la fermentación, respectivamente.

El análisis discriminante (PLS-DA) se realizó con estos tres grupos y fue validado. Se utilizó el criterio de índice VIP para reducir las variables de estudio a 56 iones. De estas, se identificaron tentativamente 20 iones, predominando tres naturalezas químicas: azúcares, aminoácidos/oligopéptidos y flavanoles. Algunos metabolitos que destacaron fueron la (epi) catequina, dímeros y trímeros de procianidinas, antocianinas y oligopéptidos de entre 3 y 12 aminoácidos.

En general, los metabolitos propuestos como azúcares y flavanoles mostraron una tendencia a la reducción continua durante el tiempo de fermentación. Caso contrario se presentó en los oligopéptidos, los cuales presentaron un aumento continuo o un máximo en la etapa de fermentación intermedia (3-4 días).

Esta investigación es una base para futuros estudios de caracterización de cacao, que permitirán optimizar procesos en la cadena de valor del cacao. Además, considerando la relevancia nutricional de los flavanoles, es posible plantear nuevas fermentaciones que contemplen criterios no sólo asociados con el desarrollo de precursores de aromas y sabores (como aminoácidos y oligopéptidos), sino también con la retención de polifenoles.

## ABSTRACT

Fermentation is a critical step in cocoa processing, because of its impact in quality. In this step flavor precursors are developed by means of biochemical reactions in the beans. Although several cocoa compounds have been reported, there is still a need of more research. This study applied a non targeted metabolomic approach to identify the metabolites that had the bigger changes in three stages: a) 0-2 days, b) 3-4 days and c) 5 to 6 days of fermentation.

Three different cocoa types were subjected to box fermentation. Clon IMC-67 was fermented in three independent assays, in a polystyrene box (23x22x17 cm) with 40 kg of load. Cocoa variety (genetically undefined) from Trinidad and Tobago, was fermented in a wood box with 30 kg of load. Finally, genetically undefined cocoa from Upala, Costa Rica, was fermented in a wood box (1,0x1,0x0,5 m), with approximated 100 kg of load.

Sampling was done daily, and the samples were frozen until the end of fermentation. Afterwards the cocoa was dried to 7% of moisture; husk was removed, and beans were grounded and defatted. Next, they were extracted with acetone:water:acetic acid (70:29,5:0,5%). The extracts were filtered and evaporated, diluted with water, and frozen.

These extracts were separated and measured in and UPLC-ESI-Q-ToF-MS, using a C18 column, and applying positive and negative modes. The data was preprocessed with MZmine software, and 2979 ions were obtained. These ions were employed as variables in statistical multivariate models, after their normalization via Pareto algorithm and logarithmic transformation.

The principal component analysis (PCA) showed that in advanced fermentation stages, the samples clustered, independently of their origin. Additionally, samples were grouped in function of their fermentation day: 0-2 days, 3-4 days and 5-6 days, which was coherent with the biological steps: beginning-anaerobic fermentation, aerobic fermentation and fermentation ending, respectively.

The discriminant analysis (PLS-DA) was constructed from these three groups, and was validated. VIP index was used for variable reduction, to obtain 56 ions. From these, 20 of them were tentatively identified, showing three predominating nature: sugars, aminoacids/oligopeptides and flavanols. Some of the main metabolites were (epi) catechin, dimers and timers of procyanidins, anthocyanins and oligopeptides with 3 to 12 aminoacids.

In general, the metabolites annotated as sugars and flavanols, showed a tendency to a continuous reduction along fermentation time. A contrasting pattern was showed by oligopeptides which increased continuously or presented a maximum in 3-4 days of fermentation.

This research is a base for future studies of cocoa characterization that could contribute to optimization of processes in favor of cocoa value chain. Besides, considering the nutritional importance of flavanols, it is possible to propose new fermentations that include criteria related not only to the development of flavor precursors (as aminoacids and oligopeptides), but also related to polyphenol retention.

## 4. LISTA DE CUADROS

	<i>Página</i>
Cuadro I. Composición química de la pulpa o mucílago de las semillas de cacao. ....	12
Cuadro II. Composición de semillas de cacao no fermentadas y secas. ....	13
Cuadro III. Enzimas más relevantes en la calidad y que intervienen en la fermentación de cacao. ....	24
Cuadro IV. Composición química de semillas de cacao fermentadas y secas. ....	25
Cuadro V. Comparación de algunas generalidades asociadas a las principales técnicas empleadas en el análisis de metabolitos. ....	32
Cuadro VI. Ejemplos de técnicas multivariadas supervisadas y no supervisadas. ....	33
Cuadro VII. Niveles de confianza para la identificación de metabolitos (Sumner et al., 2007).....	34
Cuadro VIII. Características generales asociadas a la CLAR y la CLUR.....	39
Cuadro IX. Fuentes de iones que operan a presión atmosférica. ....	40
Cuadro X. Clasificación de polifenoles, ejemplos de compuestos y forma química en la que se presenta en la naturaleza (Manach et al., 2004). ....	47
Cuadro XI. Localización de la investigación según el tipo de actividad desarrollada.....	52
Cuadro XII. Descripción de las materias primas a emplear durante los ensayos de fermentación.....	53
Cuadro XIII. Condiciones de fermentación de cacao según el tipo de materia prima.....	56
Cuadro XIV. Estándares elegidos como controles de calidad en los ensayos de metabolómica. ....	60
Cuadro XV. Condiciones instrumentales cromatográficas para la medición de los extractos de cacao.....	61
Cuadro XVI. Condiciones instrumentales en el espectrómetro de masas para la medición en modo continuo de los extractos de cacao.....	61
Cuadro XVII. Parámetros para la identificación de $m/z$ , construcción de cromatogramas, y alineamiento de picos hecho con el software MZMine.....	62
Cuadro XVIII. Lista de los 10 iones con mayor capacidad discriminante y características cromatográficas, espectrométricas y estadísticas asociadas. ....	76
Cuadro XIX. Lista de los iones clasificados tentativamente como flavonoides y características cromatográficas, espectrométricas y estadísticas asociadas.....	77

Cuadro XX. Lista de los iones clasificados tentativamente como oligopéptidos y características cromatográficas, espectrométricas y estadísticas asociadas..... 78

Cuadro XXI. Relación de los oligopéptidos identificados de manera tentativa e hipótesis de su relación con las principales proteínas de cacao..... 92

## 5. LISTA DE FIGURAS

	<i>Página</i>
Figura 1. <i>Theobroma cacao</i> , árbol de cacao (fuente propia). .....	8
Figura 2. Mazorca de cacao unida al árbol (izquierda) y semillas del fruto rodeadas de pulpa (derecha) (fuente propia). .....	11
Figura 3. Procianidinas destacadas por su concentración en cacao. ....	14
Figura 4. (-)-epicatequina, catequina más abundante en cacao. ....	15
Figura 5. Ejemplos de antocianinas presentes en cacao. ....	15
Figura 6. Tipos de fermentadores para cacao (a: apilado, b: en bandejas, c: en cajas o cajones y c: en escalera) (elaboración propia). ....	18
Figura 7. Evolución de los microorganismos durante una fermentación típica de cacao (E. Afoakwa, 2014; Schwan & Wheals, 2004). ....	19
Figura 8. Flujo de proceso general empleado en estudios metabolómicos. ....	27
Figura 9. Componentes principales de un cromatógrafo líquido (a: reservorios de solventes, b: sistema de bombeo, c: mezcladores, d: inyector de muestra, e: precolumna y f: columna analítica en un horno, g: detector y h: sistema computarizado de control y almacenamiento de información; elaboración propia). ....	36
Figura 10. Partículas de fase estacionaria desarrolladas con tecnología híbrida de puentes de etileno (BEH), desarrollada por Waters® (elaboración propia, basada en (Waters <sup>(R)</sup> , n.d.)). ....	38
Figura 11. Ejemplo de un espectrómetro de masas de tipo ESI-QToF (elaboración propia). ....	39
Figura 12. Formación de iones con la técnica de electrospray (ESI) (elaboración propia, adaptada de (Kearle & Tang, 1993)). ....	41
Figura 13. Analizador cuadrupolo-tiempo de vuelo (Q-TOF, por sus siglas en inglés). ....	44
<b>Figura 14.</b> Clasificación de polifenoles (Manach et al., 2004). ....	46
Figura 15. Ejemplos de mecanismos de fragmentación de una proantocianidina en modo de ionización negativo (Demarque, Crotti, Vessecchi, Lopes, & Lopes, 2016; Friedrich, Eberhardt, & Galensa, 2000). ....	48
Figura 16. Enlace peptídico entre dos aminoácidos. ....	49
Figura 17. Principales vías de fragmentación en péptidos (Banerjee & Mazumdar, 2012). ....	50

Figura 18. Principales rutas de fragmentación de los azúcares (Banerjee & Mazumdar, 2012).....	51
Figura 19. Resumen del número de ensayos de fermentación –representados por caja unitaria-, y los puntos de muestreo por ensayo –representados por estrellas. ....	53
Figura 20. Secuencia de operaciones empleada para el desarrollo de las fermentaciones de cacao.....	55
Figura 21. Flujo del proceso para la obtención de la fracción orgánica de interés, a partir de granos de cacao fermentados. ....	58
Figura 22. Evolución de la temperatura de fermentación en el tiempo, para los ensayos piloto del cacao clon IMC-67. ....	65
Figura 23. Evolución de la temperatura de fermentación en el tiempo, para el ensayo del cacao con genética desconocida, de Trinidad y Tobago. ....	66
Figura 24. Ejemplo de cromatogramas en modo positivo (arriba) y negativo (abajo) de una muestra de cacao con un día de fermentación.....	67
Figura 25. Pool de estándares externos empleados como control de calidad. ....	68
Figura 26. Análisis de componentes principales con coloración según tipo de híbrido: el color azul corresponde al clon IMC-65 de Costa Rica, el verde a las muestras de genética desconocida provenientes de Trinidad y Tobago, y la roja a las muestras de genética desconocida provenientes de Upala, Costa Rica.....	69
Figura 27. Análisis de componentes principales con coloración según días de fermentación: las muestras color azul (No. 1) corresponden a aquellas fermentadas de 0 a 3 días, y las de color verde (No. 2) se fermentaron entre 4 y 6 días. ....	70
Figura 28. Análisis de componentes principales con coloración según día de fermentación: cada número en la leyenda del gráfico corresponde al día de fermentación de cada muestra. ....	70
Figura 29. Distribución de las muestras de un modelo no supervisado de tipo análisis de componentes principales (PCA, arriba) (las muestras color azul corresponden a los días de fermentación de 0 a 2 días, las muestras en verde a los 3 y 4 días, y las muestras en rojo a los días 5 y 6). ....	71
Figura 30. Distribución de las muestras de un modelo supervisado de tipo discriminante (PLS-DA, abajo) (las muestras color azul corresponden a los días de fermentación de 0 a 2 días, las muestras en verde a los 3 y 4 días, y las muestras en rojo a los días 5 y 6). ...	72

<b>Figura 31.</b> $R^2$ y $Q^2$ obtenidos al realizar una prueba con 100 permutaciones y 2 componentes principales.....	72
<b>Figura 32.</b> Variables (iones medidos; color verde y color morado) del PLS-DA y variables con mayor poder discriminante, según el índice VIP (color rojo). .....	74
Figura 33. Sección del cromatograma asociado a los dos iones con mayor poder discriminante ( $m/z$ 343.1 y 325.1) de una muestra de cacao.....	79
Figura 34. Sección del cromatograma asociado a los dos iones con mayor poder discriminante ( $m/z$ 343.1 y 325.1) de un pool de estándares de sacáridos: sacarosa $[M+H]^+=343.1$ , maltosa $[M+H]^+=343.1$ , lactosa $[M+H]^+=343.1$ y glucosa $[M+H]^+=181.0$ . ...	80
<b>Figura 35.</b> Barrido del espectro de masas experimental en los tiempos de retención (0,7-0,8 min) correspondientes a los dos iones con mayor poder discriminante ( $m/z$ 343.1 y 325.1) en un extracto de cacao.....	81
<b>Figura 36.</b> Espectro de masas MS/MS en los tiempos de retención 0,7-0,8 min de un pool de estándares de sacáridos, a 20 eV.....	81
Figura 37. Espectro de masas en modo negativo asociado a la región de tiempo de los dos iones con mayor poder discriminante: $[M-H]^-$ $m/z$ 341.1 y 323.1, determinado experimentalmente (arriba) y espectro reportado para la sacarosa en la base de datos Metlin a 10 V (abajo). .....	82
Figura 38. Comportamiento en el tiempo de fermentación del ión $[M+H]^+=343.1$ (azul: 0-2 días de fermentación, verde: 3-4 días de fermentación y rojo: 5-6 días de fermentación). Ver nomenclatura en sección 11.4. ....	83
Figura 39. Espectro de masas MS/MS experimentales obtenidos en modo positivo en a) la presente investigación para el ión $[M+H]^+$ $m/z$ 291,0 con RT de 8,9 a 15 eV (arriba) y b) de la epicatequina a 20 eV según la base de datos MetLine (abajo). .....	84
Figura 40. Espectro de masas de una muestra de cacao con dos días de fermentación, en modo positivo y en una ventana de 70 a 1000 Da. ....	85
Figura 41. Ejemplo de descomposición por eliminación Retro Diels Alder del ión $[M+H]^+$ 579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX). .....	86
Figura 42. Ejemplo de descomposición por fisión heterocíclica del anillo del ión $[M+H]^+$ 579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX). .....	86

Figura 43. Ejemplo de descomposición por ruptura directa del enlace interflavano del ión $[M+H]^+$ 579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX). .....	86
Figura 44. Espectro de masas MS/MS del ión $[M+H]^+$ 579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX). .....	87
Figura 45. Espectro MS/MS de la procianidina B2, a 20 eV en modo positivo, reportado en la base MetLin. ....	87
Figura 46. Ión $m/z$ 739.1 con TR de 10.2 min en modo positivo y con una energía de colisión de 15 eV (arriba) y espectro predicho por la base de datos HMDB en modo positivo a 20 eV para 3-O- $\beta$ -D-galactopiranosil proantocianidina A5'.....	89
Figura 47. Ejemplo del comportamiento en el tiempo de fermentación del ión identificado tentativamente como (epi) catequina (azul: 0-2 días de fermentación, verde: 3-4 días de fermentación y rojo: 5-6 días de fermentación). Ver nomenclatura en sección 11.4. ....	90
Figura 48. Espectros de masas, en modo positivo y a 15 eV, del ión anotado tentativamente como PVN (arriba), y espectro de masas experimental reportado por PVN en la base MetLin a 20 eV (abajo). ....	91
Figura 49. Evolución a lo largo del tiempo de fermentación de los metabolitos anotados tentativamente como PVN y VSTDVN catequina (azul: 0-2 días de fermentación, verde: 3-4 días de fermentación y rojo: 5-6 días de fermentación). Ver nomenclatura en sección 11.4. ....	93

## 6. LISTA DE ABREVIATURAS

**ANDEVA/ANOVA:** análisis de variancia

**CG:** Cromatografía de gases

**CL:** cromatografía líquida

**CLAR:** cromatografía líquida de alta resolución

**CLUR:** cromatografía líquida de ultra alta resolución

**EC:** electroforesis capilar

**EM:** espectrometría de masas

**FE:** fase estacionaria

**FHA:** fisión heterocíclica de anillo

**FM:** fase móvil

**FN:** fase normal

**FR:** fase reversa

**FTIR:** espectroscopía infrarroja por transformada de Fourier

**HILIC:** cromatografía de interacción hidrofílica

**IDEM:** infusión directa en un espectrómetro de masas

**IUPAC:** unión internacional para la química pura y aplicada

***m/z:*** relación masa/carga

**PCA:** análisis de componentes principales

**Q-TOF:** analizador cuadrupolo-tiempo de vuelo

**RDA:** eliminación Retro-Diels-Alder

**RMN:** espectroscopía de resonancia magnética nuclear

**UFC:** unidades formadoras de colonias

## 7. JUSTIFICACIÓN

El cacao es un árbol perteneciente a la familia Malvaceae, género *Theobroma*. A pesar de que existen aproximadamente 22 especies, la de mayor aprovechamiento comercial es la especie *cacao*. Este árbol crece en zonas tropicales, a unos 10° de latitud con respecto a la franja del Ecuador. La altitud de siembra es por lo general menor a 1000 m.s.n.m., a temperaturas entre los 15 y 32 °C, con una cantidad anual de lluvias entre los 1250 y los 3000 mm (E. Afoakwa, 2014; Biehl & Ziegler, 2003b; Dand, 2011b; ITIS, n.d.; Nair & Nair, 2010).

Según reportes del International Cocoa Organization (ICCO), para el período 2014-2015 más de un 70 % de la producción mundial de granos de cacao provenía de África; América representó en promedio un 18 % de este total y Asia en conjunto con Oceanía aportaron un valor cercano al 10 % (Navarro, 2016). Para el período 2015-2016 se ha estimado un déficit de producción de cacao mayor a las 200 mil toneladas para poder satisfacer la demanda de los mercados mundiales (ICCO, 2016b). El impacto del cambio climático en la producción, la creciente demanda de derivados de cacao y la exigencia por garantizar sistemas productivos justos, responsables y sostenibles, son algunos de los factores que explican estas proyecciones (Gross, 2014; Läderach, Martínez-Valle, Schroth, & Castro, 2013).

Ante este panorama se han visualizado nuevas oportunidades y planteado diversas estrategias tanto en los sectores productivos como en los industriales. Un ejemplo es el caso de la empresa alemana de chocolates Ritter Sport, la cual ha apostado a tener su propia plantación de cacao en Nicaragua, con 2500 Ha, constituyéndose en una de las más grandes del mundo (RitterSport, n.d.).

A nivel comercial, los granos de cacao son clasificados, según su calidad, en dos categorías. Estas son cacao ordinario, también conocido como *bulk* (por su nombre en inglés), y el cacao fino y de aroma. Éste último representa tan sólo un 5 % del total de la producción mundial de cacao y se distingue principalmente por sabores y aromas que incluyen notas frutales, florales, herbales y otros. Algunos criterios adicionales que se consideran al realizar la clasificación son el origen genético, las características

morfológicas de la planta, la composición química, el color de los granos, el grado de fermentación y secado y el porcentaje de contaminantes (ICCO, 2016a).

Centroamérica participa aproximadamente con un 0,22% de la producción mundial de cacao (Navarro, 2016). Esta región se visualiza como un competidor por calidad y no por volumen (Orozco, Aguilar, Deheuvels, Villalobos, Rodríguez, & Somarriba, Chávez, 2015), ya que una parte importante de su producción es cacao fino y de aroma.

Con la misma distinción de la región centroamericana, el cacao costarricense destaca por su calidad y, según estadísticas de la ICCO, la totalidad de exportaciones de cacao de Costa Rica corresponden a cacao fino y de aroma (ICCO, 2016a). En cuanto a productividad, este país se posiciona en cuarto lugar en Centroamérica, superado casi el doble por Guatemala y Honduras y seis veces por Nicaragua, quien lidera la región (Navarro, 2016).

Al 2014 se reportó que la producción de cacao costarricense fue de 700 ton/año, en donde la zona Huetar Atlántica concentró más del 60% de la producción, seguida por la zona Brunca con casi un 25%, la zona Huetar Norte con un 7% y finalmente la zona Pacífico Central con menos de un 1 % (Flores-González, 2015).

Actualmente Costa Rica experimenta un proceso de reactivación del cultivo. Existen diversos proyectos de siembra y capacitación agrícola en el país, así como estrategias para fomentar el valor agregado, por lo cual se pretende que estas cifras aumenten. Lo anterior ha tenido el apoyo a nivel gubernamental, que queda expreso al emitir un decreto en el cual se "*declara de interés nacional, la producción, procesamiento, comercialización y exportación del cacao y sus derivados*" (Solís-Rivera, Arauz-Cavallini, & Durán-Salvatierra, 2015).

Uno de los principales problemas asociados con el sector cacaotero costarricense es la variabilidad en la calidad de los granos de cacao fermentados y secos que entregan los productores a las industrias. La calidad es un conjunto de características intrínsecas que se ven influenciadas por factores como el tipo de variedad empleada, las técnicas agronómicas aplicadas y los procesos de transformación poscosecha utilizados (Aprotosoai, Luca, & Miron, 2016). Es por esto por lo que los esfuerzos para garantizar la calidad deben ser abordados de manera integral.

El proceso tradicional para la obtención de granos de cacao comercializables combina dos operaciones básicas: la fermentación y el secado. Posteriormente, estos granos son transformados en plantas industriales que permiten agregar valor en forma de cacao en polvo, manteca de cacao, chocolates y otros derivados. Es en la fermentación donde se producen la mayor parte de precursores de aromas y sabores en los granos como péptidos, aminoácidos, ácidos orgánicos y azúcares reductores (Großmann-Kühnau, 2015), los cuales en la operación de tostado se traducen en notas características a chocolate. Es usual que los granos comercializables se obtengan en la zona de siembra mientras que los distintos derivados se elaboren en industrias en zonas distintas, por lo cual es necesaria la existencia de controles a lo largo de la cadena que permitan conservar la calidad (Aprotosoia et al., 2016; Biehl & Ziegleder, 2003b).

Durante la fermentación las mazorcas se quiebran para extraer los granos y, en el proceso, estos son inoculados de manera espontánea por el ambiente que los rodea. El inóculo, el cual está conformado por una diversidad de bacterias y hongos, permite que se lleven a cabo una serie de reacciones secuenciales en dos etapas distinguibles: en primer lugar una fermentación anaerobia, seguida por fermentación aerobia. Algunos de los cambios que ocurren en los granos son: la degradación de la pulpa, la producción de ácido acético, la muerte del embrión y el subsecuente despliegue de reacciones *post-mortem* como hidrólisis proteica, oxidación de polifenoles, hidrólisis de azúcares, cambios de color y textura del grano, entre otros (E. O. Afoakwa, Quao, Takrama, Budu, & Saalia, 2012; Biehl & Ziegleder, 2003a).

La fermentación puede realizarse con distintas técnicas, como, por ejemplo, en cajas de madera, en sacos, baldes, o realizando cúmulos de granos sobre hojas de plátano (Großmann-Kühnau, 2015; Papalexandratou, Vrancken, De Bruyne, Vandamme, & De Vuyst, 2011). A pesar de que todas estas técnicas son frecuentemente empleadas, los controles durante el proceso son pocos, generando mucha variabilidad en el producto final y, por ende, pérdidas por producto no conforme. Es por esto por lo que es necesario desarrollar protocolos de fermentación óptimos, de la mano con la implementación de técnicas de control de calidad apropiadas.

Desde un punto de vista bioquímico, se han realizado una serie de investigaciones que pretenden comprender los distintos fenómenos que ocurren durante la fermentación. Por ejemplo, se ha estudiado el papel de diversos microorganismos, como levaduras y

bacterias ácido lácticas y ácido acéticas, en esta etapa (Ardhana & Fleet, 2003; Mozzi, Ortiz, Bleckwedel, De Vuyst, & Pescuma, 2013). En contraposición, otros autores han demostrado que la formación de precursores puede ocurrir en ausencia de microorganismos, al incubar con alcohol, ácido acético y pectinasas (Großmann-Kühnau, 2015).

Se ha estudiado la secuencia de metabolitos generados o transformados durante esta etapa, ciertas rutas asociadas y su impacto en la calidad. Algunos de los avances incluyen la identificación de productos de la hidrólisis de albúminas y globulinas de tipo vicilina (7S), como por ejemplo oligopéptidos hidrofílicos y amino ácidos hidrofóbicos (Kongor et al., 2016), los cuales se han visto implicados en la generación de sabores y aromas durante el tostado de cacao (Amin, Jinap, & Jamilah, 1997; Biehl & Ziegleder, 2003a; Kongor et al., 2016; Marseglia et al., 2014; J. Voigt, Heinrichs, Voigt, & Biehl, 1994). Los flavonoides también son afectados durante la fermentación, ya que suelen oxidarse, hidrolizarse, polimerizar o formar distintos complejos (E. O. Afoakwa et al., 2012; Nazaruddin, Seng, Hassan, & Said, 2006). Estos metabolitos contribuyen con el perfil sensorial de los granos de cacao al proporcionar notas astringentes y diversidad de colores. A pesar del conocimiento generado hasta el momento, todavía existen vacíos que permitan comprender desde una perspectiva global los cambios y rutas que ocurren a nivel molecular durante la fermentación; por tanto, la investigación permanece abierta.

Otra perspectiva de la calidad del cacao podría abarcar la dimensión del valor nutricional, un tema que ha atraído la atención de consumidores que desean cuidar su salud de manera preventiva a través de la alimentación. El cacao ha llamado la atención en este sentido, principalmente por su elevado contenido de flavonoides, que podría alcanzar un 13% del grano seco sin fermentar (Kim, Lee, & Lee, 2011), siendo de particular importancia algunos flavonoides como la epicatequina, la catequina, los oligómeros de procianidinas y las cianidinas glicosidadas (Kim et al., 2011; Niemenak, Rohsius, Elwers, Omokolo Ndoumou, & Lieberei, 2006; Ortega et al., 2008). Incluso la Autoridad Europea de Inocuidad Alimentaria (EFSA, por sus siglas en inglés), aprobó una declaración nutricional que expresa que los flavanoles de cacao poseen efectos fisiológicos positivos, en cuanto contribuyen con el mantenimiento normal de la vasodilatación dependiente del endotelio (EFSA, 2014). El consumo de flavanoles se ha correlacionado con distintos beneficios en la salud; por ejemplo la regulación de la presión, el mejoramiento de la

función vascular, efectos antiinflamatorios y otros (Balzer et al., 2008; Corti, Flammer, Hollenberg, & Lüscher, 2009). Por tanto, las implicaciones de la fermentación sobre la funcionalidad del producto final podrían considerarse como un criterio para evaluar la calidad.

La alimentómica es una disciplina que estudia los dominios de los alimentos y la nutrición, por medio de la aplicación de técnicas *ómicas* avanzadas para mejorar el bienestar, salud y confianza del consumidor (Cifuentes, 2009). Una de estas técnicas *ómicas* es la metabolómica, cuya apuesta es el determinar de manera integral los metabolitos presentes en un sistema biológico y describir sus interrelaciones o rutas de conexión metabólica. Esto da una ventaja distintiva con respecto a las metodologías tradicionales, las cuales se enfocan en la determinación de un metabolito o grupos de metabolitos similares de manera no específica.

Bajo este contexto se propone la presente investigación, la cual permitirá aproximarse al fenómeno de la fermentación de cacao desde una perspectiva molecular. En esta se pretende analizar el comportamiento de una fracción del metaboloma de cacao a lo largo del tiempo de fermentación, combinando técnicas de análisis instrumental de tipo cromatografía líquida de ultra alta resolución acoplada a espectrometría de masas (CLUR-ES) y análisis estadístico multivariado. Esto se realizará con miras a proponer moléculas que funjan como marcadores de proceso, que permitan comprender con mayor profundidad la fermentación y que eventualmente contribuyan con el desarrollo de nuevos protocolos de fermentación y/o control de calidad.

## 8. OBJETIVOS

### 8.1. Objetivo general

Analizar la evolución de metabolitos de cacao (*Theobroma cacao* L.) a lo largo del tiempo de fermentación, con un abordaje metabolómico no dirigido, en extractos de agua-acetona acidificados medidos con la técnica de cromatografía líquida de ultra alta resolución acoplada a espectrometría de masas (CLUR-EM).

### 8.2. Objetivos específicos

1. Analizar y seleccionar los iones discriminantes más relevantes en la distinción de tiempos de fermentación de cacao, en extractos de agua-acetona acidificada, medidos con CLUR-EM y clasificados con la aplicación de estadística multivariada.
2. Asignar una identificación tentativa de los iones con mayor poder discriminante entre tiempos de fermentación y definirlos como marcadores biológicos del proceso.

## 9. MARCO TEÓRICO

### 9.1. Cacao (*Theobroma cacao* L.)

#### 9.1.1. Perspectiva histórica del cacao

El *Theobroma cacao* o cacao es un árbol de la familia Malvaceae (ITIS, n.d.), atractivo por sus frutos, los cuales son la materia prima para producir importantes derivados como el chocolate.

La primera publicación referente al cacao fue hecha en 1605 por Clusius, el cual lo llamó con el nombre de *Cacao fructus* (Clusius, 1605). Posteriormente, en el año 1700, Tournefort propuso, para efectos de clasificación taxonómica, el género *Cacao* con una única especie (Tournefort, 1700). Fue hasta el año 1737 que Linnaeus cambió el nombre del género por *Theobroma*, cuyo significado es “alimento para los dioses”. Además, planteó la existencia de dos especies: *T. cacao* y *T. guazuma* (Linnaeus, 1737). A lo largo de la historia se fueron proponiendo distintas clasificaciones de los tipos de cacao; esto se profundizará en la sección 10.3. del presente documento.

Se considera que el cacao tiene su origen en América del Sur, y que posteriormente fue introducido en América Central a través de rutas comerciales indígenas (Dand, 2011a; Motamayor et al., 2002).

El cacao fue un elemento importante que contribuyó al desarrollo de las poblaciones. Por ejemplo, existen reportes de que los mayas y aztecas empleaban las semillas de cacao como moneda de intercambio, y también solían preparar una bebida a base de cacao, vainilla, maíz y chile que se ofrecía en ceremonias religiosas (Dand, 2011a).

En el siglo XV el encuentro de poblaciones americanas y europeas provocó la creación de nuevas rutas y zonas mercantiles. Este hito histórico facilitó la dispersión del cacao en diferentes países, como por ejemplo Jamaica, Trinidad, Martinica, Santa Lucía, Java, Surinam (Dand, 2011a), y otros.

### 9.1.2. Cultivo

El *Theobroma cacao* L. es mayormente cultivado en zonas tropicales húmedas, 10 ° al norte y al sur del Ecuador. Es un cultivo que se siembra en zonas de poca altura, por lo general debajo de 1000 m.s.n.m., con temperaturas entre 18 y 32 °C. El nivel de precipitación de la zona debe encontrarse preferiblemente entre los 1500 y los 2500 mm anuales. Este cultivo se adapta a diferentes tipos de suelo, con pH entre 5 y 7,5 (Emmanuel Ohene Afoakwa, 2014).

La altura del árbol es de 4 a 10 m (Emmanuel Ohene Afoakwa, 2014; Dand, 2011b), que en estado salvaje puede llegar a alcanzar alturas de hasta 25 m (Ph. Lachenaud, Paulin, Ducamp, & Thevenin, 2007) (Figura 1).



Figura 1. *Theobroma cacao*, árbol de cacao (fuente propia).

Un árbol puede producir en promedio entre 10 y 50 mazorcas de cacao al año (Dand, 2011b), que normalmente se cosechan en dos períodos, determinados según la región específica (Vega & Kwik-Urbe, 2012). El rendimiento de la producción de cacao ronda los 300-600 kg/Ha, a pesar de que se ha demostrado de que el potencial en distintas variedades puede ser mucho mayor, inclusive superando los 6000 kg/Ha (Dand, 2011c; van Vliet & Giller, 2017; Wessel & Quist-Wessel, 2015).

Se ha planteado que algunas de las principales causas directas de los bajos rendimientos de producción son el inadecuado mantenimiento de las plantaciones en cuanto a control de plagas y enfermedades, mal manejo de sombra, uso de fertilizantes, carencia de nutrientes, así como también se le atribuye a la antigüedad de los árboles (van Vliet & Giller, 2017; Wessel & Quist-Wessel, 2015); así como también es determinado por el tipo de material genético.

Dentro de las principales enfermedades se señalan la mazorca negra, escoba de bruja, virus del brote hinchado y moniliasis; causadas respectivamente por los microorganismos *Phytophthora* spp., *Moniliophthora perniciosa*, *Badnavirus* y *Moniliophthora roreri* (O. A. Gutiérrez, Campbell, & Phillips-Mora, 2016).

### 9.1.3. Clasificaciones

En 1882, Morris propuso formalmente por primera vez una nomenclatura para el *Theobroma cacao*, planteando dos grandes clases: cacao Criollo y cacao Forastero (Morris, 1882), división que para el año 1944 se mantuvo apoyada por Cheesman (Cheesman, 1944). En 1964 se establecieron dos subespecies *Theobroma cacao* subsp. *cacao* y *Theobroma cacao* subsp. *sphaerocarpum*; en donde la primera subespecie se hipotetizó como originaria de Centroamérica y México y recibió el nombre de cacao criollo, mientras que la segunda se asoció a una mayor ocurrencia en América del Sur, conocido comúnmente como cacao Calabacillo, Amelonado y Forastero, entre otros nombres (Cuatrecasas, 1964). Subsiguientemente esta teoría de origen fue substituida por aquella que propone que el cacao se originó en América del Sur (Motamayor et al., 2002).

El cacao Criollo fue asociado con una calidad distinguida y valorada en el mercado. En general, éste presentaba algunas características, como, por ejemplo, mazorcas puntiagudas con corteza delgada, superficie rugosa y surcada, semillas blancas o con poca pigmentación que requieren de una fermentación corta (2-4 días), entre otros (Cuatrecasas, 1964; G.A.R., Wood, & Lass, 1985; Lockwood & Gyamfi, 1979; Soria-V, 1970).

El cacao Trinitario se refiere a un híbrido entre un Criollo y un Amelonado, proveniente de Trinidad (P. Lachenaud & Motamayor, 2017). Este híbrido en principio ofrece características intermedias de rendimiento, resistencia y calidad, entre sus padres.

Estas clasificaciones consideraban criterios limitados, y frecuentemente no describían de forma comprensiva muchos tipos de cacao; era común encontrar una alta variabilidad dentro de los grupos propuestos. Además, era frecuente encontrar diferentes nombres para una misma variedad, o un mismo nombre para diferentes genotipos (Bartley, 2005).

En contraposición con la clasificación tradicional, para el año 2001 (Motamayor, 2001) se señala que el cacao llamado “Criollo” no conforma una subespecie y para el 2008 se realiza por primera vez una clasificación del cacao que incorporaba criterios genéticos. Este estudio propone 10 grupos genéticos: Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañón, Nacional, Nanay y Purús (Motamayor et al., 2008). A pesar de esto, todavía es muy común encontrar en la literatura actual el uso de las clasificaciones clásicas de Criollo, Forastero y Trinitario.

De la mano con los avances científicos, es común obtener variedades de cacao mejoradas, lo que permite potenciar características genéticas como los son la resistencia a enfermedades, productividad y calidad (Lieberei, Kadow, & Seigler, 2013).

Algunos de los genotipos empleados tradicionalmente en las estrategias de mejoramiento provienen de las series Scavina (SCA), Pound, Nanay (NA), Parinari (PA), United Fruit (UF), Iquitos Mixed Calabacillo (IMC) e Imperial College Selection (ICS), los cuales se derivan de diferentes procesos de selección (Phillips-Mora, Arciniegas-Leal, Mata-Quirós, & Motamayor-Arias, 2012).

Fruto del Programa de Mejoramiento Genético del CATIE, en el 2007 se eligió un paquete tecnológico de 6 clones CATIE-R1, CATIE-R4, CATIE-R6, CC-13, ICS-95 T1 y PMCT-8, que es actualmente distribuido en Centroamérica, incluyendo a Costa Rica (Phillips-Mora et al., 2012). Cada uno de estos híbridos posee perfiles de resistencia, productividad y calidad distintos, pero la siembra en forma conjunta permite compensar las debilidades de cada uno y potenciar las fortalezas.

#### **9.1.4. Morfología y composición del fruto de cacao**

El fruto del cacao puede medir de 15 a 25 cm de longitud y pesar en promedio 500 g (Emmanuel Ohene Afoakwa, 2014; Ph. Lachenaud et al., 2007). Recibe comúnmente el nombre de mazorca y alberga en su interior un total de 30 a 50 semillas rodeadas de un mucílago o pulpa, las cuales están sujetas a una placenta (Nair & Nair, 2010) (Figura 2).



**Figura 2.** Mazorca de cacao unida al árbol (izquierda) y semillas del fruto rodeadas de pulpa (derecha) (fuente propia).

La cáscara de la mazorca es gruesa, y puede llegar a alcanzar desde el 50 % y hasta el 80% del peso total del fruto (de Souza, Moreira, Sarmiento, & da Costa, 2018; Dias, Schwan, Freire, & Seródio, 2007). Su color está en función de la variedad y su estado de madurez. Por ejemplo, si en forma inmadura presenta coloraciones verdes, se espera que al madurar se tornen amarillas, mientras que aquellas que son originalmente rojizas, suelen alcanzar tonalidades amarillas y anaranjadas al madurar (Nair & Nair, 2010).

La pulpa o mucilago que rodea a las semillas es el endospermo y puede constituir 0,8-1,2 g/semilla (Biehl & Ziegleder, 2003a). Es de carácter ácido –atribuido principalmente al ácido cítrico- y dulce, dado el alto contenido de azúcares, en donde destacan la glucosa y la fructosa (Cuadro I) (E. Afoakwa, 2014; Anvoh, Bi, & Gnakri, 2009; Igbinador, 2009). Además, posee una serie de compuestos inhibidores de la germinación, que previenen que las semillas germinen dentro de la mazorca durante el período de madurez (Wood & Lass, 2001).

La pulpa y las semillas de cacao se encuentran divididas por una capa llamada testa, que posee altos niveles de fibra dietética -mayoritariamente insoluble-, y diversidad de polifenoles (Okiyama, Navarro, & Rodrigues, 2017; Vega & Kwik-Urbe, 2012).

**Cuadro I.** Composición química de la pulpa o mucílago de las semillas de cacao.

<b>Propiedad</b>	<b>Concentración</b>
Humedad (g/100 g)	79,2-87 <sup>eg</sup>
Proteína (g/100 g b.s.)	5,47-5,56 <sup>a</sup>
Cenizas (g/100 g b.s.)	7,51-7,68 <sup>a</sup>
Grasa (g/100 g b.s.)	1,91-1,92 <sup>a</sup>
Carbohidratos (g/100 g b.s.)	67,99-68,35 <sup>a</sup>
Azúcares totales (g/100 g)	10-17,4 <sup>bg</sup>
Azúcares reductores (g/100 g)	10,7 <sup>b</sup>
Almidón (g/100 g)	0,16 <sup>b</sup>
Pectina total (g/100 g)	1-2,5 <sup>bg</sup>
Pectinas (mg/semilla)	12-13 <sup>c</sup>
Hemicelulosas y celulosas (mg/semilla)	23-31 <sup>b</sup>
pH	3,2-3,8 <sup>bcd</sup>
Ácido cítrico (g/100 g)	1-2 <sup>g</sup>
Acidez titulable (mEq ácido málico/L)	1,0-1,84 <sup>bf</sup>
Sólidos solubles (°Brix)	14,81-20,5 <sup>bdf</sup>

Nota: b.s. = base seca; a=(Martínez et al., 2012), considerado ; b=(Dias et al., 2007); c=(Biehl & Ziegleder, 2003a), d=(Anvoh et al., 2009), e=(Igbinalodor, 2009), f=(de Souza et al., 2018), g=(E. Afoakwa, 2014).

Las semillas de cacao están conformadas por dos cotiledones unidos por un eje embrionario. Usualmente tienen una longitud entre los 2 y 3 cm y, dependiendo de la variedad, pueden presentar coloraciones que van desde blanco hasta morado (de Souza et al., 2018). Su composición es variable y depende en gran parte de su genética; algunos reportes se detallan en el Cuadro II.

Cerca de un 80% de las células presentes en las semillas almacenan lípidos y proteínas. Los primeros son almacenados en cuerpos lipídicos de 2  $\mu\text{m}$  de tamaño, que rodean los amiloplastos y las vacuolas donde se almacenan proteínas (Biehl & Ziegleder, 2003a; Lieberei et al., 2013; Martini, Figueira, Lenci, & Tavares, 2008).

La fracción lipídica es la fracción mayoritaria en las semillas de cacao, que es cerca de un 95% triacilgliceroles, 2% diacilgliceroles y el porcentaje restante lo conforman monoacilgliceroles, lípidos polares y ácidos grasos libres. Los ácidos oleico, esteárico y palmítico se encuentran en porcentajes similares y constituyen cerca del 90% de los ácidos grasos (E. Afoakwa, 2014; Biehl & Ziegleder, 2003a).

**Cuadro II.** Composición de semillas de cacao no fermentadas y secas.

<b>Propiedad</b>	<b>Concentración</b>
Humedad (g/100 g)	4,2-5,0 <sup>1</sup> ab
Grasa (g/100 g)	55,2-62,9 <sup>ab</sup>
Proteínas (g/100 g)	17,5-21,6 <sup>ab</sup>
Carbohidratos (g/100 g)	15,5 <sup>a</sup>
Fibra (g/100 g)	5,9 <sup>b</sup>
Cenizas (g/100 g)	3,5-4,4 <sup>ab</sup>
Polifenoles totales (mg equivalentes de catequina/g)	77,31 <sup>c</sup>
Antocianinas totales (mg/g)	3,01 <sup>c</sup>
HCN (mg/100 g)	7,1 <sup>b</sup>
Oxalatos totales (mg/100 g)	194,3 <sup>b</sup>
Oxalatos solubles (mg/100 g)	29,3 <sup>b</sup>
Teobromina (g/100 g)	2,9 <sup>b</sup>

*Nota:* a=(Emmanuel Ohene Afoakwa, Quao, Takrama, Budu, & Saalia, 2013); se analizaron muestras deshidratadas, b=(Aremu, Agiang, & Ayatse, 1995), reportados en base seca, c=(do Carmo Brito, Campos Chisté, da Silva Pena, Abreu Gloria, & Santos Lopes, 2017); se analizaron muestras liofilizadas.

Las proteínas del cacao son principalmente de reserva: albúminas y globulinas. Una albúmina de 21 kDa, soluble en agua, alcanza cerca de un 50% del total proteico, mientras que cerca de un 40% lo conforman las globulinas tipo vicilina (7S) -solubles en soluciones salinas- que poseen tres subunidades (Biehl & Ziegleder, 2003a; Kongor et al., 2016; Kratzer et al., 2009). Además, se ha reportado la presencia de diversas enzimas que tienen papeles preponderantes en etapas como la fermentación de cacao, como proteasas, polifenol oxidasas, glicosidas, lipasas, peroxidasas, catalasas, fosfatasas y otras (E. Afoakwa, 2014; Vega & Kwik-Urbe, 2012).

