



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS**

INSTITUTO DE BIOLOGÍA
BIOLOGÍA EVOLUTIVA

**RELACIONES ENTRE MARCADORES MOLECULARES Y LAS
CARACTERÍSTICAS DE CALIDAD DE GRANO DE *Theobroma cacao*
L. cv. CRIOLLO CULTIVADO EN EL SOCONUSCO, CHIAPAS**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

JOSÉ ALFREDO VÁZQUEZ OVANDO

TUTOR PRINCIPAL DE TESIS

DR. MIGUEL SALVADOR FIGUEROA, Instituto de Biología

COTUTOR DE TESIS

DR. FRANCISCO ELIZANDRO MOLINA FREANER, Instituto de Ecología

COMITÉ TUTOR

DR. JUAN SERVANDO NÚÑEZ FARFÁN, Instituto de Ecología

MÉXICO, D.F. *****I @C

2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión del Subcomité por Campo de Conocimiento de Biología Evolutiva y Sistemática del Posgrado en Ciencias Biológicas, celebrada el día 23 de marzo de 2015, se aprobó el siguiente jurado para el examen de grado de DOCTOR EN CIENCIAS del alumno VÁZQUEZ OVANDO JOSÉ ALFREDO con número de cuenta 511012026 con la tesis titulada: "Relaciones entre marcadores moleculares y las características de calidad del grano de *Theobroma cacao* L. cv. criollo cultivado en el Soconusco, Chiapas", realizada bajo la dirección del DR. MIGUEL SALVADOR FIGUEROA:

Presidente:	DR. DANIEL IGNACIO PIÑERO DALMAU
Vocal:	DR. FRANCISCO JAVIER ESPINOSA GARCÍA
Secretario:	DR. JUAN SERVANDO NÚÑEZ FARFÁN
Suplente:	DR. HÉCTOR BERNARDO ESCALONA BUENDÍA
Suplente	DR. ALEJANDRO NETTEL HERNANZ

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F. a 11 de junio de 2015.

M. del Coro Arizmendi Arriaga

DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



c.c.p. Expediente del (la) interesado (a).

AGRADECIMIENTOS

Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México por permitir el espacio para desarrollar estudios de doctorado y fomentar a través de su gente la generación de conocimiento.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT, México) por el apoyo económico bajo el esquema de beca (becario número 200343) otorgado para la realización de estudios de doctorado.

A los integrantes del Comité Tutoral, quienes guiaron el curso de la investigación y de mi formación como estudiante posgraduado.

Dr. Miguel Salvador Figueroa

Dr. Francisco Molina Freaner

Dr. Juan Núñez Farfán

AGRADECIMIENTOS A TÍTULO PERSONAL

Este documento es el resultado del trabajo colaborativo de muchas personas e instituciones. A TODAS GRACIAS POR SU APORTACIÓN TÉCNICA, INTELECTUAL Y MORAL.

Agradezco a la Universidad Nacional Autónoma de México, a su gente y a sus programas por fomentar la formación de seres humanos sensibles a las problemáticas del país y por aceptarme como estudiante en el Posgrado en Ciencias Biológicas.

Al Consejo Nacional de Ciencia y Tecnología de México por el apoyo económico.

A la Universidad Autónoma de Chiapas, especialmente al Centro de Biociencias (CenBio) que aportó los recursos materiales y humanos para el desarrollo experimental del proyecto y me permitió convivir en un espacio laboral de trabajo en equipo.

Agradezco inmensamente al director de la tesis, Dr. Miguel Salvador Figueroa. Agradezco su incisivo empeño en que realizara los estudios de doctorado, por revisar mi progreso académico, compartir su amistad conmigo e incorporarme en su grupo de trabajo.

A mi cotutor, Dr. Francisco Molina Frenaner y mi sinodal, Dr. Juan Núñez Farfán por escuchar y encausar de manera adecuada el curso de la investigación, por sus valiosos y atinados comentarios y por compartir conmigo su conocimiento y experiencia como investigadores.

A los revisores de la tesis, Dres. Daniel Piñero Dalmau, Juan Servando Núñez Farfán, Francisco Espinosa García, Héctor Bernabé Escalona Buendía y Alejandro Nettel Hernanz, por su cuidadosa aportación y sus críticas que enriquecieron a la misma.

A la Dra. María de Lourdes Adriano Anaya, por su apoyo sostenido y la enseñanza continua en todas áreas del quehacer de la investigación. También por su amistad. A los muchísimos colegas, estudiantes y amigos del CenBio que acompañaron mi etapa de estudiante y aportaron mucho a mi naciente carrera en la investigación. Al Dr. Isidro Ovando Medina, agradecimiento especial por compartir conmigo su experiencia, amistad y conocimiento en áreas para mí novedosas y también por sus aportaciones al manuscrito. A los IBT's Adriana Sánchez Gutiérrez, Julio Magaña Ramos, Lisbeth Chacón Martínez, Nidia Zulema Gutiérrez, Andy Erick Villarreal Cruz, Karen Dafne Castillo López y Julio Cesar Coutiño Montes por su invaluable apoyo técnico en las actividades de colecta de hojas y frutos y en el laboratorio de Genética de Poblaciones. Al MB Gamaliel Velázquez Ovalle por su apoyo en las colectas de campo y su desinteresado apoyo y amistad.

A la MB María Guadalupe de Gyves Córdova, por su apoyo en aspectos laborales y administrativos dentro del CenBio y su valiosa amistad.

A los colegas, con quienes compartimos historias propias del quehacer cotidiano, MB Sonia Ruiz González, MB Rito Coronel Niño, IBT Gemelli López Martínez, MB Victor Albores Flores, Dr. Rodolfo Torres de los Santos, MC Manuel Rincón Rabanales.

A la MC Nancy Gálvez Reyes por su invaluable apoyo en las gestiones administrativas en la coordinación y por su amistad.

Al Dr. Héctor Escalona Buendía de la Universidad Autónoma Metropolitana Iztapalapa, por su asesoría y apoyo en los estudios sensoriales y recibirme en su laboratorio.

Al Dr. David Betancur Ancona de la Universidad Autónoma de Yucatán, por sus revisiones y observaciones a los manuscritos, por recibirme en su facultad y por su amistad.

Agradezco también a la Dra. Alma Rosa González Esquinca de la Universidad de Ciencias y Artes de Chiapas por su asesoría y apoyo en los procesos administrativos como coordinadora del programa en Chiapas. En este mismo sentido, agradezco a la MC Christian Anabi Riley Saldaña por su invaluable apoyo.

Agradezco a cada uno de los cacocultores que nos permitieron ingresar a sus parcelas, extraer hojas y frutos y en muchos casos hasta nos invitaron a tomar pozol o agua de mango. De manera especial al Sr. Samuel Guillén Díaz, cacaocultor apasionado, quien nos proporcionó material vegetal y quien nos acompañó a los recorridos por las parcelas de Soconusco.

Agradecimiento especial a mis familias, por su incondicional apoyo moral. A mi madre, Paula Ovando Domínguez con toda la admiración por la familia que creaste y el ser humano que me ayudaste a ser. A mis hermanos, cuñados y sobrinos todos, por ser aliciente al preguntar frecuentemente de mi progreso como estudiante de posgrado. A Marisol Martínez Sánchez mi compañera de vida, por su valioso apoyo en tantas cosas, por las palmaditas en la espalda, por ser mamá y papá cuando no he estado y por nunca cuestionar mi pasión por la investigación. A mi hija Arlette Valeria Vázquez Martínez por crecer conmigo, entender mis ausencias y enseñarme otra forma de ser feliz.

ÍNDICE

	Página
Resumen	1
Abstract	3
CAPÍTULO 1. Introducción	5
CAPÍTULO 2. Potencial de los marcadores moleculares para el rescate de individuos de <i>Theobroma cacao</i> L. de alta calidad	10
Resumen, Abstract	11
Introducción	12
Situación actual del cultivo de cacao	14
La utilidad de los marcadores moleculares para comprender los secretos del cacao	16
Conclusiones	25
Referencias	25
CAPÍTULO 3. De la estimación de la diversidad y estructura genética de árboles de cacao con fenotipo criollo cultivados en la región Soconusco, al sur de Chiapas México	32
Genetic identification of <i>Theobroma cacao</i> L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico	33
Abstract	33
Introduction	34
Material and Methods	35
Results	38
Discussion	40
Acknowledgments	42
References	42
CAPÍTULO 4. Importancia de alcaloides y polifenoles en el sabor y aroma del cacao	44
Alcaloides y polifenoles del cacao: mecanismos que regulan su síntesis y sus implicaciones en el sabor y aroma	45
Resumen, Summary, Resumo	45
Introducción	46
Factores que afectan la calidad del cacao	47
Alcaloides del cacao	48
Polifenoles del cacao	51
Aroma y sabor del cacao, efecto del procesamiento poscosecha	56
Conclusión	61
Agradecimientos	62
Referencias	62
CAPÍTULO 5. Agrupamiento de muestras de semillas de cacao empleando métodos de estadística multivariada	67
Classification of cacao beans (<i>Theobroma cacao</i> L.) of Southern Mexico	68

based on chemometric analysis with multivariate approach	
Abstract	68
Introduction	68
Materials and Methods	69
Results and discussions	70
Conclusions	75
Acknowledgments	77
References	78
CAPÍTULO 6. Descriptores de olor y sabor en semillas de cacao de Soconusco, Chiapas, México	80
Sensory descriptors of cocoa beans from cultivated trees of Soconusco, Chiapas, Mexico	81
Abstract	81
Introduction	81
Materials and methods	82
Results and discussions	83
Conclusions	85
Acknowledgments	85
References	85
CAPÍTULO 7. De las relaciones entre los marcadores moleculares SSR y características químicas y sensoriales de semillas de cacao	87
Asociación entre marcadores moleculares SSR y características químicas y sensoriales de muestras de cacao cultivado en Soconusco, Chiapas; México	88
Introducción	88
Materiales y métodos	89
Resultados y discusión	91
Referencias	97
CAPÍTULO 8. Discusión General y Conclusiones	101
Literatura citada	104

RESUMEN

Por su gran importancia industrial, el cacao (*Theobroma cacao* L.) ha aumentado su cultivo, sobre todo en países de África y Asia. Esta planta nativa de América fue llevada al Viejo Mundo por los colonizadores españoles. La clasificación convencional de esta especie, basada en características morfo-geográficas, reconoce a tres grupos o variedades: Criollo, Forastero y Trinitario, siendo el primero el menos producido a nivel mundial, pero del que se obtienen las almendras de mayor calidad sensorial. El registro más antiguo de uso del cacao por los seres humanos data de hace casi 4000 años, en Soconusco, en el estado de Chiapas, México. Considerando que las variedades Forastero y Trinitario fueron introducidas a Mesoamérica hace pocos siglos, se hipotetiza que el germoplasma inicialmente domesticado era Criollo y que en las actuales plantaciones de Soconusco persisten genotipos ancestrales de cacao Criollo. Con el objetivo de caracterizar el germoplasma representado por los árboles que fenotípicamente son identificados por los cacaocultores de Soconusco como tipo Criollo se realizaron estudios moleculares, fisicoquímicos y sensoriales de las semillas.

Para el estudio molecular se muestrearon 108 árboles (de 7 demos o subpoblaciones) en plantaciones de la región Soconusco y se analizó su diversidad y estructura genética empleando 12 marcadores microsatélites (SSR por su abreviación de *short sequence repeat*). Además se incluyeron como controles, 10 accesiones de la variedad Criollo de bancos de germoplasma. Los SSRs amplificaron un total de 74 alelos. Los mismos revelaron baja heterocigosidad observada ($H_o = 0.28$) y deficiencia de heterocigotos con un valor del índice de fijación $F = 0.50$, aunque moderada diversidad alélica estimada por el índice de Shannon, ($I = 0.97$ para todas las poblaciones, todos los *loci*). No se encontró evidencia de agrupamiento genético por localización geográfica, de acuerdo a la prueba de Mantel ($R_{xy} = 0.54$, $P = 0.120$). Empleando el software *Structure* se determinó la formación de 2 grupos genéticos ($K = 2$). Un análisis posterior de asignación, agrupó a 37 árboles (incluidos 10 controles) en un grupo genético (Criollo), 54 individuos en un segundo grupo (presumiblemente Forasteros), y las 27 plantas restantes (entremezclados) que no correspondieron (<90% asignación) a ninguno de los otros dos grupos. Estos individuos posiblemente corresponden al genotipo Trinitario (híbrido Criollo x Forastero). Los individuos agrupados junto con los controles (con alta ancestría de Criollo) podrían investigarse posteriormente como potenciales reservorios de genes para programas de mejoramiento centrado en la calidad sensorial.

Por otro lado, se analizaron las variables físicas (peso, longitud, ancho y diámetro) y químicas (humedad, cenizas, grasa, proteínas, composición de ácidos grasos), además de contenido de polifenoles y capacidad antioxidante de las semillas de 45 plantas y una planta control (*T. bicolor*, conocida en Soconusco como "pataste"). Con estos datos se realizó un agrupamiento como pre-requisito para realizar la evaluación sensorial. Todas las variables se analizaron

mediante análisis multivariados (análisis de componentes principales-ACP, K-medias y análisis discriminante-AD). Las variables morfológicas fueron de poca utilidad y dieron lugar a errores de clasificación de la muestra control, por lo que fueron descartados para los análisis posteriores. Entre las variables químicas, las que más contribuyeron y explicaron la varianza total del análisis de componentes principales (ACP) fueron los ácidos grasos palmitoleico, palmítico, oleico y esteárico, así como el contenido de humedad y de polifenoles. Con este enfoque de análisis, seis caracteres fenotípicos fueron útiles para clasificar las muestras en siete grupos homogéneos. El origen geográfico de las muestras no se correlacionó con la agrupación quimiométrica, lo que muestra posible dispersión de materiales altamente relacionados (como también demostró la alta homocigosis del estudio molecular).

Con base en el agrupamiento anterior, mediante un panel de jueces entrenados se caracterizó el olor y sabor de las mismas 45 muestras de cacao (conglomeradas en siete grupos G1-G7). Se evaluaron cuatro atributos de gusto (dulce, amargo, ácido y astringente) y nueve de olor (chocolate, nueces, avellana, dulce, ácido, tostado, picante, húmedo y mal olor). Se detectó una muestra (G7) con mayores puntuaciones en gusto dulce y olores dulce y nuez, así como alta asociación entre estos descriptores y la muestra, analizada mediante ACP. Del mismo modo se identificaron muestras que presentaron altas puntuaciones para olores no deseables en cacaos como mal olor y humedad, y relacionadas mediante ACP con olor a tostado y gusto astringente (G2 y G4). Con base en esta puntuación, las muestras fueron catalogadas en orden descendente por su calidad sensorial en G7 > G5 > G6 > G3 > G1 > G4 > G2.

Finalmente, se analizó la utilidad de los *loci* marcadores obtenidos del estudio molecular, y su relación con los caracteres fenotípicos. Mediante análisis de regresión multivariada se encontraron correlaciones altamente significativas ($p < 0.05$) entre bandas obtenidas de los marcadores SSR y los contenidos de ácidos grasos (mirístico, palmitoleico, palmítico, esteárico, eicosanoico) así como los contenidos de grasa y proteína. Del mismo modo se hallaron asociaciones con los caracteres sensoriales amargor, acidez, olor a nuez, olor ácido, olor a moho y olor a tostado. En todos los casos se anotan los modelos matemáticos que permiten emplear entre tres y diez bandas SSR en conjunto con fines predictivos de caracteres fenotípicos.

La investigación mostró que: a) menos de la mitad de los árboles fenotípicamente Criollos tienen alta ancestría Criollo, b) que el germoplasma de cacao Criollo tiene de moderada a baja diversidad genética, c) que físicoquímica y sensorialmente está diferenciado, con solo algunos grupos de árboles de alta calidad de grano, y que d) existen asociaciones entre algunos marcadores moleculares microsatélites y características fenotípicas, las cuales podrían aprovecharse en estudios futuros de mejoramiento del ancestral cultivo de cacao.

ABSTRACT

In the past, Spaniards carried cocoa (*Theobroma cacao* L.) to the Old World, and nowadays the acreage this crop is increasing in Africa and Asia due to their industrial importance. The conventional classification of cocoa assumes three races or varieties: Criollo, Forastero and Trinitario. Although Criollo is not widely grown, produces high quality beans which are used only for luxury confectionery industry. In the Soconusco, Chiapas its use dates from nearly 4000 years; however a few centuries ago the varieties Forastero and Trinitario were introduced in Mesoamerica. Several supported hypotheses state that the original cocoa germplasm was Criollo, thus, in the current Soconusco plantations these ancestors could persist. Therefore, this study aimed to characterize the germplasm present in trees phenotypically identified as Criollo, thus we performed molecular, physico-chemical and sensorial studies on seeds.

For molecular research, we used 108 cocoa trees (from 7 demes) Soconusco plantations and ten accessions (from germplasm bank) as control. Diversity and genetic structure was analyzed by using 12 SSR. SSRs amplified 74 alleles in total, which showed low heterocigosity (observed $H_o=0.28$) and deficiency of heterocigotes (Fixation index $F=0.50$) and moderate allelic diversity with the Shannon index ($I=0.97$ in all populations). Our results showed no evidence of genetic grouping by geography (Mantel, $R_{xy} = 0.54$, $P = 0.120$). However, the *Structure* software defined two genetic groups ($K=2$). Then by assignment test, 37 trees (controls included) were grouped in one cluster (Criollo), 54 individuals in a second cluster (presumably Forastero) and the other 27 plants (admixed) did not fit (" $<90\%$ assignment) in any of the former groups. We considered the admixed group as Trinitario (hybrid Criollo x Forastero). Individuals grouping with controls (Criollo high ancestry) could be investigated further as potential gene reservoirs in breeding programs focused on sensory quality.

On the other hand, we studied physical (weight, length, width and bean circumference) and chemical (moisture, ash, lipids, proteins, fatty acids), polyphenol content and antioxidant activity in seeds from 45 plants and one control sample (patate *T. bicolor*). We analyzed our results with multivariate analysis (principal component analysis-PCA, K-means, and discriminant analysis-DA). The physical traits were not useful when incorporated into the analysis altogether with chemical variables and led to wrong grouping; furthermore, PCA did not separate *T. bicolor* of the cocoa samples, so they physical traits were discarded for further analysis. The chemical variables that best differentiate clusters were moisture content, polyphenols content and palmitoleic, palmitic,

oleic and stearic fatty acids. With this analysis approach, these variables separated our samples in seven homogeneous groups. The geographical origin of the samples did not correlate with chemometric grouping, which shows the possible dispersion of highly related materials (as also demonstrated with the high level of homozygosity in molecular study).

Based on this last grouping, panels of trained judges characterized odor and flavor in 45 cacao samples (grouped in seven groups G1 – G7). Four flavor variables were evaluated (sweet, sour, acid and astringent) and nine for odor (chocolate, nutty, hazelnut, sweet, acid, toasted, spicy, musty and off-odor). G7 was the sample with higher scores in sweet flavor and sweet and nutty odor, besides a high association between descriptors and the samples (PCA analysis). We also identified samples with high scores for undesirable odors (G2 and G4, based on PCA analysis). Based on sensorial quality, samples were characterized in descending order $G7 > G5 > G6 > G3 > G1 > G4 > G2$.

Finally, with our results, we also assayed the utility of related *loci* markers with phenotypic characters, by using a multivariate regression analysis (MRA). We found significant correlations ($p < 0.05$) between SSR and fatty acids (myristic, palmitoleic, palmitic, stearic, eicosanoic), fatty acid content and protein. We also registered associations between SSR and sensorial characters, such as bitter, acidity, and nutty, acid, musty, toasted odors, in order to predict phenotypical characters in the future. In all cases the mathematical models to use a set of three to ten bands SSR are proposed.