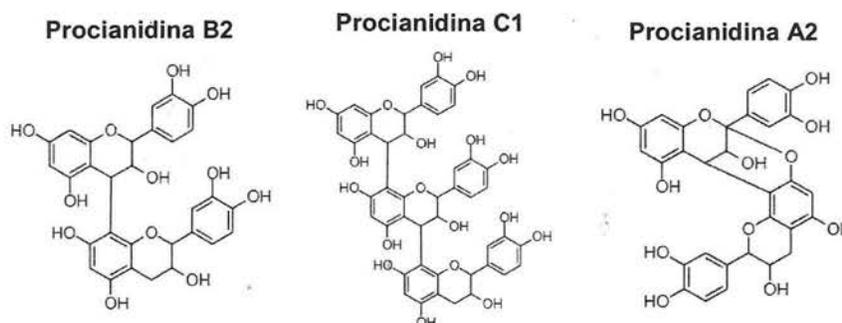
La sacarosa es el azúcar más concentrado en las semillas de cacao, puede llegar a representar un 2 % del total de la semilla, y cerca de un 90 % de los azúcares totales. Además, existen otros carbohidratos presentes como: fructosa, glucosa, almidón, fibra dietética (celulosa, pentosanos y otros), estaquiosa, rafinosa y verbascotetrosa, entre otros (E. Afoakwa, 2014; Hii, Law, Suzannah, Misnawi, & Cloke, 2010)

<sup>1</sup> El grano fresco posee un porcentaje cercano al 30 % de humedad (Hii et al., 2010; Kongor et al., 2016).

Alrededor de una quinta parte de las células de las semillas de cacao son un tipo de idioblastos que almacenan metabolitos secundarios; cerca de un 15 % son purinas y más de un 60 % polifenoles (Biehl & Ziegleder, 2003a; Lieberei et al., 2013; Martini et al., 2008).

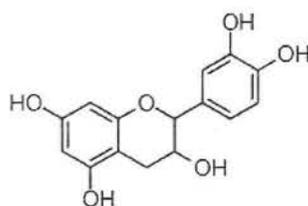
Las tres principales xantinas son la teobromina, la cafeína y la teofilina. La teobromina se encuentra por lo general en porcentajes entre 1 y 3 % (b.s.) mientras que la cafeína usualmente está por debajo del 2 % (b.s.). La teofilina por su lado y otras metilxantinas se encuentran en cantidades mucho menores (Belščak, Komes, Horžić, Ganić, & Karlović, 2009; Biehl & Ziegleder, 2003a; Vega & Kwik-Urbe, 2012).

Los flavonoides conforman la mayor parte de los polifenoles de cacao. Se ha reportado que del total de polifenoles más de la mitad corresponden a proantocianidinas. Estas son principalmente procianidinas formadas con dos o hasta más de diez unidades monoméricas. Algunos ejemplos relevantes en cacao son la B2, C1 y A2 (D'Souza et al., 2017; Dorenkott et al., 2014; Hii et al., 2010; Wollgast & Anklam, 2000) (Figura 3).



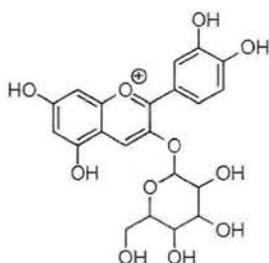
**Figura 3.** Procianidinas destacadas por su concentración en cacao.

Alrededor de una tercera parte de los polifenoles son catequinas, en particular de (-)-epicatequina (Figura 4). Las antocianinas son el tercer grupo más abundante de los flavonoides del cacao, con valores promedio entre 1,7 y 4 % de los polifenoles totales; algunas leucoantocianinas y cianidinas son representativas de esta fracción (Biehl & Ziegleder, 2003a) (Figura 5).

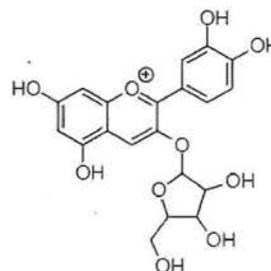


**Figura 4.** (-)-epicatequina, catequina más abundante en cacao.

**Cianidina-3-galactósido**



**Cianidina-3-arabinósido**



**Figura 5.** Ejemplos de antocianinas presentes en cacao.

Existe una amplia gama de otros compuestos presentes en semillas de cacao, por citar, ácido (iso) cítrico, hexósido de apigenina, ácido cafeoilquínico, clovamida, kaempferol-3-O-rutinósido, aspartato de *p*-cumaroil, luteolina, hexósido de naringenina y otros (D'Souza et al., 2017).

Adicionalmente, se ha reportado la presencia de algunos compuestos anti nutricionales como HCN, oxalatos y teobromina, los cuales se ven reducidos durante la fermentación (Aremu et al., 1995).

### 9.1.5. Procesamiento de cacao

Una vez cosechado, el cacao es sometido a etapas poscosecha de fermentación y deshidratación; obteniéndose así los granos que son la principal materia prima para elaborar chocolates, manteca y cacao en polvo.

Estas etapas poscosecha, o de primera transformación, se realizan mayoritariamente en los países productores, mientras que la segunda transformación, o transformación

industrial de los granos fermentados y secos, suelen ocurrir principalmente en países no productores.

En cada etapa se generan alteraciones en la composición, y por tanto en la calidad. Por esta razón, es importante comprender el impacto de los parámetros en cada etapa, y velar por seleccionar aquellos que permitan potenciar la calidad a lo largo de la cadena de valor.

#### **9.1.5.1. Primera transformación**

##### **Cosecha**

El cacao debe cosecharse en el momento adecuado y uno de los indicadores más usados por los productores es el color. En términos generales, las mazorcas inmaduras que presentan coloraciones verdes suelen tornarse amarillas cuando alcanzan su madurez, mientras que las que son de color morado o rojizas en estadios tempranos, alcanzan tonos amarillos o anaranjados al madurar (de Souza et al., 2018; Nair & Nair, 2010).

La cosecha de mazorcas inmaduras puede incidir negativamente en la calidad del cacao, debido a que presenta un menor contenido de azúcares y requieren mayores tiempos de secado, lo cual aumenta la probabilidad de contaminación con hongos de deterioro y patógenos. Contrariamente, las mazorcas sobre maduras pueden presentar una pulpa con baja humedad, y sus semillas inician los procesos de germinación, en donde ocurren daños estructurales en la semilla, facilitándose la contaminación del grano por acción de mohos, insectos y otros contaminantes (CAOBISCO, ECA, & FCC, 2015).

Las mazorcas, además, no deben presentar síntomas de enfermedades o daños visibles, como, por ejemplo, aquellos producto de la acción de plagas. Al cortar el fruto del árbol, se deben emplear herramientas limpias y en buen estado, con la finalidad de no dañar los cojines florales y evitar la transmisión de enfermedades entre plantas.

Una vez cosechadas las mazorcas se recomienda no almacenarlas por más de una semana, ya que pueden empezar a desarrollarse hongos patógenos; aunque es necesario establecer los días según variedad de cacao, ya que se conoce que el pre almacenamiento influye en la calidad (CAOBISCO et al., 2015).

Las mazorcas deben abrirse para continuar con la etapa de fermentación de los granos. La apertura debe realizarse de forma cuidadosa con la finalidad de no quebrar o dañar los granos, lo que provocaría un incremento en su susceptibilidad a ser contaminados.

El índice de fruto o de mazorca, permite conocer cuántas mazorcas son necesarias para obtener 1 kg de cacao fermentado y seco. Comúnmente, éstos índices rondan los 15 a 31 frutos (de Souza et al., 2018; Phillips-Mora et al., 2012).

## **Fermentación**

Actualmente, la fermentación de cacao es el método tradicional que permite la producción de precursores de aromas -ej. amino ácidos libres, péptidos y azúcares reductores-, el mejoramiento de sabor y el desarrollo de color, característicos del cacao (Aprotosoie et al., 2016; Kongor et al., 2016). De forma alternativa, se ha planteado que la presencia de microorganismos fermentativos no es esencial, pero sí sus metabolitos, como lo son el alcohol, ácido láctico, ácido acético y algunas enzimas, acompañados de perfiles específicos de temperatura (Kadow, Niemenak, Rohn, & Lieberei, 2015).

Por definición, una fermentación consiste en un conjunto de procesos bioquímicos que permite obtener energía a partir de azúcares y otros compuestos orgánicos y que mayoritariamente se realiza en ausencia de oxígeno (Tortora, Funke, & Case, 2010).

La fermentación que experimenta el cacao es espontánea, y se debe al desarrollo de microorganismos que son favorecidos por la composición de la pulpa y por sus variaciones en el tiempo (Cuadro I). Existe una gran variedad de microorganismos que pueden estar presentes durante la fermentación, algunos de estos son bacterias lácticas, ácido acéticas, levaduras y mohos (Ozturk & Young, 2017).

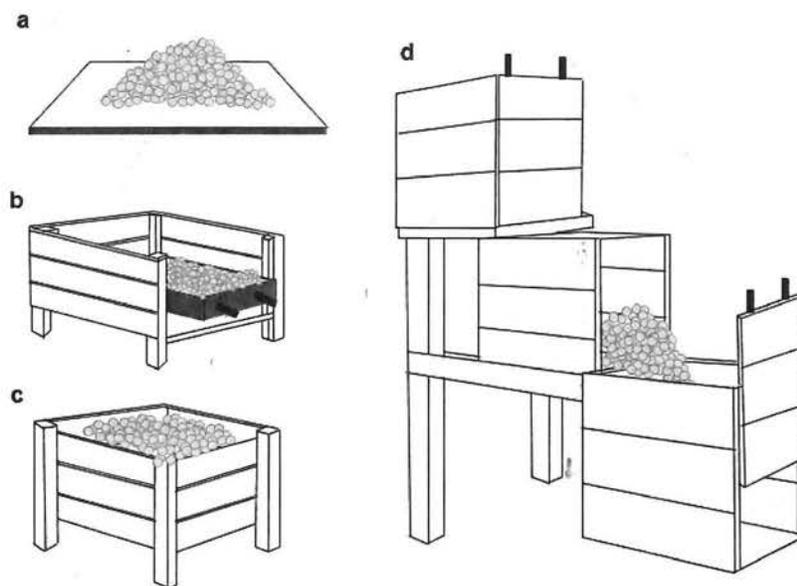
Las principales fuentes de inoculación de los granos estériles son las mazorcas de cacao, los instrumentos y recipientes utilizados durante la cosecha, apertura y almacenamiento, las manos de los agricultores, insectos, hojas de plátano o banano con la que se cubre la masa de fermentación, y otros (D. S. Nielsen, Craffack, Jespersen, & Jakobsen, 2013; Ozturk & Young, 2017; Schwan & Wheals, 2004).

Existen diferentes técnicas para fermentar, que se diferencian mayoritariamente por la configuración del sistema donde se coloca el cacao en baba. En términos generales, la

configuración de un fermentador debe poseer drenajes para la salida de lixiviados, aislar el producto del ambiente, conservar la temperatura, facilitar el volteo de los granos y considerar aspectos de salud ocupacional e inocuidad.

Algunos ejemplos son (Figura 6) (D. S. Nielsen et al., 2013):

- *Apilado*: el cacao se coloca apilado sobre hojas de plátano y banano y se cubren con estas mismas hojas.
- *En bandejas*: se colocan bandejas de madera con aproximadamente 10 cm de alto cada una, unas sobre otras. Se colocan hojas de plátano o banano en la bandeja superior.
- *En cajas o cajones*: se utilizan cajas de madera donde se contiene el cacao en baba, y se cubre la superficie con hojas de plátano o banano.
- *En escalera*: son cajas de madera acomodadas verticalmente escalonadas. Se cubre la superficie de cada cajón con hojas de plátano o banano.



**Figura 6.** Tipos de fermentadores para cacao (a: apilado, b: en bandejas, c: en cajas o cajones y d: en escalera) (elaboración propia).

Esta etapa puede durar en promedio 3 a 5 días, lo cual depende de diversos aspectos desde la genética del cacao hasta las condiciones ambientales en las cuales se desarrolla la fermentación (CAOBISCO et al., 2015), y el tiempo en el que un consorcio complejo de

microorganismos evoluciona y genera profundos cambios físicos, químicos y sensoriales en los granos (Figura 7).

La diversidad de microorganismos es muy amplia (Figura A.2, en anexos); no obstante, se destaca de manera general la acción de algunos como *Hanseniaspora opuntiae*, *Saccharomyces cerevisiae*, *Lactobacillus fermentum* y *Acetobacter pasteurianus* (Pereira, Soccol, & Soccol, 2016).

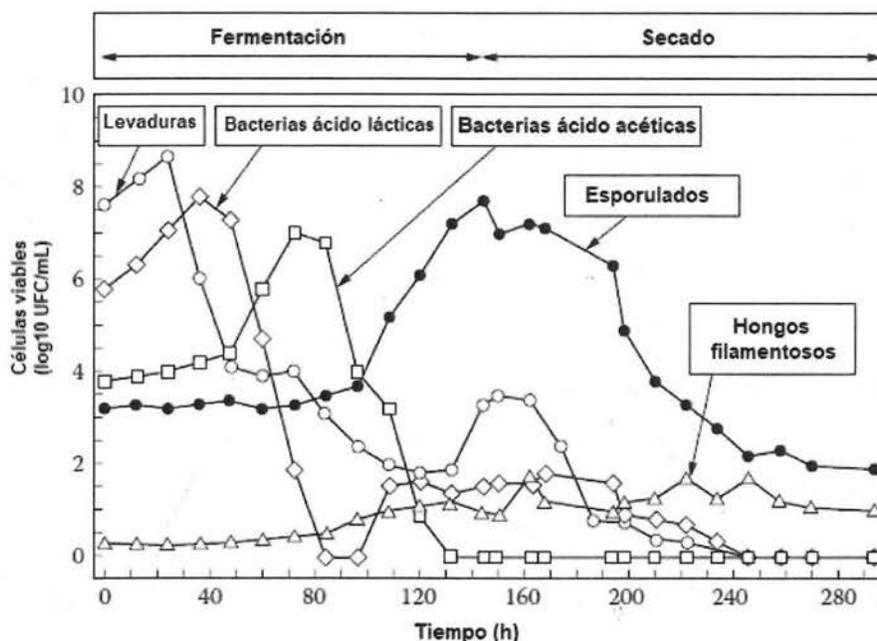


Figura 7. Evolución de los microorganismos durante una fermentación típica de cacao (E. Afoakwa, 2014; Schwan & Wheals, 2004).

La fermentación de cacao inicia por lo general con el desarrollo de levaduras facultativas, y que alcanzan sus niveles máximos -entre  $10^7$  y  $10^8$  UFC/g- durante las primeras 36 h. Estos son los principales responsables de metabolizar los azúcares como sacarosa, glucosa y fructosa para producir etanol y  $\text{CO}_2$ , conversión de naturaleza exotérmica que eleva la temperatura hasta cerca de  $35^\circ\text{C}$ . Además, estos hongos producen enzimas pectinolíticas que degradan la pulpa y favorecen la difusión de oxígeno en la masa fermentativa, eliminando la anaerobiosis. Se ha reportado que las levaduras también producen compuestos como alcoholes de alto peso molecular, ésteres, cetonas, aldehídos y ácidos orgánicos tales como acético, oxálico, fosfórico y succínico, que pueden favorecer el perfil sensorial (Ardhana & Fleet, 2003; D. S. Nielsen et al., 2013; Ozturk & Young, 2017; Pereira et al., 2016; Schwan & Wheals, 2004).

Los mohos usualmente se presentan en bajas cantidades, de  $10^2$ - $10^3$  UFC/g, y se desarrollan principalmente en superficies aireadas. Estos se caracterizan por tener una importante actividad enzimática que se refleja en la producción de poligalacturonasas, amilasas y proteasas; lo cual puede favorecer la hidrólisis de la pulpa y parte de la testa. Algunos reportes señalan picos máximos de mohos entre las 24-36 h, mientras otros apuntan a que aumentan conforme se acerca el final de la fermentación (Ardhana & Fleet, 2003; Schwan & Wheals, 2004).

El ácido cítrico es requerido en el metabolismo de las levaduras, por lo que su desarrollo va acompañado de un incremento en el pH. El contenido de etanol puede llegar a alcanzar niveles entre 5 y 6%, parte de este difunde a través de la testa y hacia el endospermo. La difusión de oxígeno, la reducción de ácido cítrico y la concentración de etanol son factores que limitan el posterior desarrollo de las levaduras y favorecen el crecimiento de las bacterias lácticas microaerofílicas (Aprotosoai et al., 2016; Ardhana & Fleet, 2003; Pereira et al., 2016).

Se ha reportado que las bacterias ácido lácticas pueden alcanzar niveles entre  $10^8$  y  $10^9$  UFC/g, con máximos entre las 36-48 h. Estas bacterias pueden consumir la glucosa y ácidos orgánicos para la producción de diversos metabolitos como ácido láctico (Nigam & Singh, 2014; Pereira et al., 2016; Schwan & Wheals, 2004).

Las metodologías de fermentación involucran volteos de la masa de fermentación, siendo el primero normalmente luego de 48 h; no obstante, esto varía según el cacao y las condiciones ambientales. Esta etapa implica el mezclado del cacao de forma manual o con ayuda de algún instrumento. El objetivo es aumentar la uniformidad de la distribución del aire y de la temperatura; ya que es común, en particular para el método de la caja o cajón, que la superficie y el fondo de la caja posean condiciones diferentes (E. Afoakwa, 2014; Biehl & Ziegler, 2003b). Los volteos, al incorporar  $O_2$  en la masa, favorecen el desarrollo de bacterias aerobias, de tipo ácido acéticas.

Las bacterias aerobias ácido acéticas poseen el rol de oxidar el etanol hacia ácido acético, y algunas especies logran oxidar este último a  $CO_2$  y agua. La naturaleza exotérmica de la oxidación del etanol provoca que la temperatura se incremente a valores superiores a  $45\text{ }^\circ\text{C}$ , lo cual perjudica a otros microorganismos sensibles a la temperatura. También se ha reportado que estas bacterias producen metabolitos de importancia

sensorial, como aldehídos, cetonas y una variedad de compuestos volátiles (Ardhana & Fleet, 2003; D. S. Nielsen et al., 2013; Schwan & Wheals, 2004).

El ácido acético penetra en las semillas, lo cual es evidenciado por las variaciones de pH. Por ejemplo, un estudio reportó la reducción del pH de las semillas de cacao desde 6,3-6,5 al inicio de la fermentación hasta 5,0-5,1 en su término. Incluso, se ha sugerido que las semillas que poseen un pH entre 4,75 y 5,19 han sido fermentadas apropiadamente, mientras que si se encuentran en rangos entre 5,5 y 5,8 experimentaron una fermentación incompleta (E. Afoakwa, 2014; Ardhana & Fleet, 2003). Otros autores señalan que fermentaciones que conlleven a obtener un pH en los granos entre 4,8 y 5,2 permiten alcanzar el mayor potencial de aroma de cacao (Jürgen Voigt, Textoris-Taube, & Wöstemeyer, 2018).

Las bacterias esporuladas aerobias, aunque intervienen en menor medida desde los inicios de la fermentación, predominan al acercarse al término de la fermentación y pueden llegar a alcanzar niveles entre  $10^7$  y  $10^8$  UFC/g entre las 8 y 72 h. Estas bacterias son enzimáticamente activas y tienen amilasas, proteasas, pectinasas y lipasas; y producen una diversidad de compuestos como pirazinas, 2,3-butanediol, ácidos acético y láctico, ácidos grasos de cadena corta y otros (Ardhana & Fleet, 2003; Ozturk & Young, 2017; Schwan & Wheals, 2004).

Los mohos son otros microorganismos que pueden desarrollarse al final de la fermentación o bien en condiciones de sobrefermentación, así como también durante el secado. Estos se ven favorecidos por bajas temperaturas y atmósferas aerobias. Pueden provocar la aparición de ácidos grasos de cadena corta y notas conocidas como "sabor a moho", que son negativas. Además, algunas cepas son patógenas producen micotoxinas termoestables, como aflatoxinas y ocratoxina A (Copetti, Iamanaka, Nester, Efraim, & Taniwaki, 2013; D. S. Nielsen et al., 2013).

Durante la fermentación, el embrión de las semillas de cacao muere por factores como el incremento en la temperatura y la presencia de metabolitos como ácido acético, láctico y etanol en la semilla, provocando cambios estructurales y de composición (Ozturk & Young, 2017; Schwan & Wheals, 2004). Además, se ha observado que el metabolismo de las levaduras y las bacterias lácticas y ácido acéticas puede inhibir el crecimiento de hongos productores de ocratoxina A (CAOBISCO et al., 2015; Pereira et al., 2016).

La difusión de diversos metabolitos en los granos y la muerte del embrión provoca una serie de reacciones enzimáticas y no enzimáticas que desembocan en la producción de precursores de aromas y sabores del cacao, como aminoácidos libres, péptidos, azúcares reductores, taninos, y compuestos volátiles como alcoholes, ácidos orgánicos, ésteres y aldehídos, entre otros (Kongor et al., 2016).

En particular, la acción enzimática inducida por la reducción del pH tiene un papel preponderante en el desarrollo de precursores de aromas y sabores (Janek, Niewienda, Wöstemeyer, & Voigt, 2016). Además, la difusión de metabolitos en la semilla propicia la ruptura de estructuras intracelulares, favoreciendo el contacto enzima-sustrato. Esta actividad es determinada por la genética, el momento de la cosecha, el tiempo de almacenamiento previo a la fermentación y las condiciones de procesamiento (Hansen, Mañez, Burri, & Bousbaine, 2000). En el Cuadro III se resumen algunas de los rasgos asociados a diferentes enzimas de interés.

La polimerización de quinonas, productos de la oxidación de catequinas, se ha sugerido que es de las principales reacciones no enzimáticas que ocurren en la semilla durante la fermentación. Adicionalmente, durante el proceso se dan pérdidas de diferentes compuestos debido a la lixiviación de la pulpa (Biehl & Ziegler, 2003a).

Una fermentación incompleta puede hacer que los granos sean color pizarroso o morados o muy amargos y astringentes. Una sobrefermentación puede provocar un incremento del pH, coloraciones oscuras, así como el desarrollo indeseado de mohos y bacterias que producen ácidos grasos de cadena corta, y sabores a moho (E. Afoakwa, 2014; CAO-BISCO et al., 2015; Schwan & Wheals, 2004).

## **Secado**

El secado se realiza con la finalidad de aumentar la estabilidad de los granos durante el almacenamiento, transporte y comercialización. No obstante, esta etapa también tiene implicaciones en la calidad del cacao.

El empleo de técnicas solares y artificiales, de forma independiente o combinada es común. Es usual alcanzar temperaturas entre los 40-60 °C, y la humedad suele disminuir desde valores cercanos al 55 % hasta menores al 7,5%. Porcentajes de humedad mayores propician el desarrollo de hongos de deterioro y patógenos, mientras que un

sobre secado hace que los granos sean frágiles y propensos a reacciones indeseadas de lipólisis (Biehl & Ziegleder, 2003b; CAOBISSCO et al., 2015; Nair & Nair, 2010).

Dada la incorporación de aire, durante el secado se prolongan las reacciones oxidativas iniciadas durante la fermentación a la vez que se reduce el contenido de ácidos volátiles, lo que favorece el perfil sensorial (Biehl & Ziegleder, 2003a; Schwan & Wheals, 2004). El contenido total de polifenoles suele reducirse ya que estos se condensan hasta formar taninos de color café (Hii et al., 2010; Hii, Law, & Suzannah, 2012).

**Cuadro III.** Enzimas más relevantes en la calidad y que intervienen en la fermentación de cacao.

Enzima	pH óptimo	Función
Endoproteasa aspártica	≤ 3,5	Hidroliza principalmente enlaces peptídicos de las globulinas tipo vicilina, en distintos rangos de pH y en sitios con residuos de aminoácidos básicos como la lisina para producir péptidos hidrofóbicos. También tiene actividad proteolítica frente a las albúminas en condiciones ácidas. Aunque es favorecida en condiciones ácidas también muestra actividad a pH cercanos a 5. Es estable durante la fermentación.
Carboxipeptidasa	5,2-6,0	Exopeptidasa que produce oligopéptidos hidrofílicos y aminoácidos hidrofóbicos libres a partir de los oligopéptidos hidrofóbicos derivados de la proteólisis de las globulinas. Su actividad es limitada a pH por debajo de pH 5. Es parcialmente estable durante la fermentación.
Invertasa	4,5	Hidroliza la sacarosa a glucosa y fructosa. Se inactiva con los procesos de fermentación.
Aminopeptidasa	6,8-7,0	Exopeptidasa que provoca la escisión en terminales amino. Se inactiva rápidamente con los procesos de fermentación.
Polifenoloxidasa	5,5-7,0	Convierte polifenoles en quinonas, que luego polimerizan. Es afectada por la temperatura y baja humedad en las etapas de fermentación y secado.
Glicosidasas	3,5-4,5	Provoca la hidrólisis de antocianinas y otros flavonoides generando variaciones en el color. Se ha reportado que son estables durante la fermentación.

Fuentes: (Emmanuel Ohene Afoakwa, 2014; Biehl & Ziegler, 2003a; Bytof, Biehl, Heinrichs, & Voigt, 1995; Hansen, Olmo, & Burri, 1998; Janek et al., 2016; John et al., 2016; SOUSA et al., 2016; J. Voigt et al., 1994; Jürgen Voigt et al., 2018).

La producción de azúcares reductores a partir de la hidrólisis de la sacarosa, sumado al incremento de oligopéptidos y aminoácidos durante la fermentación, promueven las reacciones de Maillard durante el secado, importantes para el desarrollo del perfil sensorial (Biehl & Ziegler, 2003a; Schwan & Wheals, 2004).

Un estudio analizó 81 muestras comerciales de cacao de Costa de Marfil y de diferentes regiones de Brasil, muestreadas a lo largo de un año (Hashimoto et al., 2018), obteniéndose los valores de composición mostrados en el Cuadro IV.

**Cuadro IV.** Composición química de semillas de cacao fermentadas y secas.

<b>Propiedad</b>	<b>Concentración</b>
Humedad (g/100 g)	6,0-10,3
Grasa (g/100 g)	41,0-48,7
Proteínas (g/100 g)	13,8-16,0
Polifenoles totales (g/100 g)	3,0-7,9
Acidez titulable (g/100 g)	0,7-2,1
pH	6,0-10,3
Cafeína (g/100 g)	0,04-0,26
Teobromina (g/100 g)	0,8-1,5

Los granos fermentados y secos son la principal forma de comercializar el cacao en los países productores. Estos deben cumplir con algunos requerimientos, como por ejemplo, tener una humedad menor al 7,5 %, ser uniformes y no poseer contaminaciones, además de cumplir con estándares especificados por diferentes normativas como, por ejemplo, porcentaje de granos mohosos o pizarrosos (CAOBISCO et al., 2015).

#### **10.1.6.2. Segunda transformación**

La primera transformación industrial que reciben los granos fermentados y secos es el tostado; que ocurre en temperaturas entre los 90-170 °C, con tiempos de operación variables. Esta es la etapa en la que se desarrollan los aromas, sabores y colores, a partir de los precursores formados en las etapas postcosecha del cacao. Además, la humedad

se reduce a valores cercanos al 2 %, disminuye la carga total de microorganismos y se facilita la separación posterior de la testa (E. Afoakwa, 2014; T. J. Gutiérrez, 2017).

El tostado favorece la formación de pirazinas, pirroles, pironas, lactonas, furanonas y otros (Biehl & Ziegler, 2003a), así como la reducción en el contenido de polifenoles (Ioannone et al., 2015).

Las semillas tostadas de cacao posteriormente pueden seguir diversas rutas de valorización, para la obtención de derivados como licor o pasta de cacao, cacao en polvo, manteca de cacao, chocolate y otros (Biehl & Ziegler, 2003b; Nair & Nair, 2010).

## 9.2. Metabolómica

Un metabolito se refiere a aquel compuesto obtenido del metabolismo biológico. Los metabolitos primarios son aquellos que son necesarios para que un organismo se desarrolle normalmente; ejemplos de ellos son los aminoácidos, lípidos, y carbohidratos. Por el otro lado, un metabolito secundario es aquel que no es esencial para el desarrollo, pero tiene funciones específicas como, por ejemplo, defensa o señalización, que favorece la supervivencia en determinado ambiente; por ejemplo, polifenoles, alcaloides, terpenos y otros (Wolfender, Marti, Thomas, & Bertrand, 2015).

La metabolómica se centra en el estudio de metabolitos con un tamaño molecular menor a 1000 Da y puede seguir diferentes estrategias para su análisis (O Fiehn, 2001; Scalbert et al., 2009; Wolfender et al., 2015).

- *Huella metabólica*: permite la comparación de patrones de metabolitos entre muestras, de manera no dirigida, ya que usualmente no realiza las etapas de identificación y cuantificación.
- *Análisis dirigido*: se enfoca en el estudio del efecto de una perturbación en un sistema biológico, mediante la evaluación de uno o pocos metabolitos específicos en una ruta metabólica particular.
- *Perfilamiento*: se centra en el análisis de una variedad de clases de compuestos o una variedad de metabolitos asociados a determinada ruta metabólica.
- *Metabolómica*: busca identificar y cuantificar los metabolitos sintetizados en un sistema biológico -i.e. un organismo, un órgano, tejido, célula o compartimentos

celulares-, de manera no selectiva, universal y con un abordaje analítico comprensivo; o, en otras palabras, analizar su metaboloma. Una subdivisión es la metabonomía, que se refiere a la evaluación de los metabolitos endógenos.

A diferencia del resto de las *ómicas*, una de las ventajas de la metabolómica es que permite analizar la última expresión de los genes, los metabolitos, lo que permitiría comprender los fenotipos, es decir las propiedades observables de los organismos -ej. color y sabor- y su estado fisiológico (Allwood & Goodacre, 2010; O Fiehn, 2001; Oliver Fiehn, Kloska, & Altmann, 2001; Markley et al., 2017).

### 10.8.1. Generalidades metodológicas del análisis metabolómico

El flujo de proceso de metabolómica suele iniciar con el planteamiento de una pregunta biológica, que pretende ser respondida mediante la construcción de un diseño experimental apropiado, un buen muestreo, el tratamiento óptimo de muestras, el uso de herramientas analíticas apropiadas, y un correcto procesamiento de datos (Smilde, Hendriks, Westerhuis, & Hoefsloot, 2013; Zhou, Xiao, Tuli, & Resson, 2012) (Figura 8).

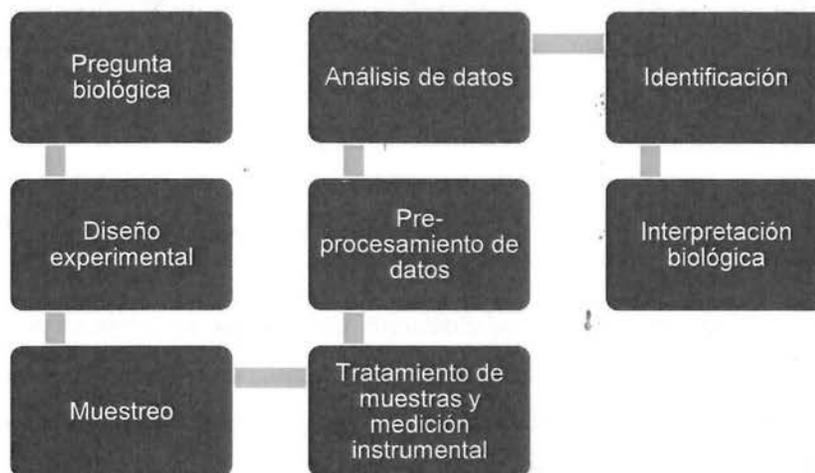


Figura 8. Flujo de proceso general empleado en estudios metabolómicos.

#### ***Pregunta biológica***

Las preguntas biológicas son de muchos tipos, y pueden ir con miras a evaluar el efecto de tratamientos sobre un sistema, comprender la relación entre metabolitos para proponer

rutas metabólicas o plantear biomarcadores de procesos, entre otras (Smilde et al., 2013). En el último caso, un biomarcador puede definirse como “una característica medida objetivamente y evaluada como un indicador de un proceso biológico normal, patogénico, o como respuesta farmacológica a una intervención terapéutica” (Biomarkers Definitions Working Group, 2001).

### ***Diseño experimental***

Al construir un diseño experimental en metabolómica, es importante definir en detalle el tipo de estudio a realizar, por ejemplo, si es observacional o posee alguna intervención experimental (Friedman, Furberg, DeMets, Reboussin, & Granger, 2015; Grimes & Schulz, 2002; Machin, Campbell, & Walters, 2001). La elección debe hacerse con la finalidad de responder la pregunta biológica original. Además, es necesario establecer los factores o parámetros del sistema que se van a controlar o modificar en el estudio (Zhou et al., 2012).

Por otro lado, es necesario fundamentar la estimación del número de réplicas biológicas, ya que de esto depende en gran medida el alcance de una investigación. En ocasiones se debe recurrir a pruebas piloto para conocer la variabilidad del sistema que se desea estudiar, y así evaluar el número de muestras a elegir (Zhou et al., 2012).

### ***Muestreo y almacenamiento de muestras***

Al trabajar con sistemas vivos, se hace imperante la aplicación de métodos que detengan las reacciones metabólicas una vez obtenidas las muestras bajo protocolos estandarizados. Se pueden utilizar técnicas de reducción de temperatura o congelación en nitrógeno líquido y posterior almacenamiento a  $-80\text{ }^{\circ}\text{C}$ , que además garantizan una mayor estabilidad durante el almacenamiento (Dunn & Ellis, 2005).

El tipo de material empleado para almacenar las muestras debe ser tal que no posea compuestos que migren y contaminen la muestra; lo mismo aplica para los extractos obtenidos de las muestras (Scalbert et al., 2009).

### ***Tratamiento de muestras y medición instrumental***

Uno de los principales retos de la metabolómica es el lidiar con un amplio rango de metabolitos a la vez, los cuales difieren en sus propiedades químicas, físicas y concentración (desde pmoles hasta mmoles); es por ello que no existe una técnica de extracción y medición universal y es necesario recurrir a la aplicación de métodos complementarios (Dunn & Ellis, 2005; Scalbert et al., 2009).

Desde un abordaje metabolómico, se prefiere en primera instancia evitar la aplicación de métodos que puedan alterar la composición de las muestras, ya que pueden eliminarse o reducirse algunos compuestos, disminuyendo el alcance.

Usualmente no es posible medir muestras de forma directa, por lo que suelen establecerse protocolos que incluyan etapas de filtración, centrifugación, desproteinización, desgrasado, dilución, concentración y extracción (Hyötyläinen & Wiedmer, 2013). La definición de un protocolo debe también considerar el tipo de plataforma analítica a utilizar, su alcance y sus limitaciones.

La infusión directa en un espectrómetro de masas (IDEM) y las técnicas espectroscópicas como la resonancia magnética nuclear (RMN) y la espectroscopía infrarroja por transformada de Fourier (FTIR), son mayoritariamente empleadas para investigaciones de huella metabólica y metabonómica. Por otro lado, la cromatografía de gases (CG) y la cromatografía líquida (CL) acopadas a espectrometría de masas (EM) son preferidas cuando se hacen estudios de perfilamiento, análisis dirigido y metabolómica (Dunn & Ellis, 2005).

La RMN, como técnica espectroscópica, estudia la interacción de los núcleos magnéticamente activos con radiofrecuencias, en el ámbito de megahercios (S. S. Nielsen, 2010). Por su naturaleza, es una técnica rápida y no destructiva que permite cuantificar y obtener estructuras. Comúnmente se requieren concentraciones que van desde los micromoles hasta los milimoles, lo que se constituye en su principal desventaja (Keski-Rahkonen, Rothwell, & Scalbert, 2017; Larive, Barding, & Dinges, 2015; Wolfender, Glauser, Boccard, & Rudaz, 2009).

La EM por su parte, tiene como principio la formación, separación y detección de iones según su relación masa/carga ( $m/z$ ) (Dunn & Ellis, 2005). Es mucho más sensible que la

RMN, ventaja que destaca principalmente cuando los metabolitos poco abundantes son de interés. No obstante, la cuantificación se limita a la presencia de estándares, los cuales no siempre son accesibles. Es por esto que al emplear esta técnica es común cuantificar en términos relativos (Keski-Rahkonen et al., 2017).

Otras técnicas, como la electroforesis capilar (EC) acoplada a espectrometría de masas, es aplicable principalmente para el estudio de metabolitos polares o cargados, como por ejemplo péptidos, proteínas y nucleótidos. A pesar de su alta resolución y bajo costo, se pueden medir volúmenes reducidos de muestra, lo que reduce la sensibilidad, excepto cuando se acopla con EM (Hyötyläinen & Wiedmer, 2013; Wolfender et al., 2015).

Entre los criterios más relevantes a tomar en cuenta para elegir el tipo de plataforma instrumental y los parámetros de operación se encuentran el tipo de aplicaciones, la selectividad, el rango dinámico, el límite de detección, la exactitud, la precisión y los costos (Scalbert et al., 2009). El Cuadro V resume algunas generalidades asociadas a las técnicas más usadas en análisis metabólico; no obstante, dependiendo de la combinación de distintas técnicas y las especificaciones de configuración, así será el potencial para ser empleado en este tipo de investigaciones.

### ***Preprocesamiento de datos***

En esta etapa es necesario depurar la información obtenida según el tipo de instrumento o instrumentos empleados durante la medición, con la finalidad de facilitar el análisis estadístico y la posterior interpretación.

Algunas acciones comunes implican la reducción del ruido, la corrección de líneas base, la eliminación de información redundante, o segmentos que no son de interés, la detección de señales, el alineamiento de señales entre muestras y la normalización (Castillo, Gopalacharyulu, Yetukuri, & Orešič, 2011). El producto final del preprocesamiento puede ser analizado vía algoritmos matemáticos y estadísticos que faciliten su interpretación.

### ***Análisis de datos***

Un primer paso para llevar a cabo el análisis de los datos, es realizar aquellas transformaciones que permitan facilitar la interpretación y aplicar correctamente las distintas técnicas. Algunos ejemplos de transformaciones básicas son el centrado de datos y el escalamiento (Goodacre et al., 2007). Existe una amplia gama de algoritmos de transformación y, en la mayoría de los casos, la elección depende del estudio particular de cada conjunto de datos.

**Cuadro V.** Comparación de algunas generalidades asociadas a las principales técnicas empleadas en el análisis de metabolitos.

Técnica	Aplicaciones	Dificultad en la elaboración de muestras	Tamaño de muestras en disolución	Sensibilidad	Selectividad	Reproducibilidad	Identificación y cuantificación
CL-MS	Lípidos, ácidos orgánicos, alcaloides, aminoácidos, compuestos fenólicos, esteroides, vitaminas.	Moderada	5-100 $\mu$ L	pmol-mmol	Alta	Moderada-alta	Limitada, se restringe al uso de estándares
CG-MS	Volátiles, aceites esenciales, ácidos grasos, esteroides, aminoácidos, azúcares, alcoholes y metabolitos no volátiles o termolábiles derivatizados.	Alta	5-100 $\mu$ L	pmol-nmol	Alta	Alta; tiempos de retención muy repetibles	Buena
RMN	Aminoácidos, ácidos grasos, colesterol, citrato, ácido hipúrico, entre otros. Mezclas líquidas, semisólidas o sólidas.	Baja	0,07-0,5 mL	$\mu$ mol-mmol	Baja	Alta	Buena

**Fuentes:** (Hyötyläinen & Wiedmer, 2013; Keski-Rahkonen et al., 2017; Markley et al., 2017; Priego-Capote et al., 2012; Putri, Yamamoto, Tsugawa, & Fukusaki, 2013; Wolfender et al., 2009, 2015; Zhang, Sun, Wang, Han, & Wang, 2012).

Posterior a la transformación, se pueden aplicar distintas técnicas estadísticas para la interpretación y visualización de la información. Por ejemplo, en investigación metabolómica, un abordaje univariado permite el estudio de un metabolito en específico, mientras que el uso de estadística multivariada analiza todas las variables de forma simultánea. Ambos abordajes pueden emplearse, pero por la naturaleza de los datos metabolómicos, la segunda es la más empleada (Goodacre et al., 2007).

Ejemplos de análisis univariado son la prueba de *t*-Student y su forma generalizada, así como el análisis de varianza (ANDEVA). Mientras que el análisis de componentes principales (PCA, del inglés) es una de las técnicas multivariadas más populares (Smilde et al., 2013).

El análisis estadístico multivariado tiene la característica de reducir la dimensionalidad en bases de datos en donde el número de variables supera el número de muestras. Esto facilita el estudio e interpretación de datos, y puede aplicarse con fines exploratorios o de control, para identificar agrupaciones o identificar biomarcadores entre grupos, o para establecer correlaciones múltiples (Wheelock & Wheelock, 2013).

La estadística multivariada puede ser clasificada en métodos supervisados y no supervisados, cuya diferencia radica en que en el primer caso se conoce *a priori* información sobre la identidad de las muestras (Goodacre et al., 2007; Wheelock & Wheelock, 2013). Algunos cuantos ejemplos de estas técnicas se muestran en el Cuadro VI, pero existe una amplia variedad de otras técnicas.

**Cuadro VI.** Ejemplos de técnicas multivariadas supervisadas y no supervisadas.

Clasificación	Técnica
Supervisadas	Análisis discriminante (LDA, PLS-DA),
	Regresión (MLR, PCR, PLS)
	Redes neuronales artificiales
No supervisadas	Análisis de componentes principales (PCA)
	Análisis de conglomerados
	Redes neuronales de Kohonen

Fuentes: (Goodacre et al., 2007; Wheelock & Wheelock, 2013)

Distintos parámetros asociados a los modelos permiten valorar su calidad. Por ejemplo, el parámetro  $R^2$  permite comprender el porcentaje de los datos que son explicados por el modelo, mientras que el  $Q^2$  se asocia a la capacidad predictiva del modelo. Ambos

parámetros deben mantenerse cercanos, de lo contrario se puede sospechar de un posible sobreajuste de los datos al modelo. Por otro lado, en el caso de los métodos supervisados, el parámetro  $p$  generado del análisis CV-ANOVA se relaciona con la significancia de la separación entre grupos (Wheelock & Wheelock, 2013).

Independientemente de la técnica y como última etapa en el análisis de datos, es fundamental evaluar y validar los modelos construidos, para lo cual existen distintas técnicas. Esta validación puede ser interna, es decir, empleando una parte de los datos originales para efectos de predicción (ej. validación cruzada o permutaciones), o bien, externa, incorporando nuevos datos en el modelo (Goodacre et al., 2007; Smilde et al., 2013).

### ***Interpretación biológica***

La interpretación biológica debe hacerse a la luz de la pregunta biológica originalmente planteada. Esta, a su vez, debe estar acompañada de procesos de identificación de compuestos, que pueden eventualmente ser incorporados o clasificados dentro de una ruta metabólica.

El grupo de trabajo de análisis químico de la Iniciativa de Estándares Metabolómicos, propuso distintos niveles para la identificación de metabolitos, que se detallan en el Cuadro VII.

**Cuadro VII.** Niveles de confianza para la identificación de metabolitos (Sumner et al., 2007).

<b>Nivel</b>	<b>Descripción</b>
I	Identificación basada en la comparación contra estándares.
II	Identificación tentativa de compuestos de acuerdo con la comparación de propiedades físicoquímicas e información espectral; sin estándares.
III	Identificación tentativa de clases de compuestos de acuerdo con la comparación de propiedades físicoquímicas e información espectral de una clase de compuestos.
IV	Compuestos que, a pesar de tener identidad desconocida, pueden ser diferenciados y cuantificados con base en información espectral.

Algunas de las propiedades que ayudan en los procesos de identificación son las  $m/z$ , propuestas de fórmulas moleculares, tiempos de retención y comparación de información espectral contra bases de datos, entre otros.

## **10.8.2. Aplicación de la CL-MS en análisis de metabolitos**

### **10.8.2.1. Cromatografía líquida**

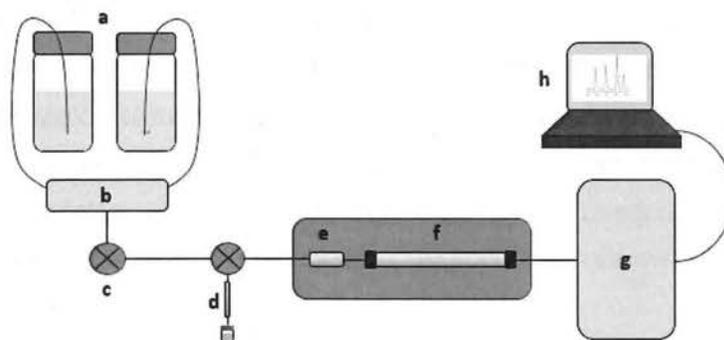
La CL-MS es una de las técnicas más populares en análisis de metabolitos, que combina la selectividad de la cromatografía con la sensibilidad de la espectrometría de masas.

La cromatografía líquida consiste en una técnica de separación de compuestos presentes en una mezcla, que es provocada por las diferencias de afinidades de los compuestos entre dos fases inmiscibles: una fase móvil líquida que eluye a lo largo de una fase estacionaria líquida o sólida. Esto genera variaciones en la velocidad de elución y, por tanto, en los tiempos de retención (Allwood & Goodacre, 2010; S. S. Nielsen, 2010).

La afinidad de un metabolito por una u otra fase depende de diferentes propiedades como, por ejemplo, su estructura, tipos de grupos funcionales, peso molecular,  $pK_a$  y solubilidad, entre otras (Snyder, Kirkland, & Glajch, 1997).

En aplicaciones metabolómicas, la cromatografía líquida de columna es la más empleada. Una columna es un tubo que contiene la fase estacionaria; estas varían en cuanto a longitud, diámetro, tamaño de partícula y composición, y contienen la fase estacionaria.

Este tipo de cromatógrafo consiste en un equipo conformado como mínimo por reservorios de solventes, una bomba, un inyector, una serie de tuberías, una precolumna y una columna, un detector y un sistema computarizado que controla el equipo (Figura 9) (Allwood & Goodacre, 2010).



**Figura 9.** Componentes principales de un cromatógrafo líquido a: reservorios de solventes, b: sistema de bombeo, c: mezcladores, d: inyector de muestra, e: precolumna y f: columna analítica en un horno, g: detector y h: sistema computarizado de control y almacenamiento de información (elaboración propia).

El sistema de bombas permite trasegar los eluentes desde sus reservorios y a través del sistema cromatográfico; permite proveer la presión necesaria para obtener un determinado flujo continuo. De requerirse fases móviles mixtas, es necesario utilizar válvulas de mezclado que permitan combinar eluentes. Estas válvulas pueden encontrarse antes o después del sistema de bombeo (Allwood & Goodacre, 2010; Meyer, 2010; Snyder, Kirkland, & Dolan, 2010).

El inyector permite introducir la muestra en la columna. Existen diversas configuraciones que buscan inyectar de manera exacta y precisa una muestra, de tal forma que no se interrumpa el flujo continuo de la fase móvil (Snyder et al., 2010).

El uso de precolumnas es común, y puede ser antes o después del inyector de muestras; en ambos casos se busca proteger la columna analítica. Si la precolumna se encuentra antes del inyector permite acondicionar el solvente, de forma tal que se sature con componentes propios de la fase estacionaria y evite que esto ocurra en la columna principal. Mientras que, si la columna se usa posterior al inyector, permite retener compuestos o impurezas que son retenidos con fuerza y pueden afectar la separación en la columna analítica. Las precolumnas son de menor longitud que la columna analítica; además poseen una composición similar o igual a ella (Meyer, 2010; S. S. Nielsen, 2010).

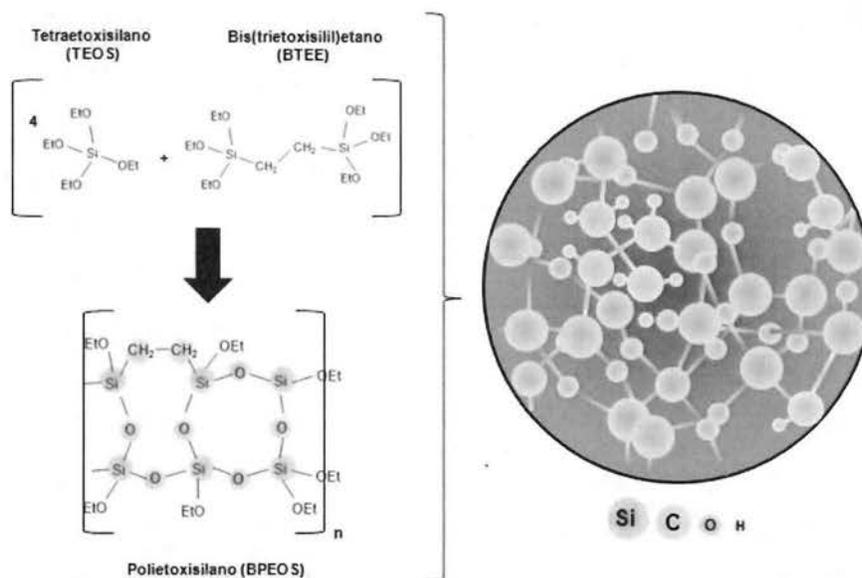
El sistema separativo se ubica dentro de un horno que controla la temperatura. Este parámetro tiene profundas implicaciones en el desempeño de los análisis, al modificar las propiedades físicoquímicas tanto de los eluentes como de los analitos. Por ejemplo, al aumentar la temperatura se modifica la viscosidad del sistema de elución, favoreciendo flujos mayores y, por ende, uso de menores presiones (Meyer, 2010).

Los detectores permiten adquirir información sobre los analitos, al medir propiedades específicas y convertirlas en señales eléctricas. Algunos detectores comunes son de absorción en la región ultravioleta-visible, de fluorescencia, de índice de refracción, de espectrometría de masas, entre otros. Finalmente, un sistema computarizado recibe y almacena las señales eléctricas del detector y las transforma en información comprensible, como lo son los cromatogramas (Allwood & Goodacre, 2010; S. S. Nielsen, 2010).

La cromatografía de adsorción se define como aquella en donde el mecanismo de separación implica la adsorción de los compuestos en una fase estacionaria sólida, a diferencia de la cromatografía de partición, en donde la fase estacionaria es líquida (Allwood & Goodacre, 2010).

Dada la diversidad de alternativas de columnas, la cromatografía de adsorción puede subdividirse en aquella de fase normal (FN), que emplea fases estacionarias polares, en contraposición con la fase reversa (FR) que hace uso de fases estacionarias no polares. Por la capacidad de analizar un amplio rango de metabolitos y por la estabilidad de las fases estacionarias, esta última es la técnica más utilizada en estudios metabolómicos (Allwood & Goodacre, 2010; Keski-Rahkonen et al., 2017; Wolfender et al., 2009).

Las columnas de tipo C8 y C18 son de FR y consisten en partículas de sílica con grupos R de naturaleza no polar, como por ejemplo  $C_8H_{17}$  o  $C_{18}H_{37}$ , correspondientemente. Hoy por hoy, las distintas casas comerciales destinan esfuerzos para desarrollar tecnologías de columnas híbridas con miras a un mejor desempeño cromatográfico y vida útil, como por ejemplo, la columna híbrida de puentes de etileno BEH, por sus siglas en inglés, mostrada en la Figura 10 (Allwood & Goodacre, 2010; Guillaume & Veuthey, 2012; Naushad & Khan, 2014).



**Figura 10.** Partículas de fase estacionaria desarrolladas con tecnología híbrida de puentes de etileno (BEH), desarrollada por Waters® (elaboración propia, basada en (Waters<sup>(R)</sup>, n.d.)).

La cromatografía líquida puede complementarse con la cromatografía de interacción hidrofílica (HILIC, del inglés) para la evaluación de compuestos polares como los carbohidratos, fosfatos, amino ácidos y otros ácidos orgánicos de cadena corta (Keski-Rahkonen et al., 2017).

Es importante optimizar los parámetros en un cromatógrafo, ya que en mezclas complejas de metabolitos, una baja resolución puede provocar la no observación de señales (Castillo et al., 2011; Keski-Rahkonen et al., 2017).

En la actualidad, los cromatógrafos líquidos de ultra alta resolución (CLUR), ofrecen una serie de ventajas con respecto a los tradicionales cromatógrafos líquidos de alta resolución (CLAR). Esto se debe en gran medida al uso de columnas con tamaño promedio de partícula inferiores a 2  $\mu\text{m}$ , lo cual va acompañado del uso de presiones de operación mayores, incrementos en la resolución, la reproducibilidad, la sensibilidad, mejoras en la relación señal-ruido y la reducción en los tiempos de análisis (Churchwell, Twaddle, Meeker, & Doerge, 2005; Guilleme & Veuthey, 2017; Nordström, O'Maille, Qin, & Siuzdak, 2006; Ortega et al., 2010; Putri et al., 2013; Swartz, 2005) (Cuadro VIII).

**Cuadro VIII.** Características generales asociadas a la CLAR y la CLUR.

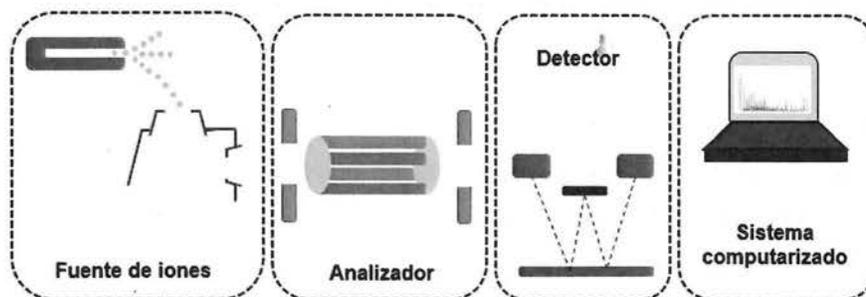
Características	CLAR	CLUR
Presiones de trabajo (psi)	2 000-6 000	15 000
Tamaño de partícula de la FE ( $\mu\text{m}$ )	3-10	$\leq 2$
Flujos (mL/min)	0,4-1,0	0,01-2,0
Volumen inyección ( $\mu\text{L}$ )	10-100	0,1-20
Longitud de columna analítica (cm)	3-25	5-15
Diámetro interno (mm)	3,0-6-5 <sup>2</sup>	2,1-4,6

Fuentes: (Guillarme & Veuthey, 2012; Naushad & Khan, 2014; S. S. Nielsen, 2010; Putri et al., 2013).

Al realizar análisis cromatográficos, es importante el uso de controles de calidad durante las mediciones que permitan estimar la variabilidad de un experimento (Zhou et al., 2012). Algunos controles consisten en medir una mezcla de muestras o *pool*, estándares de referencia y blancos, ya sea al inicio, en intervalos regulares o al final de cada lote de muestras (Dunn et al., 2017).

### 10.8.2.2. Espectrometría de masas

Un espectrómetro de masas es un tipo de detector, que se distingue por ionizar los compuestos de previo a su medición. Este destaca por su alta sensibilidad y rapidez de medición (Allwood & Goodacre, 2010). El instrumento consta de tres partes principales: una fuente de ionización, un analizador, un detector y un dispositivo como una computadora que permite controlar los parámetros de medición, así como almacenar la información (Figura 11).



**Figura 11.** Ejemplo de un espectrómetro de masas de tipo ESI-QToF (elaboración propia).

<sup>2</sup> Las columnas capilares y de microboro poseen diámetros internos menores a 2 mm.

### **Fuente de iones**

En la actualidad, las fuentes de ionización empleadas en metabolómica funcionan a presión atmosférica, por lo cual son llamadas técnicas de ionización suave. El cuadro IX resume algunas generalidades de las fuentes más empleadas.

**Cuadro IX.** Fuentes de iones que operan a presión atmosférica.

<b>Fuente de iones</b>	<b>Principio</b>	<b>Alcance</b>
Ionización electrospray (ESI, del inglés)	Ionización suave que funciona con la aplicación de alto voltaje; permite formación de gotas cargadas.	Compuestos ionizables en solución, principalmente de alto peso molecular y polares.
Ionización química a presión atmosférica (APCI, del inglés)	Ionización química que se provoca al ampliar la temperatura y una descarga corona.	Bajo peso molecular y polaridad baja o media.
Fotoionización a presión atmosférica (APPI)	Un haz de fotones de 10 eV, que puede ser generado por una lámpara de xenón o kriptón, interacciona con una solución líquida nebulizada.	Tamaño molecular y polaridad intermedios.

Fuente: (Keski-Rahkonen et al., 2017; Steinmann & Ganzera, 2010; Wolfender et al., 2009).

La ionización de metabolitos vía ESI es provocada por un diferencial de voltaje entre el capilar metálico donde se encuentra la muestra y el cono de muestreo; que se encuentra entre los 3-5 kV (Wolfender et al., 2009)

Según el modo de ionización -positiva o negativa- se favorecen reacciones de protonación y desprotonación, formación de aductos o reacciones electroquímicas de oxidación o reducción del sistema de elución o de los analitos (Banerjee & Mazumdar, 2012; Kebarle & Verkerk, 2012).

El campo eléctrico conformado provoca una redistribución de estas cargas presentes en la solución, generando una repulsión en la punta del capilar, que se visualiza como un cono llamado cono de Taylor. Finalmente la repulsión, más el uso de corrientes de gases inertes como N<sub>2</sub>, provoca la formación de un spray de gotas cargadas que se dirigen hacia el contraelectrodo o cono de muestreo (Kebarle & Peschke, 2000) (Figura 12).

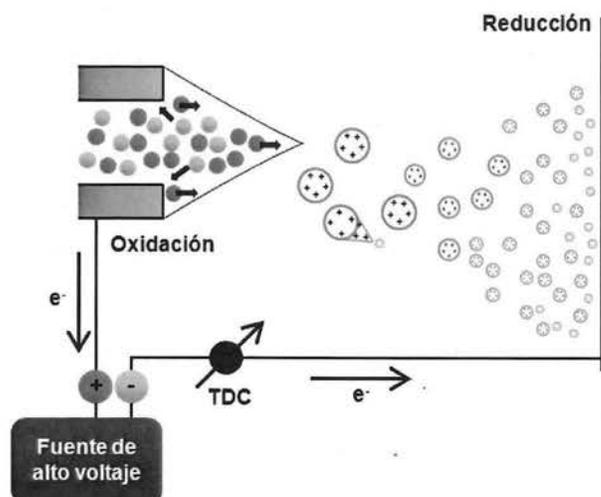


Figura 12. Formación de iones con la técnica de electrospray (ESI) (elaboración propia, adaptada de (Kearle & Tang, 1993)).

La elección del modo de ionización es determinada en gran medida por la  $pK_a$  o  $pK_b$  de los analitos de interés. Según su tendencia a protonarse o desprotonarse, puede ser preferible realizar el análisis en modo positivo o negativo. Por ejemplo, metabolitos básicos como alcaloides y aminas suelen ionizar mejor en modo positivo, mientras que algunos ácidos como compuestos fenólicos y ácidos carboxílicos tienden a ionizar mejor en modo negativo (Kearle & Verkerk, 2012; Wolfender et al., 2009). En el Cuadro A.I., en la sección de anexos, se ejemplifican algunos valores de  $pK_a$  o  $pK_b$  de compuestos orgánicos de distinta naturaleza.

Adicionalmente, para favorecer la ionización en ESI es deseable tener iones preformados, lo cual se obtiene empleando fases móviles con pH superior a la  $pK_a$  de compuestos ácidos, e inferior en el caso de compuestos básicos. Esto puede perjudicar la separación cromatográfica, al afectar la retención en la fase estacionaria cuando se emplean columnas de fase reversa, ya que las formas iónicas son más polares y solubles por lo que su retención es menor. No obstante, independientemente del tipo de analito, por lo general es posible ionizar en un amplio margen de pH (Heftmann, 2004; Kearle & Verkerk, 2012; Niessen, 2006).

Por esta razón es que cuando se emplean estrategias metabolómicas para estudiar un sistema biológico, es recomendado utilizar ambos tipos de voltaje, tanto positivo como negativo de forma complementaria (Wolfender et al., 2009).

Conforme las gotas del electrospray se acercan al cono de muestreo, van reduciendo su tamaño. Esto es producto de la acción del gas inerte -como  $N_2$ - que favorece la evaporación del solvente, por efecto de las fuerzas de repulsión que se generan en la superficie de las gotas. Cuando la tensión superficial de cada gota es superada por la repulsión entre las cargas superficiales, se generan procesos de fisión o explosiones Coulómbicas (Kearle & Peschke, 2000).

La ionización ocurre principalmente por protonación o desprotonación de los analitos, pero en algunos casos se da por formación de aductos con compuestos orgánicos polares, cationes -ej. sodio y potasio- o aniones -ej. amonio y acetato. La protonación y desprotonación múltiple tienen la ventaja de que permiten ionizar muestras de alto peso molecular (Glish & Vachet, 2003; Kearle & Verkerk, 2012).

Con la finalidad de aportar protones y favorecer la ionización, es común la adición de ácidos volátiles, como el ácido fórmico y acético, a los eluentes, en concentraciones de 0,5-0,1 %. Esta estrategia es utilizada tanto en modo positivo y negativo, pero para este último caso también se emplean bases volátiles como hidróxido de amonio o trietilamina. Entre los principales cuidados se debe considerar que altas concentraciones pueden provocar la supresión de señales, como también ocurre al usar ácidos fuertes no volátiles (Kearle & Verkerk, 2012).

Existen distintos modelos que abordan los mecanismos de ionización. El modelo de carga residual se resume en que la reducción de tamaño de las gotas es tal que llega a ser constituida únicamente por el analito con su carga. Por otro lado, el modelo de evaporación de iones propone la expulsión o emisión de los iones del analito solvatado fuera de la gota, producto del campo eléctrico generado dentro de la gota. Finalmente, el modelo Fenn de formación de iones, plantea la unión de analitos en el núcleo neutro de la gota, con cargas en la superficie respectiva. El movimiento Browniano en conjunto con la vibración y rotación internas proveen la energía suficiente para alejar al analito cierta distancia de la superficie de la gota, que en última instancia termina siendo repelido, llevándose consigo otras cargas (Banerjee & Mazumdar, 2012; Kearle & Verkerk, 2012).

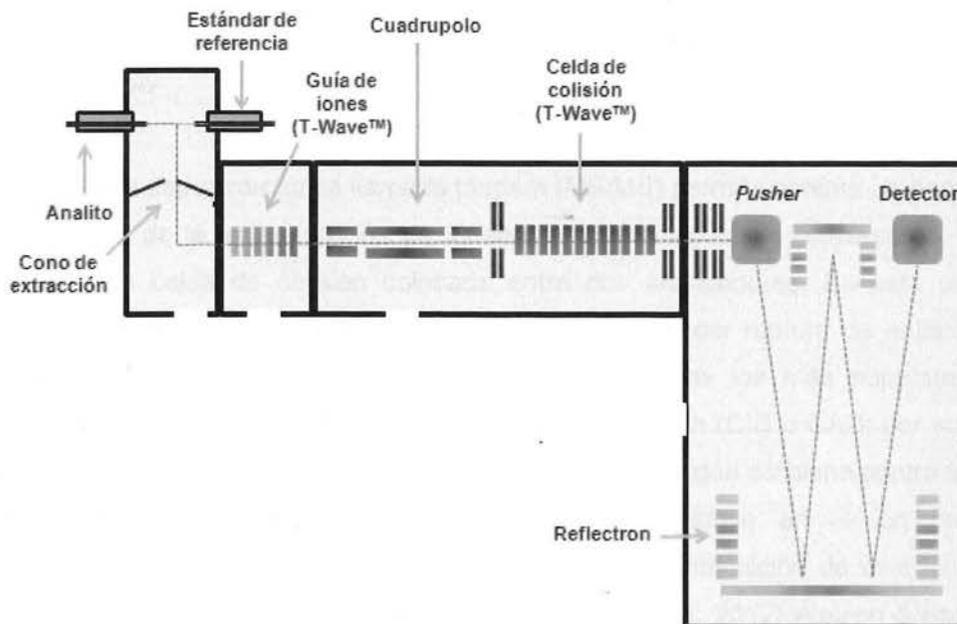
Al ingresar al cono de muestreo, los iones atraviesan un capilar a alta temperatura, el cual tiene la función de eliminar por completo el solvente residual, así como también en separar aglomerados (Kearle & Verkerk, 2012).

## ***Analizador***

Existen distintos tipos de analizadores, cuya función es separar los iones provenientes en la fuente de iones para su posterior medición. Algunos ejemplos son a) cuadrupolos (Q, por sus siglas en inglés), b) sectores, c) tiempos de vuelo (TOF, por sus siglas en inglés), d) trampa de iones orbital (Orbitrap™) y e) de resonancia de ciclotrón (i.e. transformada de Fourier; FTIR). Estos últimos tres se han planteado como los más apropiados para utilizar en estudios metabolómicos en conjunto con combinaciones como el Q-TOF (Ardrey, 2003; Glish & Vachet, 2003; Keski-Rahkonen et al., 2017).

Un cuadrupolo lineal (Q, por sus siglas en inglés), consiste en un espacio conformado por cuatro barras metálicas que se distribuyen de forma paralela en el espacio, formando pares. Cada par está conectado eléctricamente y cargado con un potencial generado por una corriente directa (positiva o negativa); a la vez se aplica un voltaje de radio frecuencia de corriente alterna, generando un campo. Este último es el que selecciona los iones que oscilan mientras atraviesan el cuadrupolo; los que no son seleccionados son filtrados en los espacios entre las barras (Allwood & Goodacre, 2010; Glish & Vachet, 2003; Niessen, 2006).

Un cuadrupolo puede estar unido a un analizador TOF (Figura 13). Existen distintas configuraciones de los TOF, pero su principio es la separación de iones con base en su velocidad o energía cinética. El sistema se basa en un acelerador de iones que favorece la movilidad a través de un tubo que conecta con un detector. La medición de tiempo en que un ión dura recorriendo este trayecto, tiene relación inversamente proporcional con el valor  $m/z$  (Glish & Vachet, 2003)



**Figura 13.** Analizador cuadrupolo-tiempo de vuelo (Q-TOF, por sus siglas en inglés) (elaboración propia, adaptada de Waters).

Un estudio realizó una comparación entre UPLC acoplados a dos tipos de detectores: Q-TOF y Orbitrap, para investigación metabolómica con aplicación en plantas. Éste concluyó que ambos tuvieron elevada exactitud y desempeño global. Algunas leves diferencias fueron que Q-TOF permitió una menor variabilidad mientras que el Orbitrap permitió alcanzar menores límites de detección para algunos compuestos (Glauser, Veyrat, Rochat, Wolfender, & Turlings, 2013).

### **Detector**

Los detectores permiten medir características de los iones que atravesaron el analizador, y convertir esta información en señales interpretables. Muchos de ellos se basan en la producción de una cascada de electrones una vez que un ión llega al detector. Esta cascada puede generar una corriente eléctrica que posteriormente es interpretada por una computadora (Allwood & Goodacre, 2010).

### **10.8.2.3. Espectrometría de masas tándem (MS/MS) y mecanismos de fragmentación**

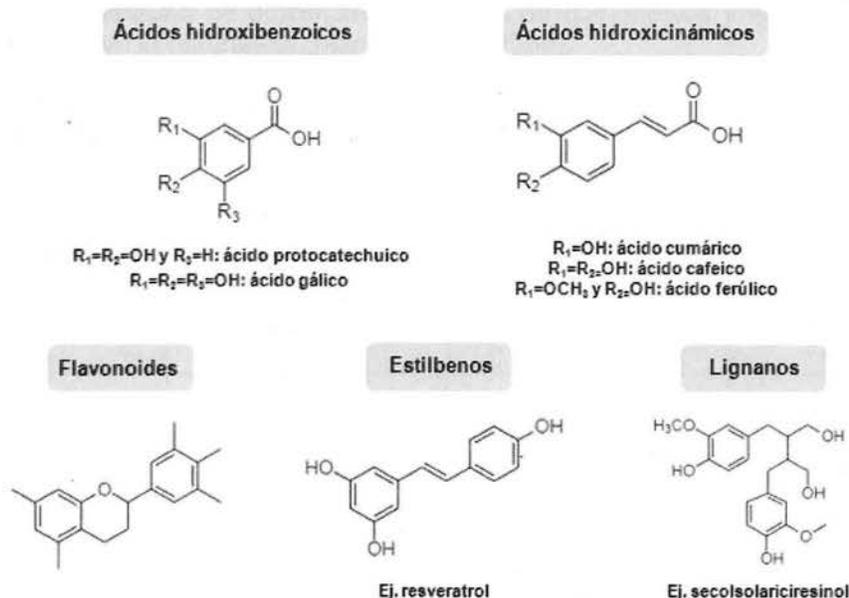
La metodología espectrométrica llamada tándem (MS/MS) permite obtener un espectro de masas a partir de la fragmentación de iones seleccionados -iones precursores. Esto se realiza en una celda de colisión colocada entre dos analizadores. En esta celda, se propicia la activación de los iones precursores, para provocar ruptura de enlaces. Esta activación puede darse con distintos métodos, en uno de los más populares es la disociación por colisión inducida o por activación por colisión (CID o CAD, por sus siglas en inglés), en donde un gas inerte como nitrógeno, helio o argón colisiona contra los iones en fase gaseosa, transfiriendo energía que se redistribuye en el ión precursor. Consecuentemente, se generan subproductos de la descomposición de este ión, dando paso a un nuevo espectro de masas (Banerjee & Mazumdar, 2012; Watson & Sparkman, 2007; Wysocki, Resing, Zhang, & Cheng, 2005).

Para comprender los principales mecanismos de fragmentación que pueden ocurrir en distintas moléculas es necesario conocer su química. A continuación, se describen algunas generalidades sobre la química de los polifenoles, oligopéptidos y azúcares, así como vías comunes de fragmentación.

#### ***Fragmentación de polifenoles***

Los polifenoles son metabolitos secundarios que se encuentran ampliamente distribuidos en las plantas. Con frecuencia, están involucrados en mecanismos de defensa contra la radiación ultravioleta, microorganismos, insectos y otros predadores. Estos también juegan un papel relevante en los alimentos, ya que han sido asociados a beneficios a nivel de salud y son determinantes en atributos sensoriales como por ejemplo color y sabor (Manach, Scalbert, Morand, Rémésy, & Jiménez, 2004; L. de L. de Oliveira, Carvalho, & Melo, 2014).

Dada su diversidad, existen distintas formas de clasificar estos compuestos. Una de ellas propone los siguientes principales grupos (Figura 14): 1) ácidos fenólicos (hidroxibenzoicos o ácidos hidroxicinámicos) 2) flavonoides (Figura A.II), 3) estilbenos y 4) lignanos. Estos pueden encontrarse naturalmente de forma individual o conjugados con carbohidratos, ácidos orgánicos u otros polifenoles (Cuadro X) (Manach et al., 2004).



**Figura 14.** Clasificación de polifenoles (Manach et al., 2004).

Los taninos son también compuestos polifenólicos complejos. Estos pueden ser I) hidrolizables, es decir formados a partir de ácidos hidroxibenzoicos (galotaninos y elagitaninos) o II) condensados (también llamados proantocianidinas), que son polímeros de catequinas (Manach et al., 2004).

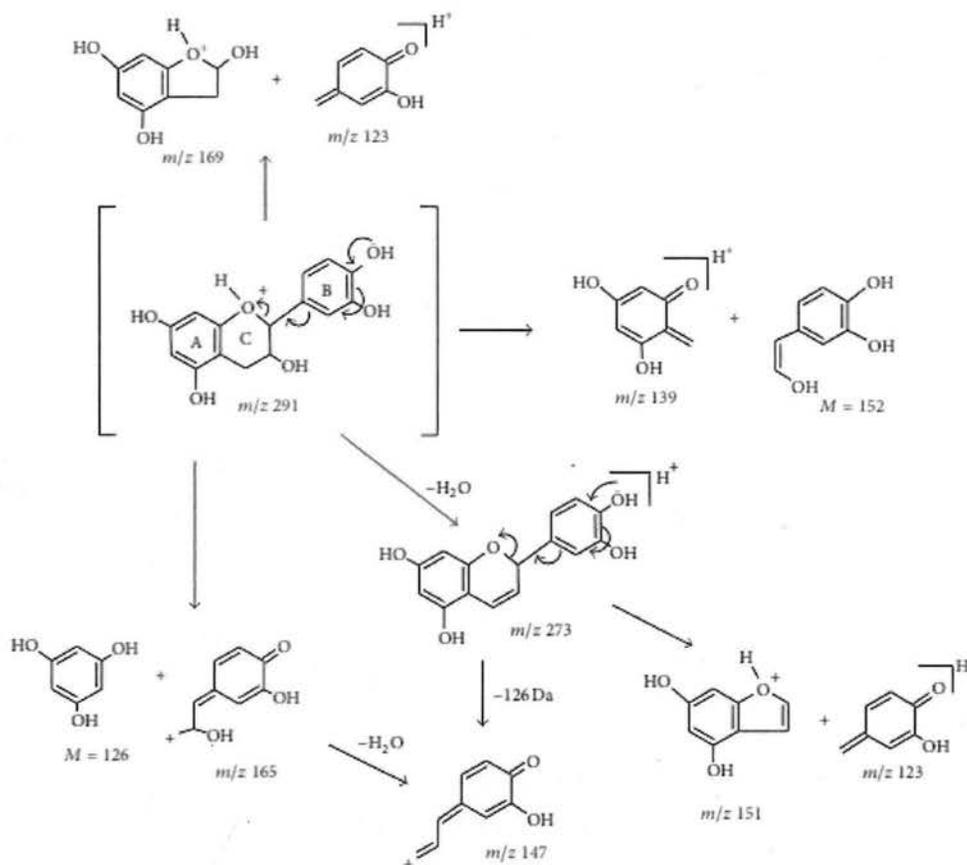
Los taninos condensados o proantocianidinas también son nombrados según la unidad monomérica que los conforma (Gu et al., 2003):

- *Procianidinas*: formadas a partir de (epi) catequinas
- *Propelargonidinas*: tienen como unidad base las (epi) afzelequinas
- *Prodelfininas*: conformadas por (epi) galocatequinas

Producto de la diversidad de estructuras polifenólicas, también existen diferentes mecanismos de fragmentación. Algunas reacciones importantes son la eliminación retro-Diels-Alder (RDA), la fisión heterocíclica de anillos (FHA) y la ruptura directa del enlace interflavano (Gu et al., 2003). Algunas de estas reacciones se ejemplifican en la Figura 15.

**Cuadro X.** Clasificación de polifenoles, ejemplos de compuestos y forma química en la que se presenta en la naturaleza (Manach et al., 2004).

Grupo de polifenoles	Subgrupos	Ejemplos	Forma química natural
Ácidos fenólicos	Ácidos hidroxibenzoicos	Ácido gálico Ácido protocatechuico	Libres o ésteres
	Ácidos hidroxicinámicos	Ácido cafeico Ácido ferúlico Ácido <i>p</i> -coumárico	Glicósidos o ésteres de ácidos quínico, shiquímico y tartárico
Flavonoides <i>(ver estructuras en anexo Figura A.I)</i>	Flavonoles	Quercetina Kaempferol	Glicósidos
	Flavonas	Luteolina Apigenina	Glicósidos
	Flavanonas	Naringenina Hesperetina	Glicósidos de disacáridos
	Isoflavonas	Genisteína Daidzeína Gliciteína	Agliconas o glicósidos
	Flavanoles monoméricos (catequinas o leucoantocianidinas)	Catequina Epicatequina	Libres
	Flavanoles poliméricos (proantocianidinas o taninos condensados)	Procianidina B2 Procianidina A1	Libres
	Antocianinas	Cianidina Pelargonidina Peonidina	Glicósidos o ésteres de ácidos orgánicos y ácidos fenólicos.
Lignanós	-	Secoisolariciresinol Matairesinol	-
Estilbenos	-	Resveratrol	-



**Figura 15.** Ejemplos de mecanismos de fragmentación de una proantocianidina en modo de ionización positivo (Flamini, 2013).

### **Fragmentación de oligopéptidos**

Según la IUPAC, un péptido se define como la unión de dos o más ácidos amino carboxílicos, ya sean iguales o diferentes, mediante el enlace peptídico (Figura 16). Según la cantidad de estas unidades, o residuos, pueden recibir diferentes nombres. Por ejemplo, un oligopéptido puede poseer entre 3 y 10 unidades, mientras que un polipéptido posee 10 o más unidades. Por otro lado, un polipéptido que posea más de 10 000 residuos aminoácidos se cataloga como una proteína. Los péptidos también pueden unirse con diversos tipos de compuestos, como es el caso de los glicopéptidos (unión con carbohidratos) (Moss, Smith, & Tavernier, 1995).



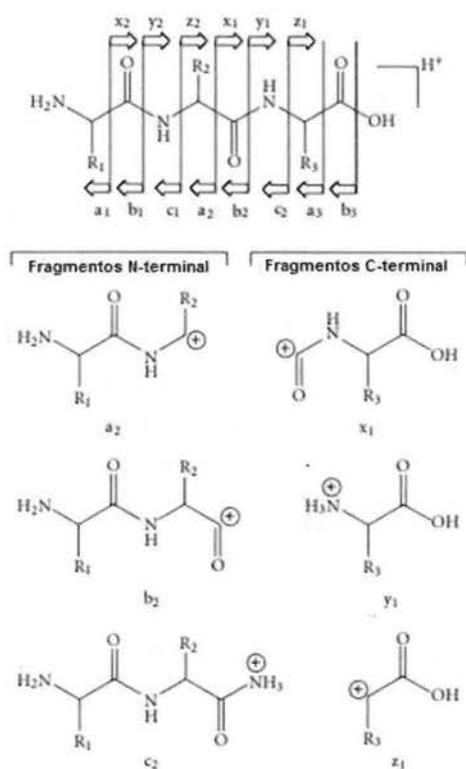


Figura 17. Principales vías de fragmentación en péptidos (Banerjee & Mazumdar, 2012).

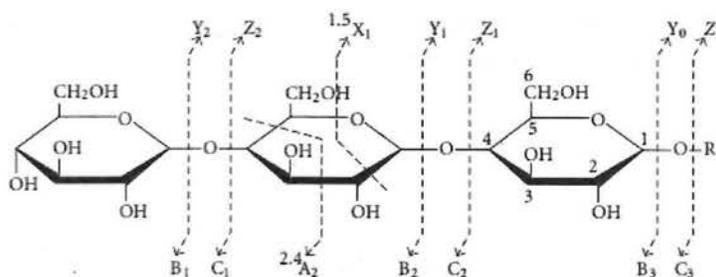
En condiciones de baja energía de colisión, se forman principalmente fragmentos A, B e Y. Estos fragmentos pueden seguir posteriormente otras rutas de fragmentación y generar una diversidad de fragmentos, como los iones iminio<sup>3</sup> (Cuadro A.II) (Banerjee & Mazumdar, 2012).

<sup>3</sup> Se refiere a aquellos cationes que poseen la estructura  $R_2C=N^+R_2$ . Según IUPAC el nombre correcto es iminio, no obstante, popularmente se utilizan los sinónimos imonio e imonio (Moss et al., 1995).

### **Fragmentación de azúcares**

Los azúcares pueden ser medidos tanto en modo de ionización positivo como negativo; además, con frecuencia forman aductos (Black & Fox, 1996). Su análisis es complejo dada la diversidad de enlaces, posiciones y configuraciones anoméricas (Banerjee & Mazumdar, 2012).

Estos por lo general siguen rutas de fragmentación que reciben distintos nombres. En la Figura 18 se pueden distinguir los puntos con mayor propensión a ocurrir en un espectrómetro de masas y la nomenclatura asociada.



**Figura 18.** Principales rutas de fragmentación de los azúcares (Banerjee & Mazumdar, 2012).

## 10.MATERIALES Y MÉTODOS

### 10.1. Localización

El Cuadro XI resume los lugares en donde se desarrollaron las distintas actividades de la investigación.

**Cuadro XI.** Localización de la investigación según el tipo de actividad desarrollada.

<b>Actividad</b>	<b>Lugar</b>
Ejecución de las fermentaciones de cacao	Cocoa Research Center (CRC), Trinidad y Tobago. Centro Nacional de Ciencia y Tecnología de Alimentos (CITA), Costa Rica. Finca productora de cacao, Upala, Costa Rica.
Obtención los extractos de cacao	CITA
Determinación instrumental mediante cromatografía líquida de alta resolución acoplada a espectrometría de masas (CLUR-EM)	Centro de Investigación en Productos Naturales (CIPRONA), Costa Rica.
Procesamiento de megadata metabólica	CITA y CIPRONA

### 10.2. Descripción de la materia prima

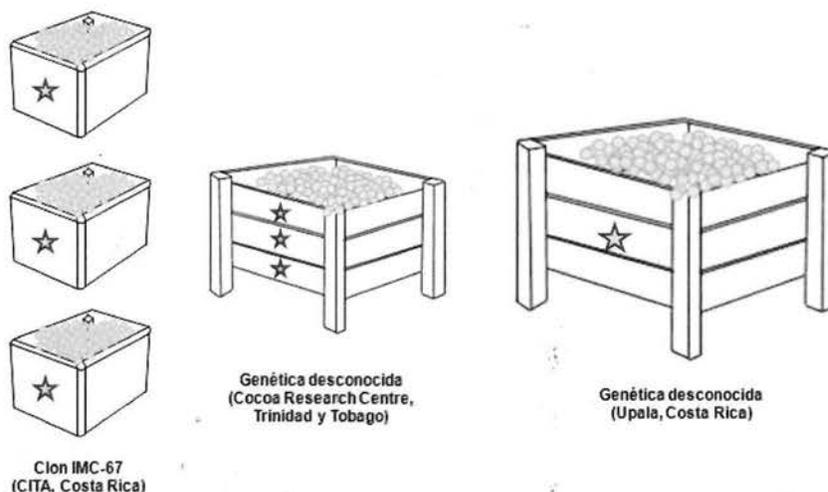
Se analizaron tres tipos de cacao, los cuales se fermentaron empleando distintas metodologías. La descripción de estas materias primas y el número de muestras se presentan en el Cuadro XII.

**Cuadro XII.** Descripción de las materias primas a emplear durante los ensayos de fermentación.

Tipo	Número de muestras*	Origen
Clon IMC-67	3	Centro Agronómico Tropical de Investigación y Enseñanza (CATIE).
Genética desconocida	3**	Variedades autóctonas de una finca de Trinidad y Tobago, con genética desconocida. Las muestras fueron provistas por el Centro de Investigación en Cacao (CRC) de Trinidad y Tobago.
Genética desconocida	1	Variedades autóctonas, con genética desconocida de una finca de Upala, Costa Rica.

Nota: \*Corresponden a lotes completos. \*\*Bajo la hipótesis de no uniformidad en la caja de fermentación, se tomarán tres muestras de un mismo lote de fermentación.

En la figura 19 se resumen los ensayos realizados, y los puntos de muestreo en cada caja:



**Figura 19.** Resumen del número de ensayos de fermentación –representados por caja unitaria–, y los puntos de muestreo por ensayo –representados por estrellas.

Se empleó el siguiente código de identificación de muestras, con base en el tipo de híbrido (primer carácter), día de fermentación (segundo carácter) y ensayo/repetición (tercer carácter). Los últimos dos caracteres indican que los extractos eran de semillas (Sp):

- *Clon IMC-67*: el primer carácter es una C; los ensayos independientes se representan con letras: A, B y C. Ej. C1ASp.
- *Muestras de Trinidad y Tobago*: el primer carácter es una T; la posición de muestreo en caja se representa con letras: A (abajo), B (en medio) y C (arriba). Ej. T4CSp.
- *Muestras de Upala*: el primer carácter es una W, sólo posee un ensayo (A). Ej. W6ASp.

### 10.3. Metodología

#### 10.3.1. Procesamiento de cacao

Para obtener los granos fermentados y secos, se siguió el proceso mostrado en la Figura 20. A continuación, se describe cada una de las operaciones del proceso:

- **Selección:** las mazorcas se seleccionaron manualmente, eliminando aquellas que presentaran síntomas de enfermedad, como crecimiento extendido de hongos, malformaciones o daño mecánico.
- **Despulpado:** los frutos de cacao se abrieron manualmente, empleando un cuchillo o mazo, con el cuidado de no dañar las semillas. Seguidamente se seleccionaron únicamente las semillas en buen estado, descartando aquellas con evidencia o sospecha de daño fisiológico; es decir, las que presentaron con pulpa color café oscuro, textura gelatinosa, presencia de larvas o estaban germinadas.
- **Fermentado:** las condiciones de fermentación espontánea se efectuaron según el tipo de cacao (Cuadro XIII). En términos generales, se colocaron los granos en cajas de fermentación, y durante las primeras 48 horas el cacao se mantuvo cubierto en la superficie con hojas de plátano. Esto se hizo para evitar la difusión de oxígeno y así promover la fermentación anaerobia, que produce etanol. Posterior, a este tiempo se realizó un volteo manual de los granos (a las 48 y 96 h), para promover la incorporación de oxígeno en la masa de fermentación. Lo anterior tuvo como objetivo favorecer el desarrollo de microorganismos aerobios, y la subsecuente conversión de etanol en ácido acético. Así como también el incremento en la temperatura, hasta valores cercanos a los 50 °C.
- **Secado:** los granos congelados se deshidrataron combinando técnicas de secado solar y secado en cabina con aire caliente a una temperatura de 40 °C. El punto final de secado se determinó mediante evaluaciones *in situ* como facilidad de fracturar manualmente los granos, aroma y color característicos, definidos con base en la experiencia, para obtener un contenido de humedad no superior al 7 %.

- **Almacenamiento:** las muestras deshidratadas se almacenaron en bolsas metalizadas y de polietileno de alta densidad, en condiciones de congelación menores a 10 °C, hasta ser extraídas para análisis en laboratorio



Figura 20. Secuencia de operaciones empleada para el desarrollo de las fermentaciones de cacao.