The research showed that: a) less than half of Criollo-like trees have high Criollo ancestry, b) Criollo germplasm showed low genetic diversity, c) cacao germoplasm are differentiated by physical-chemical and sensorial traits, and d) molecular markers and phenotypical characteristics are highly correlated, which could be useful for cacao management improvement and breeding.

CAPÍTULO 1

INTRODUCCIÓN

Se estima que la industria del chocolate genera ganancias por 110 000 millones de dólares al año. La principal materia en la elaboración de chocolates es el cacao, que se obtiene de la planta del cacaotero (*Theobroma cacao* L.), y es la especie de mayor importancia del género *Theobroma* (Schnell et al., 2005). Linneo asignó este el nombre al género *Theobroma* posiblemente al conocer las leyendas Amerindias del origen de esta planta, las cuales argumentan que fue un regalo directo de los dioses. A este género de la familia Malvacea, subfamilia Sterculioidae (antes Stercualicea) pertenecen 22 especies descritas. Otras especies de este género (*T. bicolor*, *T. angustifolium*, *T. grandiflorum*, *T. mamosum*) tienen cierta importancia a nivel local en algunos países de Centro y Sudamérica, mientras que el resto se encuentran en estado silvestre en el sur del mismo continente. Esta diversidad de especies aunada a estudios moleculares que apuntan en ese sentido, sugieren que la Amazonia es el centro de origen de *T. cacao*. Aún con lo anterior, en esa región de América no existe evidencia del uso de las semillas de esta planta. Por el contrario, en la región geográfica que ocupó Mesoamérica (desde la parte central de México hasta Nicaragua) existen vastos vestigios del uso del cacao con fines alimentarios y/o medicinales (Henderson et al., 2007). Lo anterior da cuenta de que las civilizaciones Mesoamericanas domesticaron esta planta. En la región conocida como Paso de Amada en la zona sur de México (ocupada por los pueblos Mokayas e Izapeños), además de otras regiones de Guatemala y del sur de lo que hoy es Veracruz se han datado tiestos con restos de teobromina (el principal alcaloide del cacao) hacia los años 1500-1900 a.C. (Powis et al., 2011).

De forma tradicional, las plantas de cacao están clasificadas en tres grupos morfo-geográficos, Criollo, Forastero y Trinitario (Cheesman, 1944). La variedad Forastero [que actualmente se divide en varios grupos genéticos, (Motamayor et al., 2008)], incluye plantas que presentan tolerancia o resistencia a enfermedades (moniliasis, escoba de bruja, mancha negra), además de alta productividad agronómica. Los frutos de estas plantas presentan surcos poco pronunciados y coloraciones generalmente verdes que tienden a amarillo en la madurez. Las semillas de la variedad Forastero son consideradas de baja calidad sensorial, pues carecen de sabores y aromas agradables, además que su contenido de polifenoles y teobromina es mayor que la variedad Criollo.

Por el contrario, se ha documentado que las plantas de la variedad Criollo son menos tolerantes a las enfermedades y de menor productividad agronómica. Los frutos Criollo, tienden a presentar coloración roja, con surcos muy pronunciados, de forma alargada y terminación en punta, pulpa más dulce, semillas blancas como consecuencia del menor contenido de polifenoles y teobromina, aunque con mayor contenido de cafeína comparado con los Forasteros. Sin embargo, a diferencia de la variedad Forastero, las semillas de cacao Criollo son de alta calidad sensorial, pues son más dulces, presentan aromas y sabores atractivos para mercados gourmet y por consecuencia son más cotizados. Sobre el aspecto sensorial, se hipotetiza como la principal razón que alentó la domesticación de la variedad Criollo por los habitantes de Mesoamérica (Motamayor et al., 2002).

Después del encuentro cultural, los españoles documentaron las principales áreas de producción de cacao en Mesoamérica (Figura 1). Es importante anotar que en México las áreas de mayor producción corresponden a una región de lo que hoy es el estado de Tabasco y la región Soconusco al sur del actual estado de Chiapas. Otras regiones de alta producción se localizaban en Guatemala, El Salvador y Honduras. Actualmente, la producción de cacao en México sigue un patrón similar al descrito hace 500 años, las áreas de mayor producción (99.7 %) están concentrados en los estados de Tabasco y Chiapas (Soconusco), y con una contribución casi simbólica, los estados de Guerrero y Oaxaca.



Figura 1. Distribución del cultivo de cacao Criollo en el año 1500 (tomado de Crown y Hurst, 2009).

La situación de producción en la época reciente en México ha pasado por un período de producción continua desde 1993 hasta 2008 con un dramático descenso en el año 2000.

Sin embargo a partir de 2008 se ha registrado un significativo aumento en la producción. En 2012, México produjo 83000 ton de almendras de cacao, con los que ocupó el 8vo lugar mundial (FAO, 2014). Los cacaos que actualmente se cultivan en esta región de México, son los resultantes de los cacaos Criollo ancestrales cultivados por los pobladores de Mesoamerica, de los cuales existen reportes de su presencia (Whitkus et al. 1998; Motamayor et al. 2002) y de las introducciones (no documentadas) que posiblemente ocurrieron a lo largo de 500 años. Se tiene conocimiento que a mediados del siglo XX el gobierno mexicano financió programas para la introducción de plantas con mayores ventajas agronómicas, aunque se desconocen los genotipos empleados para tal efecto. El producto de tales recombinaciones, las prácticas agronómicas (propagación por semilla) y la naturaleza altamente alógama promovida por autoincompatibilidad, aunado a los microclimas de la parte costera del estado de Chiapas, pueden ejercer efecto sobre la composición genética de los individuos plantados, contribuyendo con la diversidad mundial del cacao, la cual se ha estimado en análisis predictivos que puede llegar a ser de hasta 14 000 genotipos (Turnbull *et al.*, 2003) y esto, repercutir sobre las características químicas y sensoriales de las almendras obtenidas del procesado de los frutos maduros. Derivado de esto, actualmente se observa un fenómeno similar a lo que ocurre en la mercadotecnia de bebidas, es decir, algunas plantaciones adquieren valor por su “origen” o características “varietales”, asociados con características sensoriales muy particulares y del cual el sector cacaocultor de la región Soconusco no está totalmente ajeno, como demuestra la adquisición de la cosecha 2011 de cacao por un sector de comerciantes italianos (Vázquez-Ovando *et al.*, 2012).

La presente investigación sustenta que la variabilidad fenotípica presente en las plantaciones del estado de Chiapas puede estar relacionada directamente con variación genética, y ésta resultar útil para localizar material vegetal con características superiores, las cuales pueden ser caracterizadas mediante marcadores moleculares. El uso de estos marcadores permitirá caracterizar accesiones, otorgar identidad y coadyuvar en los programas de mejoramiento genético. Para caracterizar tal calidad se debe recurrir a descriptores sensoriales (sabor, aroma), los cuales pueden ser asociados con caracteres químicos de las semillas o marcadores moleculares de ADN.

El objetivo de la presente investigación fue establecer las relaciones entre los marcadores moleculares microsatélites y las características de calidad de *Theobroma cacao* L. Criollo de plantas cultivadas en la región Soconusco, al sur del estado de Chiapas. Para alcanzar

este objetivo se establecieron tres tipos de análisis experimentales, el primero con un enfoque poblacional se basó en analizar la diversidad y estructura genética poblacional de las plantas colectadas con fenotipo Criollo. El segundo, para analizar caracteres fenotípicos (físicos y químicos) relacionados de manera directa con la calidad final de semillas, obtenidas de las accesiones muestreadas. El tercer análisis se centró en caracterizar la calidad de las mismas semillas desde la perspectiva de la evaluación sensorial.

Con esta estructura, se inició revisando el conocimiento actual acerca de la aplicación y utilidad de los marcadores moleculares en diversos estudios que tienen como modelo a la planta del cacaotero. Se enfatizó en la revisión bibliográfica en las ventajas del uso de marcadores moleculares codominantes del tipo secuencias simples repetidas (SSR) o microsatélites y de la baja aplicación que han tenido para ser asociados con caracteres fenotípicos de calidad (química o sensorial). Lo anterior como consecuencia de que la mayoría de los trabajos se centran en asociarlos con caracteres fenotípicos de rendimiento o tolerancia a enfermedades. Esta revisión se incluye en el Capítulo posterior a esta Intriducción (Capítulo 2).

Empleando los marcadores moleculares microsatélites, se obtuvo el perfil de bandas de los individuos incluidos en el estudio y con éstas se analizó la diversidad y estructura genética poblacional. El Capítulo 3 de esta tesis, incluye el análisis y discusión de los resultados con base en el pasado de cultivo de esta planta en la región y se hipotetiza sobre los grupos genéticos revelados por este análisis.

El Capítulo 4 del documento se centra en analizar desde la perspectiva teórica la contribución de dos tipos de moléculas responsables del sabor del cacao, los compuestos polifenólicos y los alcaloides. Se enfatiza en las rutas de biosíntesis de éstos compuestos, la regulación genética de éstas, de los factores que pueden influir en su contenido y, de los cambios químicos a los que están expuestas durante el procesamiento de las semillas del cacao.

Para caracterizar la calidad de las semillas de cacao, se realizaron dos tipos de análisis, fisicoquímicos y sensoriales. Primeramente se caracterizó la morfología y composición química de las semillas. Estos datos fueron analizados con un enfoque multivariado y, derivado de esto se formaron grupos de muestras que comparten características adaptativas. En el Capítulo 5 del documento se aborda este análisis y se discute sobre la baja contribución de factores genéticos, geográficos y/o ambientales sobre el

agrupamiento. En cambio, se incorpora el término “terroir” para hacer referencia a cacaos locales con características particulares.

La caracterización sensorial de las semillas de cacao se plasma en el Capítulo 6. En este análisis se generaron descriptores de sabor y olor y se cuantificó la magnitud asignada para cada muestra por medio de jueces previamente entrenados. Se discute sobre la presencia de descriptores poco reportados para muestras de cacao de otras latitudes.

En el Capítulo 7 de esta tesis se presentan los resultados del análisis de asociación realizado para integrar los resultados del genotipado de marcadores SSR y las características fenotípicas. Se hace especial énfasis en el uso de correlaciones multivariadas para modelar y predecir contenidos de moléculas químicas y valores de características sensoriales con base en el perfil de bandas polimórficas.

Finalmente, en el Capítulo 8 se integra una discusión general de los planteamientos generales abordados en los capítulos así como de los nuevos cuestionamientos que surgen a partir de los hallazgos de la investigación.

CAPÍTULO 2

Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad

En este capítulo se muestran los aspectos generales del cultivo de cacao, los retos que enfrenta la producción y de la importancia económica y social que representa como sistema de producción. Se discute sobre el empleo de los marcadores moleculares (principalmente basados en ADN) con diversos enfoques, a decir, en genética poblacional, asociados a caracteres cuantitativos (fundamentalmente relacionados con resistencia a enfermedades y rendimiento). Se hace énfasis particular en el potencial de los marcadores microsátelites (SSR) para ser empleados como herramientas en los programas de mejoramiento con interés en características relacionadas con la calidad sensorial.

Esta revisión bibliográfica se publicó en la revista BioTecnología (Sociedad Mexicana de Biotecnología y Bioingeniería, ISSN 0188-4786). La cita es: Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J, Salvador-Figueroa M (2012) Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad. BioTecnología 16(1): 36-56.

Artículos

Potencial de los Marcadores Moleculares para el Rescate de Individuos de *Theobroma cacao* L. de Alta Calidad

Alfredo Vázquez-Ovando^{*1,4}, Francisco Molina-Freaner², Juan Nuñez-Farfán³, Miguel Salvador-Figueroa¹

¹Centro de Biociencias, Universidad Autónoma de Chiapas. Carretera a Puerto Madero km. 2. Tapachula, Chiapas 30700. Tel (fax) 52 962 6427972. E-mail: jose.vazquez@unach.mx

²Departamento de Ecología de la Biodiversidad, Instituto de Ecología, Unidad Hermosillo. Universidad Nacional Autónoma de México. Hermosillo, Sonora.

³Departamento de Ecología Evolutiva, Instituto de Ecología. Universidad Nacional Autónoma de México. Ciudad Universitaria, México D.F.

⁴Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México. Ciudad Universitaria, México D.F.

RESUMEN

Aunque originario de América tropical, en la actualidad el cacao (*Theobroma cacao* L.), planta alógama, se cultiva en muchas regiones del mundo. Originalmente, los programas para obtener variedades "mejoradas" estuvieron basados en marcadores morfológicos los cuales, poco a poco, se han sustituido cada vez más por marcadores moleculares. Se discuten aquí las principales aplicaciones de las herramientas de la biología molecular en la obtención de mapas de ligamiento de este importante cultivo, loci de caracteres cuantitativos implicados en mayor rendimiento agronómico y resistencia a enfermedades, así como en la secuenciación del genoma cuya reciente publicación apertura una importante oportunidad de trabajo dirigido. Se hace énfasis en la falta de información y de investigación en el uso de estas herramientas para la búsqueda de marcadores asociados de manera directa con las características de calidad sensorial de las almendras, siendo ésta una de las razones principales que incentivaron su visionaria domesticación hace unos 4 000 años y su amplia distribución alrededor del mundo hace unos 500 años. Conocer a los individuos que muestran estas características resulta importante, ya que su posible extinción por cuestiones sociales o enfermedades, conllevaría pérdida de información científica y tecnológica valiosa que puede ser empleada con fines de conservación y explotación.

Palabras clave: calidad sensorial, chocolate, variedad Criollo, Soconusco

ABSTRACT

Cacao (*Theobroma cacao* L.) is an allogamous plant native to tropical America, but nowadays it is grown in tropical regions across the world. Originally, programs to obtain "improved" cultivars were based on morphological markers. However, there is a trend to use molecular markers. Here we discuss the main applications of molecular biology tools to obtain linkage maps of this important

Artículos

crop, to identify quantitative trait loci involved in major agronomic performance and disease resistance, as well as genome sequencing whose recent publication opens an important opportunity for specific investigations. We emphasize on the lack of information and research on the use of these tools to find markers associated directly with the sensory quality characteristics of almonds, which is one of the main reasons that encouraged its visionary domestication 4 000 years ago and its wide distribution around the world 500 years ago. Knowledge of the individuals who exhibit these characteristics is important, because it's possible extinction by social or diseases entail loss of valuable scientific and technological information that can be used for conservation and exploitation.

Key words: sensory quality, chocolate, Criollo cultivar, *Soconusco*

INTRODUCCIÓN

Theobroma cacao L. (cacao) es el cultivo de mayor importancia comercial a nivel mundial de las 22 especies que comprenden el género *Theobroma* (Schnell *et al.*, 2005); otras que también se comercializan pero solo a nivel local son *T. bicolor* Humb. y Bonpl. (pataste o pataxte), *T. grandiflorum* Schum. (cupuazú o copuazú) y *T. angustifolium* Moçño y Sessé (cacao de montaña o de mono) (Cervantes-Martínez *et al.*, 2006). Además de la relevancia industrial, el cacao es medular también desde la perspectiva social, pues su cultivo está vinculado, por sus requerimientos edafoclimáticos, a países, básicamente en desarrollo localizados en la franja ecuatorial de América, Asia y África y del que dependen una gran cantidad de agricultores (Efombagn *et al.*, 2007).

La importancia del cacao radica en que, de éste árbol se obtienen frutos (mazorcas) de las cuales, se extraen de 30 a 50 semillas (almendras) por mazorca, que son utilizadas ampliamente en las industrias alimentaria, farmacéutica y cosmética; las cuales demandan manteca de cacao (grasas) y torta de cacao o cocoa (obtenido al extraer las grasas) o bien licor de cacao (pasta de

cacao) para la elaboración de chocolates. Esta industria genera divisas por unos 73 000 millones de dólares (Ploetz, 2007) y de esta actividad dependen unos 80 000 empleos en todo el mundo (Lanaud *et al.*, 2009).

Desafortunadamente su permanencia como sistema de producción, sobre todo en el Neotrópico, se ve amenazada por enfermedades (Brown *et al.*, 2007; Clement *et al.*, 2003a; Queiroz *et al.*, 2003) que están onllando a quienes practican esta actividad a sustituir su cultivo o, en el mejor de los casos establecerlos con nuevos individuos que ofrezcan resistencia a los embates de los antagonistas. Esta característica que parece ser ventajosa desde la perspectiva de la producción, conlleva dos marcados aspectos negativos: 1). la pérdida del ecosistema existente en los cultivos con la consecuente disminución en la diversidad biológica, y aumento en el potencial de erosión de suelos y, 2). demerita la aceptación sensorial en los mercados más exigentes, puesto que los genotipos con mayor tolerancia o resistencia a enfermedades generalmente están limitados en características de calidad sensorial de las almendras. Enfatizando en el segundo aspecto, los criterios de selección

Artículos

más difundidos en los programas de mejoramiento son, precisamente tolerancia o resistencia a enfermedades (principalmente a moniliasis [*Moniliophthora roreri* Evans, Stalpers, Samson & Benny], mancha negra [*Phytophthora* spp.] y escoba de bruja [*Moniliophthora perniciosa* Stahel]) (Phillips-Mora *et al.*, 2005), el rendimiento agronómico (Irizarry & Coenaga, 2000) y solo en menor medida, la calidad de grano.

Las almendras de cacao de calidad, provienen de la variedad Criollo que, a diferencia del cultivar Forastero y del Trinitario (obtenido por la recombinación de los dos primeros), fue domesticada (Whitkus *et al.*, 1998) y empleada como materia prima en la alimentación de los pueblos precolombinos de Centroamérica hace unos 3 800 años (Powis *et al.*, 2011). El pueblo Olmeca (1800 a.C.-100 a.C), ubicado en lo que hoy es el estado de Veracruz fue el que originalmente cultivo esta planta (Powis *et al.*, 2011), después los Mokayas e Izapeños (1100 a.C.-1200 d.C.) en la región Soconusco (Chiapas) (INAH, 2007) y Mayas en la selva Lacandona, (Chiapas) Tabasco y península de Yucatán (Whitkus *et al.*, 1998). El empleo como moneda de las semillas de cacao da cuenta de la importancia social, religiosa y económica que llegó a tener este cultivo.

El hallazgo de descendientes aislados de cacaos criollos en la región poblada por los pueblos precolombinos, hasta nuestros días (Motamayor *et al.*, 2002; Whitkus *et al.*, 1998) muestra que, pueden haber fungido como reservorio genético, y al recombinarse con

las accesiones introducidas a mediados del siglo pasado, dieron origen a gran parte de la variabilidad presente en los cultivos de Mesoamérica. Lo anterior convierte a las plantaciones asentadas en esta parte del mundo, junto con las características edafoclimáticas, en candidatas a conservar y exhibir individuos con la calidad requerida por los mercados más exigentes. Daste como muestra el reciente interés de la comunidad europea al adquirir parte importante de la cosecha de cacao programada para el año 2012 y reconocerlo como de uno de los mejores cacaos del mundo (Chiapas.gob, 2010).

Como anteriormente se comentó, los programas de mejoramiento del cacao enfocaron su atención en características fenológicas de interés agronómico las cuales, están fuertemente influenciados por el ambiente (Azofeita-Deigado, 2006). A partir de la década de los ochenta del siglo pasado se empezaron a incorporar marcadores moleculares como herramientas auxiliares en los estudios y programas relacionados con esta planta. Se han empleado isoenzimas (Lanaud, 1986), AFLP (Lerceteau *et al.*, 1997; Perry *et al.*, 1998; Queiroz *et al.*, 2003), RAPD (Lerceteau *et al.*, 1997; Moreno *et al.*, 2004; Whitkus *et al.*, 1998), y SSR ó microsatélites (Trognitz *et al.*, 2011; Zhang *et al.*, 2006a; 2006b) con el objetivo de asociarlas a características de interés comercial (detallado más adelante) como el rendimiento agronómico y resistencia a enfermedades. Dichos marcadores también se han empleado para reducir la redundancia

Artículos

y etiquetado erróneo, conocer la estructura de las poblaciones tanto en bancos de germoplasma como *ex situ*, así como para analizar la genealogía de las accesiones.