**Cuadro XIII.** Condiciones de fermentación de cacao según el tipo de materia prima.

<b>Materia prima</b>	<b>Descripción</b>
Clon IMC-67	<ul style="list-style-type: none"> <li>• Temperatura de 30 °C en cabina, con humedad relativa propia de la cabina.</li> <li>• Cajas de poliestireno de 23 x 22 x 17 cm y con una carga promedio de 40 kg.               <ul style="list-style-type: none"> <li>• Se monitoreó la temperatura interna de la masa fermentativa en el tiempo.</li> </ul> </li> </ul>
Cacao con genética desconocida (Trinidad y Tobago)	<ul style="list-style-type: none"> <li>• Temperatura y humedad relativa propias del ambiente, a realizarse en las instalaciones del CRC.               <ul style="list-style-type: none"> <li>• Caja de madera, con una carga de 30 kg.</li> </ul> </li> <li>• Se midió una vez al día la temperatura interna en la masa de fermentación.</li> </ul>
Cacao con genética desconocida (Upala)	<ul style="list-style-type: none"> <li>• Temperatura y humedad relativa propias del ambiente de la finca.</li> <li>• Caja de madera de 1,0 x 1,0 x 0,5 m, con una carga aproximada de 100 kg.</li> <li>• Se midió una vez al día la temperatura interna en la masa de fermentación.</li> </ul>

### 10.3.2. Muestreo

El muestreo de los granos de cacao se hizo diariamente. Se tomó una muestra de tamaño variable, en donde la menor fue de 10 granos promedio, para las fermentaciones en cajas de poliestireno y cerca de 100 g en el caso de la caja de madera, con carga de 100 kg.

Las muestras de los clones se congelaron inmediatamente con nitrógeno líquido y se almacenaron en congelación hasta completar los muestreos. Finalmente fueron deshidratadas según las condiciones anteriormente especificadas. En el caso de los otros dos materiales de cacao, se secaron directamente luego de ser muestreadas.

### 10.3.3. Extracción de muestras

Se siguió el protocolo indicado en la Figura 21 para obtener los extractos orgánicos de interés. El protocolo se basa en lo expuesto por varios autores, con algunas modificaciones (Ortega et al., 2010).

A continuación, se describen las distintas etapas utilizadas para obtener los extractos de cacao:

- **Cribado:** se separó manualmente la cascarilla que rodea los granos, previo a su molienda.
- **Molienda:** la molienda se realizó en un procesador de cuchillas, hasta obtener un producto en polvo, en donde visualmente no se evidenciaban trozos de granos.
- **Desgrasado:** esta etapa se hizo a temperatura ambiente, empleando una mezcla de hexanos<sup>4</sup> (grado ACS, Avantor Performance Materials). Para ello se pesaron masas entre 1 a 3 gramos de semillas de cacao molidas en tubos cónicos para centrífuga. Luego, se les adicionó la mezcla de hexanos en una proporción de 8,3 veces la masa exacta del material molido. Seguidamente, los tubos con la mezcla se mantuvieron durante 20 min en una plataforma *Orbital Shaker* a 250 rpm para luego ser sometidos a separación forzada en una centrífuga a 5000 rpm por 15

---

<sup>4</sup> Mezcla de pentano, 3-metilpentano, metilciclopentano, 2-metilpentano y hexano.

min. El supernatante restante se descartó y se repitió la operación de desgrasado hasta completar 4 ciclos.

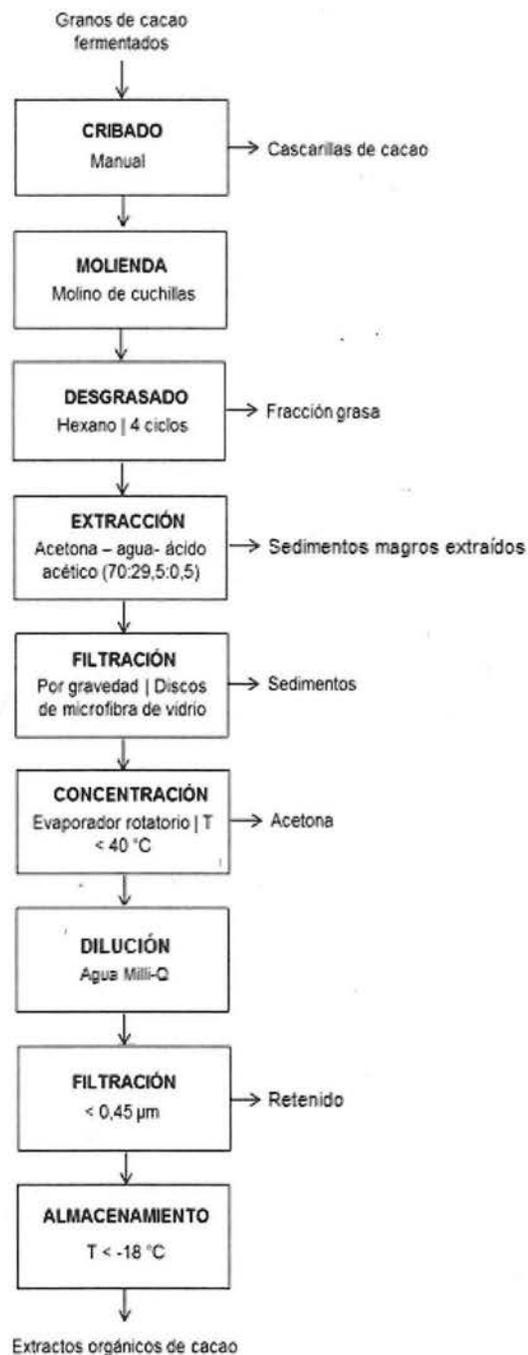


Figura 21. Flujo del proceso para la obtención de la fracción orgánica de interés, a partir de granos de cacao fermentados.

- **Extracción:** el material desgrasado se extrajo con una solución acetona-agua-ácido acético (70:29,5:0,5) (grado ACS de Avantor Performance Material y el agua será tipo Milly-Q). La solución se adicionó en una relación de 5 veces la masa inicial de muestra sin desgrasar. Una vez mezclado el material con la solución, se agitó en un Vortex por 3 min para posteriormente ser centrifugado a 5000 rpm por 15 min. El protocolo de extracción se repitió hasta completar 5 ciclos.
- **Filtración:** las fracciones extraídas se filtraron por gravedad en filtros de microfibra de vidrio (Boeco, Alemania).
- **Concentración:** se empleó un rotavapor para separar por vacío y a menos de 40 °C la acetona de los extractos.
- **Filtración:** los extractos libres de acetona se diluyeron con agua Milly-Q en una relación de 21 veces la masa de material sin desgrasar. Finalmente, estos se filtraron a presión con filtros microporo de celulosa regenerada con un tamaño promedio de 0,45  $\mu\text{m}$  (Minisart-R 25).
- **Almacenamiento:** los extractos se almacenaron a temperaturas inferiores a los -18 °C hasta su medición instrumental.

#### 10.3.4. Preparación de controles de calidad

Como controles de calidad se utilizaron los siguientes:

- *Pool de muestras:* se tomó un volumen constante de cada una de las muestras y se mezcló en un vial.
- *Pool de estándares:* se eligieron por ser representantes de clases de compuestos de interés en cacao (Cuadro XIV).

**Cuadro XIV.** Estándares elegidos como controles de calidad en los ensayos de metabolómica.

Clase de compuesto	Descripción
Aminoácidos	L-fenilalanina ( $\geq 98$ % Sigma-Aldrich) y L-triptofano ( $\geq 98$ % Sigma-Aldrich)
Ácidos orgánicos	Ácido cítrico (99,5 % Sigma-Aldrich) y ácido málico (99 % Sigma-Aldrich)
Azúcares	Sacarosa ( $\geq 98$ % Sigma-Aldrich)
Flavanoles	(+)-catequina (98 % Sigma-Aldrich)

#### 10.3.5. Determinación instrumental

Los extractos se inyectaron de forma aleatorizada en un equipo de cromatografía líquida de alta resolución, acoplada a un espectrómetro de masas (CLUR-EM). Cada 10 muestras se inyectó un pool de estándares y un pool de muestras. Se eligió una ventana de 70-1000 Da para la observación de los iones, bajo las condiciones de medición que se detallan en los Cuadros XV y XVI.

Para las pruebas de MS/MS realizadas posterior a la obtención de modelos estadísticos (ver sección 11.3.7.), se programó el equipo para analizar  $m/z$  específicos en tiempos de retención específicos.

Se midió en condiciones de voltaje tanto positivo como negativo, empero los modelos se construyeron con base en la información generada en modo positivo, dada la mejor calidad de los cromatogramas y espectros asociados.

**Cuadro XV.** Condiciones instrumentales cromatográficas para la medición de los extractos de cacao<sup>5</sup>.

Característica	Descripción
Pre columna (Acquity, Waters)	1,7 µm; CSH C <sub>18</sub>
Columna (Acquity, Waters)	2,1 * 100 mm; 1,7 µm; CSH C <sub>18</sub>
Eluentes	Acetonitrilo con 0,1 % de ácido fórmico (A) Agua grado UPLC con 0,1 % de ácido fórmico (B)
Gradiente (min, % de eluyente A)	(7; 10); (22; 95); (22,1; 0) y (26; 0).
Flujo	0,4 mL/min
Tiempo total por corrida	26 min
Temperatura de columna	30 °C
Volumen de inyección	6 µL
Temperatura de muestras antes de inyección	5 °C

**Cuadro XVI.** Condiciones instrumentales en el espectrómetro de masas para la medición en modo continuo de los extractos de cacao<sup>6</sup>.

Característica	Descripción
Rango de medición	70-1000 Da
Modo de ionización	Positivo y negativo <sup>7</sup>
Tipo de módulo	Waters SYNAPT ESI-QToF Premier (Milford, Massachusetts, USA) ®
Voltaje en el capilar	1,4 kV
Voltaje en el cono	35 V
Voltaje en el cono de extracción	4,0 V
Gas para desolvatar	Nitrógeno con un flujo de 650 L/h
Gas en el cono	Nitrógeno con un flujo a 40 L/h
Solución para calibrar	Yoduro de sodio a una concentración de 2 µg/µL
Masa de referencia externa	Leucina encefalina con una concentración de 2 ng/mL
Tiempo de escaneo	0,2 s
Tiempo inter escaneo	0,020 s
Modo de medición	Continua

<sup>5</sup> Método adaptado a partir de lo propuesto por (Llorach et al., 2014).

<sup>6</sup> Método adaptado a partir de lo propuesto por (Llorach et al., 2014).

<sup>7</sup> El modelo estadístico construido, sólo se realizó en modo positivo.

### 10.3.6. Procesamiento de la megadata

La información obtenida fue preprocesada en el software MZMine. Los parámetros se eligieron con base en distintas pruebas que involucraban el análisis de los cromatogramas y espectros originales, los controles de calidad (blancos, pool de muestras y pool de estándares) (Cuadro XVII). Se obtuvieron 2979 iones al final del preprocesamiento.

**Cuadro XVII.** Parámetros para la identificación de  $m/z$ , construcción de cromatogramas y alineamiento de picos hecho con el software MZMine.

Etapa	Parámetro	Detalle
Filtrado de cromatogramas	Rango de tiempo de cromatogramas	0-20 min
Detección de picos	Nivel de ruido	20
	Tiempo	0,08 min
	$m/z$	0,003 ó 5 ppm
Construcción de cromatogramas	Tiempo	0,08 min
	Altura	5
	$m/z$	0,003 ó 5 ppm
Deconvolución de cromatogramas	Algoritmo	Savitzky-Golay
	Altura	5
	Tiempo	0,06-0,08 min
Alineamiento	Algoritmo	RANSAC
	$m/z$	0,0035 ó 5 ppm
	Tiempo	0,05 min
	Corrección	0,5
	Iteraciones	1000
	Puntos	30
	Umbral	0,05
Modelo lineal	Sí	
Llenado de vacíos	$m/z$	0,0045 ó 8 ppm
	Tiempo	0,4 min ó 8 ppm
	Porcentaje	50
	Corrección	Sí

### 10.3.7. Análisis estadístico

Se exploró el espacio construido por los 2979 iones detectados en las distintas muestras de cacao mediante un análisis de componentes principales (PCA, por sus siglas en inglés) con el programa SIMCA 13.0.2.0 (Umetrics, Umeå, Suecia). El tratamiento de datos incluyó la transformación logarítmica y normalización con el algoritmo de Pareto, de amplio uso en estudios metabolómicos. Luego de explorar el espacio de todas las muestras se definieron tres grupos de interés, caracterizados por los tiempos de fermentación: 0-2 días; 3-4 días y 5-6 días.

Estos grupos fueron la base para el análisis discriminado de mínimos cuadrados parciales (PLS-DA, por sus siglas en inglés). Se seleccionó el número de componentes principales y se evaluaron los parámetros  $R^2$  y  $Q^2$ . Como método de validación se usaron únicamente técnicas de validación interna: validación cruzada y permutaciones. Finalmente se hizo un análisis de CV-Anova para evaluar el modelo.

Se empleó el índice de proyección según importancia (VIP, por sus siglas en inglés) como criterio de selección de las variables (iones) con mayor poder discriminante entre los grupos elegidos. Para ello, se consideró como límite un VIP mayor a 2,5 para un total inicial de 56 iones, que fueron estudiados posteriormente.

Finalmente, se realizó un análisis de variancia (ANOVA, del inglés), por ión discriminante y se obtuvo el valor  $p$  con la finalidad de clasificarlos. Estos análisis se hicieron con los programas JMP y R.

### 10.3.8. Clasificación de la información

Los 56 iones fueron estudiados mediante una combinación de técnicas, con la finalidad de dirigir el análisis:

- *Construcción de histogramas:* se realizó un gráfico en donde el eje Y incluía todas las muestras estudiadas, y el eje X la intensidad del ión. Con esta información se proveyó información adicional para la interpretación biológica de lo observado.
- *Conformación de conglomerados:* se formaron conglomerados de iones considerando principalmente dos parámetros, el valor  $p$  y el tiempo de retención:

- *Clasificación por valor-p*: se clasificaron los iones discriminantes de menor a mayor valor de probabilidad originado del análisis de variancia. Se le asignó un rango de valor  $p$  a cada ión.
- *Clasificación por tiempo de retención*: se empleó como un segundo filtro para ver la posible asociación entre iones con valores  $p$  cercanos.
- *Relación entre  $m/z$* : se observó la correspondencia entre iones con valores  $p$  y tiempos de retención similares, para ver la posibilidad de la existencia de relación entre sus  $m/z$ .

Mediante estas técnicas se lograron conformar 37 conglomerados, los cuales fueron estudiados posteriormente para su identificación tentativa.

### **10.3.9. Identificación tentativa de los iones**

La identificación de los iones se realizó combinando las siguientes técnicas:

- Comparación de tendencias de intensidad a lo largo del tiempo de fermentación.
- Evaluación de los patrones de fragmentación a distintas energías de colisión (experimentos tándem MS/MS).
- Correspondencia con patrones de fragmentación y/o masas exactas en bases de datos.
- Reportes en la literatura de metabolitos en cacao.
- Comparación con bases de datos libres disponibles en Internet.
- Empleo de estándares.

Se propuso una lista de metabolitos asociados a los iones discriminantes. Se debe tomar en cuenta que son propuestas con argumentos que lo justifican, mas no deben tomarse como confirmatorias. Para esto se requieren patrones o evaluación vía técnicas complementarias como resonancia magnética nuclear, que escapan de los objetivos de esta investigación.

## 11. Resultados y discusión

### 11.1. Fermentaciones de cacao

En las Figuras 22 y 23 se observa la evolución de las temperaturas durante las fermentaciones ejecutadas<sup>8</sup>. En el caso de los clones de cacao IMC-67, fue posible medir la temperatura de forma continua, mientras que, para el caso de la mezcla de Trinidad y Tobago, se midió una vez al día. Para el ensayo de fermentación en Upala, se controló la temperatura, no obstante no fue posible acceder al registro correspondiente.

En todos los casos, se observó una tendencia al alza de la temperatura en los primeros tres días de fermentación. Se ha señalado que el metabolismo de conversión de azúcares a etanol y CO<sub>2</sub>, por parte de las levaduras es de carácter exotérmico y puede llegar a incrementar la temperatura hasta los 35 °C durante las primeras horas del proceso (Ardhana & Fleet, 2003; D. S. Nielsen et al., 2013; Nigam & Singh, 2014; Ozturk & Young, 2017; Pereira et al., 2016; Schwan & Wheals, 2004).

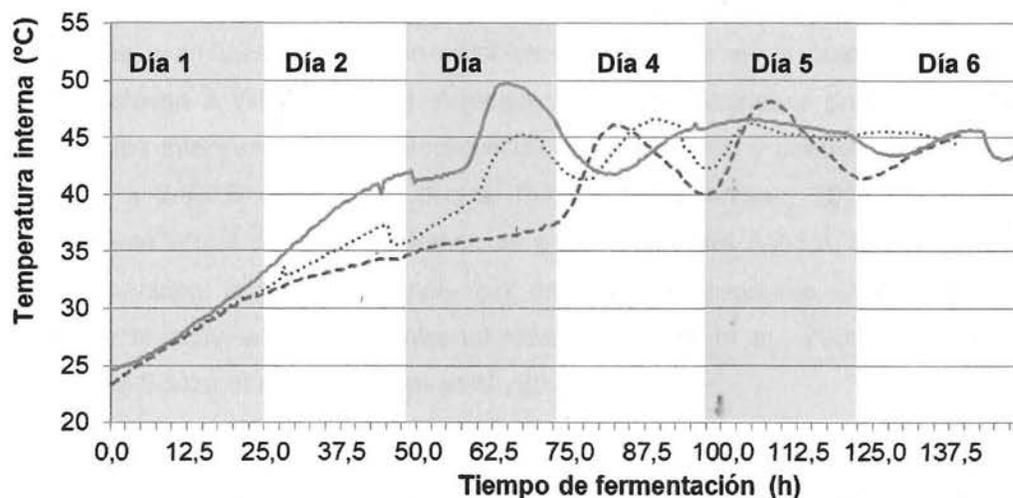
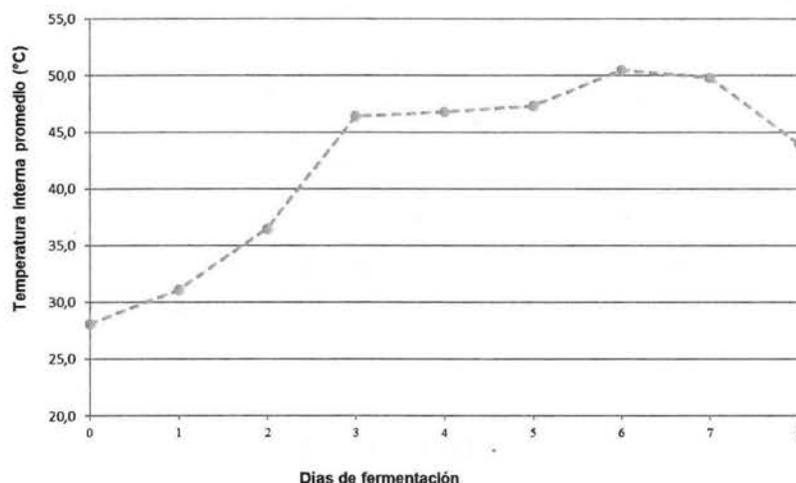


Figura 22. Evolución de la temperatura de fermentación en el tiempo, para los ensayos piloto del cacao clon IMC-67 (- - - ensayo 1, ···· ensayo 2, — ensayo 3).

<sup>8</sup> El monitoreo de los ensayos realizados con la mezcla de cacao de Costa Rica se realizó bajo un protocolo estándar; no obstante, no se hicieron los registros correspondientes.



**Figura 23.** Evolución de la temperatura de fermentación en el tiempo, para el ensayo del cacao con genética desconocida de Trinidad y Tobago.

Al avanzar la fermentación, se dio una reducción en el contenido de pulpa y un cambio en su coloración, tornándose café. Existe evidencia de que diferentes mohos y levaduras, contribuyen a la liberación de enzimas como, por ejemplo, pectinasas, poligalaturonasas, proteasas y amilasas, causando cambios estructurales en la pulpa (Ardhana & Fleet, 2003; Schwan & Wheals, 2004). Además, las polifenoloxidasas presentes naturalmente en la pulpa intervienen en los procesos de oscurecimiento y pueden tener un pH óptimo cercano a 6,5-6,6 (T. B. de Oliveira, Rogero, & Genovese, 2015). A pesar de que inicialmente el pH de la pulpa fresca se encuentra entre 3,2-3,8, el consumo de ácido cítrico -principal ácido de la pulpa- por parte de las levaduras, disminuye la acidez y favorece la actividad de las polifenoloxidasas (Anvoh et al., 2009; Biehl & Ziegleder, 2003a; de Souza et al., 2018; Dias et al., 2007).

Luego de realizar volteos manuales a las 48 h (día 2) y 96 h (día 4), el incremento de la temperatura fue más pronunciado. La incorporación de oxígeno durante el volteo manual favorece el desarrollo de microorganismos aerobios, como las bacterias ácido-acéticas, que oxidan principalmente el etanol a ácido acético. Esta conversión es de carácter exotérmico, lo cual explica el alza de las temperaturas por encima de 45 °C (Ardhana & Fleet, 2003; D. S. Nielsen et al., 2013; Schwan & Wheals, 2004).

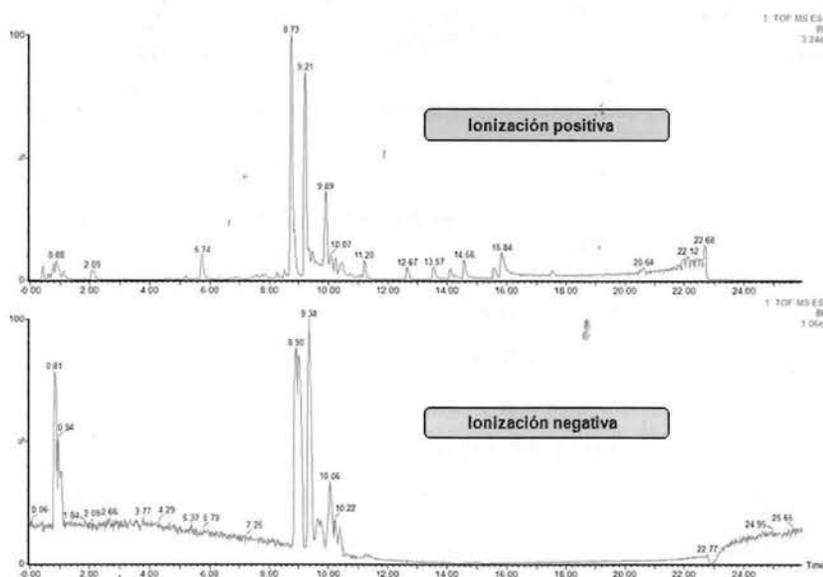
El ensayo 3 del clon IMC-67 y el ensayo con la mezcla de Trinidad y Tobago reflejaron las temperaturas más altas de todos los ensayos: 50 °C. En el caso particular del ensayo 1 del clon IMC-67, los picos de temperatura se presentaron con un retraso, en comparación

con el resto de los ensayos. A pesar de ser fermentaciones controladas, existen diferencias propias de la materia prima y de las condiciones ambientales donde se procesó el cacao, que pueden incidir en el perfil de temperaturas final.

La simulación de fermentaciones en cajas de poliestireno (Figura 22), controladas a nivel de planta piloto, se acerca al comportamiento esperado de fermentaciones clásicas en cajones de madera (Figura 23). De optimizarse, pueden servir como modelo a escala piloto para estudiar futuras fermentaciones.

## 11.2. Evaluación de las corridas cromatográficas

En términos generales, se obtuvo un mayor número de señales y estabilización de la línea base en modo positivo, comparado con modo negativo (Figura 24); adicionalmente, el perfil en modo negativo indicó un problema en la sensibilidad, comparado con el modo positivo. Por esta razón, el análisis quimiométrico se realizó con los datos obtenidos en modo positivo. La información generada en los ensayos en modo negativo se empleó como herramienta adicional para la identificación de los distintos compuestos.



**Figura 24.** Ejemplo de cromatogramas en modo positivo (arriba) y negativo (abajo) de una muestra de cacao con un día de fermentación.

El análisis del pool de estándares (Figura 25), medido de forma repetitiva a lo largo de las mediciones de las muestras, permitió observar que los corrimientos de tiempos de

retención estuvieron dentro de un rango promedio de 0,2 min. Por otro lado, algunas intensidades de picos se vieron afectadas en cantidades de hasta un 50%, pero la gran mayoría se mantuvieron en rangos similares. Estos insumos se emplearon para definir los parámetros de preprocesamiento en el programa MZmine.

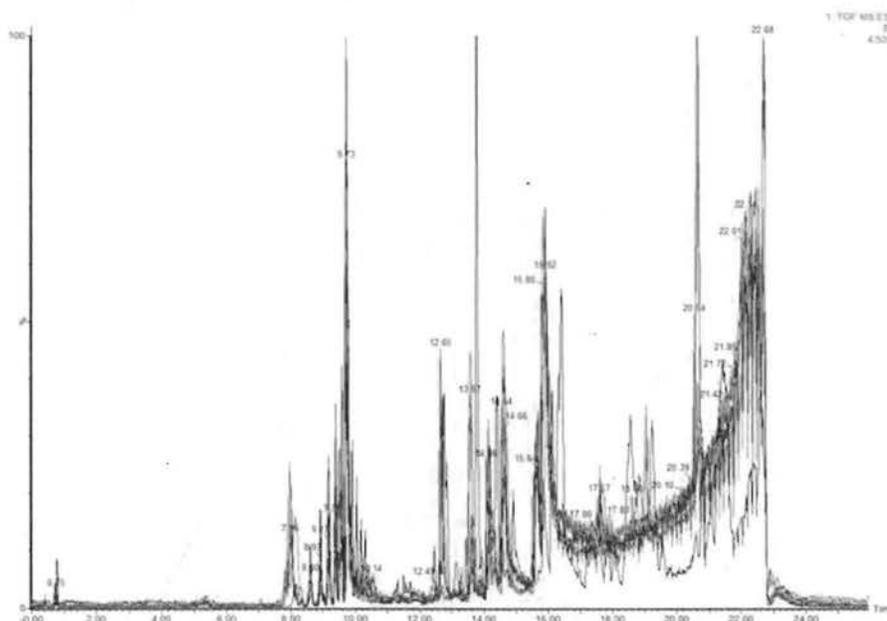


Figura 25. Pool de estándares externos empleados como control de calidad.

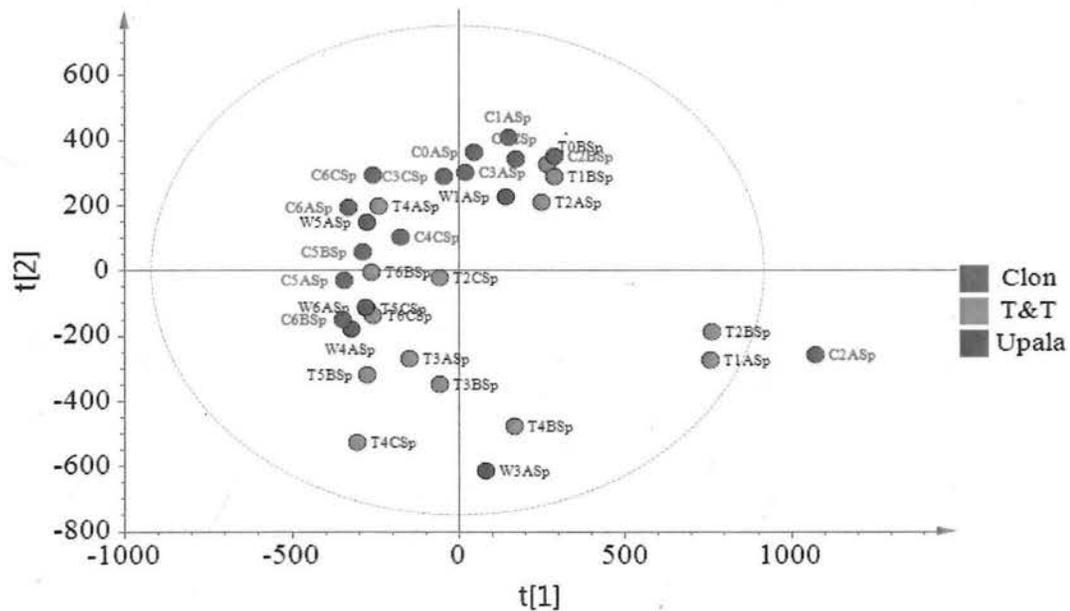
### 11.3. Evaluación del modelo estadístico

Las muestras preprocesadas mediante el programa MZmine, fueron modeladas con estadística multivariada. Se aclara que se obtuvo un menor número de puntos experimentales que los planteados en el diseño experimental debido a errores analíticos durante la extracción de las muestras. Entre estos destacan el uso de tubos Falcom con fisuras no visibles, que provocaron pérdida de muestra durante las etapas de centrifugación, así como también errores involuntarios en la dilución analítica.

Con la finalidad de explorar el comportamiento de las muestras, en función de los iones medidos, se construyó un modelo de dos componentes principales, a partir de los datos preprocesados, y normalizados con el algoritmo de Pareto. En las Figuras 26-29, las muestras se colorearon según diferentes categorías, como, por ejemplo, tipo de híbrido de

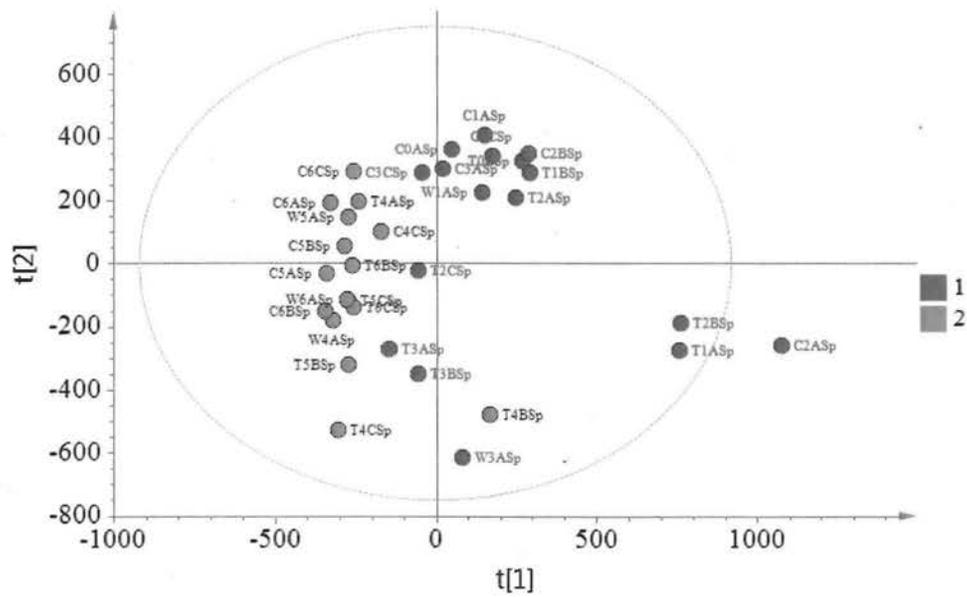
cacao, o bien los días de fermentación. Esto se realizó con la finalidad de explorar agrupamientos.

Se observan indicios de separación de las muestras, en función del tipo de híbrido evaluado (Figura 26). Las principales diferencias se evidenciaron al comparar los clones con las muestras provenientes de Trinidad y Tobago. Las muestras fermentadas en Upala, se dispersaron en el espacio ocupado por las muestras de los otros dos híbridos.



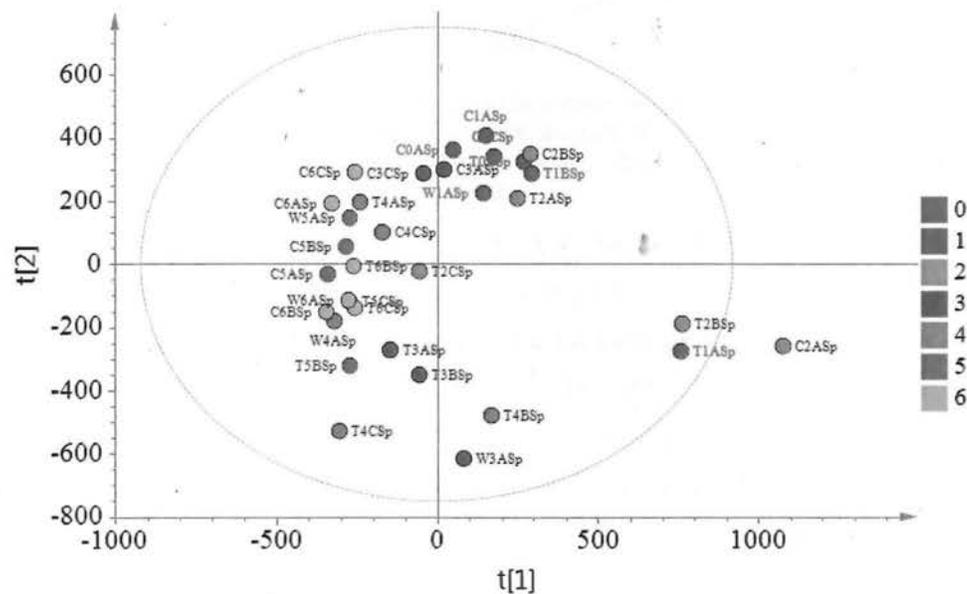
**Figura 26.** Análisis de componentes principales con coloración según tipo de híbrido: el color azul corresponde al clon IMC-65 de Costa Rica, el verde a las muestras de genética desconocida provenientes de Trinidad y Tobago, y la roja a las muestras de genética desconocida provenientes de Upala, Costa Rica.

Es importante subrayar que esta tendencia a separarse por variedad podría ser relevante para la materia prima sin fermentar, y para el cacao apenas iniciando su fermentación. Esto es claro en la Figura 27, puesto que hay una tendencia general, independientemente del híbrido, a estar mayoritariamente asociado al eje negativo del componente principal 1.



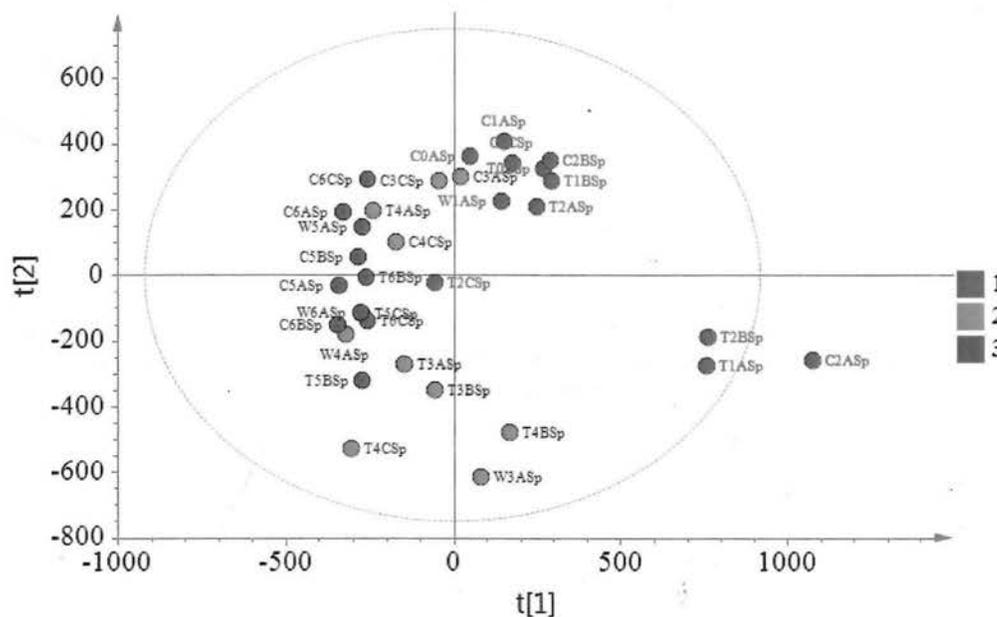
**Figura 27.** Análisis de componentes principales con coloración según días de fermentación: las muestras color azul (No. 1) corresponden a aquellas fermentadas de 0 a 3 días, y las de color verde (No. 2) se fermentaron entre 4 y 6 días.

En la Figura 28 se confirma este comportamiento, ya que en etapas avanzadas de la fermentación, como por ejemplo en los días 5 y 6, la mayor parte de las muestras se agrupó.



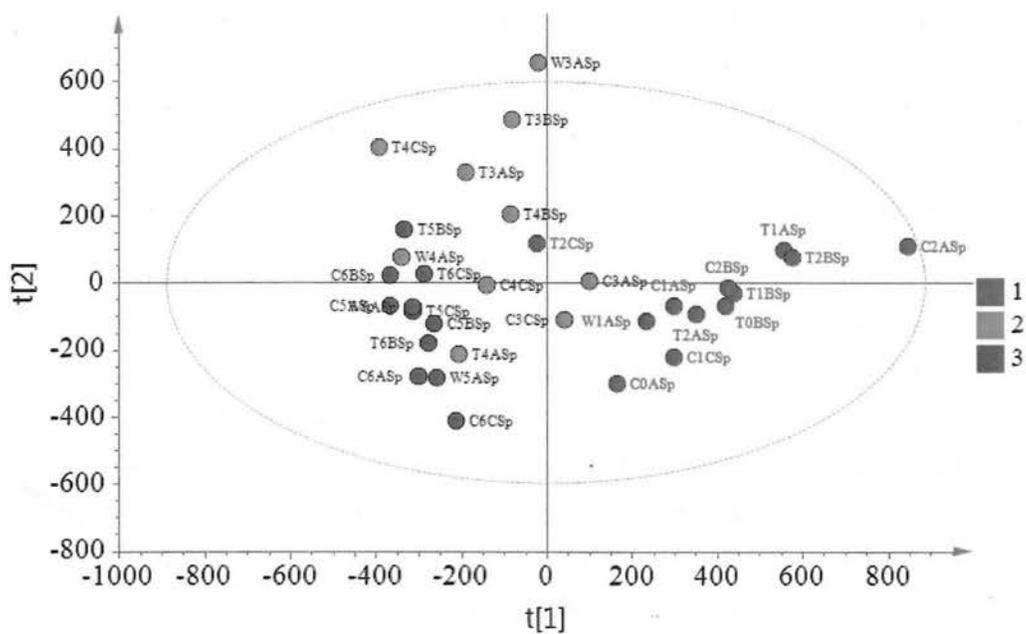
**Figura 28.** Análisis de componentes principales con coloración según día de fermentación.

Considerando los fenómenos biológicos detrás de la fermentación, se logró encontrar una tendencia a la agrupación de muestras en función de la etapa en la que se encontraba: sin fermentar, etapa de la fermentación anaerobia, etapa de fermentación aerobia (Figura 29). Esto, aplicó para todos los casos, sin importar su naturaleza genética. Dada la relevancia científica de comprender con mayor profundidad las diferencias subyacentes en estas etapas –desde un punto de vista de metabolitos- se tomó la decisión de estudiar este modelo.

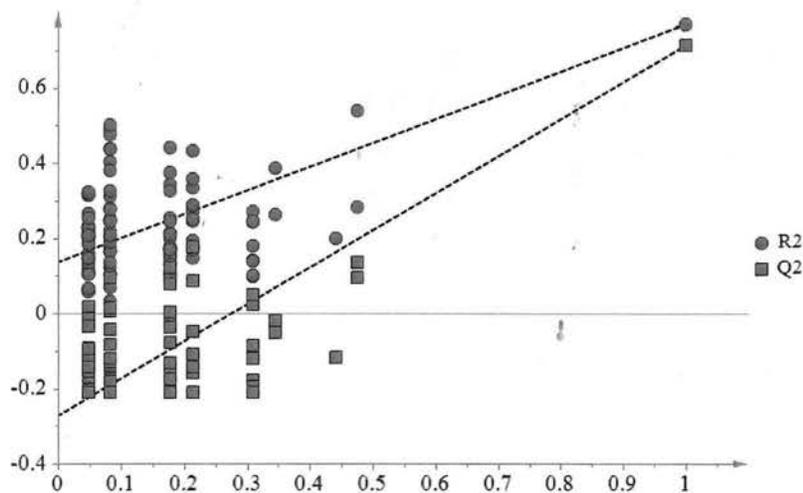


**Figura 29.** Distribución de las muestras de un modelo no supervisado de tipo análisis de componentes principales (PCA, arriba) (las muestras color azul corresponden a los días de fermentación de 0 a 2 días, las muestras en verde a los 3 y 4 días, y las muestras en rojo a los días 5 y 6).

Se construyó un modelo multivariado tipo PLS-DA, de dos componentes, que permitió discriminar apropiadamente entre tres grupos según días de fermentación: 0-2 días, 3-4 días y 5-6 días (Figura 30). Este modelo explicó un 52% de la variación de los datos y mostró una capacidad predictiva aceptable con un  $Q^2Y$  de 0,71. El modelo mostró significancia pues se obtuvo un CV-ANOVA con una  $p$  asociada de  $2,4 \times 10^{-5}$ . La realización de permutaciones permitió concluir que el modelo supera un modelo nulo y reduce la probabilidad de sobreajuste con un  $R^2Y = (0,0; 0,0143)$  y un  $Q^2Y = (0,0, -0,267)$  (Figura 31). Dada la cercanía al modelo, no se eliminó ninguna muestra con compartimento diferenciado (*outlier*), si no que se procedió a estudiarlas con el resto de puntos experimentales.



**Figura 30.** Distribución de las muestras de un modelo supervisado de tipo discriminante (PLS-DA, abajo) (las muestras color azul corresponden a los días de fermentación de 0 a 2 días, las muestras en verde a los 3 y 4 días, y las muestras en rojo a los días 5 y 6).



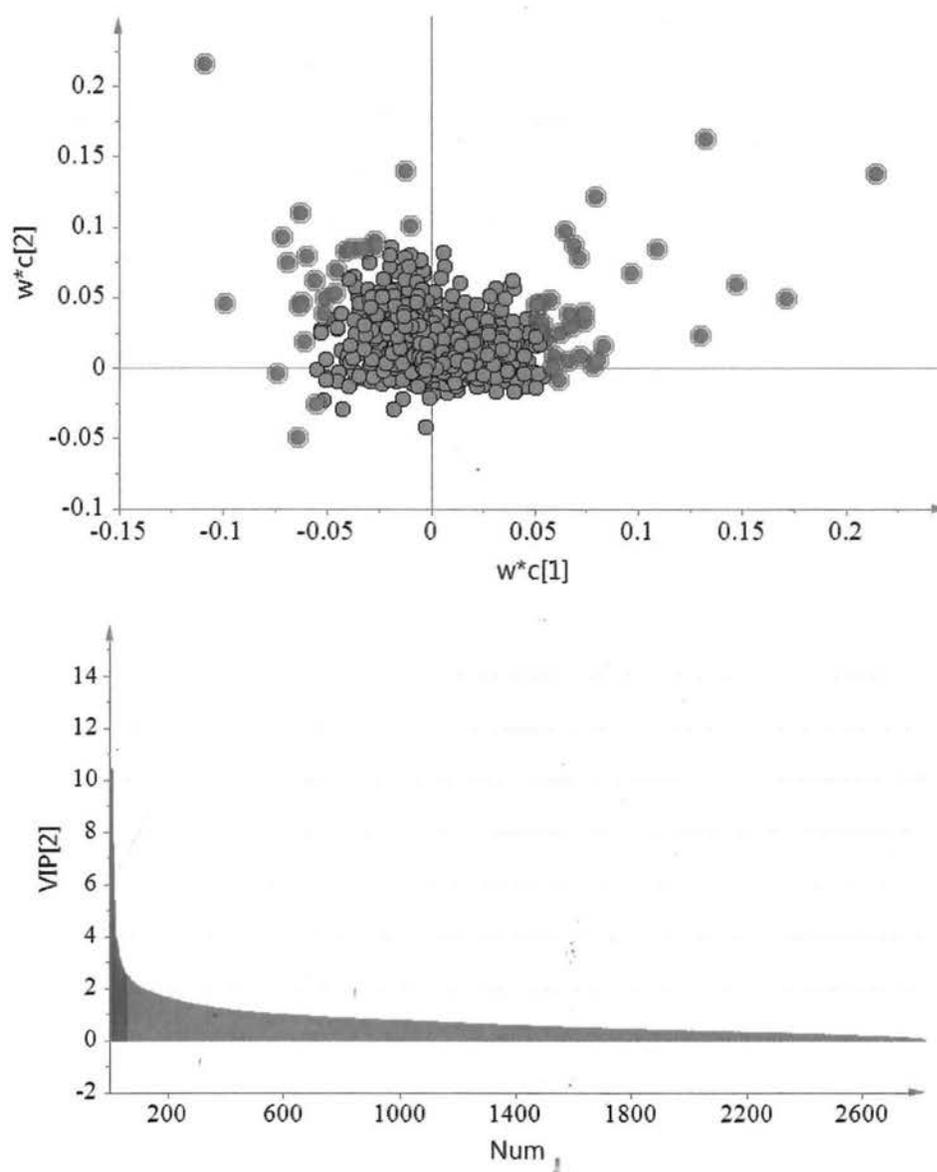
**Figura 31.**  $R^2$  y  $Q^2$  obtenidos al realizar una prueba con 100 permutaciones y 2 componentes principales.

Las similitudes de las muestras, independientemente de la variedad, en etapas avanzadas de la fermentación coincide parcialmente con lo observado en una investigación donde se analizaron 86 híbridos de cacao de 6 distintos países (Brasil, Ecuador, Indonesia, Malasia, Tanzania y Costa de Marfil), destacándose, con dos excepciones, las similitudes entre los cacaos fermentados indistintamente de su naturaleza (D'Souza et al., 2017)

De manera particular, en el grupo con muestras de 4-5 días de fermentación se observó una mayor dispersión con tendencia a separarse según la variedad. Esto podría responder a las diferencias inherentes en el tiempo requerido para alcanzar la máxima temperatura durante la etapa aerobia (Figuras 22 y 23). Es justamente en esta etapa donde se provoca la muerte del embrión, y se propician las reacciones *post-mortem* facilitadas por la deslocalización de enzimas y sustratos (E. Afoakwa, 2014; Biehl & Ziegleder, 2003a). Sustentado en la ecuación de Arrhenius, la temperatura también juega un papel fundamental en la velocidad de reacción de procesos enzimáticos y no enzimáticos y, por ende, de formación de subproductos (Tucker & Woods, 2012).

#### **11.4. Análisis de los iones discriminantes**

El modelo PLS-DA contenía 2979 variables (iones medidos en el espectrómetro de masas), los cuales se tamizaron con el índice de proyección según importancia (VIP, del inglés). Esto permitió obtener una lista de 56 iones con el mayor poder discriminante entre los tres grupos de muestras estudiados (Figura 32).



**Figura 32.** Variables (iones medidos; color verde y color morado) del PLS-DA y variables con mayor poder discriminante, según el índice VIP (color rojo).

De los 10 iones con mayor poder discriminante, 6 de ellos no pudieron ser identificados y 4 de ellos se identificaron tentativamente: 3 flávanoles y 1 oligopéptido. La gran mayoría mostraron una tendencia a la reducción en el tiempo de fermentación, excepto el oligopéptido identificado como PVN y dos iones de identidad desconocida (Cuadro XVIII).

El resto de metabolitos discriminantes se ubicaron dentro de dos clases de compuestos: flavanoles (Cuadro XIX) y oligopéptidos (Cuadro XX). Todos los anotados tentativamente como flavanoles mostraron una reducción al avanzar la fermentación, contrario a los oligopéptidos, en donde se observaron dos tendencias generales: a) un aumento en el tiempo de fermentación o b) un pico de concentración a la mitad de la fermentación.

Cuadro XVIII. Lista de los 10 iones con mayor capacidad discriminante y características cromatográficas, espectrométricas y estadísticas asociadas.

Identidad tentativa	Aducto	m/z (ordenado por valor p)	TR (min)	Valor p (Anova)	VIP	Fragmentos MS/MS	Energía MS/MS (eV)	Intensidad		
								0-2 días	3-4 días	5-6 días
Desconocido 1	[M±H] <sup>+</sup>	343.1237(1)	0.8	9.9E-9	2.7	232.1296 50 150.0889 100 127.0340 65	20	916	313	37
	-	145.0503(13)		1.4E-5	2.6					
Desconocido 2	[M±H] <sup>+</sup>	325.1121(2)	0.9	1.8E-7	3.6	163.0646 30 145.0478 50 127.0378 65 85.0303100	20	1725	561	157
	[M+H-anhidrohexosa] <sup>+</sup>	163.0578(18)		4.8E-5	3					
Desconocido 3	-	396.0336(3)	1.0	3.4E-7	3.3	276.0 80 204.0 60 187.0 100 174.0 70	20	1564	599	188
Desconocido 4	[M±H] <sup>+</sup>	431.2215(4)	4.8	4.6E-7	3.2	342.1842 20 198.1 60 141.1 65 132.1 100; 86.1 90	20	74	626	1297
Dímero de (epi) catequina (procianidina)	[[M+H]+2( <sup>13</sup> C) o + ( <sup>18</sup> O)] <sup>+</sup>	581.1652(5)	8.8	5.6E-7	3.2	409.1 80 287.1 60 139.0 100 127.0 80	20	21559	14862	7983
	[M±H] <sup>+</sup>	579.1487(6)		6.3E-7	10.5					
	[[M+H]+( <sup>13</sup> C)]	580.1565(8)		1.4E-6	6.9					
	[M+H] <sup>+</sup> → RDA - H <sub>2</sub> O	409.0922(24)		2.4 E-4	2.9					
	[M+H] <sup>+</sup> → RDA	427.1027(35)		7.5 E-4	4					
Desconocido 5	[M±H] <sup>+</sup>	632.0664(7)	8.8	1.0E-6	2.5	-	-	1477	1073	670
(epi) catequina	[M±H] <sup>+</sup>	291.0857(9)	8.9	1.6E-6	6	165.1 25 139.0 100 123.0 60	10	5739	2583	1064
Desconocido 6	[M±H] <sup>+</sup>	399.2200(10)	5.3	3.5E-6	4.7	197.1 85 158.1 95 129.1 85 84.0 100	30	554	2977	3121
	[[M+H]+( <sup>13</sup> C)] <sup>+</sup>	400.2392(14)		1.8E-5	2.5					
PVN	[M±H] <sup>+</sup>	329.1832(11)	0.8	4.3E-6	3	197.1 35 169.1 100 70.1 90	15	283	1307	1089
Trímero de (epi) catequina (procianidina)	-	577.1326(12)	9.3	5.7E-6	3.5	579.2 100 409.1 40 291.1 45 247.1 45	20	14087	11405	5061
	[[M+H]+2( <sup>13</sup> C) o + ( <sup>18</sup> O)] <sup>+</sup>	869.2264(19)		5.1E-5	3.2					
	[M+H] <sup>+</sup> → QM	579.1460(39)		1.3 E-3	4.8					
	[[M+H]+( <sup>13</sup> C)] <sup>+</sup>	868.2142(49)		6.3 E-3	4.9					
	[M±H] <sup>+</sup>	867.2109(52)		7.5 E-3	7.5					

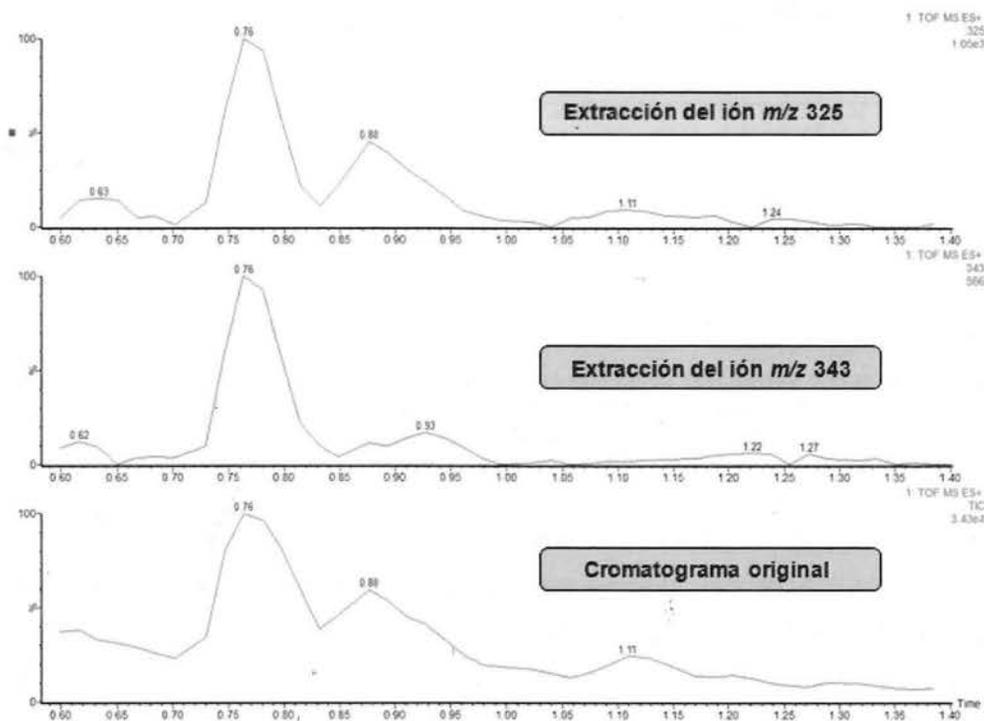
**Cuadro XIX.** Lista de los iones clasificados tentativamente como flavonoides y características cromatográficas, espectrométricas y estadísticas asociadas.

Identidad tentativa	Fórmula	Aducto	<i>m/z</i> (ordenado por valor <i>p</i> )	TR (min)	$\Delta m/z$ (ppm)	Valor <i>p</i> (Anova)	VIP	Fragmentos MS/MS	Energía MS/MS (eV)	Intensidad		
										0-2 días	3-4 días	5-6 días
Dímero de procianidina tipo B	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	[M+H] <sup>+</sup>	579.1468 (17)	10	5.0	4.3E-5	7.9	409.1 100 287.1 85 139.0 60 127.0 75	20	9763	4051	756
Dímero de procianidina tipo A	C <sub>30</sub> H <sub>24</sub> O <sub>12</sub>	[M+H] <sup>+</sup>	577.1315 (33)	11.2	4.5	5.9 E-4	2.7	425.1 65 245.0 100 123.0 30	25	1609	754	464
O-hexósido de proantocianidina A5'	C <sub>36</sub> H <sub>34</sub> O <sub>17</sub>	[M+H] <sup>+</sup>	739.1798 (50)	10.2	4.5	6.6 E-3	2.6	587.1 60 577.1 25 425.1 100 299.1 40	20	1933	1282	703
Dímero de procianidina tipo B	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>	[M+H] <sup>+</sup>	867.2157 (47)	9.8	3.0	5.5 E-3	3.9	409.0 90 289.1 40 287.1 55 247.1 100 139.0 80 127.0 70	30	3143	2279	710
O-pentósido de proantocianidina A5'	C <sub>35</sub> H <sub>32</sub> O <sub>16</sub>	[M+H] <sup>+</sup>	709.1755 (53)	10.3	1.1	0.028	2.6	557.1 100 539.1 35 437.1 30 425.1 90 299.6 45 287.1 30	20	2260	1690	906

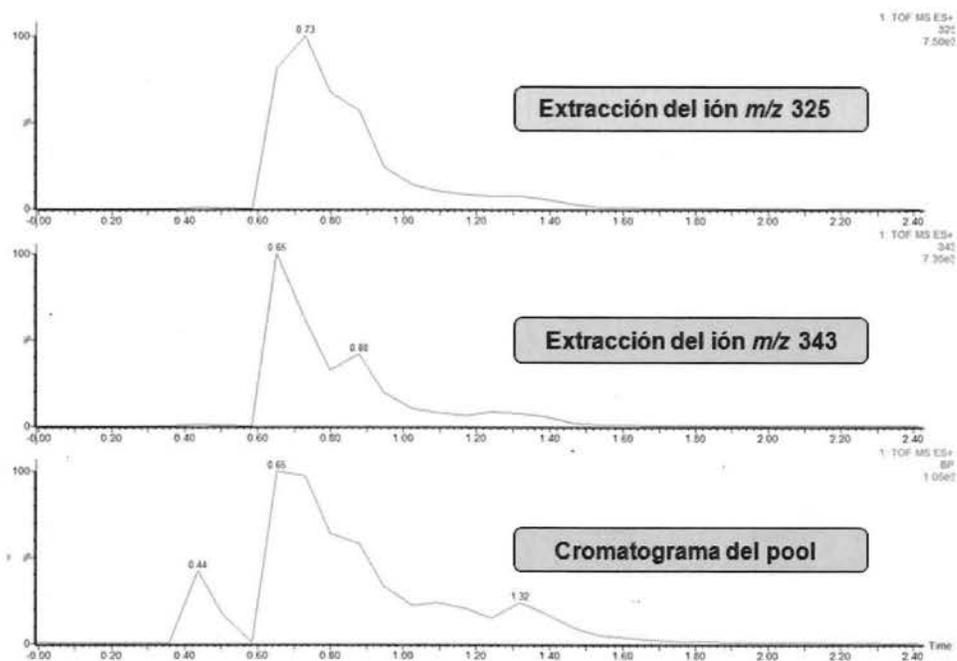
Cuadro XX. Lista de los iones clasificados tentativamente como oligopéptidos y características cromatográficas, espectrométricas y estadísticas asociadas.

Identidad tentativa	Fórmula	Aducto	m/z (ordenado por valor p)	TR (min)	$\Delta$ m/z (ppm)	Valor p (Anova)	VIP	Fragmentos MS/MS	Energía MS/MS (eV)	Intensidad		
										0-2 días	3-4 días	5-6 días
VSTDVN	C <sub>25</sub> H <sub>44</sub> N <sub>7</sub> O <sub>12</sub>	[M+H] <sup>+</sup>	634.3020 (15)	5.4	3.5	2.2E-5	2.5	403.2 100 187.1 80 72.1 100	25	82	857	798
Desconocido	-	[M+Na] <sup>+</sup>	475.3249 (16)	9.7	-	2.4E-5	2.6	197.1 65 169.1 100 70.1 90	20	1447	1979	2404
DNEW	C <sub>24</sub> H <sub>31</sub> N <sub>6</sub> O <sub>10</sub>	[M+H] <sup>+</sup>	563.2111 (20)	8.1	2.7	1.3E-4	3.1	342.1 60 270.1 40 230.1 100 130.1 40	20	187	1408	1339
SPGDVF	C <sub>28</sub> H <sub>41</sub> N <sub>6</sub> O <sub>10</sub>	[M+H] <sup>+</sup>	621.2860 (22)	8.8	3.1	1.2E-4	7.7	357.1 90 270.1 100 70.1 90	30	1283	7826	4859
GAGGGGL	C <sub>19</sub> H <sub>34</sub> N <sub>7</sub> O <sub>8</sub>	[M+H] <sup>+</sup>	488.2443 (30)	5.2	4.1	4.6 E-4	3.3	215.1 70 198.1 75 132.1 65 86.1 100 70.1 65	25	161	1435	1933
NGTPVIF	C <sub>35</sub> H <sub>55</sub> N <sub>8</sub> O <sub>10</sub>	[M+H] <sup>+</sup>	747.4064 (34)	9.8	3.7	6.4 E-4	2.8	475.3 85 197.1 100 169.1 80 70.1 60	30	101	1129	1273
SSISGAGGGGLA	C <sub>37</sub> H <sub>65</sub> N <sub>12</sub> O <sub>16</sub>	[M+H] <sup>+</sup>	933.4640 (36)	7.9	0.43	8.8 E-4	2.7	539.3 80 452.2 85 86.1 90	40	90	913	444
SSISGAGGGGL	C <sub>34</sub> H <sub>60</sub> N <sub>11</sub> O <sub>15</sub>	[M+H] <sup>+</sup>	862.4262 (37)	8.0	0.35	0.0010	2.8	369.2 40 312.1 40 300.1 40 284.6 35	40	182	1204	849
APLSPGDVF	C <sub>42</sub> H <sub>64</sub> N <sub>9</sub> O <sub>13</sub>	[M+H] <sup>+</sup>	902.4612(38)	9.9	0.66	0.0010	3.9	534.2 75 341.2 65 270.1 100	40	457	1535	351
ANSPVLD	C <sub>30</sub> H <sub>51</sub> N <sub>8</sub> O <sub>12</sub>	[M+H] <sup>+</sup>	715.3627 (40)	7.4	0.84	0.0014	2.9	443.2 95 197.1 100 169.1 100	30	161	1120	634
VPIR	C <sub>22</sub> H <sub>42</sub> N <sub>7</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	484.3266 (46)	4.8	5.0	0.0043	2.7	368.2 40 183.1 55 70.1 100	30	91	1179	991

La identificación de los dos iones con mayor poder discriminante ( $m/z$  343.1 con TR de 0,8 min y  $m/z$  325.1 con TR 0,9 min) fue dificultada porque eluyeron en tiempos menores a 1 min, en donde se obtuvo una baja resolución. La sección del cromatograma correspondiente (Figura 33), se comparó con la misma sección de un cromatograma de un pool de estándares de sacáridos, compuesto por sacarosa, maltosa, lactosa y glucosa (Figura 34). Se encontraron similitudes en cuanto a la forma y tiempos de retención.



**Figura 33.** Sección del cromatograma asociado a los dos iones con mayor poder discriminante ( $m/z$  343.1 y 325.1) de una muestra de cacao.



**Figura 34.** Sección del cromatograma asociado a los dos iones con mayor poder discriminante ( $m/z$  343.1 y 325.1) de un pool de estándares de sacáridos: sacarosa  $[M+H]^+=343.1$ , maltosa  $[M+H]^+=343.1$ , lactosa  $[M+H]^+=343.1$  y glucosa  $[M+H]^+=181.0$ .

Los espectros de masas correspondientes muestran señales repetidas al comparar el pool de estándares y los extractos de cacao; como, por ejemplo,  $m/z$  85, 127, 145, 163 y 325 ( $[M+H]^+$ ) (Figuras 35 y 36). Adicionalmente, se lograron detectar señales coherentes con el espectro de masas reportado para sacarosa ( $[M+H]^+=343.1$ ) en la base de datos Metlin (Figura 37), pero en modo negativo. Es necesario mejorar la resolución de la zona, y emplear estándares individuales y no mezclas, para una confirmación completa.

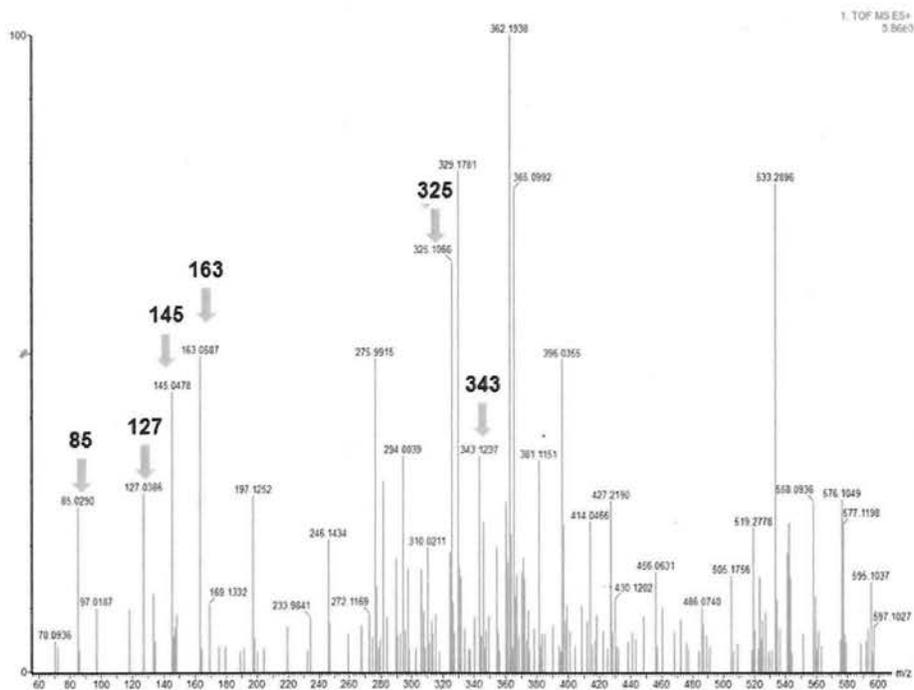


Figura 35. Barrido del espectro de masas experimental en los tiempos de retención (0,7-0,8 min) correspondientes a los dos iones con mayor poder discriminante ( $m/z$  343.1 y 325.1) en un extracto de cacao.

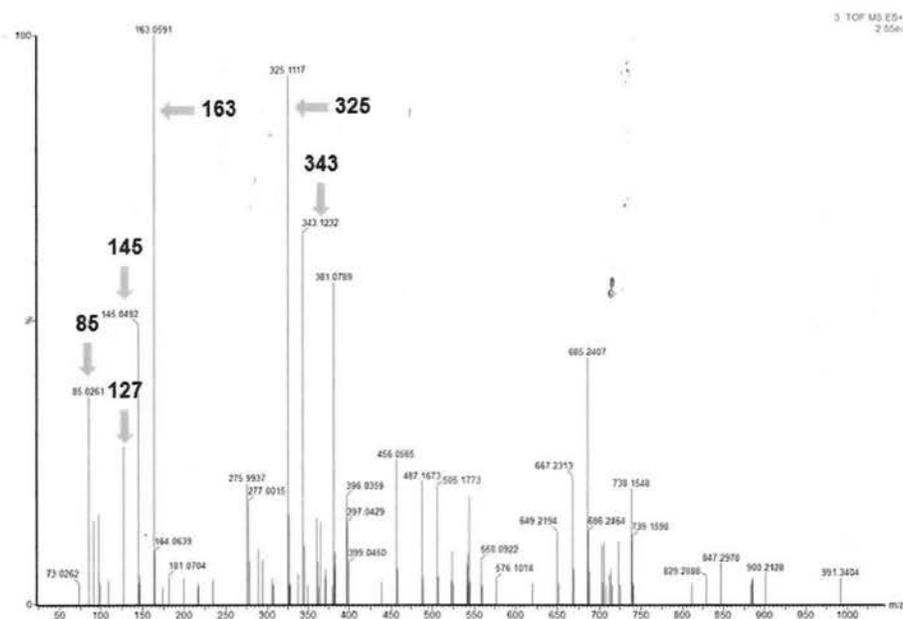
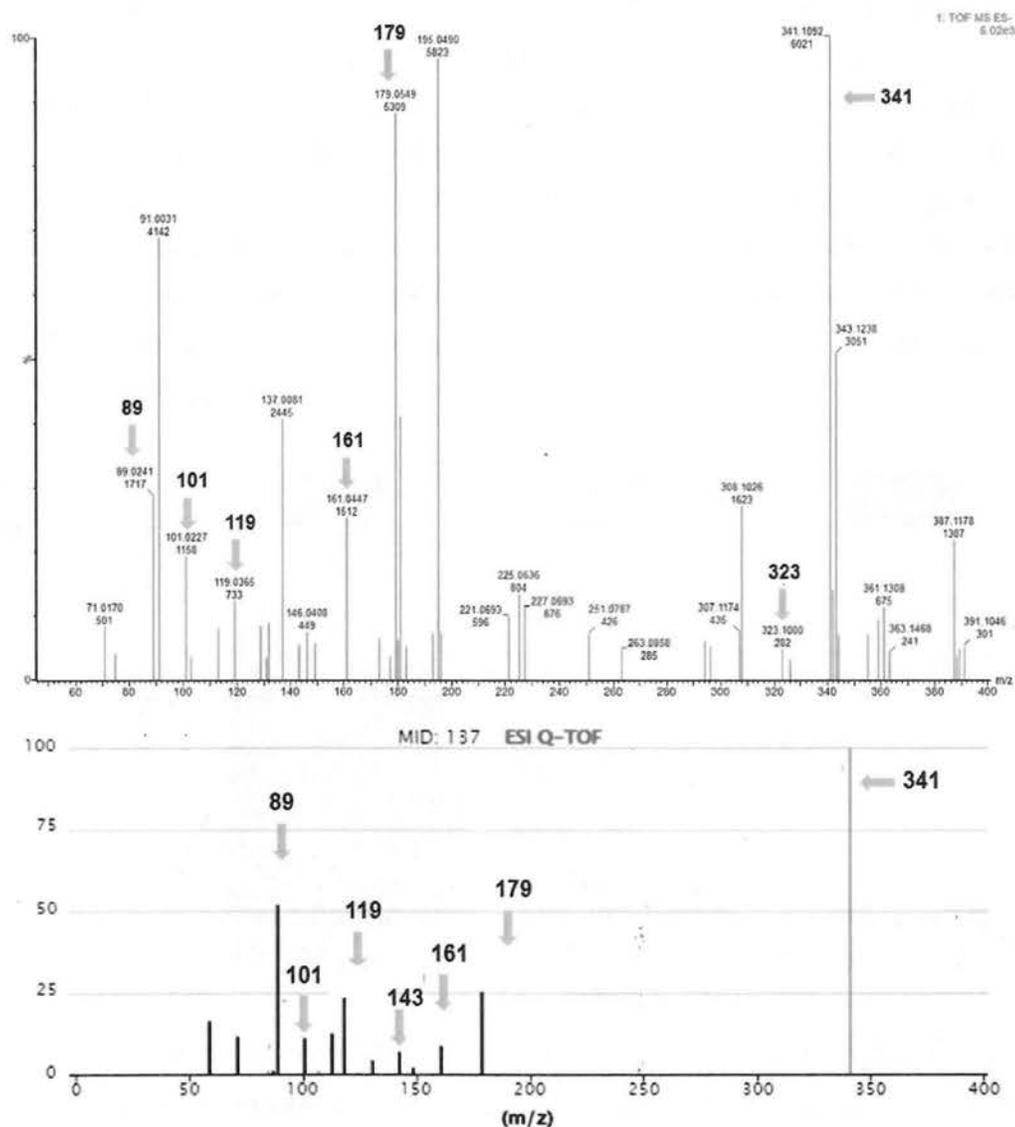


Figura 36. Espectro de masas MS/MS en los tiempos de retención 0,7-0,8 min de un pool de estándares de sacáridos, a 20 eV.



**Figura 37.** Espectro de masas en modo negativo asociado a la región de tiempo de los dos iones con mayor poder discriminante:  $[M-H]^-$   $m/z$  341.1 y 323.1, determinado experimentalmente (arriba) y espectro reportado para la sacarosa en la base de datos Metlin a 10 V (abajo).

Se ha reportado que la sacarosa ( $[M+H]^+$  343.1235) es el azúcar más concentrado en las semillas de cacao, y puede llegar a representar un 2% de la semilla (E. Afoakwa, 2014; Hii et al., 2010); por lo que existe la probabilidad que éste sea un ión discriminante. Es razonable que, de ser sacáridos estos dos iones, mostraran una tendencia a la disminución en el tiempo de fermentación, como se confirma en la Figura 38. Esto se fundamenta en el hecho de que la sacarosa es extensivamente hidrolizada por invertasas; lo que puede provocar la degradación casi completa para el tercer día de fermentación (Biehl & Ziegler, 2003a; Reineccius, Andersen, Kavanagh, & Keeney, 1972; Rohan & Stewart 1967).

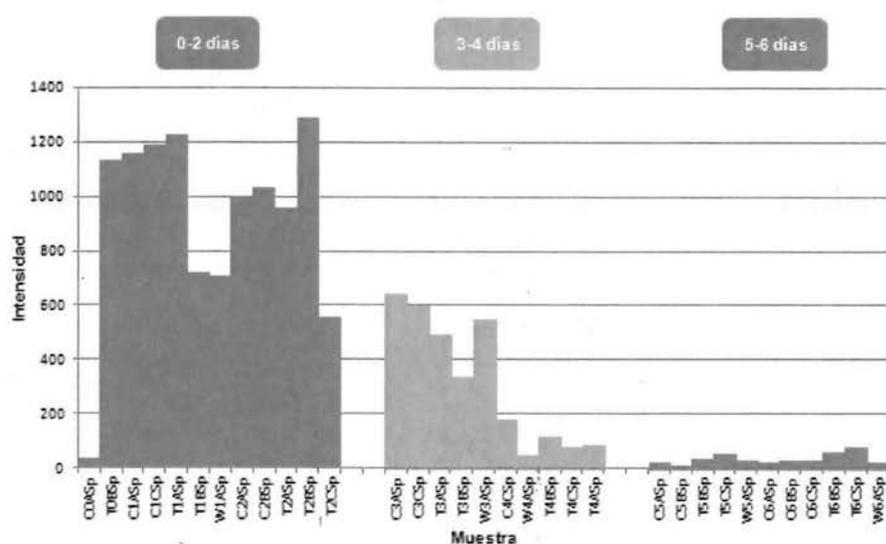


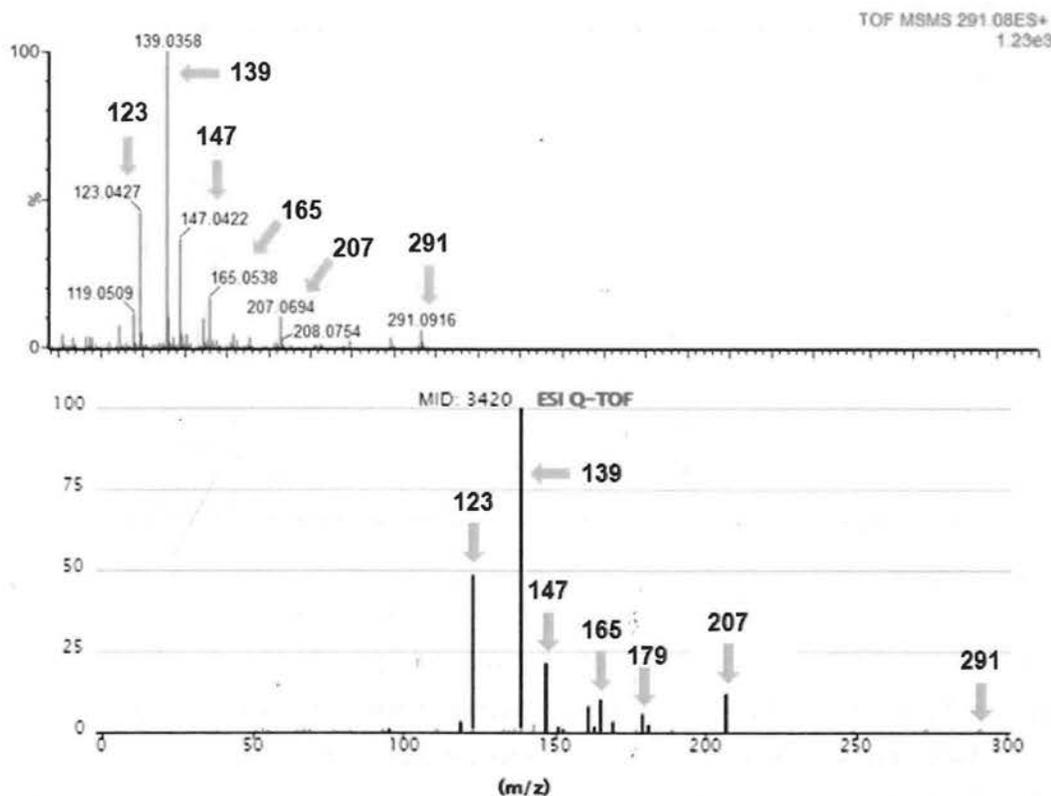
Figura 38. Comportamiento en el tiempo de fermentación del ión  $[M+H]^+$  = 343.1.

En resumen, se planteó que los dos iones discriminantes ( $m/z$  343.1 con TR de 0,8 min y  $m/z$  325.1 con TR 0,9 min) poseen naturaleza de sacáridos por: a) el traslape de tiempos de retención y la similitud de señales con respecto a los estándares externos de azúcares y, b) el comportamiento esperado de este tipo de metabolitos durante la fermentación. Se reitera en la necesidad de mejorar la resolución del cromatograma y emplear estándares individuales para confirmar su identidad.

El ión  $m/z$  291,0 con TR de 8,9 fue identificado como (epi) catequina y se ubicó dentro de los 10 iones de mayor poder discriminante. Además de la coincidencia por masa exacta ( $\Delta$  2 ppm), el espectro de este ión coincidió con el espectro de masas reportado por la

base de datos MetLin, y generado a partir de la medición de epicatequina a 20 eV con un ESI-QTOF (Figura 39).

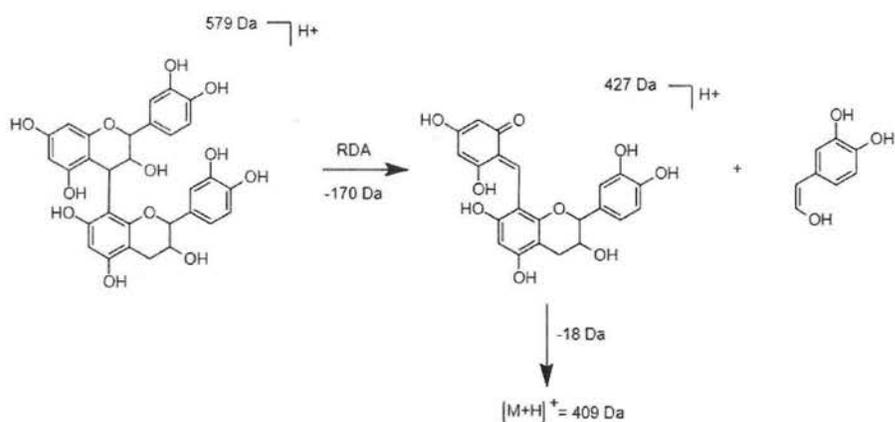
Se ha reportado que la (-)-epicatequina es la catequina más abundante y se encuentra principalmente de forma libre o como elemento estructural de las procianidinas de cacao, que conforman más de la mitad de los polifenoles del cacao (D'Souza et al., 2017; Dorenkott et al., 2014; Hii et al., 2010; Wollgast & Anklam, 2000).



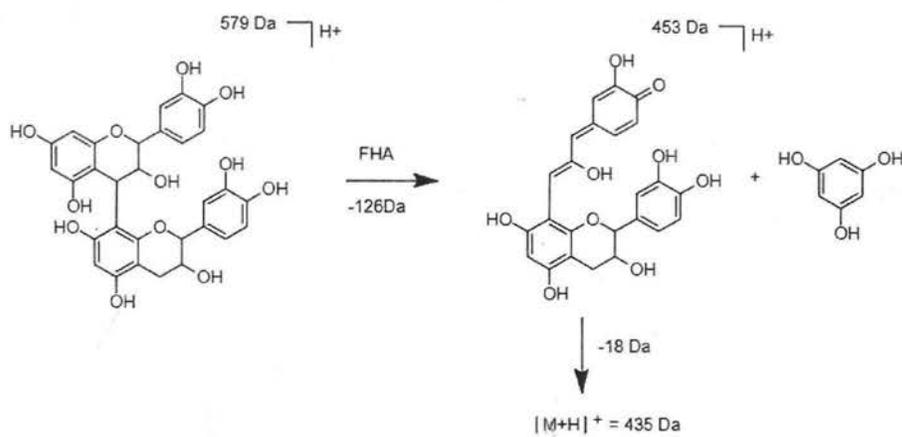
**Figura 39.** Espectro de masas MS/MS experimentales obtenidos en modo positivo en a) la presente investigación para el ión  $[M+H]^+$   $m/z$  291,0 con RT de 8,9 a 15 eV (arriba) y b) de la epicatequina a 20 eV según la base de datos MetLine (abajo).

También se lograron identificar de forma tentativa distintas procianidinas, con dos y tres monómeros (Cuadros XVIII y XIX). Entre las procianidinas más relevantes en cacao se destacan los dímeros B2 y B5, el trímero C1, el tetrámero A2 y otros pentámeros, como el A3 (Esatbeyoglu, Wray, & Winterhalter, 2015), y la gran mayoría se ubica en el rango de dos a diez unidades. Sin embargo, muchas de estas no se determinaron debido a la ventana definida en el espectrómetro de masas: de 70 a 1000 Da. Por ejemplo, al analizar experimentalmente un extracto de cacao en el rango de 70 a 3000 Da y comparar las  $m/z$

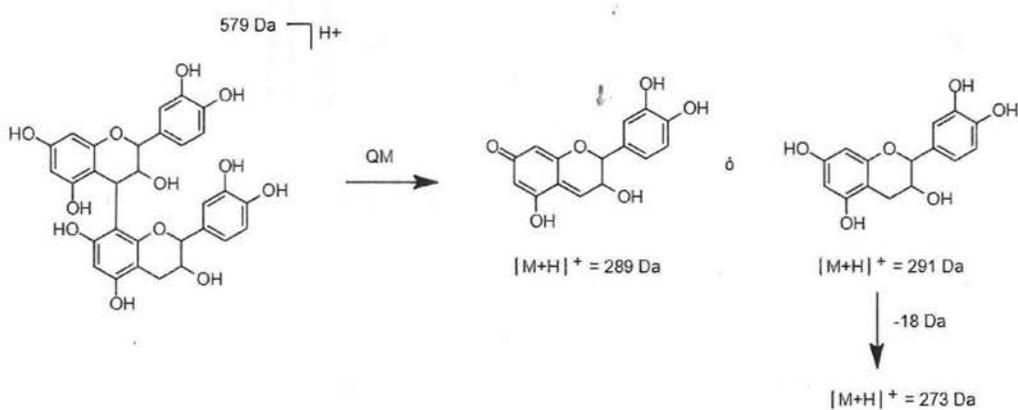




**Figura 41.** Ejemplo de descomposición por eliminación Retro Diels Alder del ión  $[M+H]^+$  579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX).



**Figura 42.** Ejemplo de descomposición por fisión heterocíclica del anillo del ión  $[M+H]^+$  579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX).



**Figura 43.** Ejemplo de descomposición por ruptura directa del enlace interflavano del ión  $[M+H]^+$  579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX).

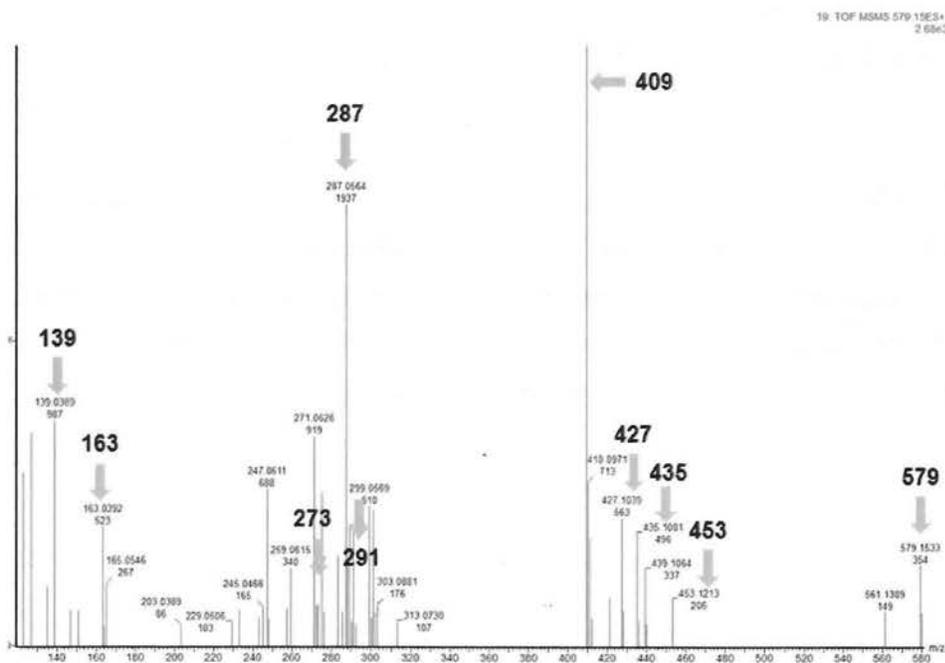


Figura 44. Espectro de masas MS/MS del ión  $[M+H]^+$  579.1468 con TR de 10 min, identificado tentativamente como procianidina con enlace tipo B (Cuadro XIX).