Pese al gran número de trabajos que emplean marcadores moleculares, muy poco se ha abordado sobre su uso y la asociación de éstos con la calidad que ostentan las almendras de cacao. En este trabajo se hace un repaso sobre la importancia del cacao, los grandes avances para descifrar los secretos guardados en su genoma, se particulariza en los hallazgos con el uso de los marcadores moleculares asociados con diversas características de interés en el cacao y se discute acerca del potencial de éstos para ser correlacionados con las características de calidad.

SITUACIÓN ACTUAL DEL CULTIVO DE CACAO

Para el año 2009, se produjeron, a nivel mundial, poco más de 4.2 millones de toneladas de grano de *T. cacao*, reporte muy similar al obtenido los dos años anteriores (2007-2008). México ocupó el lugar trece en producción con 22 680 toneladas (en 2009), lo que representó un ingreso de 23.5 millones de dólares, producción por arriba de Venezuela, Malasia, Uganda e India (FAO, 2011), y mostrando un descenso gradual respecto de años anteriores (Tabla 1). En México, el 95% de cacao se obtiene de las plantaciones en los estados sureños de Tabasco y Chiapas; tan solo en la región Soconusco (Chiapas) se cultivan 11 500 ha (Molina & Córdova, 2006) y de donde se han reportado morfotipos del cultivar Criollo.

Tabla 1. Producción de almendras de cacao seco en México en la última década.

Año	Producción (Ton)
1999-2001	38 613
2003-2005	43 435
2007	29 910
2008	27 548
2009	22 680

Fuente: FAO, 2011.

De las tres variedades cultivadas, los Trinitarios y Forasteros [de los cuales se hipotetiza que existen dos grupos genéticos (los Forasteros del Alto Amazonas UA y los del bajo Amazonas LA) Efrombagn *et al.*, 2009] ocupan cerca del 95% del total

cultivado, en países del Oeste de África, Brasil y Ecuador, y del cual se obtiene igual porcentaje del chocolate que se comercializa. De la variedad Forastero, se obtiene el cacao a granel o "básico" y contribuyó como parental en la generación de los Trinitarios

Artículos

hace unos 250 años (Motamayor *et al.*, 2003).

La variedad Criollo, clasificada como subespecie cacao (Cuatrecasas, 1964), y del cual se obtiene el cacao "fino" contribuye a la producción mundial con el 5% (Afcakwa *et al.*, 2008); aunque su consumo aún está revalorándose, se encuentra en expansión y es mucho más exigente respecto a la calidad de la materia prima y de los productos derivados de ésta. Este mercado *gourmet* paga sobrepagos por almendras obtenidas del genotipo Criollo y destina el grano a la elaboración de chocolates "finos" altamente cotizados en Estados Unidos y Europa. Esta situación, sumada a otras características de mercadeo, que premia los productos con denominación de origen y simplifica los requerimientos de comercio, estimula la conservación, cultivo y comercialización de razas o genotipos de calidad "supenor".

Se ha reportado, que los individuos Criollos "nativos" o "antiguos" (Motamayor *et al.*, 2002; Whitkus *et al.*, 1998) estarían asociados con menor rendimiento agronómico, al poseer mazorcas de menor tamaño y menos número de almendras por fruto, además de una marcada susceptibilidad a *Moniliophthora roreri* y *Phytophthora palmivora*. La moniliasis, provocada por *M. roreri*, desde su primer reporte en México (Phillips-Mora *et al.*, 2008) ha provocado el abandono del cultivo de cacao, o la reducción, tanto del área cultivada como de la producción de grano (cerca del 50% en los últimos seis años). De no atenderse esta situación y establecer programas serios de rescate de genotipos de

interés (pero no solo desde la perspectiva del rendimiento), se vislumbra un horizonte poco alentador, ya que los cacaocultores optan por cambiar el sistema a otro que demande menores esfuerzos y mayores rendimientos, siendo muchas veces la elección poco acertada. Los cacaotales, además del beneficio social y económico que representan, contribuyen como la parte soporte de nichos ecológicos para una gran diversidad de especies vegetales y animales, encontrando desde especies maderables, florales y animales como mamíferos, aves y ardillas.

La existencia de individuos criollos en Mesoamérica, es importante por varios motivos, muestra evidencia inequívoca de su cultivo y domesticación, además que ha contribuido con la diversidad genética presente en las plantaciones intensivas del sur de México. Pese a que los individuos de cacao presentan un sistema de entrecruzamiento preferentemente alogamo, se ha documentado una clara distinción entre sistemas de auto-incompatibilidad de los genotipos Forasteros y de auto-compatibilidad de los Criollos (Cheesman, 1944). Sin embargo, existen reportes de segregación hacia los dos sistemas de polinización por parte de los dos genotipos (Bartley & Cope, 1973). Hoy día y dada la fuerte recombinación de parentales por parte de algunos programas y cacaocultores, buscando aumentar el vigor híbrido (heterosis) como la principal responsable de mejoramiento, se pueden hallar plantaciones de cacao en el Soconusco, Chiapas (como en otras partes del mundo) con individuos

Artículos

que responden a ambos sistemas de entrecruzamiento tal como ha demostrado Efombagn *et al.* (2009) ocurre en la recombinación abierta de poblaciones cultivadas y de invernadero en Camerún donde ha encontrado hasta 28.3% de descendencia proveniente de polinización no controlada. Los individuos que descienden de estos sistemas de entrecruzamiento, se vuelven un importante reservorio de genes que potencialmente pueden exhibir las características de sus parentales, siendo la calidad sensorial (aroma, sabor y textura principalmente) una de éstas (cuando la heterosis y la auto-incompatibilidad en genotipos Criollos coinciden), y que mejora las características agronómicas de rendimiento y tolerancia a enfermedades, como se ha demostrado puede ocurrir en las variedades Trinitario (Johnson *et al.*, 2009).

Por otro lado, Aikpokpodion (2010), llevó a cabo una correlación entre marcadores agro-morfológicos con individuos de interés desde el punto de vista de la calidad. Reporta, que factores ajenos a la genética del cacao, tales como las condiciones edafoclimáticas del lugar (Nigeria), permiten el desarrollo de características únicas de sabor en razas "Amelonadas" nativas de ese país Africano, derivadas de Forasteros del Bajo Amazonas y del Amelonado del Oeste Africano. Este estudio, además de poner en evidencia lo que ya se ha documentado ampliamente, acerca de la influencia del ambiente sobre la expresión de los marcadores morfológicos; apoya la hipótesis de que un microclima específico puede potenciar las características sensoriales de

los productos y, en el caso particular del cacao se comentó antes de la preferencia de los consumidores más exigentes hacia productos obtenidos de las plantaciones de Soconusco (Chiapas).

LA UTILIDAD DE LOS MARCADORES MOLECULARES PARA COMPRENDER LOS SECRETOS DEL CACAO

Marcadores moleculares

Previamente se enfatizó, que hasta antes de las primeras aplicaciones de los marcadores moleculares en estudios de genética y mejoramiento de organismos, los marcadores por elección, eran aquellos regulados por genes asociados a caracteres morfológicos y de rendimiento, en general fenotipos de fácil cuantificación e identificación visual. Sin embargo, el limitado número de estos marcadores y la interferencia epistática o ambiental aortaron su expansión (Azofeita-Delgado, 2006). La revolución en este ámbito se inició con el descubrimiento y utilización de los marcadores moleculares, los cuales se desarrollaron rápidamente y su número fue ampliado, expandiendo su aplicación a prácticamente a todas las especies de organismos vivos y acelerando los avances en los programas de búsqueda de organismos de interés.

En el caso de *T. cacao*, su empleo ha resultado de gran utilidad en estudios relacionados con diversos aspectos de la biología y genealogía de esta planta tropical. Así, han sido auxiliares en estudios de etiquetado de accesiones en los bancos de germoplasma más importantes del mundo,

Artículos

donde la clasificación y muchas veces el etiquetado erróneo (Motamayor *et al.*, 2008; Motilal *et al.*, 2011; Takrama *et al.*, 2005; Zhang *et al.*, 2009, 2008) se realizó en base a caracteres morfológicos. Otras aplicaciones incluyen el estudio de la diversidad genética tanto en bancos de germoplasma como en campos de cultivo, estudios de la genealogía de individuos y poblaciones y caracterización del germoplasma, básicamente orientado a características agronómicas de interés.

Los primeros marcadores moleculares (no basados en ADN) empleados en cacao fueron las isoenzimas (Lanaud, 1986), pese a las desventajas inherentes a esta herramienta, como son el bajo número tanto de *loci* como en el polimorfismo en su momento ofrecieron excelentes resultados y encontraron gran aplicación en estudios de diversidad genética, en la identificación de genotipos, así como en la construcción de mapas de ligamiento (Lachenaud *et al.*, 2004; Ronning & Schnell, 1994; Sounigo *et al.*, 2005; Warren, 1994).

Otro tipo de marcadores que más frecuentemente se usan en cacao, son los marcadores de ADN, específicamente, los polimorfismos en la longitud de los fragmentos de restricción (RFLP), polimorfismo de ADN amplificado al azar (RAPD), polimorfismo de longitud de fragmentos amplificados (AFLP), y secuencias simples repetidas (SSR). Estos marcadores difieren en abundancia genómica, el nivel de polimorfismo detectado, la especificidad hacia el locus, la reproducibilidad, los requerimientos técnicos y el costo (Guiltinan *et al.*, 2008).

Los RFLP, se han empleado dada su alta reproducibilidad y la naturaleza alélica codominante, en la construcción de mapas de ligamiento genético y etiquetado de genes y en la localización de *loci* que afectan a caracteres cuantitativos (QTL) vinculados a fenotipos de importancia agronómica (Lanaud *et al.*, 1995; Risterucci *et al.*, 2000) y la evaluación de la diversidad genética (Motamayor *et al.*, 2002; N'goran *et al.*, 1994). Los principales inconvenientes de estas herramientas son, debido a que las sondas de RFLP son siempre específicas para un número limitado de lugares, RFLP no es una herramienta muy eficaz para la identificación de genotipos de cacao, no se puede automatizar su aplicación y la generación de datos es laboriosa y puede ser costosa (Guiltinan *et al.*, 2008).

Los RAPD's fueron los primeros marcadores basados en la reacción en cadena de la polimerasa (PCR) que se aplicaron para la caracterización genética del cacao (Wilde *et al.*, 1992). Se ha empleado dada la facilidad de su aplicación, para identificación clonal, clasificación de acuerdo al origen geográfico (Russell *et al.*, 1993), así como en estudios de etiquetado incorrecto y la duplicidad (Faleiro *et al.*, 2002; Sounigo *et al.*, 2005). Los RAPD también han sido ampliamente utilizados en los análisis de la diversidad genética (Figueira *et al.*, 1994; Lerceteanu *et al.*, 1997; N'goran *et al.*, 1994; Whitkus *et al.*, 1998). Aunque este sistema es técnicamente fácil de realizar, se obtiene una baja reproducibilidad entre experimentos y entre laboratorios. RAPD no mide directamente la heterocigosidad, lo que hace

Artículos

muestras en tiempos relativamente cortos. Debido a estas importantes ventajas, los microsatélites han sido el método de elección para conducir investigaciones en cacao en la última década.

Principales aplicaciones de los SSR en los estudios del Cacao

Por las características antes apuntadas, los microsatélites han sido empleados para analizar diversos aspectos del manejo de germoplasma de cacao, entre las que destacan:

1. En la identificación de las accesiones duplicadas y mal etiquetado en colecciones de germoplasma. Debido a las limitantes para discriminar entre accesiones, ya que sus conclusiones tenían poco rigor estadístico (Sounigo *et al.*, 2005), los AFLP's y RAPD's, fueron sustituidos en su totalidad por los SSR. Los avances en este sentido, han permitido establecer mediante la creación de un consorcio internacional, un conjunto de 15 microsatélites estándares para el etiquetado correcto de las accesiones y la identificación inequívoca, los cuales son utilizados en varios de los bancos de germoplasma alrededor del mundo, y han permitido identificar por ejemplo, de entre 105 accesiones, 10 grupos mal etiquetados (Zhang *et al.*, 2008b).

2. En estudios de diversidad genética, relaciones genealógicas y filogeografía. Los estudios buscando explicar la diversidad y estructura poblacional tanto de colecciones como de plantaciones de cacao, han mostrado un incremento pronunciado con la aparición de los microsatélites y más aún con

el apoyo de la estadística Bayesiana. Así por ejemplo Bhattacharjee *et al.* (2004) hallaron baja tasa de adopción de germoplasma mejorado en los campos agrícolas de diferentes zonas agroecológicas en Nigeria. De los pocos estudios conducidos en Mesoamérica, Ruiz *et al.* (2011), empleando diez microsatélites analizaron 70 accesiones de Nicaragua, encontraron deficiencia de heterocigosis, pero un nivel moderado de diversidad genética pese al corto historial de cultivo extensivo en ese país reportado por Trognitz *et al.* (2011) y lo atribuyen a un posible muestreo entre consanguíneos o derivados mayoritariamente de autofecundaciones o, como resultado de diferente frecuencia alélica entre poblaciones.

También con apoyo de los SSR se ha hipotetizado en varios reportes acerca de que la Amazonia Peruana posee alto nivel de diversidad genética, presentando estructura espacial, además de alta diversidad alélica en poblaciones cultivadas (Zhang *et al.*, 2008a). Del mismo modo, Johnson *et al.* (2009), encontraron una estrecha relación genética entre genotipos "Trinitarios selectos" y los posibles parentales, colectados 70 años antes. Además obtuvieron, usando 35 microsatélites una diferenciación espacial de los genotipos Trinitarios y Criollos (agrupados en un *clúster*) al compararlos con poblaciones del Alto y Dajo Amazonas y de Ecuador obtenidos de banco de germoplasma.

Por su capacidad para revelar codominancia, los microsatélites resultan muy adecuados en la búsqueda y

Artículos

esclarecimiento de relaciones genealógicas y estas investigaciones, a su vez son sumamente útiles para comprender el flujo de genes en las poblaciones naturales, la contribución de los genotipos antiguos en la estructura de las poblaciones actuales así como para rastrear la posible contribución de parentales en los programas de mejoramiento de variedades.

Otros estudios además han abordado la búsqueda de patrones de genealogía asociados con la geografía presente o pasada (filogeografía), en el supuesto de que, al conocer los patrones espaciales de biodiversidad, los procesos de flujo de genes, los efectos históricos y climáticos sobre la distribución de la diversidad genética, se tendrán más herramientas para la conservación sostenible y el uso efectivo de germoplasma de cacao (Guiltinan *et al.*, 2008).

En este sentido, Trognitz *et al.* (2011), evaluaron la composición alélica y la estructura genética a partir de 44 plantaciones y dos accesiones silvestres de Nicaragua, demuestran estructura espacial entre parcelas cultivadas la cual es atribuida por el manejo agronómico, prácticas de selección y condiciones ambientales, además reportan un genotipo único para las accesiones silvestre e hipotetizan acerca de la posible relación genética con individuos cultivados por los Mayas en la precolonia.

Sereno *et al.* (2006) compararon 94 accesiones de cuatro poblaciones naturales de la Amazonia de Brasil. Exceptuando a la población del Alto Amazonas, en las restantes encontraron bajo índice de

heterocigosidad observada, lo que revela alta tasa de homocigotos que coincide con la aseveración de Ruiz *et al.* (2011), en el sentido de estarse presentado altas frecuencias de autopolinización o el muestreo incluye alta consanguinidad. La población del Alto Amazonas exhibió la mayor diversidad genética y por lo tanto, los autores la sugieren en concordancia con otros autores como parte del centro de diversificación de la especie y, posible centro de origen (Cheesman, 1944; Motamayor *et al.*, 2002; Schultes, 1984).

En el trabajo de Motamayor *et al.* (2002) se evaluó la composición alélica en las variedades de América Central (genotipos de Guayana, la Amazonia y el Orinoco). Los resultados apoyan la hipótesis de que el cacao se originó en el Alto Amazonas y sugieren la probable ruta de dispersión mediada por humanos desde el Amazonas hasta América Central y México para establecer el cacao "Criollo" que los Olmecas domesticaron hace casi cuatro mil años. Sin embargo la baja diversidad genética reportada por estos autores para las accesiones de Criollos de Mesoamérica y que, contradice lo antes reportado (De la Cruz *et al.*, 1995; Whitkus *et al.*, 1998), debe ser corroborada y en su caso, podría ser explicada por la gran presión de selección a la que estuvo sometida la planta en los procesos de domesticación buscando características sensoriales agradables y agronómicas que pudieron haber sido de interés para los pueblos precolombinos (como pulpa más dulce, mayor número de semillas) durante varios siglos y que se

Artículos

promovió fuertemente por las características de auto compatibilidad casi exclusiva reportado para los Criollos y presuntamente observada en otras accesiones (Ruiz *et al.*, 2011; Sereno *et al.*, 2006). Respecto a la dispersión de individuos del Alto Amazonas, para dar origen a los Criollos de Mesoamérica, ésta enfrenta algunas cuestiones por ejemplo, se ha reportado que de manera natural el árbol del cacao tiene una baja o nula capacidad de dispersión, en estado silvestre, los árboles de cacao por lo general se multiplican a través de propagación vegetativa (Lachenaud & Zhang, 2008), los frutos son indehiscentes con un pericarpio espeso y duro (sobre todo los Criollos), y no presenta abscisión, lo cual significa que se mantiene unido indefinidamente al árbol. Las semillas son altamente susceptibles a la desecación (coaloitantes), sobre todo las de la variedad Criollo (Rangel, 2009), germinan dentro de la fruta y puede perderse fácilmente si las plántulas no tienen una vía de dispersión exógena. Si este papel es realizado por animales (principalmente monos), quienes pueden tragar semillas enteras sin masticarlas y terminan en el suelo en las heces, donde germinan rápidamente (endozoocoria), se han reportado principalmente a *Cebus nigrivittatus* y *C. apella* o *Alouatta seniculus* (Lachenaud & Zhang, 2000), este último género caracterizado como de baja actividad (Muñoz *et al.*, 2002) y que enfrentaría grandes retos por transportar las semillas en distancias considerables.

Si la propagación fue mediada por el hombre, como hipotetizan los estudios, - ¿cuál sería la intención de dispersar una semilla o planta que, en su centro de diversidad (Alto Amazonas) no tenía uso agronómico o social?- o al menos no hay evidencia arqueológica de esto, como si existe de manera abundante en Mesoamérica y; ¿cómo se llevaría a cabo el traslado, con semillas de la naturaleza antes descrita para recorrer los aproximadamente 2000 km de distancia entre el Alto Amazonas y accesiones de Nicaragua empleadas en el reporte (Motamayor *et al.*, 2002) en períodos breves para garantizar la no desecación y pérdida de viabilidad de las semillas?.

3. En la construcción de mapas de ligamiento y búsqueda de *loci* asociados a caracteres cuantitativos (QTL)

Se han desarrollado una gran cantidad de mapas genéticos para detectar QTL para varios rasgos sobre todo de importancia agronómica. Los principales intereses en la búsqueda de estos marcadores han sido: a) resistencia a las tres principales enfermedades causadas por hongos, caracteres de rendimiento, b) vigor de la planta, c) número de óvulos y d) características de mazorca y grano (Tabla 2). Las enfermedades son responsables de pérdidas que pueden llegar a alcanzar 100%, debido a esto, la gran mayoría de los trabajos en este sentido han dedicado esfuerzos a localizar los genes responsables de conferir resistencia a algunos genotipos y como estos pueden ser transferidos (Brown *et al.*, 2007; Cervantes-Martínez *et al.*, 2006; Clement *et al.*, 2003a; Feltus *et al.*, 2011; Lansud *et al.*,

Artículos

2009; Pugh *et al.*, 2004; Risterucci *et al.*, 2003). Un porcentaje menor de los trabajos, se enfoca en la búsqueda de genes involucrados en caracteres morfológicos de arboles, frutos y almendras (Cervantes-

Martínez *et al.*, 2006; Clement *et al.*, 2003a, 2003b; Crouzillat *et al.*, 1996; Feltus *et al.*, 2011), lo cual denota hacia donde se orientan las prioridades de la investigación en cacao.