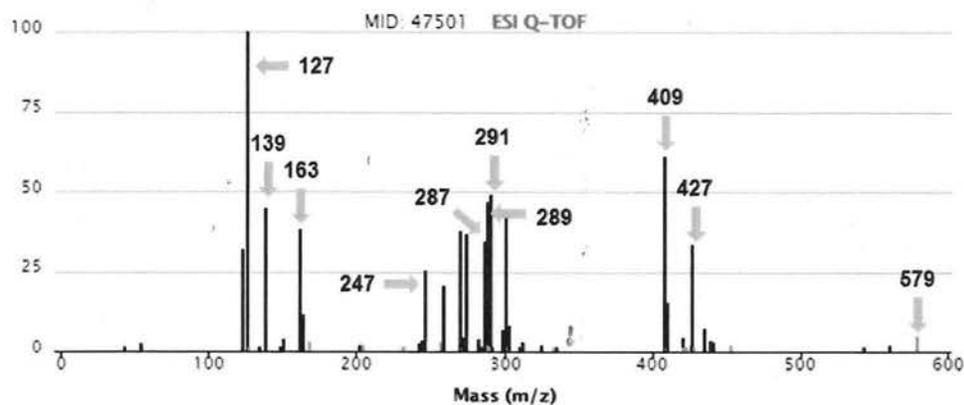
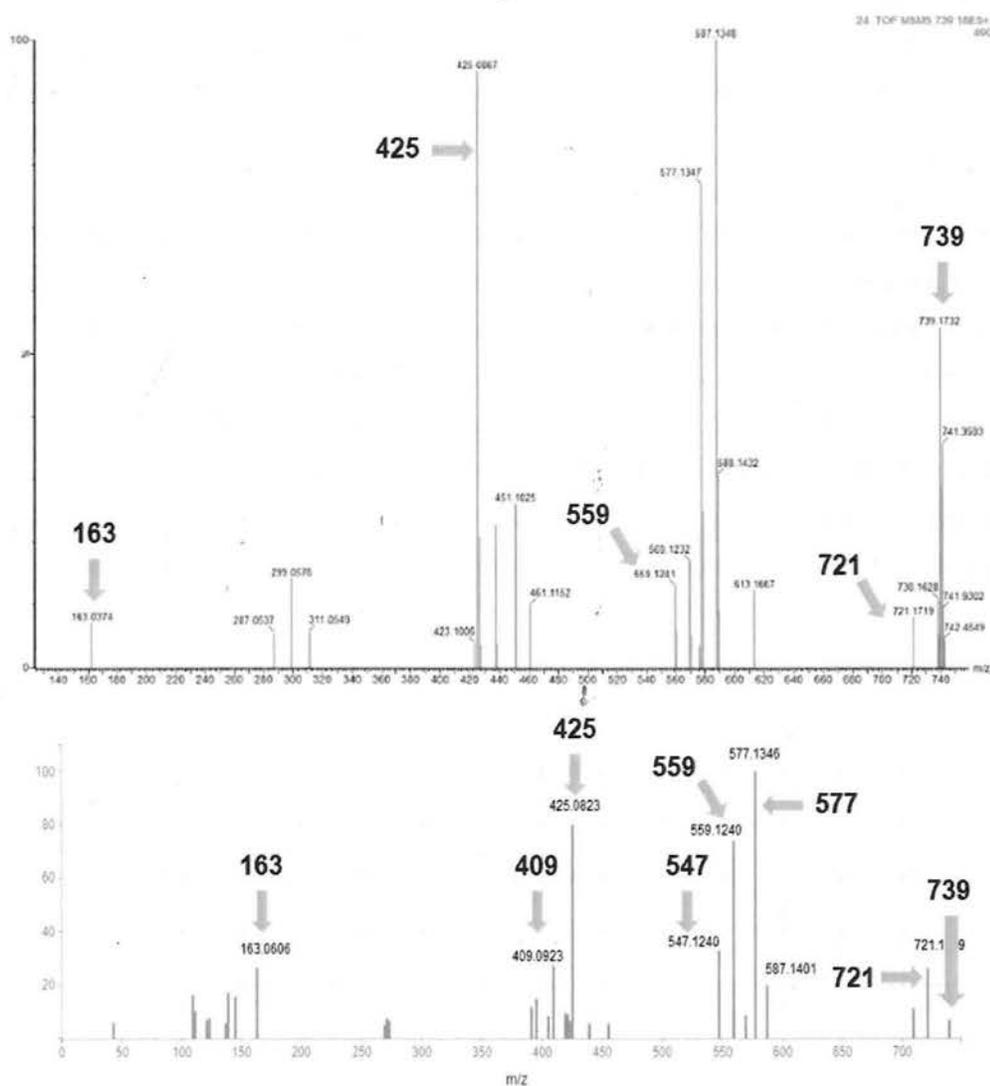


Figura 45. Espectro MS/MS de la procianidina B2, a 20 eV en modo positivo, reportado en la base MetLin.

Por otro lado, los iones  $m/z$  739 con TR 10,2 min y  $m/z$  709 con TR 10,3 min se anotaron tentativamente como antocianinas, específicamente glicósidos de dímeros de procianidinas (Cuadro XIX). Esto coincide con la literatura, ya que en cacao se han reportado compuestos de esta misma naturaleza, donde el primer ión ha sido planteado como un hexósido de un dímero procianidina A, mientras que el segundo con un

pentósido de un dímero de procianidina A (Cádiz-Gurrea et al., 2014; Calderón, Wright, Hurst, & van Breemen, 2009; D'Souza et al., 2017; Hatano et al., 2002; Porter, Ma, & Chan, 1991).

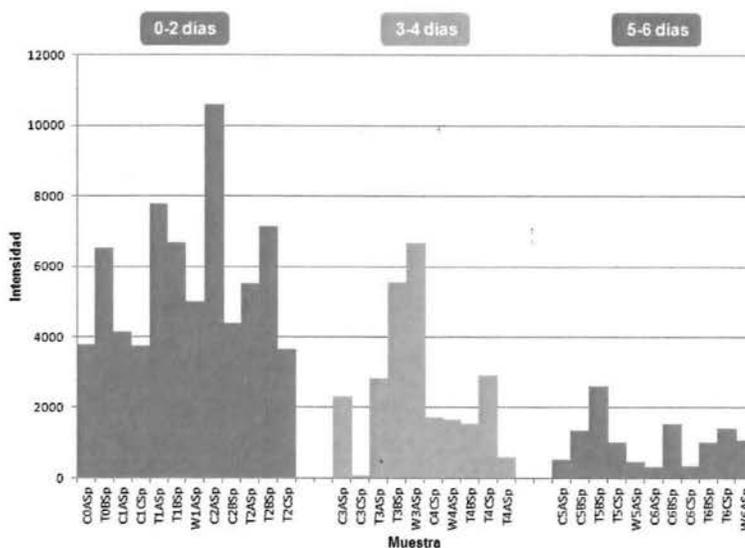
Para ambos iones se obtuvieron patrones de fragmentación similares. Por ejemplo, para el ión  $m/z$  739 la presencia de un pico en  $m/z$  577 en un espectro MS/MS sugiere la existencia de un dímero con enlace tipo A, mientras que el pico en  $m/z$  559 puede reflejar la pérdida de una hexosa (-180 Da). Otras señales evidencian la descomposición propia de una procianidina, descritas anteriormente por reacciones de RDA, FHA y QM (Figura 46).



**Figura 46.** Ión  $m/z$  739.1 con TR de 10.2 min en modo positivo y con una energía de colisión de 15 eV (arriba) y espectro predicho por la base de datos HMDB en modo positivo a 20 eV para 3-O- $\beta$ -D-galactopiranosil proantocianidina A5'.

En términos generales, los metabolitos identificados como flavanoles se redujeron en el tiempo de fermentación (ver ejemplo del comportamiento de un flavanol en la Figura 47), lo cual puede ser consecuencia de lixiviación durante los primeros días de fermentación, oxidación enzimática por polifenoloxidasas para producir quinonas, y la subsecuente polimerización de quinonas. Incluso se han reportado reducciones de entre 80 y 90 % de proantocianidinas y catequinas, posterior a la fermentación y el secado. Además las glicosidasas -estables a la fermentación- también intervienen en la formación de antocianidinas a partir de las antocianinas (E. O. Afoakwa et al., 2012; Biehl & Ziegleder, 2003a; de Brito et al., 2001; Hansen et al., 1998; Hue et al., 2014; Kim et al., 2011; Nazaruddin et al., 2006; Nigam & Singh, 2014; Payne, Hurst, Miller, Rank, & Stuart, 2010; Wollgast & Anklam, 2000).

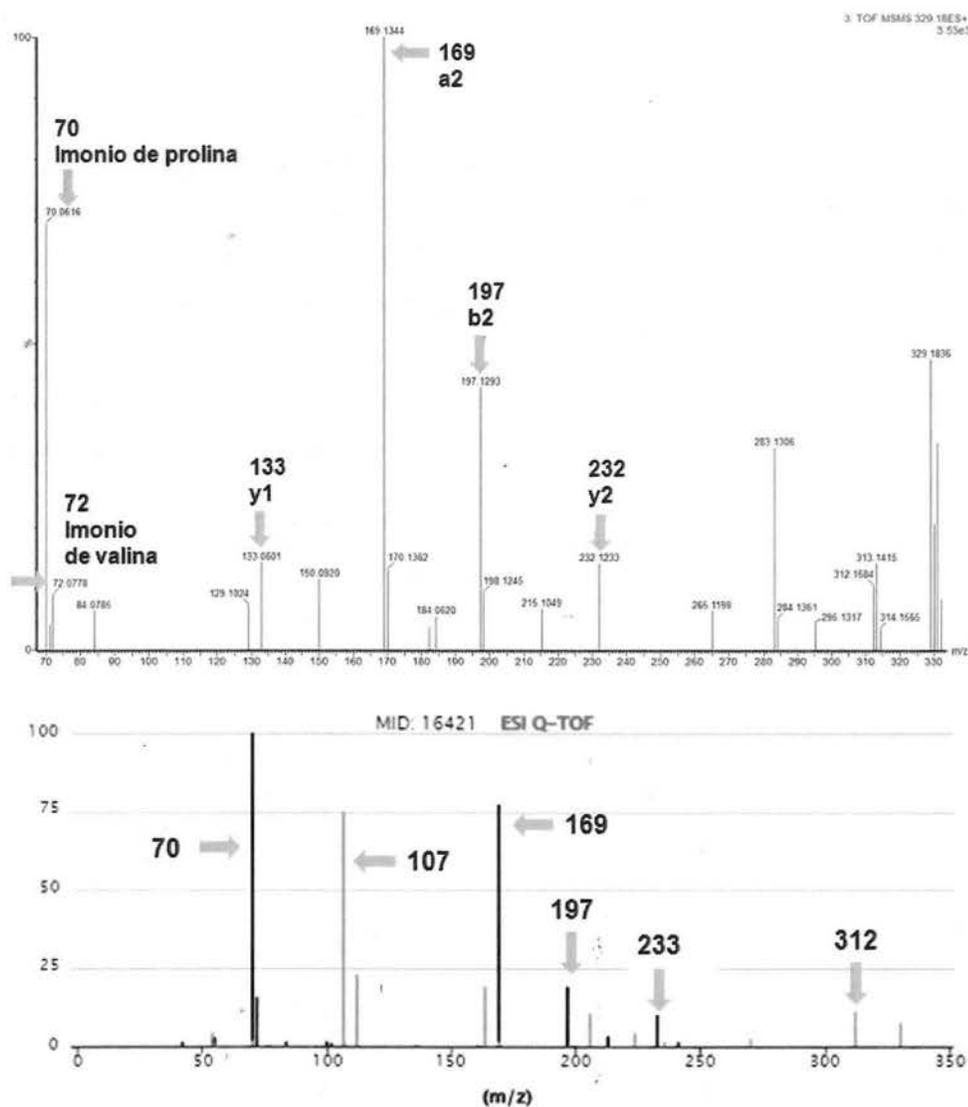
Estos resultados contrastan con los reportados en una investigación de D'Souza y colaboradores, en donde se observó un incremento del flavanol epicatequina en el tiempo de fermentación, así como de la capacidad antioxidante al comparar muestras fermentadas con no fermentadas. En el estudio se plantea que los polifenoles se ven incrementados durante la fermentación, como mecanismo de defensa de la semilla, y sugieren que las discrepancias encontradas en la literatura yacen en el hecho de que en este estudio se congelaron las muestras de cacao fresco inmediatamente después de cosechadas las mazorcas de cacao, reduciendo los niveles de estrés en las semillas (D'Souza et al., 2017).



**Figura 47.** Ejemplo del comportamiento en el tiempo de fermentación del ión identificado tentativamente como (epi) catequina.

Por otro lado, uno de los metabolitos con mayor poder discriminante, fue el oligopéptido anotado tentativamente como PVN con  $m/z$  329,1 y TR 0,8 min (Cuadro XX). Este mostró una tendencia a aumentar en el período de 3-4 días de fermentación, para posteriormente disminuir en la última etapa, es decir, 5-6 días de fermentación.

El análisis MS/MS permitió encontrar consistencias entre lo reportado en la literatura, identificándose rupturas tipo "a", "b" e "y", además de iones imonio de dos aminoácidos (Figura 48).



**Figura 48.** Espectros de masas, en modo positivo y a 15 eV, del ión anotado tentativamente como PVN (arriba), y espectro de masas experimental reportado por PVN en la base MetLin a 20 eV (abajo).

Estas rutas de fragmentación fueron consistentes en el resto de oligopéptidos identificados (Cuadro XX) así como con los mecanismos reportados previamente (Banerjee & Mazumdar, 2012; Roepstorff & Fohlman, 1984). Adicionalmente, se observó que a mayor número de residuos fue necesario emplear mayores energías de colisión para la obtención de los espectros MS/MS y que en términos generales estas fueron mayores a las requeridas durante los ensayos MS/MS de los flavonoides. Esto se explica con el hecho de que la energía de colisión tiene un efecto directo sobre la energía

vibracional de los enlaces y la subsecuente tasa de fragmentación de los oligopéptidos (Boyd & Somogyi, 2010; Wright et al., 2016).

Varios de los oligopéptidos han sido reportados previamente por otros autores y se han relacionado con secuencias específicas de proteínas del cacao, como se resume en el Cuadro XXI (US 2004/0202761 A1, 2004; Marseglia et al., 2014; Jürgen Voigt, Janek, Textoris-Taube, Niewienda, & Wöstemeyer, 2016).

**Cuadro XXI.** Relación de los oligopéptidos identificados de manera tentativa e hipótesis de su relación con las principales proteínas de cacao.

<b>Oligopéptidos identificados tentativamente en esta investigación</b>	<b>Proteína que los origina</b>
VSTDVN DNEW GAGGGGL NGTPVIF SSISGAGGGGLA SSISGAGGGG ANSPVLD VPIR	Albúmina
PVN SPGDV APLSPGDVF	Globulinas tipo vicilina

**Nota:** Ver nomenclatura de los aminoácidos en Figura A.3. de la sección de Anexos.

**Fuentes:**

En resumen, se identificaron dos patrones para los oligopéptidos en relación con su intensidad en el tiempo de fermentación: a) un aumento continuo en el tiempo o b) un aumento para alcanzar un pico a mitad de la fermentación (3-4 días) y su posterior reducción (Figura 49). En ambos casos se considera que la intervención de la enzima endoproteasa aspártica es estable durante la fermentación; mientras que la de la enzima carboxipeptidasas es parcialmente estable. La primera enzima provoca la proteólisis de

globulinas de tipo vicilina durante la fermentación, produciendo péptidos hidrofílicos, que son posteriormente hidrolizados por carboxipeptidasas para la producción de oligopéptidos hidrofóbicos y aminoácidos hidrofílicos. Estos últimos dos grupos han sido relacionados con la formación de aromas y sabores durante el tostado de cacao. Adicionalmente, las aminopeptidasas pueden tener un efecto pero en menor medida, puesto que se inactivan con la fermentación (Biehl & Ziegler, 2003a).

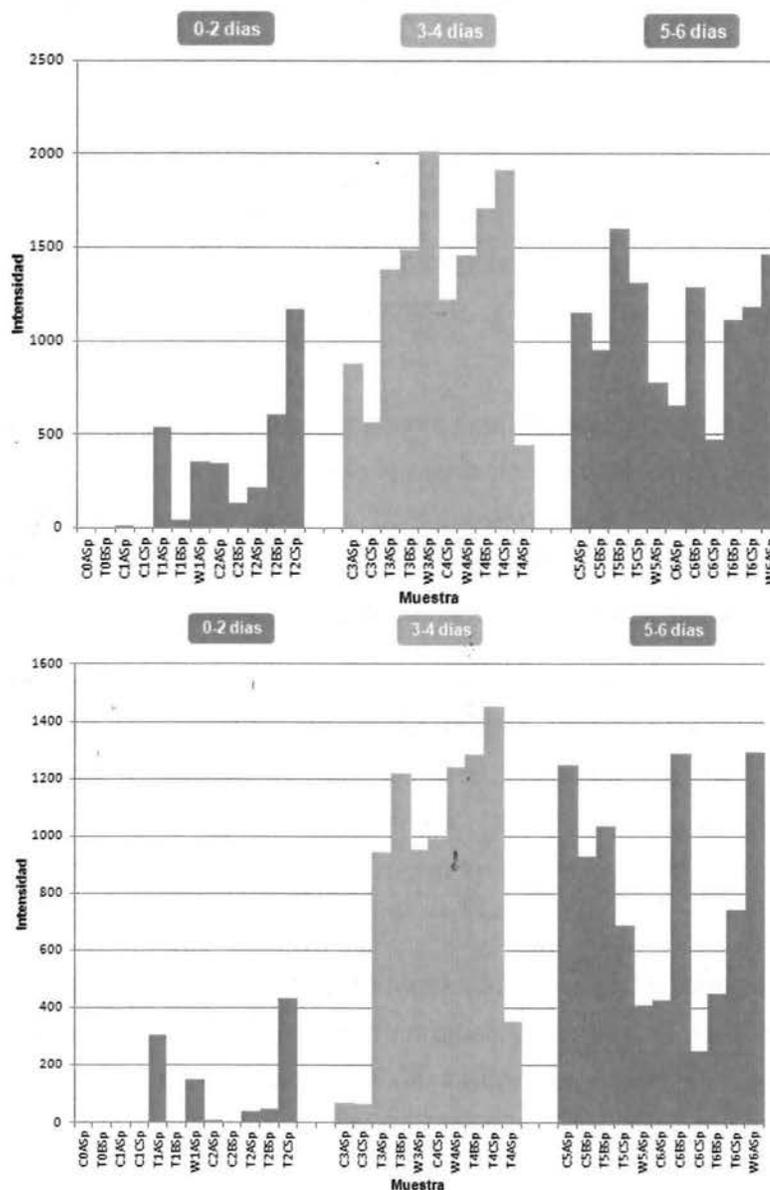


Figura 49. Evolución a lo largo del tiempo de fermentación de los metabolitos anotados tentativamente como PVN y VSTDVN catequina.

## 12. CONCLUSIONES

El presente estudio permite concluir que el empleo de metodologías metabolómicas para el análisis de la evolución de metabolitos durante la fermentación de cacao fue pertinente. A partir de una gran cantidad de información, se logró enfocar el análisis en aquellos metabolitos que experimentaron los mayores cambios en las tres etapas de la fermentación de relevancia biológica: 0-2 días, 3-4 días y 5-6 días.

Se identificaron tentativamente 20 metabolitos, clasificados en tres clases de compuestos: azúcares, flavonoides y oligopéptidos. Las tendencias que presentaron estos metabolitos, acontecieron de forma coherente con los fenómenos biológicos ocurrientes en las semillas de cacao, en las diferentes etapas de la fermentación. Muchos de estos compuestos han sido previamente reportados en cacao, y algunos de estos, a pesar de tener alto poder de discriminación, permanecieron con identidad desconocida y requieren mayor investigación.

Los azúcares y los flavonoides, principalmente (epi) catequina, dímeros y trímeros de procianidinas, se redujeron a lo largo de la fermentación, lo cual fue coherente con la gran mayoría de estudios, excepto uno. En contraposición, los oligopéptidos de 3 a 12 aminoácidos mostraron dos tendencias: un incremento continuo en el tiempo, o un incremento con un máximo a los 3-4 días de fermentación, para posteriormente disminuir.

Metabolitos asociados a las tres distintas clases de compuestos (azúcares, flavonoides y oligopéptidos) pueden proponerse como potenciales biomarcadores del proceso de fermentación. Estos pueden tener eventualmente importancia a nivel nutricional (ej. flavanoles) y sensorial (ej. aminoácidos precursores de aromas, sabores y colores del cacao). No obstante, es necesario realizar mayor investigación, que incluya cuantificación de compuestos y análisis sensorial.

Se pueden proponer nuevos protocolos de fermentación de cacao, sopesando no solo la producción de oligopéptidos precursores de aromas y sabores, si no la degradación de compuestos de interés nutricional como los flavonoides bioactivos; no obstante, es necesario realizar mayor investigación en esta línea. Se considera que los hallazgos de este trabajo pueden ser empleados como base para futuros estudios de caracterización específica.

### 13. RECOMENDACIONES

- Realizar un análisis químico proximal en conjunto con el análisis metabolómico con la finalidad de validar ciertos cambios en grupos de compuestos, así como cuantificación de clases de compuestos como por ejemplo polifenoles totales.
- Optimizar las condiciones de cromatografía y espectrometría de masas para obtener mejores resoluciones y facilitar la detección de iones y su identificación.
- Analizar otras fracciones de compuestos de carácter hidrofílico e hidrofóbico para evaluar otros compuestos que puedan tener relevancia biológica.
- Ampliar la ventana de observación de los iones en el espectrómetro de masas hasta 2000 Da para observar más compuestos como diversos polímeros de (epi) catequina, relevantes en cacao.
- Emplear estándares internos para valorar el efecto de los procesos de extracción, su estabilidad durante el almacenamiento y tiempos de espera durante la medición.
- Medir réplicas de muestras para evaluar la repetibilidad.
- Realizar recuentos e identificación de microorganismos (ej. PCR) con la finalidad de validar las buenas prácticas de fermentación y describir con mayor profundidad las distintas etapas que ocurren en el tiempo, y su relación con la microbiota.
- Estandarizar el pH de las muestras y evaluar su efecto sobre su separación cromatográfica y detección en el espectrómetro de masas.
- Realizar una investigación metabolómica en la testa de las semillas de cacao, con la finalidad de investigar los distintos procesos de transferencia de masa durante la fermentación.
- Analizar el efecto de la sobrefermentación de muestras, con la finalidad de detectar posibles marcadores biológicos de malas prácticas de fermentación.

## 14. BIBLIOGRAFÍA

- Afoakwa, E. (2014). *Cocoa Production and Processing Technology*. CRC Press.
- Afoakwa, E. O. (2014). *Cocoa production and processing technology*. CRC Press, Taylor & Francis.
- Afoakwa, E. O., Quao, J., Takrama, F. S., Budu, a. S., & Saalia, F. K. (2012). Changes in total polyphenols, o-diphenols and anthocyanin concentrations during fermentation of pulp pre-conditioned cocoa (*Theobroma cacao*) beans. *International Food Research Journal*, 19(3), 1071–1077.
- Afoakwa, E. O., Quao, J., Takrama, J., Budu, A. S., & Saalia, F. K. (2013). Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. *Journal of Food Science and Technology*, 50(6), 1097–1105.
- Allwood, J. W., & Goodacre, R. (2010). An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical Analysis*, 21(1), 33–47.
- Amin, I., Jinap, S., & Jamilah, B. (1997). Vicilin-class globulins and their degradation during cocoa fermentation. *Food Chemistry*, 59(1), 1–5.
- Anvoh, K. Y. B., Bi, A. Z., & Gnakri, D. (2009). Production and characterization of juice from mucilage of cocoa beans and its transformation into marmalade. *Pakistan Journal of Nutrition*, 8(2), 129–133.
- Aprotosoie, A. C., Luca, S. V., & Miron, A. (2016). Flavor chemistry of cocoa and cocoa products— an overview. *Comprehensive Reviews in Food Science and Food Safety*, 15(1), 73–91.
- Ardhana, M., & Fleet, G. (2003). The microbial ecology of cocoa bean fermentations in Indonesia. *International Journal of Food Microbiology*, 86(1–2), 87–99.
- Ardrey, R. E. (2003). *Liquid chromatography - Mass spectrometry: An Introduction*. West Sussex: Wiley.
- Aremu, C. Y., Agiang, M. A., & Ayatse, J. O. I. (1995). Nutrient and antinutrient profiles of raw and fermented cocoa beans. *Plant Foods for Human Nutrition*, 48(3), 217–223.
- Balzer, J., Rassaf, T., Heiss, C., Kleinbongard, P., Lauer, T., Merx, M., ... Kelm, M. (2008). Sustained benefits in vascular function through flavanol-containing cocoa in medicated diabetic patients a double-masked, randomized, controlled trial. *Journal of the American College of Cardiology*, 51(22), 2141–2149.
- Bartley, B. G. D. (Ed.). (2005). *The genetic diversity of cacao and its utilization*. Wallingford: CABI.
- Banerjee, S., & Mazumdar, S. (2012). Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International Journal of Analytical Chemistry*, 2012, 1–40.
- Belščak, A., Komes, D., Horžić, D., Ganić, K. K., & Karlović, D. (2009). Comparative study of commercially available cocoa products in terms of their bioactive composition. *Food Research International*, 42(5–6), 707–716.
- Biehl, B., & Ziegler, G. (2003a). COCOA | Chemistry of processing. In *Encyclopedia of Food Sciences and Nutrition* (Second edi, pp. 1436–1448). Elsevier.
- Biehl, B., & Ziegler, G. (2003b). COCOA | Production, products and use. In *Encyclopedia of Food Sciences and Nutrition* (Second edi, pp. 1448–1463). Elsevier.
- Biomarkers Definitions Working Group. (2001). Biomarkers and surrogate endpoints: preferred

- definitions and conceptual framework. *Clinical Pharmacology & Therapeutics*, 69(3), 89–95.
- Black, G. E., & Fox, A. (1996). Liquid chromatography with electrospray ionization tandem mass spectrometry (pp. 81–105).
- Boyd, R., & Somogyi, Á. (2010). The mobile proton hypothesis in fragmentation of protonated peptides: A perspective. *Journal of the American Society for Mass Spectrometry*, 21(8), 1275–1278.
- Bytof, G., Biehl, B., Heinrichs, H., & Voigt, J. (1995). Specificity and stability of the carboxypeptidase activity in ripe, ungerminated seeds of *Theobroma cacao* L. *Food Chemistry*, 54(1), 15–21.
- Cádiz-Gurrea, M. L., Lozano-Sanchez, J., Contreras-Gámez, M., Legeai-Mallet, L., Fernández-Arroyo, S., & Segura-Carretero, A. (2014). Isolation, comprehensive characterization and antioxidant activities of *Theobroma cacao* extract. *Journal of Functional Foods*, 10, 485–498.
- Calderón, A. I., Wright, B. J., Hurst, W. J., & van Breemen, R. B. (2009). Screening antioxidants using LC-MS: case study with cocoa. *Journal of Agricultural and Food Chemistry*, 57(13), 5693–5699.
- CAOBISCO, ECA, & FCC. (2015). *Cocoa beans: chocolate and cocoa industry quality requirements*. (M. J. End & R. Dand, Eds.). Retrieved from [http://www.cocoaquality.eu/data/Cocoa Beans Industry Quality Requirements Apr 2016\\_En.pdf](http://www.cocoaquality.eu/data/Cocoa%20Beans%20Industry%20Quality%20Requirements%20Apr%202016_En.pdf)
- Castillo, S., Gopalacharyulu, P., Yetukuri, L., & Orešič, M. (2011). Algorithms and tools for the preprocessing of LC–MS metabolomics data. *Chemometrics and Intelligent Laboratory Systems*, 108(1), 23–32.
- Cheesman, E. E. (1944). Notes on the nomenclature, classification and possible relationships of cacao populations. *Trop. Agr. (Trinidad)*, (21), 144–159.
- Churchwell, M. I., Twaddle, N. C., Meeker, L. R., & Doerge, D. R. (2005). Improving LC–MS sensitivity through increases in chromatographic performance: Comparisons of UPLC–ES/MS/MS to HPLC–ES/MS/MS. *Journal of Chromatography B*, 825(2), 134–143.
- Cifuentes, A. (2009). Food analysis and Foodomics. *Journal of Chromatography A*, 1216 (43)m 7109.
- Clusius, C. (1605). Cacao Fructus. In *Exoticorum libri decem, quibus animalium, plantarum, aromatum, aliorumque peregrinorum fructuum historiae describuntur* (pp. 55, 56). Item Petri Bellonii, eodem Carolo Clusio interprete.
- Copetti, M. V., Iamanaka, B. T., Nester, M. A., Efraim, P., & Taniwaki, M. H. (2013). Occurrence of ochratoxin A in cocoa by-products and determination of its reduction during chocolate manufacture. *Food Chemistry*, 136(1), 100–104.
- Corti, R., Flammer, A. J., Hollenberg, N. K., & Lüscher, T. F. (2009). Cocoa and cardiovascular health. *Circulation*, 119(10), 1433–1441.
- Cuatrecasas, J. (1964). Cacao and its allies, a taxonomic revision of the genus *Theobroma*. *Contributions from the United States National Herbarium*. Smithsonian Institution.
- D'Souza, R. N., Grimbs, S., Behrends, B., Bernaert, H., Ullrich, M. S., & Kuhnert, N. (2017). Origin-based polyphenolic fingerprinting of *Theobroma cacao* in unfermented and fermented beans. *Food Research International*, 99, 550–559.
- Dand, R. (2011a). 1 - History and origins of the international cocoa trade. In *The International Cocoa Trade* (pp. 1–22). Elsevier.

- Dand, R. (2011b). 2 - Agronomics of international cocoa production. In *The International Cocoa Trade* (pp. 23–64). Elsevier.
- Dand, R. (2011c). 3 - Environmental and practical factors affecting cocoa production. In *The International Cocoa Trade* (pp. 65–93). Elsevier.
- De Brito, E. S., García, N. H. P., Gallão, M., Cortelazzo, A. L., Fevereiro, P. S., & Braga, M. R. (2001). Structural and chemical changes in cocoa (*Theobroma cacao* L.) during fermentation, drying and roasting. *Journal of the Science of Food and Agriculture*, 81(2), 281–288.
- De Oliveira, T. B., Rogero, M. M., & Genovese, M. I. (2015). Polyphenolic-rich extracts from cocoa (*Theobroma cacao* L.) and cupuassu (*Theobroma grandiflorum* Willd. Ex Spreng. K. Shum) liquors: A comparison of metabolic effects in high-fat fed rats. *PharmaNutrition*, 3(2), 20–28.
- de Souza, P. A., Moreira, L. F., Sarmiento, D. H. A., & da Costa, F. B. (2018). Cacao—*Theobroma cacao*. *Exotic Fruits*, 69–76.
- Demarque, D. P., Crotti, A. E. M., Vessecchi, R., Lopes, J. L. C., & Lopes, N. P. (2016). Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Natural Product Reports*, 33(3), 432–455.
- Dias, D. R., Schwan, R. F., Freire, E. S., & Serôdio, R. dos S. (2007). Elaboration of a fruit wine from cocoa (*Theobroma cacao* L.) pulp. *International Journal of Food Science & Technology*, 42(3), 319–329.
- do Carmo Brito, B. de N., Campos Chisté, R., da Silva Pena, R., Abreu Gloria, M. B., & Santos Lopes, A. (2017). Bioactive amines and phenolic compounds in cocoa beans are affected by fermentation. *Food Chemistry*, 228, 484–490.
- Dorenkott, M. R., Griffin, L. E., Goodrich, K. M., Thompson-Witrick, K. A., Fundaro, G., Ye, L., ... Neilson, A. P. (2014). Oligomeric cocoa procyanidins possess enhanced bioactivity compared to monomeric and polymeric cocoa procyanidins for preventing the development of obesity, insulin resistance, and impaired glucose tolerance during high-fat feeding. *Journal of Agricultural and Food Chemistry*, 62(10), 2216–2227.
- Dunn, W. B., Broadhurst, D. I., Edison, A., Guillou, C., Viant, M. R., Bearden, D. W., & Beger, R. D. (2017). Quality assurance and quality control processes: summary of a metabolomics community questionnaire. *Metabolomics*, 13(5), 50.
- Dunn, W. B., & Ellis, D. I. (2005). Metabolomics: Current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry*, 24(4), 285–294.
- EFSA. (2014). *Scientific Opinion on the modification of the authorisation of a health claim related to cocoa flavanols and maintenance of normal endothelium-dependent vasodilation pursuant to Article 13 (5) of Regulation (EC) No 1924/2006 following a request in accordance with Article 14.4 of Regulation (EC) No 1825/2003*. Parma, Italy.
- Esatbeyoglu, T., Wray, V., & Winterhalter, P. (2015). Isolation of dimeric, trimeric, tetrameric and pentameric procyanidins from unroasted cocoa beans (*Theobroma cacao* L.) using countercurrent chromatography. *Food Chemistry*, 179, 278–289.
- Fiehn, O. (2001). Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics*, 2(3), 155–168.
- Fiehn, O., Kloska, S., & Altmann, T. (2001). Integrated studies on plant biology using multiparallel techniques. *Current Opinion in Biotechnology*, 12(1), 82–86.
- Flamini, R. (2013). Recent applications of mass spectrometry in the study of grape and wine polyphenols. *ISRN Spectroscopy*, 2013, 1–45.
- Flores-González, A. (2015). *Producción nacional de cacao*. San José, Costa Rica.

- Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of clinical trials, 5th ed. Journal of Chemical Information and Modeling* (5th ed., Vol. 53). Springer International Publishing.
- Friedrich, W., Eberhardt, A., & Galensa, R. (2000). Investigation of proanthocyanidins by HPLC with electrospray ionization mass spectrometry. *European Food Research and Technology*, 211(1), 56–64.
- G.A.R., Wood, & Lass, R. A. (1985). *Cocoa*. London: Longman.
- Glauser, G., Veyrat, N., Rochat, B., Wolfender, J.-L., & Turlings, T. C. J. (2013). Ultra-high pressure liquid chromatography–mass spectrometry for plant metabolomics: A systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. *Journal of Chromatography A*, 1292, 151–159.
- Glish, G. L., & Vachet, R. W. (2003). The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery*, 2(2), 140–150.
- Goodacre, R., Broadhurst, D., Smilde, A. K., Kristal, B. S., Baker, J. D., Beger, R., ... Wulfert, F. (2007). Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics*, 3(3), 231–241.
- Grimes, D. A., & Schulz, K. F. (2002). An overview of clinical research: the lay of the land. *Lancet (London, England)*, 359(9300), 57–61.
- Gross, M. (2014). Coffee and chocolate in danger. *Current Biology*.
- Großmann-Kühnau, S. (2015). Dem Kakaoaroma auf der Spur. *Internationale Literatur*, 395–397.
- Gu, L., Kelm, M. A., Hammerstone, J. F., Zhang, Z., Beecher, G., Holden, J., ... Prior, R. L. (2003). Liquid chromatographic/electrospray ionization mass spectrometric studies of proanthocyanidins in foods. *Journal of Mass Spectrometry: JMS*, 38(12), 1272–1280.
- Guillarme, D., & Veuthey, J.-L. (Eds.). (2012). *UHPLC in Life Sciences*. Cambridge: Royal Society of Chemistry.
- Guillarme, D., & Veuthey, J.-L. (2017). Theory and practice of UHPLC and UHPLC–MS. In *Handbook of Advanced Chromatography /mass Spectrometry Techniques* (pp. 1–38). Elsevier.
- Gutiérrez, O. A., Campbell, A. S., & Phillips-Mora, W. (2016). Breeding for disease resistance in cacao. In *Cacao Diseases* (pp. 567–609). Cham: Springer International Publishing.
- Gutiérrez, T. J. (2017). State-of-the-art chocolate manufacture: a review. *Comprehensive Reviews in Food Science and Food Safety*, 16(6), 1313–1344.
- Hansen, C. E., Mañez, A., Burri, C., & Bousbaine, A. (2000). Comparison of enzyme activities involved in flavour precursor formation in unfermented beans of different cocoa genotypes. *Journal of the Science of Food and Agriculture*, 80(8), 1193–1198.
- Hansen, C. E., Olmo, M. del, & Burri, C. (1998). Enzyme activities in cocoa beans during fermentation. *Journal of the Science of Food and Agriculture*, 77(2), 273–281.
- Hashimoto, J. C., Lima, J. C., Celeghini, R. M. S., Nogueira, A. B., Efraim, P., Poppi, R. J., & Pallone, J. A. L. (2018). Quality control of commercial cocoa beans (*Theobroma cacao* L.) by Near-infrared Spectroscopy. *Food Analytical Methods*, 1–8.
- Hatano, T., Miyatake, H., Natsume, M., Osakabe, N., Takizawa, T., Ito, H., & Yoshida, T. (2002). Proanthocyanidin glycosides and related polyphenols from cacao liquor and their antioxidant effects. *Phytochemistry*, 59, 749–758.
- Heftmann, E. (2004). *Chromatography: fundamentals and applications of chromatography and*

*related differential migration methods. Part A, Fundamentals and techniques.* Elsevier.

- Hii, C. L., Law, C. ., Suzannah, S., Misnawi, & Cloke, M. (2010). Polyphenols in cocoa (*Theobroma cacao* L.). *Asian Journal of Food and Agro-Industry*, 2(4), 702–722.
- Hii, C. L., Law, C. L., & Suzannah, S. (2012). Drying kinetics of the individual layer of cocoa beans during heat pump drying. *Journal of Food Engineering*, 108(2), 276–282.
- Hue, C., Brat, P., Gunata, Z., Samaniego, I., Servent, A., Morel, G., ... Davrieux, F. (2014). Near infra-red characterization of changes in flavan-3-ol derivatives in cocoa (*Theobroma cacao* L.) as a function of fermentation temperature. *Journal of Agricultural and Food Chemistry*, 62(41), 10136–10142.
- Hyötyläinen, T., & Wiedmer, S. (Eds.). (2013). *Chromatographic methods in metabolomics*. Cambridge, United Kingdom: Royal Society of Chemistry.
- ICCO. (2016a). Fine or flavour cocoa. Retrieved from <https://www.icco.org/about-cocoa/fine-or-flavour-cocoa.html>
- ICCO. (2016b). *Quarterly bulletin of cocoa statistics - august 2016. Quarterly Bulletin of Cocoa Statistics*. Retrieved from <https://www.icco.org/about-us/icco-news/324-quarterly-bulletin-of-cocoa-statistics-august-2016.html>
- Igbinadolor, R. O. (2009). Other tropical fruit vinegars. In *Vinegars of the World* (pp. 261–271). Milano: Springer Milan.
- Ioannone, F., Di Mattia, C. D., De Gregorio, M., Sergi, M., Serafini, M., & Sacchetti, G. (2015). Flavanols, proanthocyanidins and antioxidant activity changes during cocoa (*Theobroma cacao* L.) roasting as affected by temperature and time of processing. *Food Chemistry*, 174, 256–262.
- ITIS. (n.d.). *Theobroma cacao* L. 12/30/2017: Integrated Taxonomic Information System (ITIS). Retrieved from [https://www.itis.gov/servlet/SingleRpt/SingleRpt?search\\_topic=TSN&search\\_value=505487#null](https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=505487#null)
- Janek, K., Niewianda, A., Wöstemeyer, J., & Voigt, J. (2016). The cleavage specificity of the aspartic protease of cocoa beans involved in the generation of the cocoa-specific aroma precursors. *Food Chemistry*, 211, 320–328.
- John, W. A., Kumari, N., Böttcher, N. L., Koffi, K. J., Grimbs, S., Vrancken, G., ... Ullrich, M. S. (2016). Aseptic artificial fermentation of cocoa beans can be fashioned to replicate the peptide profile of commercial cocoa bean fermentations. *Food Research International*, 89, 764–772.
- Kadow, D., Niemenak, N., Rohn, S., & Lieberei, R. (2015). Fermentation-like incubation of cocoa seeds (*Theobroma cacao* L.) – Reconstruction and guidance of the fermentation process. *LWT - Food Science and Technology*, 62(1), 357–361.
- Kebarle, P., & Peschke, M. (2000). On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions. *Analytica Chimica Acta*, 406(1), 11–35.
- Kebarle, P., & Tang, L. (1993). From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry. *Analytical Chemistry*, 65(22), 972A–986A.
- Kebarle, P., & Verkerk, U. H. (2012). On the mechanism of electrospray ionization mass spectrometry (ESIMS). In *Electrospray and MALDI Mass Spectrometry* (pp. 1–48). Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Keski-Rahkonen, P., Rothwell, J. A., & Scalbert, A. (2017). Metabolomic techniques to discover food biomarkers. In D. A. Schoeller & M. S. Westerterp-Plantenga (Eds.), *Advances in the assessment of dietary intake*. Boca Raton, Florida: CRC Press, Taylor & Francis.

- Kim, J., Lee, K. W., & Lee, H. J. (2011). Cocoa (*Theobroma cacao*) seeds and phytochemicals in human health. In R. V. Preedy, R. R. Watson, & B. V. Patel (Eds.), *Nuts and Seeds in Health and Disease Prevention* (pp. 351–360). Elsevier.
- Kochhar, S., Hansen, C.-E., Juilleart, A. M., Wille, H.-J., Buyukpamukcu, E., Keely, B., & Goodall, D. M. (2004). *US 2004/0202761 A1*. United States.
- Kongor, J. E., Hinneh, M., de Walle, D. Van, Afoakwa, E. O., Boeckx, P., & Dewettinck, K. (2016). Factors influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile - a review. *Food Research International*, 82, 44–52.
- Kratzer, U., Frank, R., Kalbacher, H., Biehl, B., Wöstemeyer, J., & Voigt, J. (2009). Subunit structure of the vicilin-like globular storage protein of cocoa seeds and the origin of cocoa- and chocolate-specific aroma precursors. *Food Chemistry*, 113(4), 903–913.
- Lachenaud, P., & Motamayor, J. C. (2017). The Criollo cacao tree (*Theobroma cacao* L.): a review. *Genetic Resources and Crop Evolution*, 64(8), 1807–1820.
- Lachenaud, P., Paulin, D., Ducamp, M., & Thevenin, J.-M. (2007). Twenty years of agronomic evaluation of wild cocoa trees (*Theobroma cacao* L.) from French Guiana. *Scientia Horticulturae*, 113(4), 313–321.
- Läderach, P., Martinez-Valle, A., Schroth, G., & Castro, N. (2013). Predicting the future climatic suitability for cocoa farming of the world's leading producer countries, Ghana and Côte d'Ivoire. *Climatic Change*, 119(3), 841–854.
- Larive, C. K., Barding, G. A., & Dinges, M. M. (2015). NMR spectroscopy for metabolomics and metabolic profiling. *Analytical Chemistry*, 87(1), 133–146.
- Lieberei, R., Kadow, D., & Seigler, D. (2013). Cocoa Cultivation, Directed Breeding and Polyphenolics. In *Natural Products* (pp. 1599–1616). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Linnaeus, C. (1737). *Genera Plantarum*.
- Llorach, R., Medina, S., García-Viguera, C., Zafrilla, P., Abellán, J., Jauregui, O., ... Andrés-Lacueva, C. (2014). Discovery of human urinary biomarkers of aronia-citrus juice intake by HPLC-q-TOF-based metabolomic approach. *ELECTROPHORESIS*, 35(11), 1599–1606.
- Lockwood, G., & Gyamfi, M. M. O. (Cocoa R. I. T. (Ghana)). (1979). The CRIG [Cocoa Research Institute of Ghana] cocoa [*Theobroma cacao*] germplasm collection with notes on codes used in the breeding programme at Tafo [Ghana] and elsewhere. *CRIG Technical Bulletin (Ghana)*. Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=GH8000049>
- Machin, D., Campbell, M. J., & Walters, S. J. (2001). *Medical Statistics* (4th ed.). West Sussex: John Wiley & Sons, Ltd.
- Manach, C., Scalbert, A., Morand, C., Rémésy, C., & Jiménez, L. (2004). Polyphenols: food sources and bioavailability. *The American Journal of Clinical Nutrition*, 79(5), 727–747.
- Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34–40.
- Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G., & Caligiani, A. (2014). Extraction, identification and semi-quantification of oligopeptides in cocoa beans. *Food Research International*, 63, 382–389.
- Martínez, R., Torres, P., Meneses, M. A., Figueroa, J. G., Pérez-Álvarez, J. A., & Viuda-Martos, M. (2012). Chemical, technological and in vitro antioxidant properties of cocoa (*Theobroma cacao* L.) co-products. *Food Research International*, 49(1), 39–45.

- Martini, M. H., Figueira, A., Lenci, C. G., & Tavares, D. de Q. (2008). Polyphenolic cells and their interrelation with cotyledon cells in seven species of *Theobroma* (Sterculiaceae). *Revista Brasileira de Botânica*, 31(3), 425–431.
- Meyer, V. R. (2010). *Practical high-performance liquid chromatography* (4th ed.). Saint Gallen: John Wiley & Sons.
- Morris, D. (1882). *Cacao: how to grow and how to cure it*. Jamaica.
- Moss, G. P., Smith, P. A. S., & Tavernier, D. (1995). Glossary of class names of organic compounds and reactivity intermediates based on structure (IUPAC Recommendations 1995). *Pure and Applied Chemistry*, 67(8–9), 1307–1375.
- Motamayor, J. C., Risterucci, A. M., Lopez, P. A., Ortiz, C. F., Moreno, A., & Lanaud, C. (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity*, 89(5), 380–386.
- Motamayor, J. C. (2001). *Etude de la diversité génétique et de la domestication des cacaoyers du groupe Criollo (Theobroma cacao L.) à l'aide de marqueurs moléculaires*. Thèse de doctorat: Sciences de la vie. Université de Paris-Sud.
- Motamayor, J. C., Lachenaud, P., da Silva e Mota, J. W., Loor, R., Kuhn, D. N., Brown, J. S., & Schnell, R. J. (2008). Geographic and genetic population differentiation of the amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE*, 3(10), e3311.
- Mozzi, F., Ortiz, M. E., Bleckwedel, J., De Vuyst, L., & Pescuma, M. (2013). Metabolomics as a tool for the comprehensive understanding of fermented and functional foods with lactic acid bacteria. *Food Research International*, 54(1), 1152–1161.
- Nair, K. P. P., & Nair, K. P. P. (2010). 5 – Cocoa (*Theobroma cacao* L.). In *The Agronomy and Economy of Important Tree Crops of the Developing World* (pp. 131–180).
- Naushad, M., & Khan, M. R. (2014). *Ultra performance liquid chromatography mass spectrometry: evaluation and applications in food analysis*. CRC Press.
- Navarro, M. (2016). Organización de las cadenas de valor de cacao, Centroamérica. In *II Foro Centroamericano de cacao*. Honduras. Retrieved from [http://forocacao.com/media/activity/S1.PM1\\_Melba\\_Navarro.pdf](http://forocacao.com/media/activity/S1.PM1_Melba_Navarro.pdf)
- Nazaruddin, R., Seng, L. K., Hassan, O., & Said, M. (2006). Effect of pulp preconditioning on the content of polyphenols in cocoa beans (*Theobroma cacao*) during fermentation. *Industrial Crops and Products*, 24(1), 87–94.
- Nielsen, D. S., Crafac, M., Jespersen, L., & Jakobsen, M. (2013). The microbiology of cocoa fermentation. In *Chocolate in Health and Nutrition* (pp. 39–60). Totowa, NJ: Humana Press.
- Nielsen, S. S. (2010). *Food Analysis* (4th ed.). New York: Springer.
- Niemenak, N., Rohsius, C., Elwers, S., Omokolo Ndoumou, D., & Lieberei, R. (2006). Comparative study of different cocoa (*Theobroma cacao* L.) clones in terms of their phenolics and anthocyanins contents. *Journal of Food Composition and Analysis*, 19(6–7), 612–619.
- Niessen, W. M. A. (Wilfried M. A. . (2006). *Liquid chromatography-- mass spectrometry*. CRC/Taylor & Francis. Retrieved from [https://books.google.co.cr/books?id=XHGoVSeW\\_rYC&dq=liquid+chromatography&source=gs\\_navlinks\\_s](https://books.google.co.cr/books?id=XHGoVSeW_rYC&dq=liquid+chromatography&source=gs_navlinks_s)
- Nigam, P. S., & Singh, A. (2014). Cocoa and coffee fermentations. In *Encyclopedia of Food Microbiology* (pp. 485–492). Elsevier.
- Nordström, A., O'Maille, G., Qin, C., & Siuzdak, G. (2006). Nonlinear data alignment for UPLC-MS

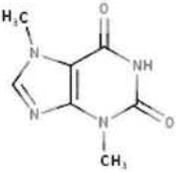
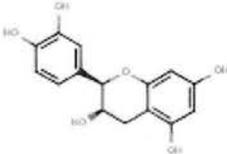
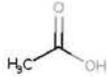
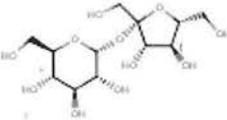
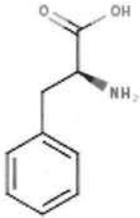
- and HPLC-MS based metabolomics: quantitative analysis of endogenous and exogenous metabolites in human serum. *Analytical Chemistry*, 78(10), 3289–3295.
- Okiyama, D. C. G., Navarro, S. L. B., & Rodrigues, C. E. C. (2017). Cocoa shell and its compounds: applications in the food industry. *Trends in Food Science & Technology*, 63, 103–112.
- Oliveira, L. de L. de, Carvalho, M. V. de, & Melo, L. (2014). Health promoting and sensory properties of phenolic compounds in food. *Revista Ceres*, 61(suppl), 764–779.
- Orozco, Aguilar, L., Deheuvels, O., Villalobos, Rodríguez, M., & Somarriba, Chávez, E. (2015). *El sector cacao en Centroamérica - Estado de desarrollo en el año 2007*. Turrialba, Costa Rica.
- Ortega, N., Romero, M.-P., Macià, A., Reguant, J., Anglès, N., Morelló, J.-R., & Motilva, M.-J. (2008). Obtention and characterization of phenolic extracts from different cocoa sources. *Journal of Agricultural and Food Chemistry*, 56(20), 9621–9627.
- Ortega, N., Romero, M.-P., Macià, A., Reguant, J., Anglès, N., Morelló, J. R., & Motilva, M. J. (2010). Comparative study of UPLC–MS/MS and HPLC–MS/MS to determine procyanidins and alkaloids in cocoa samples. *Journal of Food Composition and Analysis*, 23(3), 298–305.
- Ozturk, G., & Young, G. M. (2017). Food Evolution: The impact of society and science on the fermentation of cocoa beans. *Comprehensive Reviews in Food Science and Food Safety*, 16(3), 431–455. <https://doi.org/10.1111/1541-4337.12264>
- Papalexandratou, Z., Vrancken, G., De Bruyne, K., Vandamme, P., & De Vuyst, L. (2011). Spontaneous organic cocoa bean box fermentations in Brazil are characterized by a restricted species diversity of lactic acid bacteria and acetic acid bacteria. *Food Microbiology*, 28(7), 1326–1338.
- Payne, M. J., Hurst, W. J., Miller, K. B., Rank, C., & Stuart, D. A. (2010). Impact of fermentation, drying, roasting, and dutch processing on epicatechin and catechin content of cacao beans and cocoa ingredients. *Journal of Agricultural and Food Chemistry*, 58 (19), 10518-10527.
- Pereira, G. V. de M., Soccol, V. T., & Soccol, C. R. (2016). Current state of research on cocoa and coffee fermentations. *Current Opinion in Food Science*, 7, 50–57.
- Phillips-Mora, W., Arciniegas-Leal, A., Mata-Quirós, A., & Motamayor-Arias, J. C. (2012). *Catálogo de clones de cacao seleccionados por el CATIE para siembras comerciales* (1st ed.). Turrialba, Costa Rica: CATIE.
- Porter, L. J., Ma, Z., & Chan, B. G. (1991). Cacao procyanidins: major flavanoids and identification of some minor metabolites. *Phytochemistry*, 30(5), 1657–1663.
- Priego-Capote, F., Luque de Castro, M. D., Álvarez-Sánchez, B., Priego-Capote, F., Luque de Castro, M. D., & Álvarez-Sánchez, B. (2012). Mass spectrometry of biological molecules. In *Encyclopedia of Analytical Chemistry*. Chichester, UK: John Wiley & Sons, Ltd.
- Putri, S. P., Yamamoto, S., Tsugawa, H., & Fukusaki, E. (2013). Current metabolomics: technological advances. *Journal of Bioscience and Bioengineering*, 116(1), 9–16.
- Reineccius, G. A., Andersen, D. A., Kavanagh, T. E., & Keeney, P. G. (1972). Identification and quantification of the free sugars in cocoa beans. *Journal of Agricultural and Food Chemistry*, 20(2), 199–202.
- RitterSport. (n.d.). Product responsibility: El cacao - Ritter Sport's road to its own sustainable cocoa. Retrieved November 7, 2016, from <http://www.ritter-sport.de/en/Family-business-values/sustainability/el-cacao-00002/>
- Roepstorff, P., & Fohlman, J. (1984). Letter to the editors. *Biological Mass Spectrometry*, 11(11), 601–601.

- Rohan, T. A., & Stewart, T. (1967). The precursors of chocolate aroma: production of reducing sugars during fermentation of cocoa beans. *Journal of Food Science*, 32(4), 399–402.
- Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., ... Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics*, 5(4), 435–458.
- Schwan, R. F., & Wheals, A. E. (2004). The microbiology of cocoa fermentation and its role in chocolate quality. *Critical Reviews in Food Science and Nutrition*, 44(4), 205–221.
- Smilde, A. K., Hendriks, M. M. W. B., Westerhuis, J. A., & Hoefsloot, H. C. J. (2013). Data processing in metabolomics. In *Metabolomics in Practice* (pp. 261–284). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.
- Snyder, L. R., Kirkland, J. J. (Joseph J., & Glajch, J. L. (1997). *Practical HPLC method development*. Wiley.
- Snyder, L. R., Kirkland, J. J., & Dolan, J. W. (2010). *Introduction to modern liquid chromatography* (3rd ed.). New Jersey, United States of America: John Wiley & Sons, Inc.
- Solis-Rivera, L. G., Arauz-Cavallini, L. F., & Durán-Salvatierra, S. (2015). Decreto N° 39230-C-MAG. *La Gaceta N° 204*, p. 5. Retrieved from [https://www.imprentanacional.go.cr/pub/2015/10/21/COMP\\_21\\_10\\_2015.pdf](https://www.imprentanacional.go.cr/pub/2015/10/21/COMP_21_10_2015.pdf)
- Soria-V, J. (1970). Principal varieties of cocoa cultivated in tropical America. *Agric Soc Trinidad Tobago J.* Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=US201302393844>
- SOUSA, L. S., ROCHA, F. de S., SILVEIRA, P. T. de S., BISPO, E. da S., SOARES, S. E., SOUSA, L. S., ... SOARES, S. E. (2016). Enzymatic activity of proteases and its isoenzymes in fermentation process in cultivars of cocoa (*Theobroma cacao* L.) produced in southern Bahia, Brazil. *Food Science and Technology (Campinas)*, 36(4), 656–663.
- Steinmann, D., & Ganzera, M. (2010). Recent advances on HPLC/MS in medicinal plant analysis. *Journal of Pharmaceutical and Biomedical Analysis*, 55, 744–757.
- Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., ... Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics: Official Journal of the Metabolomic Society*, 3(3), 211–221.
- Swartz, M. E. (2005). UPLC™: An Introduction and review. *Journal of Liquid Chromatography & Related Technologies*, 28(7–8), 1253–1263.
- Tortora, G. J., Funke, B. R., & Case, C. L. (2010). *Microbiology: an introduction* (10th ed.). United States of America: Pearson.
- Tournefort, J. P. (1700). *Institutiones Rei Herbariae*. 1:600; 3: Tab. 444.
- Tucker, G. A., & Woods, L. F. J. (2012). *Enzymes in food processing*. Springer Science & Business Media.
- van Vliet, J. A., & Giller, K. E. (2017). Mineral Nutrition of Cocoa: A Review. *Advances in Agronomy*, 141, 185–270.
- Vega, C., & Kwik-Urbe, C. (2012). *Theobroma cacao*—an Introduction to the plant, its composition, uses, and health benefits. In *Cocoa Butter and Related Compounds* (pp. 35–62). Elsevier.
- Voigt, J., Heinrichs, H., Voigt, G., & Biehl, B. (1994). Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chemistry*, 50(2), 177–184.

- Voigt, J., Janek, K., Textoris-Taube, K., Niewianda, A., & Wöstemeyer, J. (2016). Partial purification and characterisation of the peptide precursors of the cocoa-specific aroma components. *Food Chemistry*, 192, 706–713.
- Voigt, J., Textoris-Taube, K., & Wöstemeyer, J. (2018). pH-Dependency of the proteolytic formation of cocoa- and nutty-specific aroma precursors. *Food Chemistry*.
- Waters(R). (n.d.). BEH (ethylene bridged hybrid) technology. Retrieved February 14, 2018, from [http://www.waters.com/waters/es\\_CR/BEH-%28Ethylene-Bridged-Hybrid%29-Technology/nav.htm?cid=134618172&locale=es\\_CR](http://www.waters.com/waters/es_CR/BEH-%28Ethylene-Bridged-Hybrid%29-Technology/nav.htm?cid=134618172&locale=es_CR)
- Watson, T. J., & Sparkman, D. O. (2007). *Introduction to mass spectrometry* (4th ed.). West Sussex: Wiley.
- Wessel, M., & Quist-Wessel, P. M. F. (2015). Cocoa production in West Africa, a review and analysis of recent developments. *NJAS - Wageningen Journal of Life Sciences*, 74–75, 1–7.
- Wheelock, Å. M., & Wheelock, C. E. (2013). Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems*, 9(11), 2589.
- Wolfender, J.-L., Glauser, G., Boccard, J., & Rudaz, S. (2009). MS-based plant metabolomic approaches for biomarker discovery. *Natural Product Communications*, 4(10), 1417–1430. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19911581>
- Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. (2015). Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*, 1382, 136–164.
- Wollgast, J., & Anklam, E. (2000). Review on polyphenols in *Theobroma cacao*: changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Research International*, 33(6), 423–447.
- Wood, G., & Lass, R. (2001). *Cocoa*. John Wiley & Sons.
- Wright, P., Alex, A., & Pullen, F. (2016). Predicting collision-induced dissociation mass spectra: understanding the role of the mobile proton in small molecule fragmentation. *Rapid Communications in Mass Spectrometry*, 30(9), 1163–1175.
- Wysocki, V. H., Resing, K. A., Zhang, Q., & Cheng, G. (2005). Mass spectrometry of peptides and proteins. *Methods*, 35(3), 211–222.
- Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *The Analyst*, 137(2), 293–300.
- Zhou, B., Xiao, J. F., Tuli, L., & Ressom, H. W. (2012). LC-MS-based metabolomics. *Molecular BioSystems*, 8(2), 470–481.

## 15.ANEXOS

Cuadro A.I. pK<sub>a</sub> de diferentes metabolitos obtenidos de la base HMDB.

Metabolito	Estructura	pK <sub>a</sub> del grupo más ácido	pKa del grupo más básico
Teobromina		9,28	-0,91
Epicatequina		9	-3,3
Ácido acético		4,5	-
Sacarosa		11,84	-3
L-fenilalanina		2,47	9,45

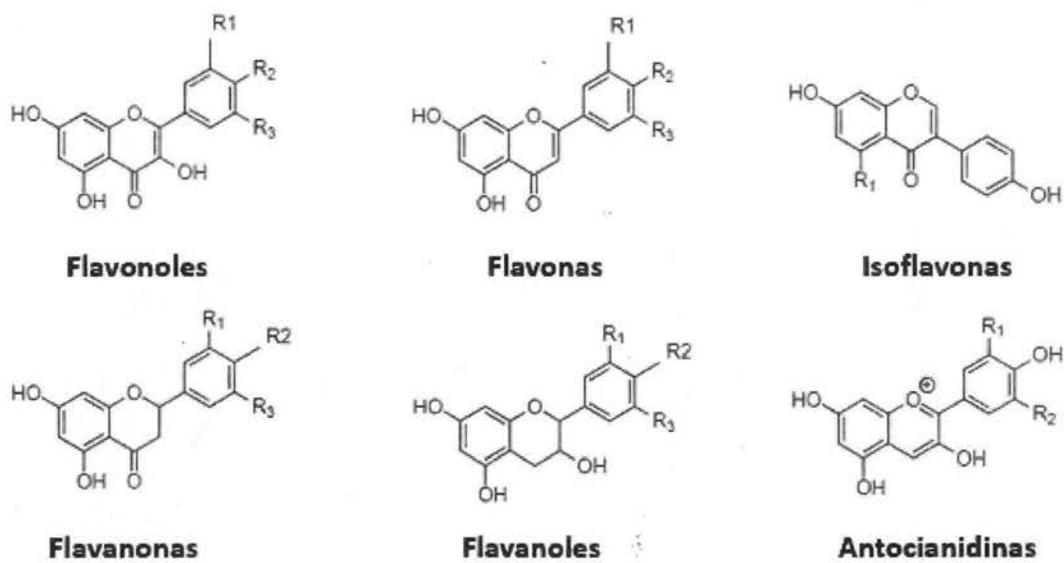


Figura A.1. Clases de flavonoides (Manach et al., 2004).



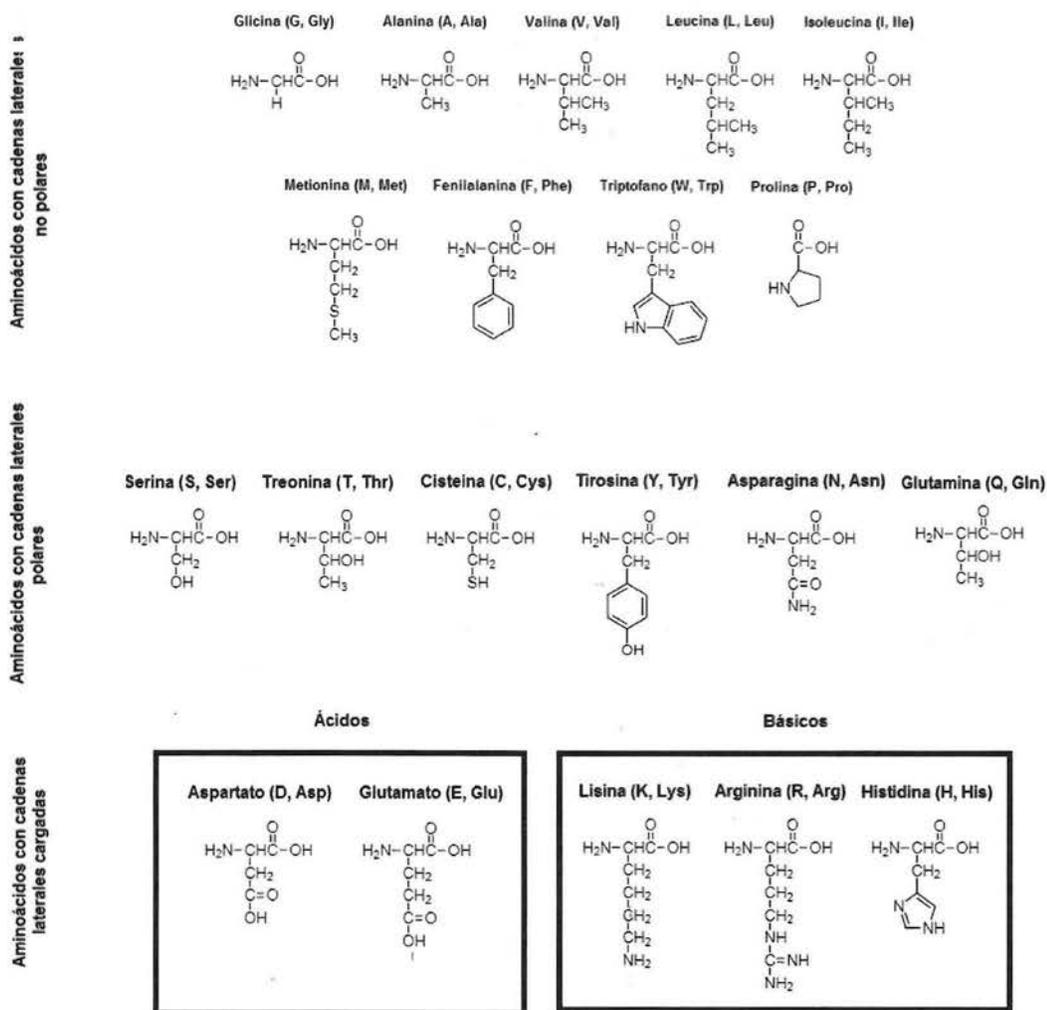


Figura A.3. Nomenclatura de distintos tipos de aminoácidos.

**Cuadro A.II.** Iones imonio asociados a distintos aminoácidos y su valor de  $m/z$  (Wysocki et al., 2005).

<b>Ión imonio</b>	<b>Residuo de aminoácido</b>	<b>Pico mayor (M) y menor (m)</b>
60.0	S	M
70.07	R o P	M
72.08	V	M
73.00	R	m
74.06	T	M
84.08	K o Q	M
86.1	I o L	M
87.09	N o R	M
88.04	D	M
100.09	R	m
101.11	K o Q	M
102.06	E	M
104.05	M	M
110.07	H	M
112.09	R	M
120.08	F	M
126.06	P	M
129.1	K o Q	m
136.08	Y	M
138.07	H	m
159.09	W	M



## An untargeted metabolomic assessment of cocoa beans during fermentation



A.L. Mayorga-Gross<sup>a</sup>, L.M. Quirós-Guerrero<sup>b</sup>, G. Fourny<sup>c,d</sup>, F. Vaillant<sup>a,d</sup>

<sup>a</sup> Centro Nacional de Ciencia y Tecnología de Alimentos, Universidad de Costa Rica, San José, P.O. Box: 11501-2060, Costa Rica

<sup>b</sup> Escuela de Química, Universidad de Costa Rica, San José 2060, Costa Rica/Centro de Investigación en Productos Naturales (CIPRONA), Universidad de Costa Rica, San José 2060, Costa Rica

<sup>c</sup> Cocoa Research Center, Sir Frank Stockdale Building, Ground Floor, North Wing, The University of the West Indies, St. Augustine, Trinidad and Tobago

<sup>d</sup> CIRAD, UMR QualiSud, 1101 avenue Agropolis, CS 24501, 34093 Montpellier Cedex 5, France

### ARTICLE INFO

#### Article history:

Received 30 October 2015

Received in revised form 25 February 2016

Accepted 14 April 2016

Available online 20 April 2016

#### Keywords:

Cocoa

Fermentation

Metabolomics

Flavonoids

Oligopeptides

Mass spectrometry

### ABSTRACT

Fermentation is a critical step in the processing of high quality cocoa; however, the biochemistry behind is still not well understood at a molecular level. In this research, using a non-targeted approach, the main metabolomic changes that occur throughout the fermentation process were explored. Genetically undefined cocoa varieties from Trinidad and Tobago ( $n = 3$ ), Costa Rica ( $n = 1$ ) and one clone IMC-67 ( $n = 3$ ) were subjected to spontaneous fermentation using farm-based and pilot plant controlled conditions. Samples were collected daily, and acetone/water/acetic acid (70/29.5/0.5%) extracts were obtained and measured using a UPLC-ESI<sup>+</sup>-Q-Tof-MS system. Analysis of pre-processed data, applying a PLS-DA model, highlighted significant differences between the three fermentation periods, observed at 0–2, 3–4 and 5–6 days, respectively. Among the most discriminating metabolites, sucrose, flavanols such as (epi) catechin, procyanidin dimers and trimers, anthocyanins and oligopeptides ranging in size from 3 to 12 amino acids were tentatively identified. In general, flavanols and sugars tended to decrease throughout fermentation, whereas most oligopeptides increased to maximal levels after 3 to 4 days of fermentation and decreased thereafter. These results shed new insight into cocoa fermentation optimization strategies and may also contribute to the development of novel alternatives for cocoa processing based on biochemical and functional values.

© 2016 Elsevier Ltd. All rights reserved.

### 1. Introduction

Traditionally, cocoa beans are subjected to fermentation and drying on farms prior to their acquisition by the chocolate industry. Fermentation is recognized as a key step, which determines final cocoa quality (Biehl & Ziegler, 2003b; Kongor et al., 2016).

During fermentation, microbial development triggers a series of changes as pulp degradation, diffusion of metabolites across the testa and diverse *postmortem* reactions that take place inside the bean. All of these phenomena result in the development of color, flavor and aroma precursors (Biehl & Ziegler, 2003b; Buyukpamukcu et al., 2001; Nigam & Singh, 2014; Saltini, Akkerman, & Frosch, 2013). After fermentation, cocoa beans are dried. This step contributes to cocoa quality, by easing the loss of volatile organic acids, and oxidizing polyphenols. The overall result of this two steps has a profound impact in the color and flavor potential of the beans (Kongor et al., 2016).

The pulp surrounding the cocoa nibs contains approximately 11% of sugars, mainly hexoses, and has a pH of 3.5–3.8 due to the high concentration of organic acids, such as citric acid (Biehl & Ziegler, 2003a; Dand, 2011). These conditions boost the spontaneous development of

microorganisms that ferment the pulp. A sequential development of yeasts (e.g., *Kloekera*, *Saccharomyces*, *Candida*, *Pichia* and *Kuyveromyces spp.*), lactic acid bacteria (e.g., *Lactobacillus*, *Leuconostoc* and *Lactococcus*), and acetic acid bacteria has been described (e.g., *Acetobacter* and *Gluconobacter spp.*) (Ardhana & Fleet, 2003; Mozzi, Ortiz, Bleckwedel, De Vuyst, & Pescuma, 2013).

Generation of microbial metabolites such as ethanol, lactic and acetic acid, combined with an increase in the cocoa mass temperature due to exothermic reactions, leads to the death of the embryo within the seed. Consequently, post mortem changes take place, for example cell membrane disruption, followed by the release of the cell's storage components. This facilitates the occurrence of biochemical reactions mainly catalyzed by glycosidases, invertases, polyphenol oxidases, peroxidases and proteases (Afoakwa, Quao, Takrama, Budu, & Saalia, 2012; Biehl & Ziegler, 2003a).

In cocoa beans, proteins are the second most abundant macronutrients and constitute approximately 10–15% (dry weight) of the bean (Marseglia et al., 2014). The most abundant proteins in cocoa are albumins and vicilin-like globulins, which undergo changes during fermentation, mediated by enzymes, such as cysteine endoprotease, leucine-*p*-nitroanilide cleaving seryl exopeptidase, aspartic endoprotease, and carboxypeptidase (seryl-exopeptidase) (Biehl & Ziegler, 2003a). Meanwhile albumin degradation is limited during fermentation,

E-mail address: [analucia.mayorga@ucr.ac.cr](mailto:analucia.mayorga@ucr.ac.cr) (A.L. Mayorga-Gross).

vicilin-type (7S) globulin is highly broken down, in derivatives such as hydrophobic oligopeptides which then turn into hydrophilic oligopeptides and hydrophobic free amino acids (Kongor et al., 2016). Both play an important role in the development of cocoa flavor, as contributors in Maillard reactions during the roasting step (Amin, Jinap, & Jamilah, 1997; Biehl & Ziegler, 2003a; Kongor et al., 2016; Marseglia et al., 2014; J. Voigt, Heinrichs, Voigt, & Biehl, 1994). Several oligopeptides have been identified in fermented cocoa beans with sequences that allows identification of the source proteins vicilin or albumin (Buyukpamukcu et al., 2001; Marseglia et al., 2014). Some oligopeptides have been putatively classified as aroma-related peptides as for example FILPH and IFVPH (Voigt, Janek, Textoris-Taube, Niewianda, & Wöstemeyer, 2016). However, the dynamic of peptides formation in the course of fermentation is not well known.

From a health functional perspective, certain peptides may also improve cocoa quality, as specific peptide classes have been reported to have antioxidant and angiotensin-converting enzyme inhibitory activities (Sarmadi, Ismail, & Hamid, 2011).

On the other hand, cocoa beans are rich in metabolites such as polyphenols (Biehl & Ziegler, 2003a) mainly flavonoids, which account nearly 13% of the dried unfermented cocoa beans mass (Kim, Lee, & Lee, 2011). Proanthocyanidins, catechins and anthocyanins represent 58–65%, 29–38% and 1.7–4%, of the total polyphenols found in cocoa, respectively (Biehl & Ziegler, 2003a). Important metabolites include epicatechin, catechin, procyanidin oligomers and some glycosides such as cyanidin-3-*O*-galactoside and cyanidin-3-*O*-arabinoside (Kim et al., 2011; Niemenak, Rohsius, Elwers, Omokolo Ndoumou, & Lieberei, 2006; Ortega et al., 2008). Flavonoids have been shown to decrease during fermentation, caused by lixiviation, oxidation, hydrolysis and polymerization to form complex tannins (Afoakwa et al., 2012; Nazaruddin, Seng, Hassan, & Said, 2006).

Flavonoids also influence cocoa quality, as they impart astringent flavors and drive a wide range of colors, from violet to dark brown. In addition, the bioactivity of the cocoa flavonoids include antioxidant, anti-inflammatory, immune modulating, and antimicrobial activities (Hii, Law, Suzannah, Misnawi, & Cloke, 2010; Rusconi & Conti, 2010).

From a global perspective, the biochemical changes that occur during cocoa fermentation have been hypothesized; yet, the comprehension of the dynamics at a molecular level still has some gaps to be fulfilled. This study attempts to contribute in this direction, by trying to identify the main metabolites (<1000 Da) of a cocoa extract, that enhance the differences between three fermentation stages.

## 2. Materials and methods

### 2.1. Samples

Three varieties of cocoa were used in this study. Two genetically undefined varieties were provided from the Cocoa Research Centre of The University of The West Indies, Trinidad and Tobago ( $n = 3$ ) and from a local producer from Upala, Costa Rica ( $n = 1$ ). The clone IMC-67 was provided by CATIE (The Tropical Agricultural Research and Higher Education Center), Costa Rica ( $n = 3$ ).

### 2.2. Chemicals

Hexanes and acetone ACS grade were acquired from Avantor Performance Materials. The following standards were employed for UPLC–MS analysis: citric acid (99.5% ACS reagent, Sigma-Aldrich, St. Louis, Missouri, USA), malic acid (99%, Sigma-Aldrich), L-tryptophan ( $\geq 98\%$ , Sigma-Aldrich) and L-phenylalanine ( $\geq 98\%$ , Sigma-Aldrich), sucrose ( $\geq 98\%$ , Sigma-Aldrich), and (+)-catechin (98%, Sigma-Aldrich). Standards were prepared in Milli-Q water. Acetonitrile (Avantor Performance Materials, Center Valley, Pennsylvania, USA) and methanol (LiChrosolv®, Merck Millipore, Billerica, Massachusetts, USA) were

acquired with HPLC grade. Formic acid was added to the solvents (98%, Fluka, Sigma-Aldrich) at a 0.1 % (v/v) concentration.

### 2.3. Fermentation trials

Cocoa pods were visually selected and those with physical damage or disease symptoms were dismissed. Selected pods were manually opened and the beans were placed in boxes following spontaneous fermentation. Internal temperature was monitored throughout the fermentation process in the central point, and mixing of the cocoa mass was done manually at 48 and 96 h. The mix of cocoa varieties from Upala (Northern Costa Rica), were fermented at a farm in a wooden box (1 \* 1 \* 0.5 m) that was divided into four equal dimensions, with approximately 100 kg of cocoa raw beans stored *per* unit. Cocoa from Trinidad and Tobago was fermented at a farm using the outlined box method, with a load of 32 kg. The clone IMC-67 was fermented in polystyrene boxes (23 \* 22 \* 17 cm) with an average of 40 kg of cocoa raw beans *per* box, that were placed inside a cabin with controlled temperature (30 °C). Nibs were manually sampled, daily, and in the center of the fermentation mass. Sampling mass ranged in between 10 and 100 g depending on the size of the batch; and preserved using flash freezing in liquid nitrogen and subsequently stored at  $-18$  °C. Afterwards, samples from Costa Rica (clon and mix of varieties) were dehydrated at the same time and for a week, with a combination of solar and artificial forced-air tray drying at less than 30 °C. The end of the drying step was determined by evaluating the following criteria: color, odor, hardness and degree of cracking of shell under manual pressure, to obtain approximated moisture below 7% (w/w).

### 2.4. Cocoa extracts

Dried samples were de-shelled manually and grinded in a blade grinder. Next, the samples were subjected to the procedure outlined in (Ortega et al., 2010) with some modifications. An exact mass ( $m$ ) of approximately 1–3 g of the grinded cocoa material was weighed and followed defatting, extraction and dilution steps. The volume of the solvents and the dilution factors employed were determined in relation to the initial grinded cocoa mass ( $m$ ), to reduce variance related to sample mass. Grinded material was defatted with  $8.3 * m$  hexane volume (mL) for 20 min in an orbital shaker at 250 rpm and followed by centrifugation at 5000 rpm for 15 min. This defatting procedure was repeated three more times. Afterwards, defatted cocoa material was extracted with a  $5 * m$  volume (mL) of acetone–water–acetic acid (70:29.5:0.5), vortexed for 3 min and centrifuged at the same conditions mentioned previously. This extraction procedure was repeated three more times. Combined supernatants were filtered through glass-microfibre discs (Boeco, Germany) and acetone was evaporated at maximum temperature of 40 °C and under vacuum. Concentrated extracts were diluted with Milli-Q water  $21 * m$  times. Extracts were filtered through a 0.45  $\mu$ m regenerated cellulose filter (Minisart-RC 25) prior to analysis.

### 2.5. Instrumental analysis

Samples were analyzed using an UPLC-ESI-QToF-MS system (Acquity, Waters, Milford, Massachusetts, USA). Chromatographic separation was performed using a 1.7  $\mu$ m CSH C<sub>18</sub> column of 2.1 \* 100 mm (Acquity, Waters) coupled to a 1.7  $\mu$ m CSH C<sub>18</sub> VanGuard pre-column in an ACQUITY Ultra Performance LC™ system equipped with an auto sampler and PDA detector (Waters). Gradient elution was carried out using water (solvent A) and acetonitrile (solvent B), both adjusted with 0.1% formic acid. The elution started at 0% of B and continued as follows (min, B %): (7, 10); (22, 95); (22.1, 0) and ended with a re-equilibration step (26, 0). The flow rate was fixed at 0.4 mL/min with a total run time of 26 min. Column temperature was fixed at 30 °C. Sample volumes of 6  $\mu$ L were injected and their temperatures were controlled at 5 °C before injection.

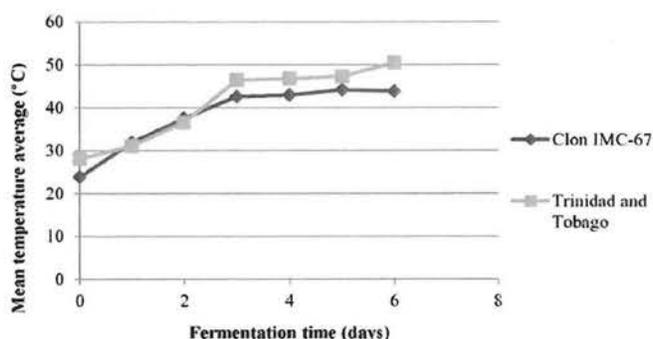


Fig. 1. Internal mass temperature of clon IMC-67 cocoa (40 kg of cocoa beans in polystyrene boxes) and genetically undefined trinitarian cocoa (32 kg in wooden boxes) during fermentation.

Positive mode ionizing chromatography was used to separate metabolites in a Waters (Milford, Massachusetts, USA) © SYNAPT ESI-QToF system under a full scan from 70 to 1000 Da in continuous mode. ESI-Qtof-MS conditions were as follows: capillary voltage at 1.4 kV, sampling cone voltage of 35 V and extraction cone of 4.0 V. Nitrogen was used as cone and desolvation gas with a flow of 40 L/h and 650 L/h, respectively. Additional MS<sup>n</sup> experiments were done at different collision energies and ramps from 4 eV to 50 eV. Scan time was set to 0.2 s and interscan time to 0.020 s. Equipment was calibrated with a sodium iodide solution (2 µg/µL, reference mass V<sup>+</sup>: 922.3552) and a solution of leucine-enkephalin (2 ng/mL) was used as lock mass (V<sup>+</sup>: 556.2771). For quality control, external standards (described in the section of *Chemicals*) and a pool of all samples were measured every 10 samples. All chromatograms and mass spectra are presented on supporting information.

## 2.6. Data processing

Raw files were collected in a continuous mode and centered with the lock mass using MassLynx V3.1 SCN639 software (Waters Inc., Milford, Massachusetts, USA). Transformed files were next converted to CDF-type files using Databridge software (Waters Inc.). Preprocessing of the chromatograms and mass spectra was done with MZmine software and spectra were aligned using the RANSAC algorithm, with 1000 iterations, with an *m/z* tolerance of 5 ppm and retention time tolerance of

0.05 min. Gap filling was done with *am/z* tolerance of 8 ppm and retention time tolerance of 0.4 min.

## 2.7. Statistical analysis

Multivariate analyses were made using SIMCA 13.0.2.0 software (Umetrics, Umeå, Sweden). Unsupervised Principal Component Analysis (PCA) was firstly applied to explore the dataset. Logarithmic transformation and Pareto scaling were applied to the data. Afterwards, a supervised Partial Least Squares discriminant analysis (PLS-DA) was applied using three fermentation time groups (0–2 days, 3–4 days and 5–6 days). The significance of this model was assessed using a cross-validated analysis of variance p-value (CV-ANOVA) and its quality was evaluated by the parameters R<sup>2</sup>X, R<sup>2</sup>Y and Q<sup>2</sup>X. Permutation test (n = 100) was carried out to evaluate data overfitting. A variable importance projection index (VIP index) was assigned to each variable (ion), to show their discrimination power between the different groups of the PLS-DA model. Overall, ions that were chosen to study displayed a VIP value higher than 2.5 and an ANOVA p-value lower than 0.05. Ions were grouped into 37 clusters by means of their retention times, intensity pattern over fermentation time, and other chemical relationships (i.e., metabolites).

## 2.8. Metabolite identification

Related ions are expected to show similar graphical tendencies; therefore, the intensity vs. fermentation time was graphed and used for pattern analysis between ions. Chromatograms and mass spectra at different collision energies were also analyzed for the identification of parent ions. MS/MS experiments were analyzed to explain the molecule's fragmentation pathways. For peptides, nomenclature proposed by (Paizs & Suhai, 2005), based on Roepstorff & Fohlmann (1984) and Biemann (1988) was employed. Different databases were consulted for matching exact mass (maximum tolerance of 5 ppm) and mass spectrum: METLIN Metabolite and Tandem MS Database (metlin.scripps.edu; Smith et al., 2005); The Human Metabolome Database (www.hmdb.ca; Wishart, Tzur, Knox, et al., 2007; Wishart, Knox, Guo, et al., 2009 & Wishart et al., 2013); Protein Prospector v 5.14.3 (prospector.ucsf.edu/prospector/mshome.htm; (Baker & Chalkley, 2014)); ReSpect for Phytochemicals (spectra.psc.riken.jp; (Sawada et al., 2012)); Mass Bank repository (massbank.jp; Nishioka, Nihei, Ojima, & Ikeda, 2012) and MetFrag fragmenter (msbi.ipb-halle.de/MetFrag; (Wolf, Schmidt, Müller-Hannemann, & Neumann, 2010)).

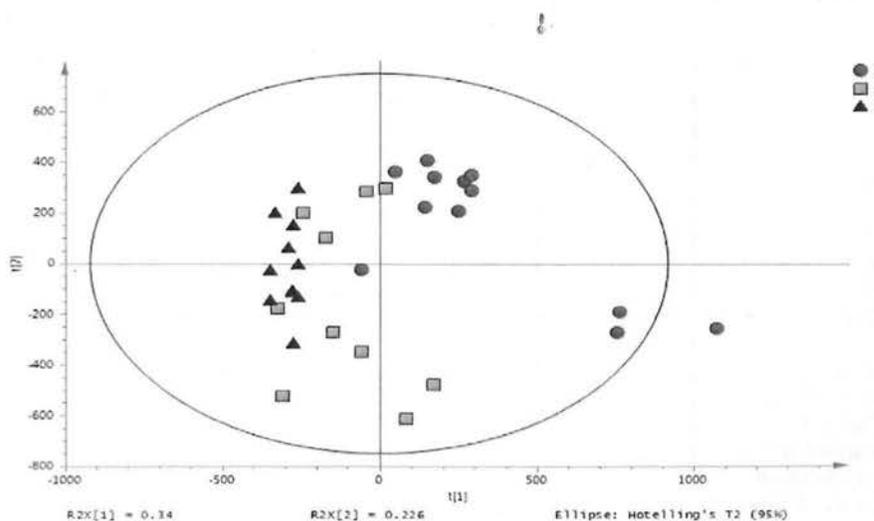
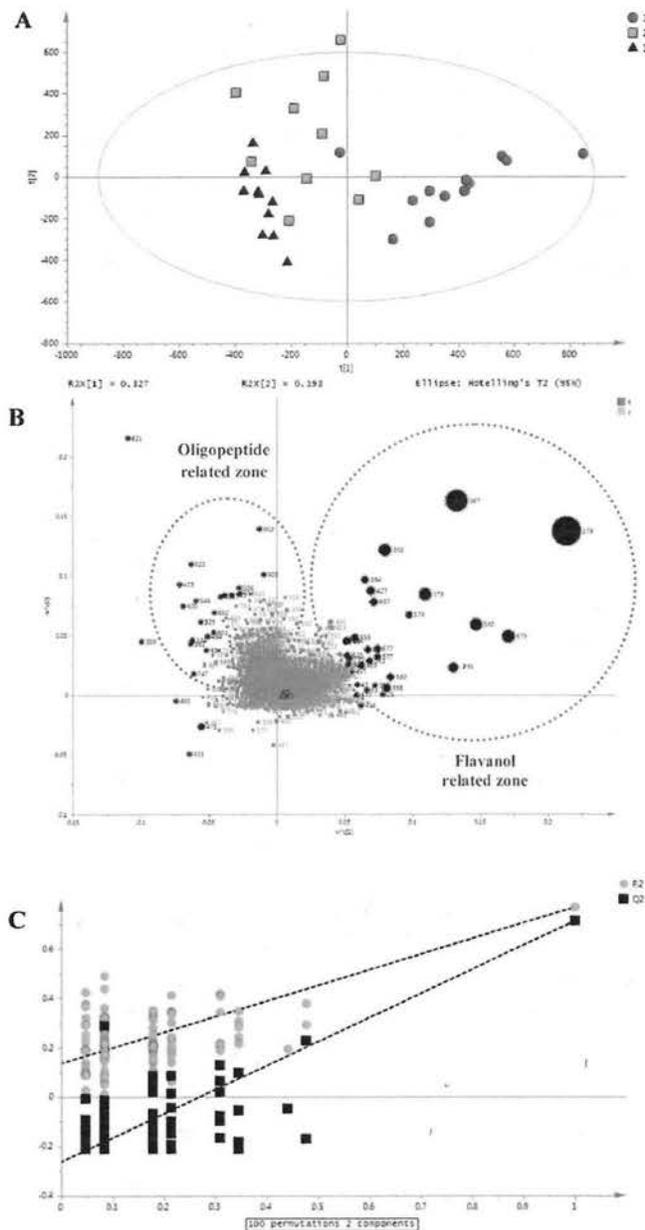


Fig. 2. Principal component analysis score plot of the dataset (1 = 0–2 days; 2 = 3–4 days and 3 = 4–6 days of fermentation).