Tabla 2. Principales investigaciones en *Theobroma cacao* L. empleando marcadores moleculares en la búsqueda de caracteres de interés agronómico

QTL's asociados con resistencia a <i>Phytophthora palmivora</i>	Brown <i>et al.</i> , 2007; Cervantes-Martínez <i>et al.</i> , 2006; Clement <i>et al.</i> , 2003a; Crouzillat <i>et al.</i> , 2000; Feltus <i>et al.</i> , 2011; Flament <i>et al.</i> , 2001; Lanaud <i>et al.</i> , 1995; Lanaud <i>et al.</i> , 2009; Pugh <i>et al.</i> , 2004; Risterucci <i>et al.</i> , 2000; Risterucci <i>et al.</i> , 2003
QTL's asociado con resistencia a <i>Crinipellis (Moniliophthora)</i> <i>perniciosa</i>	Brown <i>et al.</i> , 2005; Faleiro <i>et al.</i> , 2006; Figueira <i>et al.</i> , 2006; Lanaud <i>et al.</i> , 2009; Queiroz <i>et al.</i> , 2003
QTL's asociados con caracteres morfológicos (vigor, floración, rendimiento, número de óvulos, índice de grano y peso de mazorca)	Cervantes-Martínez <i>et al.</i> , 2006; Clement <i>et al.</i> , 2003a; Clement <i>et al.</i> , 2003b; Crouzillat <i>et al.</i> , 1996; Feltus <i>et al.</i> , 2011
Localización de QTL's asociados a resistencia a <i>Moniliophthora roerei</i>	Brown <i>et al.</i> , 2007; Cervantes-Martínez <i>et al.</i> , 2006; Lanaud <i>et al.</i> , 2009
Marcador asociado a tolerancia al estrés	Borrone <i>et al.</i> , 2004

La reciente publicación de la secuencia genómica del cacao (Argout *et al.*, 2011), ofrece importantes aportaciones a la comprensión de posibles explicaciones en la síntesis de compuestos de interés (algunos implicados en la calidad sensorial) y ofrece una oportunidad única a los investigadores para dirigir estudios buscando explicar las

interrogantes guardadas en el genoma de esta planta. De los hallazgos importantes, se estima de manera más precisa (mediante citometría de flujo) el tamaño del genoma de esta planta (430 Mb del genotipo crollo B97-81/B2), con número cromosómico $2n=20$, el cual es considerado como "pequeño" en comparación con individuos de las mismas

Artículos

características fenotípicas que el árbol del cacao y, que había sido estimado en 390 Mb por reasociación cinética (Couch *et al.*, 1993).

En este estudio, basándose en el conocimiento de los mecanismos moleculares y bioquímicos de plantas modelo y, después de realizar análisis de ortología con genes de *Arabidopsis thaliana*, *Vitis vinifera*, *Populus sp.* y *Glycine max*, muestra asociaciones entre genes posibles candidatos relacionados con la resistencia a enfermedades y, con la calidad de las almendras. Ésta última, entendida como la cantidad genes potencialmente involucrados en las síntesis de macromoléculas (grasa, proteínas, almidón) y metabolitos secundarios como flavonoides, terpenoides y alcaloides, que se sabe desempeñan un papel importante en las características sensoriales del cacao.

La calidad sensorial del cacao, su asociación con marcadores

La calidad en el cacao, como en todo producto o materia prima, está dada por un número amplio de atributos o características, muchos de ellos intrínsecos a la genética del vegetal y otros, atribuibles a las condiciones ambientales donde se desarrolla, así como a las prácticas pre y poscosecha.

Para tratar de homogenizar la terminología y atributos entendibles como "calidad" se ha acordado mediante estándares internacionales, establecer que el cacao de buena calidad comercial debe ser: a) fermentado, completamente seco, libre de almendras vanas, libre de olores anormales o

extraños y libre de cualquier evidencia de adulteración, b) de tamaño bastante uniforme, considerablemente libre de granos fragmentados, fragmentos y piezas de cáscara, y prácticamente exentas de materias extrañas (Dongo *et al.*, 2009).

Aunque esta definición no incluye de manera total, los aspectos importantes del sabor y aroma que determinan la aceptación de algunos mercados *gourmet*, se puede entender que, hay características que dependen del manejo poscosecha, como la uniformidad de las semillas, la presencia de materia extraña y otras; pero también otras en las que se involucran procesos microbianos (fermentación) y de secado que de realizarse de manera inadecuada demeritan la calidad de las almendras (Criollo y Trinitario), ya que no permiten potenciar la transformación o aparición de moléculas implicadas en algunas de las características sensoriales importantes de estos cultivares (Rodríguez-Campos *et al.*, 2011).

Algunos argumentos, atribuyen una carga sumamente importante para conservar la calidad del cacao al tiempo de cosecha y selección de frutos y semillas que deberán fermentarse, así como al mismo proceso de fermentación pues se ha reportado que, emplear frutos (mazorcas) fuera de madurez aumenta el contenido de ácido cítrico (Papalexandratou *et al.*, 2011a), lo cual ejerce efecto indeseable sobre el proceso de fermentación y el posterior secado y tostado, que conllevaría la aparición de metabolitos específicos responsables de aroma no deseados (ácidos carboxílicos), principalmente ácido butírico (Rodríguez-

Artículos

Campos *et al.*, 2011), ácido propiónico y metilpropiónico (Frauendorfer & Schieberley, 2008) que demeritan la calidad de cacaoos deficientemente fermentados, secados o tostados.

La fermentación por sí misma, indudablemente tiene un papel crucial en mantener o realzar la calidad de los cacaoos Criollos o Trinitarios, pero es cuestionable el argumento de que aplicada bajo ciertos estándares, puede mejorar la calidad del grano de aquellos considerados de "calidad inferior" en un estudio conducido con cacao Brasileño (Papalexandratou *et al.* 2011b). Estas conclusiones, son ciertas solo de manera parcial, puesto que se ha documentado de antaño que los genotipos Criollos, independientemente de los procesos poscosecha ofertan características superior u otros con procesado poscosecha similar (Afoakwa *et al.*, 2008; Ciferri & Ciferri, 1967; Clapperton *et al.*, 1994). Siendo más realista, podría pensarse que los factores ambientales y procesos practicados posteriores al corte (fermentación, secado, tostado) influyen en un 50% de la calidad final del producto, mientras que la mitad restante está guardada en la genética de cada individuo.

El otro factor que tampoco es tomado en cuenta en los estándares y que supone la mayor predisposición para obtener calidad superior es el origen genético, ya que, como antes se apuntó de la variedad Forastero se obtiene cacao de calidad inferior, del Trinitario de calidad intermedia y los granos de alta calidad son producidos por los cacaoos Criollos (Sukha & Butler, 2005). Eskes *et al.* (2007), encontraron correlación entre notas

de aroma y sabor halladas en mazorcas (pulpa) de diferente variedad y la calidad sensorial de las almendras, mostrando la variación asociada con el genotipo pues reportan la presencia de astringencia y poco dulzor en variedades de baja calidad y gustos dulce y notas de sabor afrutado en genotipos de calidad superior, además de poder diferenciar a las accesiones por su origen genético.

En Ecuador, se han empleado algunas características sensoriales para la selección de individuos que forman parte del llamado cacao fino producido en ese país (Eskes *et al.*, 2007), lo que da cuenta una vez más, que la calidad desde este punto de vista podría ser correlacionada con algún tipo de marcador y, los moleculares, se prestan como los candidatos más apropiados en esta encomienda.

La evaluación sensorial es una herramienta ampliamente utilizada en el campo de los alimentos, tanto para el desarrollo de productos, como para los estudios de clasificación y determinación de la calidad de productos de alto valor comercial. Su aplicación como herramienta de medición supone el entrenamiento de un panel de evaluadores quienes fungen como "instrumentos" de medición y emiten su veredicto inequívoco en alguna característica específica. Su correlación con mediciones químicas (acidez, pH) o físicas (viscosidad, textura) resulta útil pero, tediosa puesto que las características sensoriales evaluadas, son la resultante de una suma compleja de atributos fisicoquímicos y, evaluar una sola característica resulta insuficiente.

Artículos

De manera similar, las pruebas sensoriales pueden ser desde el punto de vista operativo costosas y tardadas, puesto que no siempre se dispone de un panel entrenado en el momento de querer evaluar la calidad de una accesión de cacao en particular.

Establecer marcadores moleculares que puedan estar asociados a regiones no codificantes (SSR) vinculadas a genes que están potencialmente involucrados en la calidad como sugieren Argout *et al.*, (2011), pero que permiten al ser detectados, segregar a individuos de diferente calidad sensorial (Smulders *et al.*, 2010), podría ser una realidad que ofrecería ventajas de tiempo y costo a los programas de mejoramiento de variedades cuando esta calidad sea el objeto de interés o, ayudar a proteger accesiones con potencial en este mismo sentido.

CONCLUSIONES

El mercado del cacao lejos de extinguirse, a pesar de las enfermedades y problemáticas sociales que le aquejan continuará en expansión en los años venideros, como también habrá de aumentar el mercado que exige productos con mayor contenido de chocolate (no cocoa o manteca de cacao) y que éste sea de calidad superior. Esto supone rescatar y aumentar el cultivo e introducción de genotipos que producen semillas de calidad superior. Debido a esto, los estudios moleculares relacionados con *T. cacao*, que cada vez se apoyan o lo harán de herramientas como los SSRs, ISSRs, búsqueda de SNP's, SCARs, deberán

extender su alcance a buscar cada vez más marcadores o QTL con la calidad; el conocimiento del genoma de esta planta resulta una herramienta útil para este propósito. Otro campo de interés que se apertura y que deberá ser explotado en su totalidad, es de la bioinformática, puesto que ahora se cuenta con un código que deberá ser descifrado con más claridad para tratar de comprender los enigmas que esta planta ha guardado por varios milenios.

REFERENCIAS

- Afoakwa E, Paterson A, Fowler M & Ryan A (2008) Flavor formation and character in cocoa and chocolate: a critical review. *Crit. Rev. Food Sci.* 48: 840-857.
- Aikpokpodion P (2010) Variation in agromorphological characteristics of cacao, *Theobroma cacao* L., in farmers' fields in Nigeria. *New Zcal. J. Crop Hort.* 38(2): 157-170.
- Argout X, Salse J, Aury J, Guiltinar M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Douge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S & Lanaud C (2011)

Artículos

- The genome of *Theobroma cacao*. *Nat. Genet.* 43(2): 101-109.
- Azofeita-Delgado A (2008) Uso de marcadores moleculares en plantas; Aplicaciones en frutales del trópico. *Agron. Mesoam.* 17(2): 221-242.
- Bartley B & Cope F (1973) Practical aspects of self-incompatibility in *Theobroma cacao* L. In: *Agricultural Genetics - Selected Topics*. Moav R (ed) John Wiley & Sons, Inc. New York. pp 109-134.
- Bhattacharjee R, Aikpokpodion P, Kolesnikova-Allen M, Badaru K & Schnell R (2004) West African Cocoa: A pilot study on DNA fingerprinting the germplasm from Cross River State of Nigeria. *INGENIC Newsl.* 9: 15-20.
- Borrone J, Kuhn D & Schnell R (2004) Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theor. Appl. Genet.* 109(3): 495-507.
- Brown J, Phillips-Mora W, Power E, Krol C, Cervantes-Martinez C, Motamayor J & Schnell J (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. *Crop Sci.* 47: 1851-1858.
- Brown J, Kuhn D, Lopez U & Schnell R (2005) Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an F2 population using SSR markers and candidate genes. *J. Am. Soc. Hort. Sci.* 130: 366-373.
- Cervantes-Martinez C, Brown J & Schnell R (2008) Combining ability for disease resistance, yield, and horticultural traits of cacao (*Theobroma cacao* L.) clones. *J. Am. Soc. Hort. Sci.* 131(2): 231-241.
- Cheesman E (1944) Notes on the nomenclature, classification and possible relationships of cocoa populations. *Trop. Agric.* 21: 144-159.
- Chiapas.gob (2010) Cacao del Soconusco, de los mejores del mundo: empresarios franceses. Instituto de Comunicación Social. Diciembre de 2010. Consultado junio de 2011. Disponible: <http://www.cocoso.chiapas.gob.mx/documento.php?id=20101212124448>.
- Ciferri R & Ciferri F (1957). The evolution of cultivated cacao. *Evolution* 11: 381-397.
- Clapperton J (1994) A review of research to identify the origins of cocoa flavour characteristics. *Cocoa Growers' Bull.* 48: 7-16.
- Clement D, Risterucci A, Motamayor J, N'Goran J & Lanaud C (2003a) Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. *Genome* 46: 204-212.
- Clement D, Risterucci A, Motamayor J, N'Goran J & Lanaud C (2003b) Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. *Genome* 46: 103-111.
- Couch J, Zintel H & Fritz P (1993) The genome of the tropical *Theobroma cacao* L. *Mol. Gen. Genet.* 237(1-2): 123-120.
- Crouzillat D, Menard B, Mora A, Phillips W & Petiard V (2000) Quantitative trait analysis in *Theobroma cacao* using molecular markers. *Euphytica* 114: 13-23.

Artículos

- Crouzillat D, Lerceteau E, Petiard V, Morera J, Rodriguez H, Walker D, Phillips W, Ronning C, Schnell R, Osei J & Fritz P (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. *Theor. Appl. Genet.* 93:205-214.
- Cuatrecasas J (1964) Cacao and its allies: a taxonomic revision of the genus *Theobroma*. *Contrib. U. S. Natl. Herb.* 35(6): 379-614.
- De la Cruz M, Whitkus R, Gomez-Pompa A & Mota-Bravo L (1995). Origins of cacao cultivation. *Nature* 375: 542-543.
- Dongo L, Aigbekaen E, Jayeola C, Emaku L & Orisajo S (2009) Influence of farmers practices on cocoa bean quality. Nigeria field experience. *Afr. Crop Sci. Conf. Proc.* 9: 299-302.
- Efombagn M, Sounigo O, Eskes A, Motamayor J, Manzanares-Dauleux M, Sohneil R & Nyassé S (2000) Parentage analysis and outcrossing patterns in cacao (*Theobroma cacao* L.) farms in Cameroon. *Heredity* 103: 46-53.
- Efombagn M, Nyassé S, Sounigo O, Kolesnikova-Allen M & Eskes A (2007) Participatory cocoa (*Theobroma cacao*) selection in Cameroon: *Phytophthora* pod rot resistant accessions identified in farmers' fields. *Crop Prot.* 26: 1467-1473.
- Eskes A, Guarda D, García L & García P (2007) Is genetic variation for sensory traits of cocoa pulp related to fine flavor cocoa traits?. *INGENIC News*. 11: 22-28.
- Faleiro F, Queiroz V, Lopes U, Guimarães C, Pires J, Yamada M, Araújo I, Pereira M, Schnell R & De Souza G (2006) Mapping QTLs for witches' broom (*Crinipellis perniciosa*) resistance in cacao (*Theobroma cacao* L.). *Euphytica* 149: 227-235.
- Faleiro F, Yamada M, Lopes U, Faleiro A, Siqueira R, Costa L, Santos R & dos Santos R (2002) Genetic similarity of *Theobroma cacao* L. accessions maintained in duplicates in the Cacao Research Center germplasm collection, based on RAPD markers. *Crop Breed. Appl. Biotechnol.* 2(3): 439-444.
- FAO (2011) Organización de las Naciones Unidas para la agricultura y la Alimentación. Anuario Estadístico 2010. Tabla B7.
- Feltus F, Saski C, Mockaitis K, Haiminen N, Parida L, Smith Z, Ford J, Staton M, Ficklin S, Blackmon B, Cheng C, Schnell R, Kuhn D & Motamayor J (2011) Sequencing of a QTL-rich region of the *Theobroma cacao* genome using pooled BACs and the identification of trait specific candidate genes. *BMC Genomics* 12: 379.
- Figueira A, Albuquerque S & Leal G (2006) Genetic mapping and differential gene expression of Brazilian alternative resistance sources to witches' broom (causal agent *Crinipellis perniciosa*). *15th Intern. Cocoa Res. Confer.* p. 166
- Figueira A, Janick J, Levy M & Goldsbrough P (1994) Reexamining the classification of *Theobroma cacao* L. using molecular markers. *J. Am. Soc. Hort. Sci.* 119: 1073-1082.
- Flament M, Kebe I, Clement D, Pieretti I, Risterucci A, N'Goran J, Cilas C, Despréaux D & Lanaud C (2001) Genetic

Artículos

- mapping of resistance factors to *Phytophthora palmivora* in cocoa. *Genome* 44(1): 79-85.
- Fraudendorfer F & Schieberle P (2008). Changes in key aroma compounds of criollo cocoa beans during roasting. *J. Agr. Food Chem.* 56:10244-10251.
- Guitinan M, Verica J, Zhang D & Figueira A (2008) Genomics of the *Theobroma cacao*, "The Food of the Gods". In: Genomics of Tropical Crop Plants. Moore P & Mings R (eds). Springer Science+Business Media. NY USA. pp. 145-170.
- INAH (2007) Instituto Nacional de Antropología e Historia. Comisión de defensa del patrimonio cultural. Considerations and alternative proposal for the encroachment on the archaeological site by the Tapachula-Talismán highway. Disponible: www.mesoweb.com/reports/tzapa.pdf.
- Irizarry H & Goenaga R (2000) Clonal selection in cacao base on early yield performance of grafted trees. *J. Agr. Univ. Puert. Rico* 84(3-4): 153-163.
- Johnson E, Bekele F, Brown S, Song Q, Zhang D, Meinhardt L & Schnell R (2009) Population structure and genetic diversity of the Trinitario cacao (*Theobroma cacao* L.) from Trinidad and Tobago. *Crop Sci.* 49: 564-572.
- Lachenaud P & Zhang D (2000) Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana. *Ann. For. Sci.* 65: 310.
- Lachenaud P, Sounigo O & Oliver G (2004) Genetic structure of Guianan wild cocoa (*Theobroma cacao* L.) described using isozyme electrophoresis. *Plant Genetic Res. Newsl.* 139: 24-30.
- Lanaud C, Fouet O, Clément D, Boccara M, Risterucci A, Surujdeo-Maharaj S, Legavre T & Argout X (2009) A meta-QTL analysis of disease resistance traits of *Theobroma cacao* L. *Mol. Breeding* 24: 361-374.
- Lanaud C, Risterucci A, Pieretti I, Falque M, Bouet A & Lagoda P (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* 8: 2141-2152.
- Lanaud C, Risterucci A, N'Goran A, Clement D, Flament M, Laurent V & Falque M (1995) A genetic linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 01(6 7): 087 003.
- Lanaud C (1988) Genetic studies of *Theobroma cacao* L. with the help of enzymatic markers. I: Genetic control and linkage of nine enzymatic markers. *Cafe Cacao The* 30: 259-270.
- Lerocteau E, Robert T, Pétiard V & Crouzillat D (1997) Evaluation of the extent of genetic variability among *Theobroma cacao* accesions using RAPD and RFLP markers. *Theor. Appl. Genet.* 95(1-2): 10-19.
- Molina J & Córdova L (2006) Recursos fitogenéticos de México para la alimentación y la agricultura: Informe Nacional 2006. Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación y Sociedad Mexicana de