**Fig. 3.** A) Partial least square discriminant analysis (PLS-DA) score and B) loading plots of the dataset, sized by VIP (1 = 0–2 days; 2 = 3–4 days and 3 = 4–6 days of fermentation -discriminant ions in dark color-) and C) permutation test for the dataset ( $n = 100$ ;  $R^2 = (0.0, 0.143)$ ;  $Q^2 = (0.0, -0.267)$ ).

### 3. Results and discussion

#### 3.1. Fermentation trials

The behavior of the internal cocoa mass temperature throughout fermentation was simulated for a small load of clon IMC-67 (40 kg) in polystyrene boxes and compared to bulk fermentations performed at a farm in Trinidad and Tobago (Fig. 1). Temperature was shown to follow the same trend with a progressive increase until the 3rd day. The mixing of the pulp that was done after the 2nd day to enhance the development of acetic acid bacteria, led to exothermic metabolic reactions rising temperature above 40 °C (Fig. 1) (Nigam & Singh, 2014; Pereira, Soccol, & Soccol, 2016; Saltini et al., 2013). After this point, the temperature remained within a range of 40–50 °C for up to 6 days after the beginning of the fermentation process. Simulation of fermentation in

polystyrene boxes with load of 40 kg of cocoa beans was considered adiabatic and was deemed to be similar to the fermentation process at farm. As a consequence, all samples collected daily were classified according to fermentation day and, in this work, analyzed independently from the place of origin and fermentation method.

#### 3.2. Construction of the model

A principal component analysis (PCA) was applied, considering the base peak intensity of the approximately 3000 ions that were detected by the instrumental system and pre-processed using MZmine. The space constructed by the two first principal components explained a total of 52% of the total variation (Fig. 2). Data grouping was observed according to the fermentation time (0–2 days, 3–4 days and 5–6 days). The effects of the cocoa origin and the fermentation procedure were not considered in this study. A partial least squares discriminant analysis (PLS-DA) was performed taking as response the different groups observed previously (Fig. 3A and B). The model obtained was deemed statistically acceptable with a predictive capacity by cross validation ( $n = 100$ ) of  $Q^2Y = 0.71$ . The analysis of variance of the cross-validation predictive residuals (CV-ANOVA) arose a low p-value ( $2.4 \cdot 10^{-5}$ ) and finally, the permutation test performed, gave a low probability of model overfitting ( $R^2Y = (0.0, 0.143)$ )  $Q^2Y = (0.0, -0.267)$  (Fig. 3C).

The highest discriminant ions with a VIP  $\geq 2.5$  and a p-value  $< 0.05$  were chosen for further analysis. The clustering of ions according to retention times and mass spectrum yielded a total of 37 highly discriminating metabolites, which were used as the basis for this research.

#### 4. Metabolite analysis

Table 1 shows information related to the 10 most discriminating metabolites, including three of which were putatively annotated as flavanols, one was assigned to a tripeptide, one hypothesized as sugar and the others remained unknown.

Metabolites at  $m/z$  343.12 at 0.8 min,  $m/z$  325.11 at 0.9 min, at  $m/z$  396.03 at 1 min, and at  $m/z$  632.07 at 8.8 min were shown to drop in intensity over the fermentation period. From these, the first showed mass spectra and retention time similarities, compared to a measured sucrose standard solution, suggesting that these metabolites are sugars. Sucrose ( $m/z$  343.1235  $[M + H]^+$ ) is reported to decrease inside the bean during fermentation, mainly because of the action of invertases; glucose and fructose produced also undergo further reactions (Biehl & Ziegler, 2003a; Camu et al., 2008). Due to their low retention time further chromatographic optimization during this time frame are recommended to resolve the different metabolites, as well as negative mode analysis.

In contrast, other metabolites at  $m/z$  of 431.21 with an RT of 4.8 min and at  $m/z$  399.22 at RT of 5.3 min increased with fermentation time, revealing a maximum intensity at the last stage of fermentation. Their nature remained unknown and requires further analysis.

The fifth highest discriminating metabolite  $m/z$  579.14  $[M + H]^+$  with a RT of 8.8 min was putatively identified as an (epi) catechin dimer (B-type procyanidin). Observed mass matched with exact mass value reported in databases. Fragmentation pattern is also in agreement with this identification, revealing Retro Diels Alder (RDA) pathway ( $m/z$  427), heterocyclic ring fission (HRF) ( $m/z$  453) and quinone methide fission (QM) ( $m/z$  291 and  $m/z$  289), which are commonly observed for flavanols (Gu et al., 2003; Lin, Sun, Chen, Monagas, & Harnly, 2014). Additionally, diode array spectrum showed absorption in the range of 260 nm to 295 nm, with a maximum at 278–279 nm, which is in line with the expected values for flavan-3-ol A-rings (de Rijke et al., 2006). Among the top ten discriminant metabolites, (epi) catechin ( $m/z$  291.08  $[M + H]^+$  with RT of 8.9 min) and a procyanidin trimer ( $m/z$  867.21  $[M + H]^+$  with RT of 9.3 min), were putatively identified. Both match exact mass and spectra reported in databases. Products derived from common flavanol fragmentation pathways

**Table 1**List of the top 10 discriminant metabolites, proposed identity and chemical and statistical parameters associated.<sup>a</sup>

Assigned identity	Formula	Adduct	m/z (p-value ranking)	RT (min)	p-value (Anova)	VIP value	Main MSMS fragments	MSMS energy (eV)	Intensity <sup>a</sup>			ID method			
									E	I	L				
Sucrose		[M ± H] <sup>+</sup>	343.1245(1)	0.8	9.9E−9	2.7	232.1 50	20	916	313	37	Standards			
			145.0503(13)		1.4E−5	2.6	150.0 100 127.0 65								
Unknown 1		[M ± H] <sup>+</sup>	325.1121(2)	0.9	1.8E−7	3.6	145.1 50	10	1725	561	157	NA			
			[M + H-anhydrohexose] <sup>+</sup>				163.0578(18)						4.8E−5	3	97.0 65 91.0 60 85.0 100 276.0 80 204.0 60 187.0 100 174.0 70 342.1 20 198.1 60
Unknown 2			396.0336(3)	1.0	3.4E−7	3.3	141.1 65 132.1 100 86.1 90	20	1564	599	188	NA			
Unknown 3		[M ± H] <sup>+</sup>	431.2215(4)	4.8	4.6E−7	3.2	141.1 65 132.1 100 86.1 90	20	74	626	1297	NA			
(Epi) catechin dimer (procyanidin)	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	[[M + H] + 2(13C) + (18O)] <sup>+</sup>	581.1652(5)	8.8	5.6E−7	3.2	409.1 80	20	21559	14862	7983	Gu et al., 2003, Lin et al. (2014).			
			[M ± H] <sup>+</sup>				579.1487(6)						6.3E−7	10.5	287.1 60
			[[M + H] <sup>+</sup> + (13C)]				580.1565(8)						1.4E−6	6.9	139.0 100
			[M + H] <sup>+</sup> → RDA − H <sub>2</sub> O				409.0922(24)						2.4 E−4	2.9	127.0 80
Unknown 4		[M ± H] <sup>+</sup>	427.1027(35)	8.8	7.5 E−4	4	165.1 25 139.0 100 123.0 60	10	1477	1073	670	NA			
(epi) catechin	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	[M ± H] <sup>+</sup>	291.0857(9)										1.6E−6	6	197.1 85 158.1 95 129.1 85 84.0 100 197.1 35
Unknown 5		[[M + H] + (13C)] <sup>+</sup>	400.2392(14)	5.3	1.8E−5	2.5	169.1 100 70.1 90	15	283	1307	1089	MetLin 22067			
PVN	C <sub>14</sub> H <sub>25</sub> N <sub>4</sub> O <sub>5</sub>	[M ± H] <sup>+</sup>	329.1832(11)	0.8	4.3E−6	3	577.1326(12) 869.2264(19)	20	14087	11405	5061	ResPect PT212070			
(Epi) catechin trimer (procyanidin)	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>	[[M + H] + 2(13C) + (18O)] <sup>+</sup>	577.1326(12)	5.7E−6	3.5	579.2 100									
			[M ± H] <sup>+</sup>	579.1460(39)	1.3 E−3	4.8	409.1 40								
			[M + H] <sup>+</sup> → QM	868.2142(49)	6.3 E−3	4.9	291.1 45								
			[[M + H] + (13C)] <sup>+</sup>	867.2109(52)	7.5 E−3	7.5	247.1 45								

<sup>a</sup> E: early fermentation (0–2 days), I: intermediate fermentation (3–4 days), L: advanced fermentation (5–6 days). RDA: Retro Diels Alder reaction and QM:quinone methide fission.**Table 2**List of putatively assigned flavonoid-type metabolites and their chemical and statistical parameters associated.<sup>b</sup>

Assigned identity	Formula	Adduct	m/z (p-value ranking)	RT (min)	Δ m/z (ppm)	p-Value (Anova)	VIP value	Main MSMS fragments	MSMS energy (eV)	Intensity <sup>a</sup>			
										E	I	L	
B type (epi) catechin dimer (procyanidin)	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	[M + H] <sup>+</sup>	579.1468 (17)	10	5.0	4.3E−5	7.9	409.1 100 287.1 85 139.0 60 127.0 75 425.1 65	20		9763	4051	756
A type (epi) catechin dimer (procyanidin)	C <sub>30</sub> H <sub>24</sub> O <sub>12</sub>	[M + H] <sup>+</sup>	577.1315 (33)	11.2	4.5	5.9 E−4	2.7	245.0 100 123.0 30 587.1 60	25		1609	754	464
O-hexoside proanthocyanidin A5'	C <sub>36</sub> H <sub>34</sub> O <sub>17</sub>	[M + H] <sup>+</sup>	739.1798 (50)	10.2	4.5	6.6 E−3	2.6	577.1 25 425.1 100 299.1 40 409.0 90 289.1 40	20		1933	1282	703
B type (epi) catechin trimer (procyanidin)	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>	[M + H] <sup>+</sup>	867.2157 (47)	9.8	3.0	5.5 E−3	3.9	287.1 55 247.1 100 139.0 80 127.0 70 557.1 100 539.1 35	30		3143	2279	710
O-pentoside proanthocyanidin A5'	C <sub>35</sub> H <sub>32</sub> O <sub>16</sub>	[M + H] <sup>+</sup>	709.1755 (53)	10.3	1.1	0.028	2.6	437.1 30 425.1 90 299.6 45 287.1 30	20		2260	1690	906

<sup>a</sup> E: early fermentation (0–2 days), I: intermediate fermentation (3–4 days), L: advanced fermentation (5–6 days).<sup>b</sup> Proposed metabolites was based on fragmentation pathways proposed by (Gu et al., 2003; Lin et al., 2014).

were observed in both cases. For example, the B-type procyanidin trimer fragmentation yielded ions at  $m/z$  715 and at  $m/z$  427 that correspond to products of the RDA reaction,  $m/z$  741 and  $m/z$  453 from HRF reactions and  $m/z$  579 and  $m/z$  291 from QM cleavage.

The (–)-epicatechin has been reported as the main catechin in cocoa (Wollgast & Anklam, 2000). On the other side, procyanidin dimers such as B1, B2 and B5 (Calderón, Wright, Hurst, & van Breemen, 2009; Dorenkott et al., 2014; Hatano et al., 2002; Karen A. Cooper et al., 2007; Porter, Ma, & Chan, 1991) and trimers such as cinnamtannin A1 have been also reported among the main flavanols in cocoa (Cádiz-Gurrea et al., 2014; Calderón et al., 2009; Dorenkott et al., 2014; Hatano et al., 2002).

Even though they were not among top ten most discriminant metabolites, other flavanols were also putatively identified (Table 2). Some of them tentatively correspond to isomers of the procyanidin dimers ( $m/z$  579.14 [M + H]<sup>+</sup>) and trimers ( $m/z$  867.21 [M + H]<sup>+</sup>) observed previously. Metabolite at  $m/z$  577.13 [M + H]<sup>+</sup> followed the flavanol pathways previously described, and was putatively annotated as a dimeric procyanidin, with the difference of presenting an A-type bond. Other two flavonoids at  $m/z$  739.18 [M + H]<sup>+</sup> and at  $m/z$  of 709.18 [M + H]<sup>+</sup> were annotated as procyanidin glycosides.

Proanthocyanidin A1 (Hatano et al., 2002), and other isomers have been reported in cocoa (Cádiz-Gurrea et al., 2014). As for anthocyanins, arabino and galactoside cyanidins have been reported; 3-T-O-L-arabinopyranosyl-ent-epicatechin-(2 $\alpha$  → 7,4 $\alpha$  → 8)-epicatechin and 3T-O- $\beta$ -D-galactopyranosyl-ent-epicatechin-(2 $\alpha$  → 7,4 $\alpha$  → 8)-epicatechin reported for cocoa might have some relation with the observed metabolites (Cádiz-Gurrea et al., 2014; Calderón et al., 2009; Hatano et al., 2002; Porter et al., 1991).

Flavanols putatively identified decreased during fermentation time (Tables 1 and 2). In Fig. 4 are represented the histograms of base peak intensities of respectively (epi) catechin, dimeric and trimeric procyanidins for the different samples collected. It can be observed a steady decrease during fermentation. Total polyphenol reductions in cocoa samples during fermentation has been observed previously, with a reported retention of only 10–20% of proanthocyanidins and catechins after fermentation and drying (Afoakwa et al., 2012; Hue et al., 2014; Nazaruddin et al., 2006; Nigam & Singh, 2014; Payne, Hurst, Miller, Rank, & Stuart, 2010). In this case, it is interesting to note that dimeric and trimeric procyanidins decreased drastically only after the fourth day of fermentation, while (epi) catechin monomer starts to decline since the very beginning of fermentation.

Several mechanisms have been proposed to explain this phenomena, including lixiviation of soluble compounds (10–40% loss), enzymatic and non-enzymatic oxidation reactions that lead to quinone formation and subsequent polymerization, and polyphenol complexing with oligopeptides, amino acids and cell wall material (Afoakwa et al., 2012; Biehl & Ziegler, 2003a; de Brito et al., 2001; Kim et al., 2011; Nigam & Singh, 2014; Wollgast & Anklam, 2000). The reduction of anthocyanins may also be attributed to glycosidases activity. These enzymes are highly stable and active during fermentation (Hansen, del Olmo, & Burri, 1998). Although increment in the aglycone's intensities was not observed, they may be more easily lixiviated or they may take part in condensation reactions, yielding complex tannins that could not be detected by the analytical method employed.

Reduction of the content of polyphenols appears to be one of the main phenomena involved during fermentation and it is associated with a decrease of astringency and bitterness in the final chocolate. Also negative correlations between these two descriptors and aroma acceptability has been described (Afoakwa et al., 2012; Albertini et al., 2015; Saltini et al., 2013).

In Table 1, among top ten discriminant ions, metabolite observed at  $m/z$  329.18 [M + H]<sup>+</sup> at RT of 0.8 min was tentatively assigned to a tripeptide PVN. This molecule may be related to PVNSPGK and PVNSPGKY oligopeptides, which are derived from vicilin-like cocoa seed protein, and classified as an aroma-related peptide (Kochhar

et al., 2004; Voigt et al., 2016). Actually, peptides appeared to be among the highest discriminant metabolites during fermentation period. In Table 3, there are reported 10 oligopeptides, ranging in length from 3 to 12 amino acid units, that were tentatively identified among the 37 most discriminant ions. Identification was based on the generated sequence observed during fragmentation of y, a and b ions, as well as internal fragments and immonium ions which were compared to in-silico fragmentation pattern.

Oligopeptide at  $m/z$  634.30 [M + H]<sup>+</sup> with an RT 5.4 min was tentatively identified as VSTDVN. Different authors (Kochhar et al., 2004; Marseglia et al., 2014) have proposed that this hexapeptide may come from cocoa albumin seed fractions. Actually, most oligopeptides tentatively identified could be related to their protein sources, cocoa vicilin or albumin.

Metabolites at  $m/z$  563.20,  $m/z$  488.24,  $m/z$  747.40,  $m/z$  933.46,  $m/z$  862.42,  $m/z$  715.36 and  $m/z$  484.32 [M + H]<sup>+</sup>, were respectively annotated as DNEW, GAGGGGL, NGTPVIF, SSISGAGGGGLA, SSISGAGGGG, ANSPVLD and VPIR. All these oligopeptides match a sequence of the albumin cocoa seed (Kochhar et al., 2004; Marseglia et al., 2014). On the other hand, metabolite at  $m/z$  621.29 and at  $m/z$  902.46 [M + H]<sup>+</sup> which were tentatively annotated as SPGDV and APLSPGDVF respectively, match a sequence of the vicilin-like cocoa protein derivatives (Kochhar et al., 2004; Marseglia et al., 2014).

Almost all the oligopeptides showed a similar trend in intensity across fermentation time, which was characterized by an initial increase that led to the highest intensity observed 3–4 days after the start of fermentation, followed by a drop in intensity (Fig. 5). During the aerobic fermentation stage, acetic acid diffuses into the bean, following the death of the embryo and cellular disruption. Several enzymes as aspartic endoproteases, aminopeptidases and carboxypeptidases are able to act on their substrates. This might explain the increase in oligopeptides at middle fermentation period. Nonetheless, a drop of these metabolites

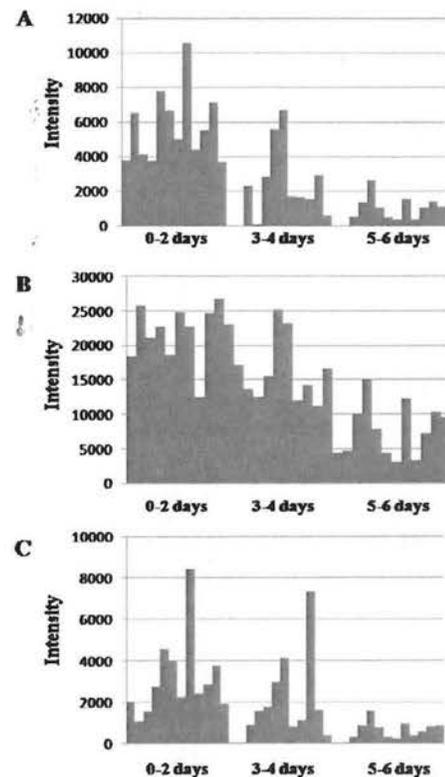


Fig. 4. Examples of metabolite intensity of different cocoa samples over fermentation time for some flavanols A) [(epi) catechin, B) (epi) catechin monomer, dimer and C) (epi) catechin trimer].

observed at advanced fermentation might be a consequence of complexing with oxidized polyphenols, quinones or condensed tannins (Biehl & Ziegler, 2003a; Jinap, Nazamid, & Jamilah, 2002; Kyi et al., 2005; Nazaruddin et al., 2006).

These results are consistent with recent findings indicating that at the beginning of the fermentation transitional peptides ranging from 3 to 10 kDa are in low concentration (Hue et al., 2016). As protein degradation advanced further, content of oligopeptides increases and reaches a peak between 3 to 4 days after the beginning of the fermentation (Fig. 5), during the development of acetic acid bacteria that led to the increase of temperature and to the intensification of *post-mortem* enzymatic reactions.

## 5. Conclusions

During fermentation of cocoa beans, and in the fraction analyzed, sugars, flavanols and oligopeptides were the most discriminant low molecular weight compounds ( $\leq 1000$  Da). In the fraction of the cocoa beans that was studied, sugars and flavonoids decreased during fermentation while almost all oligopeptides ranging from 3 to 12 amino acids reached a peak between 3 and 4 days of fermentation. All these compounds may be considered as precursors of compounds present in the final chocolate and that are most likely involved in the building of typical flavor and aroma during roasting of cocoa beans. It appears that 3 to 4 days after the beginning of the fermentation, an optimum equilibrium between the studied precursor compounds could be achieved. From the results of this study, it could be hypothesized that a shorter cocoa fermentation time of up to 3–4 days, compared to the traditional 6 days

period for the varieties analyzed, may improve cocoa quality, due to an increased abundance of flavor and aroma precursors with a slight reduction of bitter compounds. Further research has to be done, in order to understand the behavior of other fractions of compounds inside the cocoa beans, and their interactions with the studied fraction.

From a functional perspective, a shorter fermentation or no fermentation at all provides higher flavanol content in the beans with an improved antioxidant profile, associated to these compounds. Additionally, the tentative identification of the main cocoa oligopeptides released during fermentation can be used for characterization of their potential functional properties and specific contribution to overall quality. This work has to be done taking into account the rest of the fractions inside the cocoa beans which on this perspective might be of interest.

These results could contribute not only to improve current processing techniques but to promote alternatives of valorization with higher functional value for cocoa beans.

## Acknowledgments

Special gratitude is given to the following accounts: CIRAD, France; Natural Products Research Center (CIPRONA), University of Costa Rica; Escuela de Química, Universidad de Costa Rica; Cocoa Research Center (University of West Indies, Trinidad and Tobago); Tropical Agricultural Research and Higher Education Center (CATIE), Costa Rica, and University of Costa Rica within the framework of the project VI-735-B5-050 "Valorization of food matrices through the analysis of the impact of processing on the metabolomic profile".

**Table 3**

List of putatively assigned peptide-type metabolites and their chemical and statistical parameters associated<sup>a</sup>.

Assigned identity	Formula	Adduct	m/z (p-value ranking)	RT (min)	$\Delta$ m/z (ppm)	p-Value (Anova)	VIP value	Main MSMS fragments	MSMS energy (eV)	Intensity <sup>b</sup>		
										E	I	L
VSTDVN	C <sub>25</sub> H <sub>44</sub> N <sub>7</sub> O <sub>12</sub>	[M + H] <sup>+</sup>	634.3020 (15)	5.4	3.5	2.2E−5	2.5	403.2 100 187.1 80 72.1 100	25	82	857	798
DNEW	C <sub>24</sub> H <sub>31</sub> N <sub>6</sub> O <sub>10</sub>	[M + H] <sup>+</sup>	563.2111 (20)	8.1	2.7	1.3E−4	3.1	342.1 60 270.1 40 230.1 100 130.1 40	20	187	1408	1339
SPGDVF	C <sub>28</sub> H <sub>41</sub> N <sub>6</sub> O <sub>10</sub>	[M + H] <sup>+</sup>	621.2860 (22)	8.8	3.1	1.2E−4	7.7	357.1 90 270.1 100 70.1 90	30	1283	7826	4859
GAGGGGL	C <sub>19</sub> H <sub>34</sub> N <sub>7</sub> O <sub>8</sub>	[M + H] <sup>+</sup>	488.2443 (30)	5.2	4.1	4.6 E−4	3.3	215.1 70 198.1 75 132.1 65 86.1 100 70.1 65	25	161	1435	1933
NGTPVIF	C <sub>35</sub> H <sub>55</sub> N <sub>8</sub> O <sub>10</sub>	[M + H] <sup>+</sup>	747.4064 (34)	9.8	3.7	6.4 E−4	2.8	475.3 85 197.1 100 169.1 80 70.1 60	30	101	1129	1273
SSISGAGGGGLA	C <sub>37</sub> H <sub>65</sub> N <sub>12</sub> O <sub>16</sub>	[M + H] <sup>+</sup>	933.4640 (36)	7.9	0.43	8.8 E−4	2.7	539.3 80 452.2 85 86.1 90	40	90	913	444
SSISGAGGGGL	C <sub>34</sub> H <sub>60</sub> N <sub>11</sub> O <sub>15</sub>	[M + H] <sup>+</sup>	862.4262 (37)	8.0	0.35	0.0010	2.8	369.2 40 312.1 40 300.1 40 284.6 35	40	182	1204	849
APLSPGDVF	C <sub>42</sub> H <sub>64</sub> N <sub>9</sub> O <sub>13</sub>	[M + H] <sup>+</sup>	902.4612(38)	9.9	0.66	0.0010	3.9	534.2 75 341.2 65 270.1 100	40	457	1535	351
ANSPVLD	C <sub>30</sub> H <sub>51</sub> N <sub>8</sub> O <sub>12</sub>	[M + H] <sup>+</sup>	715.3627 (40)	7.4	0.84	0.0014	2.9	443.2 95 197.1 100 169.1 100	30	161	1120	634
VPIR	C <sub>22</sub> H <sub>42</sub> N <sub>7</sub> O <sub>5</sub>	[M + H] <sup>+</sup>	484.3266 (46)	4.8	5.0	0.0043	2.7	368.2 40 183.1 55 70.1 100	30	91	1179	991

<sup>a</sup> Main identification method was in-silico fragmentation patterns produced by Protein prospector (<http://prospector.ucsf.edu/prospector/mshome.htm>).

<sup>b</sup> E: early fermentation (0–2 days), I: intermediate fermentation (3–4 days), L: advanced fermentation (5–6 days).

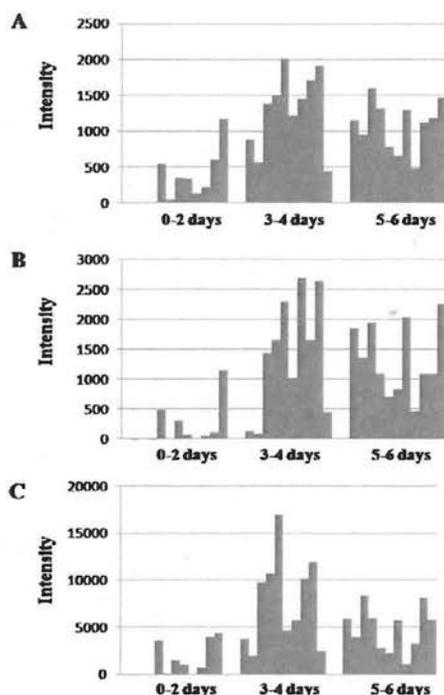


Fig. 5. Examples of metabolite intensity of different cocoa samples over fermentation time for some oligopeptides putatively annotated as A) PVN, B) DNEW and C), SPGDVF.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.foodres.2016.04.017>.

## References

- Afoakwa, E. O., Quao, J., Takrama, F. S., Budu, A. S., & Saalia, F. K. (2012). Changes in total polyphenols, o-diphenols and anthocyanin concentrations during fermentation of pulp pre-conditioned cocoa (*Theobroma cacao*) beans. *International Food Research Journal*, 19(3), 1071–1077.
- Albertini, B., Schoubben, A., Guarnaccia, D., Pinelli, F., Della Vecchia, M., Ricci, M., ... Blasi, P. (2015). Effect of fermentation and drying on cocoa polyphenols. *Journal of Agricultural and Food Chemistry*, 63(45), 9948–9953. <http://dx.doi.org/10.1021/acs.jafc.5b01062>.
- Amin, I., Jinap, S., & Jamilah, B. (1997). Vicilin-class globulins and their degradation during cocoa fermentation. *Food Chemistry*, 59(1), 1–5. [http://dx.doi.org/10.1016/0308-8146\(95\)00188-3](http://dx.doi.org/10.1016/0308-8146(95)00188-3).
- Ardhana, M., & Fleet, G. (2003). The microbial ecology of cocoa bean fermentations in Indonesia. *International Journal of Food Microbiology*, 86(1–2), 87–99. [http://dx.doi.org/10.1016/S0168-1605\(03\)00081-3](http://dx.doi.org/10.1016/S0168-1605(03)00081-3).
- Baker, P. R., & Chalkley, R. J. (2014). MS-viewer: a web-based spectral viewer for proteomics results. *Molecular & Cellular Proteomics*, 13(5), 1392–1396. <http://dx.doi.org/10.1074/mcp.0113.037200>.
- Biehl, B., & Ziegler, G. (2003a). COCOA | chemistry of processing. *Encyclopedia of food sciences and nutrition* (pp. 1436–1448) (Second ed.). Elsevier. <http://dx.doi.org/10.1016/B0-12-227055-X/00261-3>.
- Biemann, K. (1988). Contributions of mass spectrometry to peptide and protein structure. *Biomedical and Environmental Mass Spectrometry*, 16, 99–111.
- Biehl, B. and Ziegler, G., 2003b Biehl, B., & Ziegler, G. (2003b). COCOA | production, products and use. *Encyclopedia of food sciences and nutrition* (pp. 1448–1463) (Second ed.). Elsevier. <http://dx.doi.org/10.1016/B0-12-227055-X/00262-5>.
- de Brito, E. S., Garcia, N. H. P., Gallão, M., Cortelazzo, A. L., Fervereiro, P. S., & Braga, M. R. (2001). Structural and chemical changes in cocoa (*Theobroma cacao* L.) during fermentation, drying and roasting. *Journal of the Science of Food and Agriculture*, 81(2), 281–288. [http://dx.doi.org/10.1002/1097-0010\(20010115\)81:2<281::AID-JSFA808>3.0.CO;2-B](http://dx.doi.org/10.1002/1097-0010(20010115)81:2<281::AID-JSFA808>3.0.CO;2-B).
- Buyukpamukcu, E., Goodall, D. M., Hansen, C. -E., Keely, B. J., Kochhar, S., & Wille, H. (2001). Characterization of peptides formed during fermentation of cocoa Bean. *Journal of Agricultural and Food Chemistry*, 49(12), 5822–5827. <http://dx.doi.org/10.1021/jf0104127>.
- Cádiz-Gurrea, M. L., Lozano-Sanchez, J., Contreras-Gómez, M., Legeai-Mallet, L., Fernández-Arroyo, S., & Segura-Carretero, A. (2014). Isolation, comprehensive characterization and antioxidant activities of *Theobroma cacao* extract. *Journal of Functional Foods*, 10, 485–498. <http://dx.doi.org/10.1016/j.jff.2014.07.016>.
- Calderón, A. I., Wright, B. J., Hurst, W. J., & van Breemen, R. B. (2009). Screening antioxidants using LC–MS: case study with cocoa. *Journal of Agricultural and Food Chemistry*, 57(13), 5693–5699. <http://dx.doi.org/10.1021/jf9014203>.
- Camu, N., González, A., De Winter, T., Van Schoor, A., De Bruyne, K., Vandamme, P., ... De Vuyst, L. (2008). Influence of turning and environmental contamination on the dynamics of populations of lactic acid and acetic acid bacteria involved in spontaneous cocoa bean heap fermentation in Ghana. *Applied and Environmental Microbiology*, 74(1), 86–98. <http://dx.doi.org/10.1128/AEM.01512-07>.
- Cooper, K. A., Campos-Giménez, E., Alvarez, D. J., Nagy, K., Donovan, J. L., & Williamson, G. (2007). Rapid reversed phase ultra-performance liquid chromatography analysis of the major cocoa polyphenols and inter-relationships of their concentrations in chocolate. (Retrieved from <http://pubs.acs.org.ezproxy.sibdi.ucr.ac.cr:2048/doi/full/10.1021/jf063277c>).
- Dand, R. (2011). 2 - agronomics of international cocoa production. *The international cocoa trade* (pp. 23–64). Elsevier. <http://dx.doi.org/10.1016/B978-0-85709-125-3.50002-1>.
- Dorenkott, M. R., Griffin, L. E., Goodrich, K. M., Thompson-Witrick, K. A., Fundaro, G., Ye, L., ... Neilson, A. P. (2014). Oligomeric cocoa procyanidins possess enhanced bioactivity compared to monomeric and polymeric cocoa procyanidins for preventing the development of obesity, insulin resistance, and impaired glucose tolerance during high-fat feeding. *Journal of Agricultural and Food Chemistry*, 62(10), 2216–2227. <http://dx.doi.org/10.1021/jf500333y>.
- Gu, L., Kelm, M. A., Hammerstone, J. F., Zhang, Z., Beecher, G., Holden, J., ... Prior, R. L. (2003). Liquid chromatographic/electrospray ionization mass spectrometric studies of proanthocyanidins in foods. *Journal of Mass Spectrometry*, 38(12), 1272–1280. <http://dx.doi.org/10.1002/jms.541>.
- Hansen, C. E., del Olmo, M., & Burri, C. (1998). Enzyme activities in cocoa beans during fermentation. *Journal of the Science of Food and Agriculture*, 77(2), 273–281. [http://dx.doi.org/10.1002/\(SICI\)1097-0010\(199806\)77:2<273::AID-JSFA40>3.0.CO;2-M](http://dx.doi.org/10.1002/(SICI)1097-0010(199806)77:2<273::AID-JSFA40>3.0.CO;2-M).
- Hatano, T., Miyatake, H., Natsume, M., Osakabe, N., Takizawa, T., Ito, H., & Yoshida, T. (2002). Proanthocyanidin glycosides and related polyphenols from cacao liquor and their antioxidant effects. *Phytochemistry*, 59, 749–758. [http://dx.doi.org/10.1016/S0031-9422\(02\)00051-1](http://dx.doi.org/10.1016/S0031-9422(02)00051-1).
- Hii, C. L., Law, C., Suzannah, S., Misnawi, & Cloke, M. (2010). Polyphenols in cocoa (*Theobroma cacao* L.). *Asian Journal of Food and Agro-Industry*, 2(4), 702–722.
- Hue, C., Brat, P., Gunata, Z., Samaniego, I., Servent, A., Morel, G., ... Davrieux, F. (2014). Near infra-red characterization of changes in flavan-3-ol derivatives in cocoa (*Theobroma cacao* L.) as a function of fermentation temperature. *Journal of Agricultural and Food Chemistry*, 62(41), 10136–10142. <http://dx.doi.org/10.1021/jf501070d>.
- Hue, C., Gunata, Z., Breyse, A., Davrieux, F., Boulanger, R., & Sauvage, F. X. (2016). Impact of fermentation on nitrogenous compounds of cocoa beans (*Theobroma cacao* L.) from various origins. *Food Chemistry*, 192, 958–964. <http://dx.doi.org/10.1016/j.foodchem.2015.07.115>.
- Jinap, M. S., Nazamid, S., & Jamilah, B. (2002). Activation of remaining key enzymes in dried under-fermented cocoa beans and its effect on aroma precursor formation. *Food Chemistry*, 78(4), 407–417. [http://dx.doi.org/10.1016/S0308-8146\(02\)00120-6](http://dx.doi.org/10.1016/S0308-8146(02)00120-6).
- Kim, J., Lee, K. W., & Lee, H. J. (2011). Cocoa (*Theobroma cacao*) seeds and phytochemicals in human health. In R. V. Preedy, R. R. Watson, & B. V. Patel (Eds.), *Nuts and seeds in health and disease prevention* (pp. 351–360). Elsevier. <http://dx.doi.org/10.1016/B978-0-12-375688-6.10042-8>.
- Kochhar, S., Hansen, C. -E., Juilleart, A. M., Wille, H. -J., Buyukpamukcu, E., Keely, B., & Goodall, D. M. (2004). *Cocoa flavor precursor peptides*. United States.
- Kongor, J. E., Hinneh, M., Van de Walle, D., Afoakwa, E. O., Boeckx, P., & Dewettinck, K. (2016). Factors influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile – A review. *Food Research International*, 82, 44–52. <http://dx.doi.org/10.1016/j.foodres.2016.01.012>.
- Kyi, T. M., Daud, W. R. W., Mohamad, A. B., Samsudin, M. W., Kadhum, A. A. H., & Talib, M. Z. M. (2005). The kinetics of polyphenol degradation during the drying of Malaysian cocoa beans. *International Journal of Food Science and Technology*, 40(3), 323–331. <http://dx.doi.org/10.1111/j.1365-2621.2005.00959.x>.
- Lin, L. -Z., Sun, J., Chen, P., Monagas, M. J., & Harnly, J. M. (2014). UHPLC-PDA-ESI/HRMS n profiling method to identify and quantify oligomeric Proanthocyanidins in plant products. *Journal of Agricultural and Food Chemistry*, 62(39), 9387–9400. <http://dx.doi.org/10.1021/jf501011y>.
- Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G., & Caligiani, A. (2014). Extraction, identification and semi-quantification of oligopeptides in cocoa beans. *Food Research International*, 63, 382–389. <http://dx.doi.org/10.1016/j.foodres.2014.03.046>.
- Mozzi, F., Ortiz, M. E., Bleckwedel, J., De Vuyst, L., & Pescuma, M. (2013). Metabolomics as a tool for the comprehensive understanding of fermented and functional foods with lactic acid bacteria. *Food Research International*, 54(1), 1152–1161. <http://dx.doi.org/10.1016/j.foodres.2012.11.010>.
- Nazaruddin, R., Seng, L. K., Hassan, O., & Said, M. (2006). Effect of pulp preconditioning on the content of polyphenols in cocoa beans (*Theobroma cacao*) during fermentation. *Industrial Crops and Products*, 24(1), 87–94. <http://dx.doi.org/10.1016/j.indcrop.2006.03.013>.
- Niemenak, N., Rohsius, C., Elwers, S., Omokolo Ndoumou, D., & Lieberei, R. (2006). Comparative study of different cocoa (*Theobroma cacao* L.) clones in terms of their phenolics and anthocyanins contents. *Journal of Food Composition and Analysis*, 19(6–7), 612–619. <http://dx.doi.org/10.1016/j.jfca.2005.02.006>.
- Nigam, P. S., & Singh, A. (2014). Encyclopedia of food microbiology. *Encyclopedia of food microbiology* (pp. 485–492). Elsevier. <http://dx.doi.org/10.1016/B978-0-12-384730-0.00074-4>.
- Nishioka, T., Nihei, Y., Ojima, Y., & Ikeda, T. (2012). MassBank: public mass spectral database – “chemically accurate” reference MS data-. Oral presentation in 19th international mass spectrometry conference (IMSC2012). Kyoto, Japan.
- Ortega, N., Romero, M. -P., Macià, A., Reguant, J., Anglès, N., Morelló, J. -R., & Motilva, M. -J. (2008). Obtention and characterization of phenolic extracts from different cocoa sources.

- (Retrieved from <http://pubs.acs.org.ezproxy.sibdi.ucr.ac.cr:2048/doi/abs/10.1021/jf8014415>).
- Ortega, N., Romero, M. -P., Macià, A., Reguant, J., Anglès, N., Morelló, J. -R., & Motilva, M. -J. (2010). Comparative study of UPLC–MS/MS and HPLC–MS/MS to determine procyanidins and alkaloids in cocoa samples. *Journal of Food Composition and Analysis*, 23(3), 298–305. <http://dx.doi.org/10.1016/j.jfca.2009.10.005>.
- Paizs, B., & Suhai, S. (2005). Fragmentation pathways of protonated peptides. *Mass Spectrometry Reviews*, 24(4), 508–548. <http://dx.doi.org/10.1002/mas.20024>.
- Payne, M. J., Hurst, W. J., Miller, K. B., Rank, C., & Stuart, D. A. (2010). *Impact of fermentation, drying, roasting, and Dutch processing on Epicatechin and Catechin content of cocoa beans and cocoa ingredients*. (Retrieved from <http://pubs.acs.org.ezproxy.sibdi.ucr.ac.cr:2048/doi/abs/10.1021/jf102391q>).
- Pereira, G. V. d. M., Soccol, V. T., & Soccol, C. R. (2016). Current state of research on cocoa and coffee fermentations. *Current Opinion in Food Science*, 7, 50–57. <http://dx.doi.org/10.1016/j.cofs.2015.11.001>.
- Porter, L. J., Ma, Z., & Chan, B. G. (1991). Cocoa procyanidins: major flavanoids and identification of some minor metabolites. *Phytochemistry*, 30(5), 1657–1663. [http://dx.doi.org/10.1016/0031-9422\(91\)84228-K](http://dx.doi.org/10.1016/0031-9422(91)84228-K).
- de Rijke, E., Out, P., Niessen, W. M. a., Ariese, F., Gooijer, C., & Brinkman, U. A. T. (2006). Analytical separation and detection methods for flavonoids. *Journal of Chromatography A*, 1112(1–2), 31–63. <http://dx.doi.org/10.1016/j.chroma.2006.01.019>.
- Roepstorff, P., & Fohlmann, J. (1984). Proposal for common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry*, 11, 601.
- Rusconi, M., & Conti, A. (2010). *Theobroma cacao* L., the food of the gods: a scientific approach beyond myths and claims. *Pharmacological Research: The Official Journal of the Italian Pharmacological Society*, 61(1), 5–13. <http://dx.doi.org/10.1016/j.phrs.2009.08.008>.
- Saltini, R., Akkerman, R., & Frosch, S. (2013). Optimizing chocolate production through traceability: a review of the influence of farming practices on cocoa bean quality. *Food Control*, 29(1), 167–187. <http://dx.doi.org/10.1016/j.foodcont.2012.05.054>.
- Sarmadi, B., Ismail, A., & Hamid, M. (2011). Antioxidant and angiotensin converting enzyme (ACE) inhibitory activities of cocoa (*Theobroma cacao* L.) autolysates. *Food Research International*, 44(1), 290–296. <http://dx.doi.org/10.1016/j.foodres.2010.10.017>.
- Sawada, Y., Nakabayashi, R., Yamada, Y., Suzuki, M., Sato, M., Sakata, A., ... Saito, K. (2012). RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry*, 82, 38–45. <http://dx.doi.org/10.1016/j.phytochem.2012.07.007>.
- Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., ... Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. *Therapeutic Drug Monitoring*, 27(6), 747–751.
- Voigt, J., Heinrichs, H., Voigt, G., & Biehl, B. (1994). Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chemistry*, 50(2), 177–184. [http://dx.doi.org/10.1016/0308-8146\(94\)90117-1](http://dx.doi.org/10.1016/0308-8146(94)90117-1).
- Voigt, J., Janek, K., Textoris-Taube, K., Niewianda, A., & Wöstemeyer, J. (2016). Partial purification and characterisation of the peptide precursors of the cocoa-specific aroma components. *Food Chemistry*, 192, 706–713. <http://dx.doi.org/10.1016/j.foodchem.2015.07.068>.
- Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., et al. (2013). HMDB 3.0 – the human metabolome database. *Nucleic Acids Research*, 41(D1), D801–D807 Jan 1. (23161693).
- Wishart, D. S., Knox, C., Guo, A. C., et al. (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research*, 37(Database issue), D603–D610 (18953024).
- Wishart, D. S., Tzur, D., Knox, C., et al. (2007). HMDB: the human metabolome database. *Nucleic Acids Research*, 35(Database issue), D521–D526 Jan. (17202168).
- Wolf, S., Schmidt, S., Müller-Hannemann, M., & Neumann, S. (2010). In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics*, 11(1), 148. <http://dx.doi.org/10.1186/1471-2105-11-148>.
- Wollgast, J., & Anklam, E. (2000). Review on polyphenols in *Theobroma cacao*: changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Research International*, 33(6), 423–447. [http://dx.doi.org/10.1016/S0963-9969\(00\)00068-5](http://dx.doi.org/10.1016/S0963-9969(00)00068-5).