Artículos

- Fitogenética, A.C. Chapingo, México. pp. 19-20.
- Moreno Y, Melgarejo L, Hernández M, Quintero L & Vargas G (2004) Caracterización molecular de un banco de germoplasma del género *Theobroma* mediante la técnica de RAPD. *Rev. Colomb. Biotecn.* 8(2): 15-24.
- Motamayor J, Lachneaud P, da Silva e Mota J, Looer R, Kuhn D, Brown J & Schnell R (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE* 3: e3311.
- Motamayor J, Risterucci A, Heath M & Lanaud C (2003) Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity* 91: 322-330.
- Motamayor J, Risterucci A, Lopez P, Ortiz C, Moreno A & Lanaud C (2002) Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89: 380-388.
- Motilal L, Zhang D, Pathmanathan U, Mischke S, Pinney S & Meirhardt L (2011) Microsatellite fingerprinting in the International Cocoa Genebank, Trinidad: accession and plot homogeneity information for germplasm management. *Plant Genet. Resour.* 9: 430-438.
- Muñoz D, García del Valle Y, Franco B, Estrada E & Magaña M (2002) Estudio del patrón de actividad general de monoxaulladores (*Alouatta palliata*) en el parque Yumká, Tabasco, México. *Neotrop. Primate* 10(1): 11-16.
- N'Goran J, Laurent V, Risterucci A & Lanaud C (1994) Comparative genetic diversity studies of *Theobroma cacao* using RFLP and RAPD markers. *Heredity* 73: 589-597.
- Papalexandratou Z, Camu N, Falony G & De Vuyst L (2011a) Comparison of the bacterial species diversity of spontaneous cocoa bean fermentations carried out at selected farms in Ivory Coast and Brazil. *Food Microb.* 28: 964-973.
- Papalexandratou Z, Vrancken G, De Bruyne K, Vandamme P & De Vuyst L (2011b) Spontaneous organic cocoa bean box fermentations in Brazil are characterized by a restricted species diversity of lactic acid bacteria and acetic acid bacteria. *Food Microb.* 28: 1326-1338.
- Perry M, Davey M, Power J, Lowe K, Bligh H, Roach P & Jones C (1998) DNA isolation and AFLP genetic fingerprinting of *Theobroma cacao* (L.). *Plant Mol. Biol. Rep.* 16: 40-50.
- Phillips-Mora W, Castillo J, Krauss U, Rodríguez E, & Wilkinson M (2005) Evaluation of cacao (*Theobroma cacao*) clones against seven Colombian isolates of *Moniliophthora roreri* from four pathogen genetic groups. *Plant Pathol.* 54: 483-490.
- Phillips-Mora W, Coutiño A, Ortiz C, López A, Hernández J & Aime M (2006) First report of *Moniliophthora roreri* causing frosty pod rot (moniliasis disease) of cocoa in México. *Plant Pathol.* 55: 504.
- Ploetz R (2007) Cacao diseases: important threats to chocolate production worldwide. *Phytopathol.* 97: 1634-1639.
- Powis T, Cyphers A, GaiKWad N, Grivetti L & Cheong K (2011) Cacao use and the San

Artículos

- Lorenzo Olmec. *Proc. Natl. Acad. Sci. U. S. A.* 108(21): 8595-8600.
- Pugh T, Fouet O, Risterucci A, Brottier P, Abouladze M, Deletrez C, Courtois B, Clement D, Larmande P, N'Goran J & Lanaud C (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theor. Appl. Genet.* 108 (8): 1151-1161.
- Queiroz V, Guimarães C, Anherdt D, Schuster I, Daher T, Pereira M, Miranda V, Loguercio L, Barros E & Moreira M (2003) Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. *Plant Breeding* 122: 268-272.
- Rangel M (2009) Anatomía y tolerancia a la desecación de semillas de cacao (*Theobroma cacao* L.) Tesis de doctorado en Ciencias. Colegio de Posgraduados. Texcoco, Estado de México. pp.1-49.
- Risterucci A, Paulin D, Ducamp M, N'Goran J & Lanaud C (2003) Identification of QTLs related to cocoa resistance to three species of *Phytophthora*. *Theor. Appl. Genet.* 108(1): 168-174.
- Risterucci A, Grivet L, N'Goran J, Pieretti I, Flament M & Lanaud C (2000) A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 101: 948-955.
- Rodriguez-Campos J, Escalona-Buendía H, Orozco-Avila I, Lugo-Cervantes C & Jaramillo-Flores M (2011) Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Res. Intern.* 44: 250-258.
- Ronning C & Schnell R (1994) Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *J. Heredity* 85(4): 291-295.
- Ruiz J, Roa O & Marin I (2011) Molecular ecology of genetic diversity of cacao cultivated in the south-east region of Nicaragua. *Inter. Res. J. Agric. Sci.* 1(1): 6-13.
- Russell J, Hosein F, Johnson E, Waugh R & Powell W (1993) Genetic differentiation of cocoa (*Theobroma cacao* L.) populations revealed by RAPD analysis. *Mol. Ecol.* 2:89-97.
- Schnell R, Olano C, Brown J, Meerow A & Cervantes-Martínez C (2005). Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. *J. Am. Soc. Hort. Sci.* 130(2): 181-190.
- Schultes R (1984). Amazonian cultigens and their northward and westward migrations in pre-Columbian times. In: Pre-Columbian plant migration. Papers of the Peabody Museum of Archaeology and Ethnology. Vol 76: Mass. Stone D (ed). Harvard University Press. Cambridge. pp 69-83.
- Sereno M, Albuquerque P, Vencovsky R & Figueira A (2006). Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conser. Genet.* 7: 13-24.

Artículos

- Smulders M, Esselink D, Amores F, Ramos G, Sukha D, Butler D, Vosman B & Van Loo E (2010) Identification of cocoa (*Theobroma cacao* L.) varieties with different quality attributes and parentage analysis of their beans. *INGENIC Newsl.* 12: 1-13.
- Sounigo O, Umaharan R, Christopher Y, Sankar A & Ramdahin S (2005) Assessing the genetic diversity in the International Cocoa Genebank, Trinidad (ICG,T) using isozyme electrophoresis and RAPD. *Genet. Res. Crop. Evol.* 52: 1111-1120.
- Sukha D & Butler D (2005) The CFC/ICCO/INIAP Cocoa Flavour Project - Investigating the spectrum of fine flavour within genotypes and between origins. *INGENIC Newsl.* 10: 22-25.
- Takrama J, Cervantes-Martínez C, Philips-Mora W, Brown J, Motamayor J & Schnell R (2005) Determination of off-types in a cacao breeding program using microsatellites. *INGENIC Newsl.* 10: 2-7.
- Trognitz B, Scheldeman X, Hansel-Hohl K, Kuant A, Grebe H & Hermann M (2011) Genetic population structure of cacao plantings within a young production area in Nicaragua. *PLoS ONE* 6(1): e16056.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M & Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23(21):4407-4414.
- Warren J (1994) Isozyme variation in a number of populations of *Theobroma cacao* obtained through various sampling regimes. *Euphytica* 72: 121-126.
- Whitkus R, de la Cruz M, Mota-Bravo L & Gómez-Pompa A (1998) Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theor. Appl. Genet.* 96(1-2): 621-627.
- Wilde J, Waugh R & Powell W (1992) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. *Theor. Appl. Genet.* 83: 871-877.
- Zhang D, Boccara M, Motilal L, Mischke S, Johnson E, Butler D, Bailey B & Meinhardt L (2009) Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers. *Tree Genet. & Genomes* 5: 595-607.
- Zhang D, Boccara M, Motilal L, Butler D, Umaharan P, Mischke S & Meinhardt L (2008) Microsatellite variation and population structure in the 'Refractario' cacao of Ecuador. *Conserv. Genet.* 9: 327-337.
- Zhang D, Arevalo-Gardini E, Mischke S, Zuñiga-Cernades L, Barreto-Chavez A & Adiazola J (2008a) Genetic diversity and structure of managed and semi-natural populations of cocoa (*Theobroma cacao*) in the Huallaga and Ucayali Valleys of Peru. *Ann. Bot.* 98: 847-855.
- Zhang D, Mischke S, Coenaga R, Almeida A & Saunders J (2008b) Accuracy and reliability of high-throughput microsatellite genotyping for cacao clone identification. *Crop Sci.* 48: 2084-2092.

CAPÍTULO 3

De la estimación de la diversidad y estructura genética de árboles de cacao con fenotipo Criollo cultivados en la región Soconusco, al sur de Chiapas México

Bajo un enfoque de análisis poblacional, se analizó la diversidad y estructura genética de árboles cultivados en la región Soconusco, Chiapas México, empleando microsatélites (SSR) como marcadores moleculares. El muestreo de plantas fue intencionado hacia árboles con fenotipo Criollo. Se incluyeron como controles árboles con alta ancestría Criollo. Los indicadores de variabilidad revelaron de moderada a baja diversidad genética, aunque una moderada diversidad alélica (índice de Shannon $I=0.97$ para los doce *loci*). Por el tipo de muestreo intencionado hacia un grupo fenotípico, en todas las poblaciones se encontró deficiencia de heterocigotos ($H_e > H_o$). El análisis de la estructura genética reveló dos grupos ($K=2$) donde, realizando un análisis de asignación, quedaron agrupados 91 de 108 árboles con una pertenencia $>90\%$. 27 árboles cultivados y 10 controles se conglomeran en el grupo 1 (Criollo). 54 árboles pertenecieron al grupo 2. Los restantes 27 árboles se encontraron en una zona de hibridación (entremezclados) entre los dos grupos. Se hipotetiza sobre la pertenencia de los árboles del grupo 2 al llamado grupo genético Amelonado y de los individuos híbridos (entremezclados) al genotipo Trinitario.

Este capítulo muestra el artículo de investigación publicado en la revista Genetics and Molecular Research (Ribeirão Preto Foundation for Research, ISSN 1676-5680). La cita es: Vázquez-Ovando JA, Molina-Freaner F, Nuñez-Farfán J, Ovando-Medina I, Salvador-Figueroa M (2014) Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico. Genetics and Molecular Research 13(4): 10404-10414.



Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico

J.A. Vázquez-Ovando^{1,4}, F. Molina-Freaner², J. Nuñez-Farfán³,
I. Ovando-Medina⁴ and M. Salvador-Figueroa⁴

¹Graduate Program in Biological Sciences,
National Autonomous University of Mexico, Mexico City, Mexico

²Institute of Ecology, Ecology of Biodiversity Department,
National Autonomous University of Mexico, Hermosillo, Sonora, Mexico

³Institute of Ecology, Evolutionary Ecology Department,
National Autonomous University of Mexico, Mexico City, Mexico

⁴Center of Biosciences, Autonomous University of Chiapas, Tapachula,
Chiapas, Mexico

Corresponding author: M. Salvador-Figueroa
E-mail: miguel.salvador@unach.mx

Genet. Mol. Res. 13 (4): 10404-10414 (2014)

Received October 11, 2013

Accepted July 7, 2014

Published December 12, 2014

DOI <http://dx.doi.org/10.4238/2014.December.12.2>

ABSTRACT. Criollo-type cacao trees are an important pool of genes with potential to be used in cacao breeding and selection programs. For that reason, we assessed the diversity and population structure of Criollo-type trees (108 cultivars with Criollo phenotypic characteristics and 10 Criollo references) using 12 simple sequence repeat (SSR) markers. Cultivars were selected from 7 demes in the Soconusco region of southern Mexico. SSRs amplified 74 alleles with an average of 3.6 alleles per population. The overall populations showed an average observed heterozygosity of 0.28, indicating heterozygote deficiency (average fixation index $F = 0.50$). However, moderate allelic diversity was found within populations

(Shannon index for all populations $I = 0.97$). Bayesian method analysis determined 2 genetic clusters ($K = 2$) within individuals. In concordance, an assignment test grouped 37 multilocus genotypes (including 10 references) into a first cluster (Criollo), 54 into a second (presumably Amelonado), and 27 admixed individuals unassigned at the 90% threshold likely corresponding to the Trinitario genotype. This classification was supported by the principal coordinate analysis and analysis of molecular variance, which showed 12% of variation among populations ($F_{ST} = 0.123$, $P < 0.0001$). Sampled demes sites (1-7) in the Soconusco region did not show any evidence of clustering by geographic location, and this was supported by the Mantel test ($R_m = 0.54$, $P = 0.120$). Individuals with high Criollo lineage planted in Soconusco farms could be an important reservoir of genes for future breeding programs searching for fine, taste, flavor, and aroma cocoa.

Key words: Population structure; Genetic diversity; Genetic resources; Fine flavor cocoa; Simple sequence repeat markers

INTRODUCTION

The cacao tree (*Theobroma cacao* L.) is the most economically important species of the genus *Theobroma* (Cuatrecasas, 1964). Cocoa butter, cocoa powder, and chocolate can be obtained from its beans. In addition, it is a source of products for the food, pharmaceutical, and cosmetic industries (Wood and Lass, 1985). This plant, which is almost exclusively allogamous, appears to be native to South America (Cuatrecasas, 1964; Motamayor et al., 2008), although it was domesticated in Mesoamerica (Gómez-Pompa et al., 1990; De la Cruz et al., 1995; Coe and Coe, 1996). In this region of the American continent, the history of cacao dates from ancient times (Henderson et al., 2007). The earliest evidence of cacao use dates from 1500 to 1900 BC and was reported by Powis et al. (2007, 2008) in Paso de Amada, in the southern Mexican state of Chiapas. Cacao was also cultivated and used in the pre-Columbian era in countries such as Guatemala, Honduras, and Belize (Hurst et al., 2002; Powis et al., 2007). The first group classification accepted of *T. cacao* is based on morpho-geographical traits of trees and fruits. From this classification, Forastero, Trinitario (admixture of Forastero and Criollo), and Criollo, likely domesticated by Mesoamerican natives (Cheesman, 1944), are commonly known. More recently, a study by Motamayor et al. (2008) using simple sequence repeat (SSR) markers reported 10 genetic groups in cacao: Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañón, Nacional, Nanay, and Puris.

The Criollo variety produces fruits with high-quality sensory characteristics, such as sweet pulp, less bitter beans, and attractive aroma, flavor, and taste (Smith, 1999). Although the Criollo cacao spread in Mesoamerica and was socio-economically attractive, the presence of unknown diseases and other socio-cultural events caused the loss of cacao crops in Southern Mexico and Central America (Zhang et al., 2011). However, Whitkus et al. (1998) and Motamayor et al. (2002) reported that, in that region, there are still naturally grown Criollo cacao trees. Such individuals of the Criollo-type could be a useful resource for cacao breed-

ing and selection programs. Especially, gourmet markets (niche for the Criollo-type) have increased in recent years (Vázquez-Ovando et al., 2012) up to 6.8% of the current cacao bean demand in the world (ICCO, 2012).

In 2011, Mexico produced 21,400 tons of cacao beans (FAO, 2012). The Soconusco region of southern Mexico produces 30% of the national production of cacao beans (SIAP, 2013). During the second half of the twentieth century, Mexican government programs promoted the renovation of old, sick Criollo orchards with Trinitario and Forastero varieties (López-Mendoza, 1987; Ogata, 2003). Nevertheless, experienced cacao producers are still able to identify and select native Criollo trees based on phenotypic characteristics from traditional plantations.

The finding of cacao trees of near-pure Criollo lineage in the area of Soconusco suggests that seedling propagation practices by cacao farmers have disseminated this material in the area. The knowledge of the population genetic diversity and structure of that germplasm or farmer selections compared with wild Criollo individuals can provide guidelines for future selection and improvement programs of fine cacao. The objectives for this study were to identify trees with high Criollo ancestry and to compare the molecular genetic diversity of trees selected based on Criollo fruit traits in the Soconusco area in Mexico.

MATERIAL AND METHODS

Plant material and sample collection

A total of 118 leaf samples comprising 108 cultivars from farms in Soconusco, Mexico, and 10 Criollo references from wild collections were analyzed. Cultivars were selected from 19 farms in 10 municipalities in Soconusco, Chiapas, Mexico. They were *a priori* grouped in 7 demes according to their geographical proximity and similar environmental conditions (Table 1; Figure 1). The individuals were selected based on traits of fruits (pod) and seeds, which resembled those of the Criollo-type. Pods were elongated, deeply grooved, pointed at the pod end, had a lumpy surface with a warty appearance outside, white or slightly pigmented seeds, and sweet mucilage as described by Engels (1983). Samples were collected from 1 to 16 cacao trees per farm (Table 1). The average age of cacao trees was 30 years, and particular care was taken to include only sexually propagated plants (farmer's selections). Reference accessions with Criollo lineage referenced as Carmelo, Lacandón 06, Lacandón 28, Xocen, Yaxcabá, Loxicha, and Pentagona from the germplasm bank of the Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (National Research Institute of Forestry, Agriculture, and Livestock; INIFAP), Campo Experimental Rosario Izapa, Chiapas, Mexico, were kindly provided by Hugo Avendaño-Arrazate. The other reference accessions (SL1, SL2, and SL3) were collected in the Lacandon rainforest (Selva Lacandona, SL), where expeditions to collect wild Criollo cacao individuals have previously been reported (Whitkus et al., 1998; Motamayor et al., 2008). Leaves were sampled and placed in plastic bags, taken to the laboratory (4°C), and processed on the same day of sampling. Because of the long distance to the SL site, 2 leaves were collected and kept at ambient temperature in 20 mL aqueous solution [35% NaCl, 1.5% cetyltrimethylammonium bromide (CTAB), and 0.02% sodium azide] to maintain DNA integrity (Bhattacharjee et al., 2004).

Table 1. Location and number of *Theobroma cacao* L. trees considered in this study.

Socomsusco location	Geographic location	Code	Sample size (N)	a priori population assignment (deme)
Tipachula	14°52'55"N, 92°11'42"W	TASG	16	1
Cacahonán	14°59'53"N, 92°04'44"W	CAAM	14	2
Tuxtla Chico	14°56'41"N, 92°09'59"W	TCHR	8	2
Villa Comaltitlán	15°10'31"N, 92°08'06"W	VCHL	4	3
Villa Comaltitlán	15°11'17"N, 92°06'55"W	VCLB	3	3
Villa Comaltitlán	15°15'06"N, 92°06'17"W	VCAV	5	3
Mapastepec	15°28'07"N, 92°08'42"W	MAJH	3	3
Frontera Hidalgo	14°47'31"N, 92°11'11"W	FHSA	6	4
Suchiate	14°38'27"N, 92°13'47"W	SUED	7	4
Muzatán	14°38'58"N, 92°09'06"W	MAMG	13	5
Huachuacán	14°59'27"N, 92°06'34"W	HUTG	4	6
Huachuacán	14°59'28"N, 92°06'44"W	HUJF	6	6
Huachuacán	14°59'09"N, 92°06'41"W	HUCM	2	6
Huachuacán	15°00'46"N, 92°06'32"W	HUCO	3	6
Huachuacán	15°03'22"N, 92°01'54"W	HUJL	3	7
Tuzantán	15°08'10"N, 92°04'11"W	TUCA	2	7
Tuzantán	15°08'32"N, 92°05'11"W	TURA	1	7
Tuzantán	15°09'37"N, 92°03'39"W	TURH	2	7
Tuzantán	15°04'59"N, 92°04'01"W	TUMM	6	7
Severál ^a	Severál	INIFAP	7	8
Benemérito de las Américas ^b	16°64'37"N, 90°56'26"W	SL	3	8

^aTrees identified as Criollo wild cacaos (sampled from the germplasm bank of the National Research Institute of Forestry, Agriculture, and Livestock, INIFAP). ^bSamples collected in the Lacandon Forest (wild Criollo).

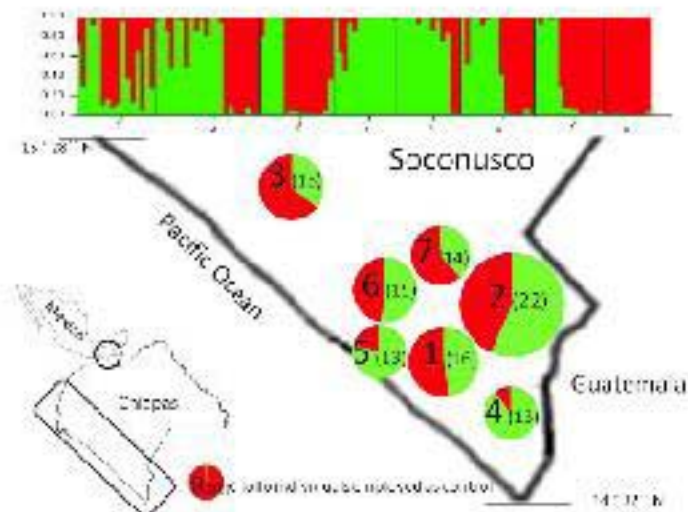


Figure 1. Genetic structure of the 7 demes of Criollo cacao trees sampled in Soconusco, Chiapas, Mexico, and one reference group with high Criollo ancestry. Results were obtained using the Structure 2.3.2 software (Pritchard et al., 2000) with 200,000 burn-in iterations, 400,000 iterations after burn-in, and 20 repetitions of each genetic population (K1-K5), admixture ancestry, and allele frequency-correlated model. The numbers in the pie charts denote the geographic deme and the number of individuals in parentheses. Red and green represent Criollo and presumably Amelonado types, respectively.

DNA extraction and SSR analysis

Total DNA extraction was performed by modifying the method described by Doyle and Doyle (1990). Leaves were washed with sterile water and 70% ethyl alcohol. Approximately

200 mg cacao leaves were ground with liquid nitrogen with 60 mg polyvinyl pyrrolidone and 1 mL CTAB buffer [2% CTAB (w/v), 20 mM ethylenediaminetetraacetic acid (EDTA), 1.4 M NaCl, 100 mM Trizma® base, pH adjusted to 8 with HCl, and 1% 2-mercaptoethanol (v/v)]. DNA extractions were performed with chloroform-isoamyl alcohol and precipitation with isopropanol. The extracted DNA was purified with a mixture of phenol:chloroform:isoamyl alcohol (25:24:1). The DNA integrity (dissolved in 60 µL Milli-Q water) was checked by 0.8% agarose electrophoresis and quantified by spectrophotometry at 260 nm (GBC Cintra 10e™ spectrophotometer, Dandenong, VIC Australia), and the purity was inferred by the 260/280 and 260/230 absorbance ratios.

Twelve highly polymorphic SSR markers were used in this test (Lanaud et al., 1999; Riju et al., 2009; Table 2). After evaluating reagent concentration and temperature conditions of several protocols, polymerase chain reaction (PCR) was performed according to Schnell et al. (2005) with modifications. These loci markers were chosen because they are distributed throughout the *T. cacao* genome (TropGene DB, Hamelin et al., 2013). They are part of the international set of SSR markers for cacao tree characterization reported by Saunders et al. (2004) with good discriminatory power (Motilal et al., 2009). Only those that showed reproducible and polymorphic bands in at least 3 trials were selected.

The PCR mixture contained 30 ng genomic DNA, 0.4 µM of each primer (Integrated DNA Technologies®, Coralville, IO USA), 0.2 mM dNTP Mix (Promega®, Madison, WI, USA), 1X PCR assay buffer ViBuffer A (Vivantis Oceanside, CA, USA), 4 mM MgCl₂, and 1.25 U Taq DNA polymerase (Vivantis, USA) in a reaction volume of 25 µL. DNA template was initially denatured at 94°C for 4 min, followed by 32 cycles of PCR amplification in a thermal cycler TC3000 (Techne, Cambridge, UK) with the following conditions: 30 s denaturation at 94°C, 1 min at appropriate annealing temperature for each primer (Table 2), and 1 min primer extension at 72°C, followed by a final extension at 72°C for 5 min. PCR products (13 samples per gel) were separated on 12% polyacrylamide gels (size 13 x 12 cm) using 0.5X Tris-borate-EDTA buffer, stained with ethidium bromide (0.6 ng/µL) for 30 min, and visualized under ultraviolet light and photographed with a Gel Doc™ EZ Imager gel documentation system (Bio-Rad, Hercules, CA, USA). Fragment sizes were estimated using the Image Lab software utility v. 4.0.1 (Bio-Rad Laboratories) option of the gel documentation system and by integrating GeneRuler™ Low Range DNA Ladder (Fermentas, Carlsbad, CA, USA). Weak bands (ghost bands) and smeared products were ignored. To confirm the reproducibility of results, DNA was extracted twice, amplified, and quantified (4 amplifications per individual).

Data analysis

Using the GenAlEx v6.5b3 software (Peakall and Smouse, 2006), genetic diversity analysis was performed by calculating global and population parameters of polymorphism, number of alleles, effective number of alleles, Shannon index, observed heterozygosity, expected heterozygosity, and fixation index. Chi-square tests were performed to test the balance-imbalance of Hardy-Weinberg equilibrium by locus in all populations. Private alleles were identified by individual and by population. Genetic and geographic linear distances were calculated for all individuals except for reference trees. Mantel correlation test (10,000 permutations) was performed using the GenAlEx v6.5b3 software. The presence of null alleles and scoring for errors were analyzed using the Micro-Checker software (Van Oosterhout et al., 2004).

The population structure and genetic ancestry of individuals were inferred using the Bayesian statistical methods from Structure v 2.3.2 (Pritchard et al., 2000). The following parameters were set up: admixture model of ancestry, 200,000 iterations during the burn-in period, 400,000 iterations after the burn-in period, and 20 repetitions for each genetic subpopulation (*K*1-*K*5) using the allele frequencies correlated model. The number of *K* was estimated following the procedure described by Evanno et al. (2005) using the Structure Harvester v0.6.93 (Earl and von Holdt, 2012). To determine the degree of differentiation within and among populations, analysis of molecular variance (AMOVA) was performed with 10,000 iterations using the GenAlEx v6.5b3.

On the basis of the coefficient of membership to 1 of 2 distinct genetic clusters generated by structure analysis, all individuals were separated into genotypic classes (assignment percentage threshold of 90%). The first class included individuals with higher than 90% membership to cluster 1 (primarily red vertical bars in Figure 1). A second class included trees with higher than 90% membership to cluster 2 (primarily green vertical bars in Figure 1). Unassigned individuals (less than 90% membership to any cluster) were considered trees with admixed ancestry. We performed an additional AMOVA for these classes. The grouping was further verified by the assignment method of Paetkau et al. (2004) using the GeaAlEx v6.5b3 software. The assignment was considered correct when 100% of individuals converged with their respective group. A principal coordinate analysis (PCoA) was conducted using a pairwise distance matrix with distance and covariance standardized. The plot was built with the 2 axes that best explain the total variation (43%) of subpopulations.

RESULTS

Twelve SSR loci detected 12 homozygous individuals (range 0-3 individuals per population) and distinguished 74 alleles among the 118 individuals analyzed, with a range of 3 to 13 alleles per locus (average of 6.2 alleles per locus). Fragment sizes were as expected, except for the mTcCIR3 locus, which was shorter than expected [previously reported by Lanaud et al. (1999); Table 2]. The Micro-Checker analysis confirmed no scoring error due to stuttering and no evidence for large allele dropout for mTcCIR3. However, the analysis suggests that null alleles may be present at the mTcCIR3 locus. Seven private alleles were detected in 5 of the 12 microsatellites. Individuals HUTG01, MAMG07, and TASG18 have a unique private allele at locus mTcCIR 6, mTcCIR 7, and mTcCIR 28, respectively. The Tapachula subpopulation had the highest proportion of individuals with private alleles detected at the mTcCIR28 locus (Table 3).

Table 2. Microsatellite primers used in molecular characterization, melting temperatures (*T_m*), and fragment sizes obtained.

Microsatellite locus	Chromosome ^a	Physical position ^b (cM)	<i>T_m</i> (°C)	Size of expected fragment ^c (bp)	Size range in study (bp)
mTcCIR1	8	0.0	64.7	143	133-155
mTcCIR3	2	19.5	57.1	249	197-227
mTcCIR6	6	10.3	54.2	231	163-265
mTcCIR7	7	27.3	56.2	160	142-194
mTcCIR8	9	52.6	51.5	301	265-319
mTcCIR11	2	92.2	49.7	298	278-370
mTcCIR12	4	45.7	54.4	188	168-274
mTcCIR15	1	18.0	53.4	254	228-296
mTcCIR19	2	10.3	56.7	376	214-310
mTcCIR21	3	25.8	61.0	157	142-170
mTcCIR25	6	45.0	54.1	153	128-175
mTcCIR28	6	55.3	57.3	336	322-362

^aRisterucci et al. (2000); ^bTropGene database (Hamelin et al., 2013); ^cLanaud et al. (1999).

Table 3. Private alleles by population and individual of *Theobroma cacao* L. trees sampled in Soconusco, Chiapas, Mexico.

Population	Locus	Private allele	Frequency	Individuals
Huehuetán (Pop 6)	mTcCIR 6	233	0.077	HUTG01
Huehuetán	mTcCIR 8	307	0.200	HUTG02, HUIP05
Mixtán (Pop 5)	mTcCIR 7	194	0.083	MAMG07
Sochiute (Pop 4)	mTcCIR 19	352	0.111	SUED02, SUED05
Tapachula (Pop 1)	mTcCIR28	328	0.036	TASG18
Tapachula	mTcCIR28	330	0.357	TASG03, TASG09, TASG10, TASG13, TASG14, TASG15, TASG16, TASG17
Tapachula	mTcCIR28	338	0.250	TASG14, TASG15, TASG16, TASG17, TASG18

Additionally, the microsatellite loci revealed polymorphism rates for subpopulations with values over 91.7%, effective number of alleles between 2.35 and 2.63, and values for the Shannon diversity index between 0.87 and 1.02 (Table 4). Looking at the expected heterozygosity ($H_e = 0.54$), observed heterozygosity ($H_o = 0.28$), and positive value of fixation index (F), all of the *a priori* determined subpopulations presented heterozygote deficiency ($H_e > H_o$). Except for mTcCIR 1 (which revealed a value of 4.29), the other SSR loci had low numbers of migrants for all subpopulations (0.284-1.784), with an average value of 1.39 for all SSR markers.

Table 4. Average values as indicators of population diversity in *Theobroma cacao* L. trees sampled from Soconusco, Chiapas, Mexico. Values were obtained using the GenAlEx v6.5b3 software (Peakal and Smouse, 2006).

Population	N	N_A	N_e	I	H_o	H_e	F
Pop 1	13.8	3.75	2.53	1.02	0.34	0.56	0.41
Pop 2	15.2	4.00	2.59	1.03	0.24	0.56	0.54
Pop 3	12.8	3.67	2.60	1.02	0.23	0.56	0.64
Pop 4	9.0	3.17	2.39	0.89	0.25	0.52	0.53
Pop 5	8.6	3.67	2.63	1.02	0.20	0.56	0.66
Pop 6	10.8	3.50	2.60	0.95	0.35	0.52	0.32
Pop 7	10.8	3.42	2.51	0.95	0.22	0.55	0.63
Pop 8	8.0	3.17	2.35	0.87	0.38	0.50	0.19
Mean	11.1	3.54	2.52	0.97	0.28	0.54	0.49
SE	0.4	0.16	0.11	0.04	0.03	0.02	0.05

N = number of individuals; N_A = number of different alleles; N_e = effective number of alleles; I = Shannon diversity index; H_o = observed heterozygosity; H_e = expected heterozygosity; F = fixation index; SE = standard error.

The effective number of alleles ($N_e = 2.52$) for all subpopulations was lower than the number of alleles ($N_A = 3.54$). The Hardy-Weinberg test revealed disequilibrium in five loci (mTcCIR 3, mTcCIR 6, mTcCIR 8, mTcCIR 25, and mTcCIR 28) in all subpopulations ($P < 0.05$). The subpopulations Tapachula (1) and Villa Comaltitlán (3) exhibited this disequilibrium for 11 of 12 tested SSR markers.

AMOVA was conducted to detect inter- and intra-population variation. The analysis revealed that the largest variation (88%) was explained by the variation within populations, while the remaining (12%) was among populations, with an F_{ST} value of 0.123 ($P < 0.0001$). Bayesian statistical analysis using an admixture model and calculating the best value of K by Evanno et al. methods (Evanno et al., 2005) indicated that the 118 individuals grouped in 2 genetic clusters ($K = 2$, $\Delta K = 294.21$). Structure analysis without *a priori* deme designations indicated that the 118 individuals were effectively grouped into 2 genetic clusters ($\Delta K = 278.89$).

Geographic demes Villa Comaltitlán (3) and Tuzantán (7) shared a high proportion of alleles (red color key in Figure 1) with Criollo references (8). The other group had individuals from demes Cacaboatán (2), Suchiate (4), and Mazatán (5), with more than 50% of non-Criollo alleles. Tapachula (1) and Huehuetán (6) shared an equal proportion of alleles from both genetic clusters. Sampled demes sites (1-7) in the Soconusco region did not show any evidence of clustering by geographic location (Figure 1), and this was supported by a non-significant Mantel test ($R_m = 0.54$, $P = 0.120$).

Assignment analysis performed with the GenAlEx v6.5b3 software defining 3 clusters *a priori* grouped 37 individuals in cluster 1 (red color key in Figure 2) that included the Criollo references, 54 individuals in cluster 2 (green color key in Figure 2), and 27 individuals in cluster 3 (blue color key in Figure 2), which was called “hybrids”, with 100% probability. We found moderate differentiation among that clusters (AMOVA $F_{ST} = 0.111$). The PCoA plot showed more clearly the differences among clusters. Coordinates 1, 2, and 3 explain 22.7, 20.01, and 17.15% of the total variation, respectively. Figure 2 shows the separation between the Criollo-type trees, non-Criollo, and “hybrids” in the overlapping region where allele frequencies are shared.

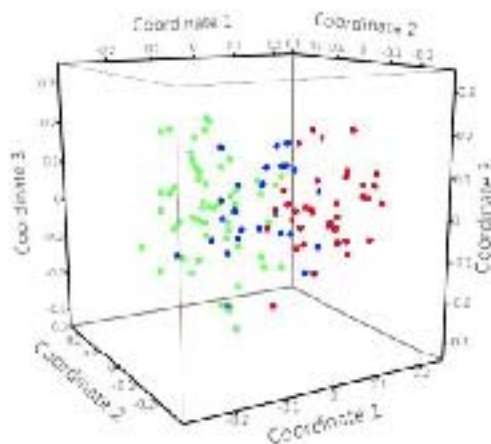


Figure 2. Principal coordinate analysis of 118 cacao trees comprising 108 cultivars sampled from 7 demes in Soconusco, Chiapas, Mexico, and 10 Criollo references belonging to 2 genetic clusters $K=2$ based on the Structure software. Criollo cluster (red) and genetic cluster 2 (green) are shown. Blue indicates hybrid individuals that likely belong to the Trinitario cultivar. Analysis was conducted with data from 12 microsatellites and 100% assignment by GenAlEx v6.5b3 (Peakall and Smouse, 2006). Coordinates 1, 2, and 3 explain 22.7, 20.01, and 17.15% of the total variation, respectively.

DISCUSSION

In this study, we investigated the genetic structure and relationships among 7 demes of Criollo-type cacao trees cultivated in Soconusco, Chiapas, Mexico, using SSR markers. To our knowledge, this is the first report on the genetic diversity of cacao grown in the Soconusco area where theobromine residues have tested positive on archaeological pottery specifically at the Mokaya site of Paso de Amada (Powis et al., 2007, 2008). There are studies on genetic diversity of the germplasm collected in southeast Mexico; however, these have focused on wild accessions. For example, Motamayor et al. (2008) analyzed, among others, wild cacaos from

the Lacandon Forest with SSRs markers. In the same way, Whitkus et al. (1998) employed random amplified polymorphic DNA (RAPD) markers to examine the genetic diversity of cultivars from the Tabasco area, which is located 300 km north of our study, and wild trees collected from the Lacandon Rainforest.

The fixation rates found in our study for all *a priori* subpopulations represents high levels of inbreeding or possible crossover between siblings as supported by a deficiency of heterozygotes ($H_E > H_O$). Such levels of homozygosity may suggest intensive cultivation and selection of Criollo genotypes for a long period of time. Selection pressure increases levels of inbreeding (Loor Solorzano et al., 2012) by 2 mechanisms: self-compatibility, as reported in ancient Criollo (Cheesman, 1944; Haddon, 1961), and selection (unintended) of homozygous trees by local cacao farmers. It is not surprising that populations had a deficiency of heterozygotes because of the effect of deliberate sampling of individuals with Criollo characteristics. Another hypothesis to explain the heterozygote deficiency is the presence of a founder effect, i.e., the populations analyzed descended from a narrow number of parents and were continuously cultivated in the Soconusco region by local cacao farmers. The values obtained in the Hardy-Weinberg test support this hypothesis.

The sampling effort in this study was influenced by the intention toward a specific group of individuals (Criollo-type trees), reflected by the $N_E < N_A$ values, setting limits for the population study. It would be desirable to assess the diversity of all individuals in a greater sampling effort in future studies. It is important to note that, despite the high homozygosity, the Shannon diversity index shows significant allelic diversity within populations, which is consistent with the total number of alleles revealed by microsatellites (total = 74; mean = 6.2/SSR) for all individuals. Johnson et al. (2009), using microsatellite markers (260 alleles, using 35 SSRs; mean 7.4/SSR), reported a genetic diversity that was similar to that of our study. This proximity in mean number of alleles could be an indication that more than 50% of the individuals analyzed in our study (particularly demes 2, 4, 5, and 6) corresponded to Trinitario types (cluster 2, green color key in Figures 1 and 2). This may have 2 explanations. First, the phenotypic characteristics utilized for collecting individuals in the farmer's field were not discriminatory enough to select only Criollo trees. Second, when selection was based on sensory characteristics, trees were also the Trinitario type.

By analyzing the genetic structure of populations and finding $K = 2$, a genetic differentiation can be observed between the 2 clusters (Figures 1 and 2). Almost 75% of the total individuals were assigned to the groups with high levels of correspondence; however, the remaining individuals (27) are located in the transition region between the pools. The genetic structure analysis clustered wild individuals that were collected in the Mayan area (INIFAP and SL Table 1; mean coefficient of membership to cluster 1 = 0.984) together with 27 cultivated individuals (mean coefficient of membership to cluster 1 = 0.974). This finding suggests that a group of Criollo-type cacao was found in the Mokaya region. In contrast, Whitkus et al. (1998), using RAPD marker analysis, did not find an association between wild Criollo (from the Lacandon Forest and Yucatan State, a zone of influence of the Mayan culture in Mexico) and accessions cultivated from Tabasco, Mexico, and South America. Individuals in the overlapping area (blue color key in Figure 2), which were identified as "hybrids," are admixed individuals that share alleles with Criollos and could be considered candidates for Trinitario. The parental contribution of Mesoamerican Criollos to the appearance of Trinitario cacao has been documented in the past (Motamayor et al., 2003).

Cluster 2 grouped 54 individuals (green in Figures 1 and 2) with a clear non-Criollo

lineage, but these individuals share alleles with “hybrids.” Possibly, these individuals belong to a particular genetic group lacking reference accessions in this study; therefore, they are now unassigned. A strong contender for this group, based on Trinitario history, is the Amelonado. More studies, including parental controls of Trinitario, are needed to address this question.

Criollo-type individuals represent a gene pool with potential to be included in breeding programs searching for fine flavor cocoa; however, is necessary to evaluate their agronomic potential, disease tolerance, and sensory qualities. Some of these materials will also serve as the basis for future selection studies to propagate and establish plantations with enhanced cocoa qualities. On the other hand, it is important to increase the heterozygosity of populations without neglecting the important features of the local cacaos. Private bands could be isolated, sequenced, and converted to sequenced-characterized amplified region or other markers, which would assist the identification of improved genotypes, because they are specific tags.

ACKNOWLEDGMENTS

We thank Dr. Hugo Avendaño-Arrazate, who kindly provided the leaf material of wild cacao accessions belonging to the germplasm bank of the National Research Institute of Forestry, Agriculture, and Livestock (INIFAP, Mexico); the Graduate Program in Biological Sciences of the National Autonomous University of Mexico (UNAM) for the academic instruction in doctoral studies for the first author; and the National Council of Sciences and Technology (CONACyT, Mexico) for financial grant awarded to the first author. Special thanks to Samuel Guillén for contact with farmers and to Carlos Amores, Julio Magaña, Dafne Castillo, Nayelly Hernández, Nidia Gutiérrez, and Julio Coutiño for technical support.

REFERENCES

- Rhattacharjee R, Kolosnikova-Allen M, Aikopolouidi P, Taino S, et al (2004). An improved semi-automated rapid method of extracting genomic DNA for molecular marker analysis in cocoa, *Theobroma cacao* L. *Plant Mol. Biol. Rep.* 22: 435-436.
- Chesman E (1944). Notes on the nomenclature, classification and possible relationships of cocoa populations. *Trop. Agr.* 21: 144-159.
- Coe SD and Coe MD (1996). *The True History of Chocolate*. 1st edn. Thames and Hudson Ltd., New York.
- Cuatrecasas J (1964). Cacao and its allies: A taxonomic revision of the genus *Theobroma*. *Contr. U. S. Natl. Herb.* 35: 379-614.
- De la Cruz M, Whitkus R, Gómez-Pompa A and Mota-Bravo L (1995). Origins of cacao cultivation. *Nature* 375: 542-543.
- Doyle JJ and Doyle JL (1990). Isolation of plantDNA from fresh tissue. *Focus* 12: 13-15.
- Earl DA and von Holdt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361.
- Engels JMM (1983). A systematic description of cacao clones. I. The discriminative value of quantitative characteristics. *Euphytica* 32: 377-385.
- Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620.
- FAO (Food and Agriculture Organization) (2011). FAOSTAT Online Statistical Service. Available at [<http://faostat.fao.org/>]. Accessed September 12, 2013.
- Gómez-Pompa A, Flores JS and Fernández MA (1990). The sacred cacao groves of the Maya. *Lat. Am. Antiq.* 1: 247-257.
- Haddon AV (1961). Variety trials of seedling cacao in Malaya. *Malayan Agric. J.* 43: 169-232.
- Hamelin C, Sempere G, Jouffé V and Ruiz M (2013). TropGensDB, the multi-tropical crop information system updated and extended. *Nucleic Acids Res.* 41: D1172-D1175.
- Henderson JS, Joyce RA, Hall GR, Hurst WJ, et al. (2007). Chemical and archaeological evidence for the earliest cacao beverages. *Proc. Natl. Acad. Sci. U. S. A.* 104: 18937-18940.

- Hurst WJ, Tarka SM Jr, Powis TG, Valdez F Jr, et al. (2002). Cacao usage by the earliest Maya civilization. *Nature* 418: 289-290.
- ICCO (International Cocoa Organization) (2012). Quarterly Bulletin of Cocoa Statistics. Vol. XXXVIII-No.2. Cocoa Year 2010/11 London (May 2012).
- Johnson ES, Bekale FL, Brown SJ, Song Q, et al. (2009). Population structure and genetic diversity of the Trinitario cacao (*Theobroma cacao* L.) from Trinidad and Tobago. *Crop Sci.* 49: 564-572.
- Lanard C, Risturacci AM, Piarotti I, Falque M, et al. (1999). Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* 8: 2141-2143.
- Loor Solórzano RG, Fouat O, Lemaître A, Pavak S, et al. (2012). Insight into the wild origin, migration and domestication history of the fine flavour Nacional *Theobroma cacao* L. variety from Ecuador. *PLoS One* 7: e48438.
- Lopez-Mendoza R (1987). El cacao en Tabasco. *Colección Cuadernos Universitarios. Serie Agronómica* no. 13. Universidad Autónoma de Chapingo, Texcoco.
- Motamayor JC, Risturacci AM, Lopez PA, Ortiz CF, et al. (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89: 380-386.
- Motamayor JC, Risturacci AM, Heath M and Lanard C (2003). Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity* 91: 322-330.
- Motamayor JC, Lachandez P, da Silva e Mota J, Loor R, et al. (2008). Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS One* 3: e3311.
- Motilal LA, Zhang D, Umaharan P, Mischke S, et al. (2009). Increasing accuracy and throughput in large-scale microsatellite fingerprinting of cacao field germplasm collections. *Trop. Plant Biol.* 2: 23-37.
- Ogata N (2003). Domestication and Distribution of the Chocolate Tree (*Theobroma cacao* L.) in Mesoamerica. In: *The Lowland Maya Area. Three Millennia at the Human-Wildland Interface* (Gómez-Pompa A, Allen MF, Fedick SL and Jiménez-Osorio JJ, eds.). The Haworth Press Inc., New York, 415-438.
- Pætkau D, Slade R, Burden M and Estoup A (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* 13: 55-65.
- Peakall R and Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6: 288-295.
- Powis T, Hurst J, Rodríguez MC, Ortiz P, et al. (2007). Oldest chocolate in the New World. *Antiquity* 81: 314.
- Powis T, Hurst J, Rodríguez MC, Ortiz P, et al. (2008). The origins of cacao use in Mesoamerica. *Mexicon* 30: 35-38.
- Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Riju A, Rajesh MK, Sherin PT, Chandrasekar A, et al. (2009). Mining of expressed sequence tag libraries of cacao for microsatellite markers using five computational tools. *J. Genet.* 88: 117-225.
- Risturacci AM, Grivet L, N'Goran JAK, Piarotti I, et al. (2000). A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 101: 948-955.
- Sanders JA, Mischke S, Leamy EA and Hemsida AA (2004). Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theor. Appl. Genet.* 110: 41-47.
- SIAP (Servicio de Información Agroalimentaria y Pesquera, SAGARPA México) (2013). Producción Agrícola por Cultivo, Reporte 2012. Available at [http://www.siap.gob.mx/index.php?option=com_wrapper&view=wrapper&layout=350]. Accessed July 30, 2013.
- Schnell RJ, Olano CT, Brown JS, Mesrow AW, et al. (2005). Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. *J. Am. Soc. Hort. Sci.* 130: 181-190.
- Smith N (1999). *The Amazon River Forest: a Natural History of Plants, Animals, and People*. Oxford University Press, New York.
- Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4: 535-538.
- Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J and Salvador-Figueroa M (2012). Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad. *BioTecnología* 16: 36-56.
- Whitkus R, de la Cruz M, Mota-Bravo L and Gómez-Pompa A (1998). Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theor. Appl. Genet.* 96: 621-627.
- Zhang D, Gardini EA, Motilal LA, Baligar V, et al. (2011). Dissecting genetic structure in farmer selections of *Theobroma cacao* in the Peruvian Amazon: implications for on farm conservation and rehabilitation. *Trop. Plant Biol.* 4: 106-116.
- Wood GAR and Lass RA (1985). *Cocoa*. Tropical Agriculture Series. 4th edn. Blackwell Science Logman Group Ltd., New York.

CAPÍTULO 4

Importancia de alcaloides y polifenoles en el sabor y aroma del cacao

Las semillas de cacao son matrices ricas en alcaloides y compuestos polifenólicos. Su concentración y proporción de cada uno de los compuestos específicos depende de varios factores. Además, su contribución en el sabor y olor tanto de la almendra como del producto que de ellas puede obtenerse, es innegable. En este capítulo se aborda una revisión sobre estos aspectos y se enfatiza en las rutas de biosíntesis de estas importantes moléculas, de los mecanismos genéticos que regulan esta síntesis y de cómo algunos factores externos a la planta pueden modificar su composición y contenido. Se aborda además como los factores poscosecha, la fermentación, el secado y el tostado de las almendras incide en la dinámica de estos compuestos, en las transformaciones que ocurren y de los cambios últimos en el sabor y aroma del cacao.

Esta revisión se envió para ser considerada a publicación a la revista *Interciencia* (Asociación Interciencia, ISSN 0378-1844). El título propuesto fue: Alcaloides y polifenoles del cacao: mecanismos que regulan su síntesis y sus implicaciones en el sabor y aroma.

ALCALOIDES Y POLIFENOLES DEL CACAO: MECANISMOS QUE REGULAN SU SÍNTESIS Y SUS IMPLICACIONES EN EL SABOR Y AROMA

COCOA ALKALOIDS AND POLYPHENOLS: MECHANISMS THAT REGULATE THEIR SYNTHESIS AND ITS IMPLICATIONS ON THE TASTE AND AROMA

ALCALÓIDES E POLIFENÓIS DO CACAU: MECANISMOS QUE REGULAM A SUA SÍNTESE E SUAS IMPLICAÇÕES PARA O SABOR E AROMA

RESUMEN

El sabor y aroma de los granos de cacao (*Theobroma cacao* L.) de la variedad Criollo, fueron las principales razones que promovieron su domesticación y uso alimentario por los pueblos precolombinos de Mesoamérica. Los alcaloides y los polifenoles, son compuestos que de manera directa inciden en el sabor de las almendras y de manera indirecta sobre los precursores de aroma. Los alcaloides, están asociados con el amargor, su concentración está relacionada con la variedad y se modifica con el procesamiento. Los polifenoles son responsables, junto con otras moléculas de la astringencia (poco deseable en chocolates), pero también de propiedades antioxidantes deseables por los consumidores. En esta revisión se abordan aspectos de la biosíntesis de estas importantes moléculas en las semillas de cacao, los cambios que ocurren durante el procesamiento de las mismas, así como del aporte biológico y a la salud de los consumidores.

Palabras clave: Cacao; alcaloides; polifenoles; antioxidantes; biosíntesis.

SUMMARY

The flavor and aroma of cocoa (*Theobroma cacao* L.) beans cv. Criollo, were the main reasons that promoted its domestication and food use by pre-Columbian peoples of Mesoamerica. Polyphenols and alkaloids are compounds that directly affect the flavor of the cocoa beans and indirectly on the flavor precursors. The alkaloids are associated with bitterness; its concentration is related to the variety and modifying with the process. The polyphenols are responsible together with other molecules of the astringency (not desirable in chocolate), but also of antioxidant properties high desirable by consumers. This review focuses on aspects of the biosynthesis of these important molecules in cocoa beans. The changes of these molecules that occur during processing is also approached, as well as the biological contribution and consumer health

RESUMO

O sabor eo aroma do cacau (*Theobroma cacao* L.) da variedade Criollo, foram as principais razões que promoveram sua domesticação e comida uso por povos pré-colombianos da Mesoamérica. Os alcalóides e polifenóis são compostos que afetam diretamente o sabor de amêndoas e, indiretamente sobre os precursores de aromas. Os alcalóides estão associados

com a amargura, a sua concentração está relacionada com a variedade e a alterações com o processamento. Os polifenóis são responsáveis, em conjunto com outras moléculas da adstringência (não desejável no chocolate), mas também das propriedades antioxidantes elevadas desejáveis por parte dos consumidores. Esta revisão enfoca aspectos da biossíntese destas moléculas importantes no cacau. As variações destas moléculas que ocorrem durante o processamento também são aproximadas, bem como a contribuição biológica e da saúde do consumidor.

Introducción

El cacao (*Theobroma cacao* L.), es una planta de reproducción preferentemente alógama, con número cromosómico $2n = 20$ (Dantas y Guerra, 2010), de la familia Malvaceae, cultivada en las regiones tropicales del mundo y de cuyos frutos se obtienen semillas (almendras) que son empleadas en la industria alimentaria, farmacéutica y cosmética. En el año 2012, se produjeron de 5 003 211 toneladas de granos secos, lo que representó un ingreso superior a los 4 000 millones de dólares (FAO, 2014). La industrialización de los granos, y la producción de derivados del cacao, ha reportado históricamente dividendos por más de 73 000 millones de dólares (Ploetz, 2007).

En el neotrópico, la importancia y demanda de los granos de cacao, data de tiempos precolombinos. Existen vestigios de su domesticación y uso con fines comestibles, hacia los años 1600 – 1200 a.C. por los Mokaya, asentados en el sitio conocido como Paso de Amada en Soconusco, Chiapas; México (Powis *et al.*, 2008) y los Olmecas de la región de San Lorenzo, Veracruz; México (Powis *et al.*, 2011). A la llegada de los españoles a América, se documentó ampliamente el uso del cacao como moneda entre los nativos, mostrando la importancia social y económica que ostentaba este cultivo. Paralelo al empleo como moneda, los Aztecas, destinaban su consumo a clases sociales de mayor *estatus* y en eventos religiosos, pues era sinónimo de poder y de divinidad. A la fecha se sabe que el cacao se consumía como restaurador del estado de ánimo, como energético, como estimulante de la libido y como estimulante digestivo. Con fines curativos se han documentado hasta 150 usos (Dillinger *et al.*, 2000), además de su empleo en eventos ceremoniales fúnebres de dignatarios (Ramiro *et al.*, 2005).

Posterior al descubrimiento de América, el cacao se extendió hacia las regiones tropicales de África y Asia, donde actualmente se produce más del 80 % del cacao mundial y del que los corporativos mundiales obtienen igual porcentaje de materia prima para producir la gran cantidad de productos derivados como la cocoa en polvo, chocolates, confites, cosméticos y medicamentos. La continua popularidad del cacao se debe tanto a los factores intrínsecos de las semillas como a los productos que posteriormente se obtuvieron a partir de ellas: chocolate, cocoa y “mantequilla” de cacao. Esta última, alcanza valores de hasta 56 % y su composición, evaluada como la relación ácidos grasos saturados: insaturados (promedio de 1 : 0.60) permite obtener una grasa sólida a 25 °C, con punto de fusión entre 33 - 37 °C

(Liendo *et al.*, 1997; Álvarez *et al.*, 2007) lo que permite a los productos derretirse al estar en contacto con el cuerpo humano, característica altamente apreciada en los chocolates y en productos farmacéuticos.

Se puede inferir que los pueblos mesoamericanos que domesticaron el cacao, hace casi 4000 años, basaron sus criterios de selección en aspectos relacionados con el sabor y el aroma tanto del fruto como de las almendras, criterios empíricos que en la actualidad se siguen empleando para calificar la “calidad” de las almendras comercializables y para obtener nuevos individuos (plantas) productores.

Por lo anterior, en esta revisión se aborda el papel de los alcaloides, particularmente metilxantinas y de los compuestos fenólicos en el sabor y aroma del cacao, la principal ruta biosintética de dichas moléculas y los mecanismos genéticos que la regulan, así como los cambios que dichas moléculas tienen durante el procesamiento de las semillas.

Factores que la afectan la calidad del cacao

La composición proximal de las almendras (Tabla I), resultado de la interacción de factores genéticos, ambientales y de manejo es, sin lugar a dudas, una característica que ha contribuido a la popularidad del cacao ya que influyen tanto en las características sensoriales del producto, junto a moléculas igualmente importantes como los alcaloides y los compuestos fenólicos. Estos últimos, actualmente tienen gran importancia, pues contribuyen a la capacidad antioxidante total del cacao y sus derivados (Figura 1).

Tabla I. Composición proximal de cacaos de diferentes variedades y regiones geográficas (g /kg)

Parámetro	Variedad de cacao		
	cv. Criollo*	cv. Criollo OC 63**	cv. MAR 4***
Humedad	525	46.2	63.7
Proteína	78.8	191.5	140
Grasa	239.2	563.7	560
Fibra cruda	32.3	42.3	3.7
Cenizas	20.7	52.4	33.2
Carbohidratos	105	103.9	199.4

* Sotelo y Álvarez (1991), resultados expresados en base húmeda y realizados con semillas frescas; ** Liendo *et al.* (1997); ***Álvarez *et al.* (2007).

Otro factor determinante en la calidad del cacao, es el procesamiento al que son sometidos, primero los frutos y posteriormente las almendras, ya que se ha demostrado ampliamente el efecto que puede ejercer el procesamiento poscosecha, principalmente la fermentación, el secado y tostado de las almendras que promoverán la transformación, disminución o síntesis de moléculas que han de contribuir con las características sensoriales deseables en los cacaos de calidad.

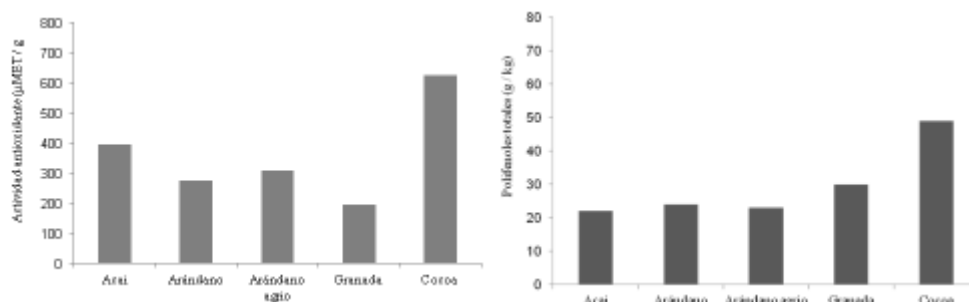


Figura 1. Actividad antioxidante evaluada por el método ORAC (IZQ) y contenido de polifenoles totales (DER), en polvo de cacao (cocoa) comparado con polvo de otras frutas (Crozier *et al.*, 2011).

Alcaloides del cacao

La teobromina (3,7-dihidro-3,7-dimetil-1*H*-purina-2,6-diona), el principal alcaloide-purina que se encuentra en los productos de los árboles de cacao, sus almendras y cáscaras, pertenece a una clase de moléculas alcaloides conocidas como metil-xantinas. Las metil-xantinas se producen de forma natural hasta en 60 diferentes especies de plantas y también incluyen a la cafeína (la principal metil-xantina del café) y teofilina (la principal metil-xantina del té). Los cotiledones de almendras maduras contienen entre 2.2 y 2.7% del peso en base seca de teobromina para la variedad Forastero y hasta 1.5 % en la variedad Criollo. La segunda metil xantina importante por su contenido es la cafeína, cuyos valores se reportan desde 0.1 hasta 0.8% (Figura 2), siendo más elevado su contenido en almendras de la variedad criollo (Sotelo y Álvarez, 1991). Otros alcaloides reportados para cacao o el género *Theobroma* pero en concentraciones más bajas son la teofilina con 0.3 % (Sotelo y Álvarez, 1991) y menos de 0.5% de teacrina (Hammerstone *et al.*, 1994).

Sobre la función que desempeñan los alcaloides en los tejidos vegetales, existen diferentes teorías. La más sólida, argumenta que su acumulación en hojas, cotiledones o semillas implica un rol biológico como químicos de defensa, ya que fungen como antiherbívoros y como compuestos alelopáticos (Bernays *et al.*, 2000; Smyth, 1992). También se ha propuesto que su presencia en los tejidos, responde a productos de desecho o

almacenamiento del nitrógeno sobrante, como una equivalencia a la del ácido úrico o de la urea en los animales o bien como reguladores del crecimiento, ya que se ha reportado su presencia incrementada en semillas en estado de germinación, sobre todo con deficiencia de nutrimentos (Evans, 2000).

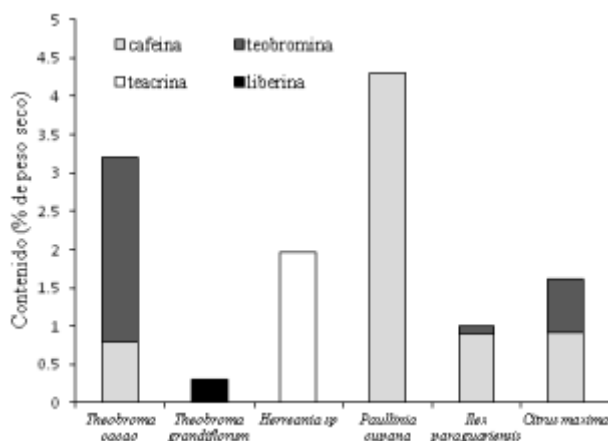


Figura 2. Contenido de alcaloides – purina en almendras de cacao comparado con otras fuentes (Ashihara *et al.*, 2011).

Desde la perspectiva del consumo, las semillas de cacao deben su sabor amargo en mayor medida a los alcaloides presentes, que no exclusivamente, ya que se ha comprobado (Stark *et al.*, 2006), que otras moléculas de naturaleza diferente a alcaloides (dicetopiperazinas, L-aminoácidos libres o péptidos) contribuyen a la percepción de amargor, o que pueden generar confusión en la percepción entre amargo-astringente (taninos de bajo peso molecular como epicatequina, catequina, procianidinas). Es indudable que algunas características de los alcaloides como estimulantes del sistema nervioso, y otros efectos farmacológicos que se les atribuyen, fueron seguramente responsables de muchas de los usos que se dio al cacao (o derivados de éste), antes incluso de conocer su composición.

El contenido de estos alcaloides durante las fases del procesamiento primario se ve disminuido dando a los productos tostados el balance adecuado de amargor deseable en cacaos finos. Con los datos antes presentados, puede notarse como los cacaos Criollos resultan ser menos amargos como resultado del mayor contenido de cafeína, pero sustancialmente menor contenido de teobromina que resulta a la percepción de un consumidor hasta once veces más amargo que la misma cafeína (Stark *et al.*, 2006).

Algunas versiones sugieren que la relación teobromina / cafeína en almendras de cacao, puede resultar un indicio de la “fineza” del cacao, toda vez que los cacaos Forasteros o “a granel” presentan valores de esta relación por arriba de 4, pudiendo superar el valor 10,

mientras que los cacaos Criollos debieran reportar valores inferiores a 4; tal como lo reporta Sotelo y Álvarez, (1991), con valores en esta relación de 11.2 y 1.6, para cacaos fermentados Forastero y Criollo, respectivamente.

Regulación genética y biosíntesis de las metil xantinas

Aunque la cafeína, que es la principal metil xantina estudiada, fue conocida desde los años 1800's no fue sino hasta el año 2000, cuando se elucidó completamente la ruta de su biosíntesis (Ashihara *et al.*, 2008). Basado en esto, se estableció que las metilxantinas, derivan de nucleótidos de purina los cuales son compuestos importantes para el metabolismo energético y como elementos esenciales para la biosíntesis de ácidos nucleicos (Stasolla *et al.*, 2003; Koyama *et al.*, 2003). Se sintetizan en las plantas tanto por la vía de *novo* como por la vía de salvamento de purinas. La vía de *novo* que produce nucleótidos de purina a partir de moléculas precursoras más pequeñas, tales como los aminoácidos y CO₂, es conservada, pero la vía de salvamento difiere entre especies (Sugiura y Takeda, 2000).

La biosíntesis de teobromina a partir de las purinas se ha estudiado mucho menos comparada con la biosíntesis de cafeína en café y otros cultivos. La teobromina, al igual que en cacao resulta ser el alcaloide predominante en plantas como *Camellia ptilophylla* (Ashihara *et al.*, 1998) y *Camellia irrawadiensis* (Ashihara y Kubota, 1987) aunque se ha reportado no se produce en todas las especies del genero *Theobroma* y relacionados de la misma familia como se puede comprobar en la Figura 2.

La síntesis de cafeína (Figura 3), con teobromina como intermediario, está catalizada predominantemente por la enzima cafeína sintasa bifuncional y fue purificada originalmente a partir de hojas de té (Kato *et al.*, 1999). Esta enzima es monomérica, con una masa molecular aparente de 41 kDa, y muestra un pH óptimo de pH 8.5. La longitud total del ADNc aislado, denominado TCS1 (AB031280), es 1 438 pb y codifica una proteína de 369 aminoácidos (Kato *et al.*, 2000).

En café, la cafeína sintasa bifuncional ha sido parcialmente purificada a partir de frutas y hojas (Mazzafera *et al.*, 1994). Esta preparación enzimática poseía actividad N-metil-transferasa asociado con los últimos pasos de biosíntesis de cafeína, y ya se han reportado los genes que codifican para los tipos de N-metil-transferasa implicadas en los mismos dos últimos pasos (Mizuno *et al.*, 2003; Mizuno *et al.*, 2001; Uefuji *et al.*, 2003). Estos son CTS1 (AB034700), CaMXMT1 (AB048796), CCS1 (AB086414) y CaDXMT1 (AB084125).

A diferencia de lo ocurre en café, en cacao y otras especies de *Camellia*, la teobromina es el alcaloide principal, lo que puede ser explicado porque la enzima cafeína sintasa posee actividad de metilación en 3-N pero no en 1-N. Ensayos realizados en este sentido, muestran que extractos de *C. ptilophylla* pueden metilar 7-metilxantina, pero no las

dimetilxantinas, teobromina, teofilina paraxantina (Ashihara *et al.*, 1998). Un gen homólogo (BTS1) para la cafeína sintasa fue clonado a partir de *Theobroma cacao* por RT-PCR. Posteriormente se construyeron plásmidos de expresión para BTS1 empleando el vector pET23d, a partir de éstos se produjeron BTS1 recombinantes en *E. coli* BL21 (BL21). Cuando un lisado bacteriano se incubó con una variedad de derivados de xantina en la presencia de [metil-¹⁴C] SAM, BTS1 catalizó la 3-N-metilación de 7-metilxantina, lo que sugiere que BTS1 era teobromina sintasa. La acumulación de alcaloides de purina es por lo tanto, dependiente de especificidad por el sustrato de la N-metil-transferasa (Uefuji *et al.*, 2003).

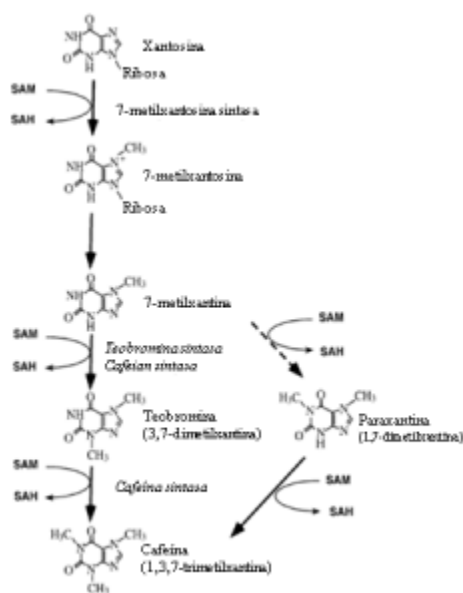


Figura 3. Vía de síntesis de alcaloides en plantas (Ishida *et al.*, 2009).

Polifenoles del cacao

Los compuestos fenólicos o polifenoles pertenecen a un amplio grupo de sustancias químicas, considerados no esenciales para la supervivencia de las plantas (metabolitos secundarios), pero que poseen gran cantidad de actividades, tanto como estructuras químicas (Martínez-Valverde *et al.*, 2000), ya que comprende más de 8 000 compuestos distintos identificados. Constituyen uno de los grupos de productos más numerosos y ampliamente distribuidos en el reino vegetal. Por citar solo algunos ejemplos, los polifenoles incluyen flavonoides, elagitaninos, cumarinas, furanocumarinas, estilbenos, ligninas y lignanos. Un grupo muy importante y sobre el que se ha estudiado ampliamente son los flavonoides, los cuales tienen diversas funciones en la bioquímica, fisiología y ecología de las plantas (Grotewold, 2006). Constituyen los pigmentos en flores y frutos,

contribuyen a la tolerancia al estrés biótico y abiótico, defensa contra radiación UV dañina y tienen papel en la fertilidad del polen (Martens y Mithöfer, 2005; Winkel-Shirley, 2001).

Químicamente, los compuestos fenólicos son sustancias químicas que poseen un anillo aromático, un anillo benceno, con uno o más grupos hidróxidos incluyendo derivados funcionales como los ésteres, metil ésteres, glicósidos, etc. (Martínez-Valverde *et al.*, 2000). La naturaleza de los polifenoles varía desde moléculas simples como los ácidos fenólicos hasta compuestos altamente polimerizados, como la lignina o los taninos (Tabla II). Se pueden encontrar en los tejidos vegetales en forma conjugada con uno o más residuos de azúcar unidos al grupo hidroxilo, aunque en algunos casos se pueden producir uniones directas entre una molécula de azúcar y un C aromático. Por ello la forma más común de encontrarlos en la naturaleza es en forma de glicósidos, siendo solubles en agua y solventes orgánicos (Herrmann y Weaver, 1999). Los azúcares asociados a los polifenoles pueden ser monosacáridos, disacáridos o incluso oligosacáridos. Los compuestos a los que se encuentran unidos con más frecuencia son: glucosa, galactosa, arabinosa, ramnosa, xilosa, y ácidos glucurónico y galacturónico. También pueden encontrarse unidos a ácidos carboxílicos, ácidos orgánicos, aminas, lípidos y a otros compuestos fenólicos (Bravo, 1998).

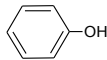

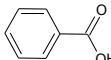
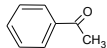
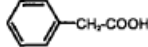
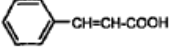
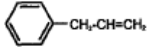
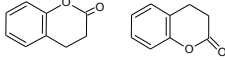
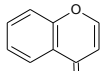
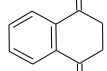
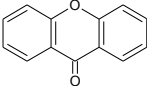
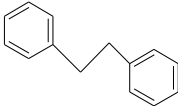
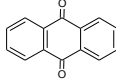
En las últimas tres décadas, se han multiplicado exponencialmente los reportes que atribuyen a los compuestos fenólicos incorporados en la dieta, efectos benéficos para la salud humana y animal, bien como antioxidantes, antitumorales, antiinflamatorios y antiaterosclerosis (Parr y Bolwell, 2000). La estructura polifenólica de flavonoides por ejemplo, resulta ideal para la actividad secuestrante de radicales libres, además han demostrado ser varias veces más eficaces como antioxidantes que por ejemplo las vitaminas E y C (Rice-Evans *et al.*, 1997), empleados convencionalmente como los más potentes antioxidantes naturales.

Por otra parte, los compuestos fenólicos están relacionados con la calidad sensorial de los alimentos de origen vegetal. Su contribución a la pigmentación de los alimentos vegetales está claramente reconocida, por ejemplo las antocianidinas, son responsables de los colores rojo, azul, violeta, naranja y púrpura de muchos frutos clasificados incluso como “superfrutos” por su notable actividad antioxidante (Crozier *et al.*, 2011) y que, en cacao se ha demostrado el contenido de polifenoles y potencial antioxidante es superior (Figura 1). Otro grupo de compuestos fenólicos, particularmente los taninos condensados o proantocianidinas se asocian con la astringencia que presentan muchos frutos, entre ellos el cacao al momento de ser cosechado.

En cacao como en muchas semillas, las células de almacenamiento de polifenoles o células pigmentarias de los cotiledones van de color blanco a violeta oscuro, dependiendo de la cantidad de antocianinas presentes en las células de almacenamiento. La degradación de las antocianinas, que ocurre como antes se explicó durante la fermentación, explica el blanqueo

del color púrpura de los cotiledones. Los polifenoles que se conservan en los cotiledones, son transformados mediante oxidación hacia quinonas, mediado por las Polifenol Oxidasas (PPO), que como antes se dijo reduce el contenido de polifenoles (astringencia) y permite el desarrollo de color marrón.

Tabla II. Esqueleto carbonado básico de los principales compuestos fenólicos (Bravo, 1998)

Clase	Esqueleto básico	Estructura básica
Fenoles simples	C ₆	
Benzoquinonas	C ₆	
Acidos fenólicos	C ₆ -C ₁	
Acetofenonas	C ₆ -C ₂	
Acidos fenilacéticos	C ₆ -C ₂	
Acidos hidroxicinámicos	C ₆ -C ₃	
Fenilpropenos	C ₆ -C ₃	
Cumarinas, isocumarinas	C ₆ -C ₃	
Cromonas	C ₆ -C ₃	
Naftoquinonas	C ₆ -C ₄	
Xantonas	C ₆ -C ₁ -C ₆	
Estilbenos	C ₆ -C ₂ -C ₆	
Antroquinonas	C ₆ -C ₂ -C ₆	
Flavonoides	C ₆ -C ₃ -C ₆	
Lignanós, neolignanós	(C ₆ -C ₃) ₂	
Ligninas	(C ₆ -C ₃) _n	

Los principales compuestos fenólicos presentes en las semillas de cacao se localizan en tres grupos, las catequinas o flavan 3-ol que representan el 37% del total; las antocianinas con cerca del 4% y las proantocianidinas con el 58% (Rusconi y Conti, 2010). Del primer

grupo, la epicatequina es casi la única molécula que compone el grupo (cerca del 98% del total de catequinas), aunque las proantocianidinas sean el principal componente fenólico en almendras de cacao (Wollgast y Anklam, 2000).

Regulación genética y biosíntesis de los compuestos fenólicos

Los polifenoles son productos de la ruta de los fenilpropanoides, sintetizados por la vía de shikimato, al igual que los compuestos fenólicos más simples, tales como ácido gálico y ácido cinámico. La biosíntesis de polifenoles complejos, como los flavonoides está relacionada con el metabolismo primario a través de productos intermediarios derivados de plástidos y mitocondria, siendo exportados en forma individual al citoplasma donde se incorporan en forma separada. Los flavonoides tienen un esqueleto carbonado base C₆-C₃-C₆, donde los dos C₆ (llamados anillos A y B, respectivamente) son de naturaleza fenólica y el C₃ llamado anillo cromano. El anillo aromático B y el anillo cromano se originan a partir del aminoácido fenilalanina (Figura 4), siendo por si mismos producto de la vía de shikimato, mientras que el anillo A partir de tres unidades de malonil-CoA (Fatland *et al.*, 2004; Tsao, 2010). Estas tres unidades de malonil-CoA se añaden a través de sucesivas reacciones de descarboxilación de condensación, que inicia la biosíntesis de flavonoides. La fenilalanina amonio liasa (PAL) es una enzima clave de la vía fenilpropanoide que cataliza la conversión de la fenilalanina en ácido cinámico (C₆-C₃). Posteriormente, el intermediario final 4-cumaril-CoA y tres moléculas de malonil-CoA se condensan para dar la primera estructura flavonoide, chalcona naringenina por acción de la enzima chalcona sintasa (CHS). La chalcona es isomerizada por la chalcona flavona isomerasa (CHI) a una flavanona. La flavanona es un intermediario fundamental, porque es básicamente de donde todas las clases de flavonoides (subgrupos incluidos), se ramifican (Tsao, 2010).

Se han reportado, los mecanismos que controlan la biosíntesis de flavonoides en diferentes tejidos. En semillas en germinación de *Arabidopsis*, se ha reportado que los niveles de ARNm que codifican a las enzimas PAL1, CHS, CHI y dihidroflavanol reductasa (DFR) se mantienen en estado estacionario en plántulas crecidas en luz blanca. En plantas crecidas en la oscuridad solamente la enzima PAL mantiene niveles altos de ARNm. Al inducir la expresión de las enzimas se puede ver que, siguen el mismo orden que las etapas de biosíntesis de los flavonoides. Además se ha comprobado que los genes de biosíntesis de flavonoides son inducidos por la luz UV-B y por la luz azul, sugiriendo la presencia de un receptor específico para ésta última (Kubasek *et al.*, 1992).

Es así que los isoflavonoides principalmente, tienen efecto protector contra la radiación UV, en cuyos niveles elevados se induce la acumulación (Beggs *et al.*, 1987). También se ha sugerido que los isoflavonoides pueden actuar como quimio-atrayentes de organismos fijadores de nitrógeno, ya que se ha comprobado que los niveles bajos de nitrógeno incrementan la concentración de éstas moléculas (Graham, 1991).

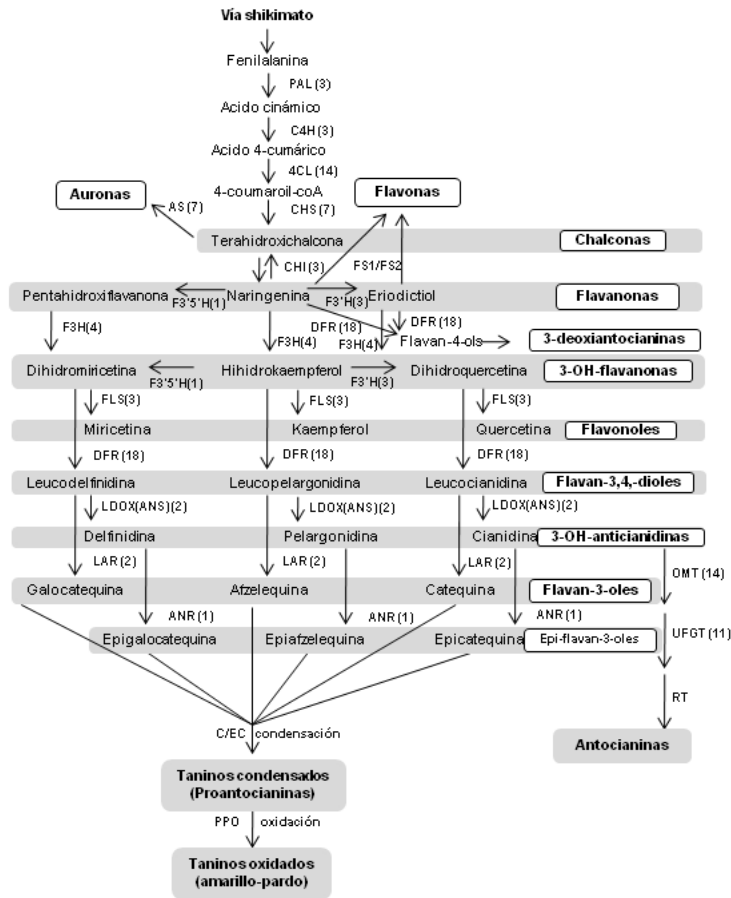


Figura 4. Ruta metabólica de la biosíntesis de flavonoides (Argout *et al.*, 2011). El número de copias de genes ortólogos a cada enzima localizados en el genoma de *Theobroma cacao* se indican entre paréntesis. PAL: fenilalanina amonioliasa, C4H: cinamato 4-hidroxilasa, 4CL: 4-cumarato-CoA ligasa, CHS: chalcona sintasa, AS: auresidina sintasa, CHI: chalcona isomerasa, FS1/FS2: flavona sintasa (número de copias no determinado en cacao), F3H: flavanona 3 hidroxilasa, F3'H: flavonoide 3'-hidroxilasa, F3'5'H: flavonoide 3'5'-hidroxilasa, FLS: flavonol sintasa, DFR: dihidroflavonol 4-reductasa, LDOX (ANS): leucoantocianidina dioxigenasa, LAR: leucoantocinidina reductasa, ANR: antocianidina reductasa, OMT: O-metiltransferasa, UFGT: UDP-glucosa:flavonoide 3-O-glucosiltransferasa, RT: ramnosil transferasa (número de copias no determinado en cacao), PPO: polifenol oxidasa. C: catequina, EC: epicatequina.

Muchos factores pueden inducir la acumulación de compuestos de naturaleza fenólica. Los flavonoles por ejemplo se acumulan en respuesta al ataque de patógenos, en el sitio de la infección suelen encontrarse en concentraciones que resultan tóxicas al patógeno (Dixon y Pavia, 1995); del mismo modo el kaempferol, se acumula en lesiones causadas por herbívoros, siendo tóxico para éstos (Hahlbrock y Scheel, 1989) y también se acumula durante el desarrollo del polen sugiriendo efecto aseptizante (Dixon y Pavia, 1995). Las flavonas y antocianinas se acumulan en respuesta a exposición intensa a la luz, actuando

como fotoprotectores, reduciendo la cantidad de luz que llega a los sistemas fotosintéticos (Beggs *et al.*, 1987).

Mediante análisis comparativo, se han identificado 96 genes ortólogos a *Arabidopsis thaliana* que están implicados en la ruta biosintética de los flavonoides, los cuales pueden revisarse a detalle en (Argout *et al.*, 2011). Son 60 genes más que en *Arabidopsis*, y de los cuales se ha evaluado en tabaco son genes funcionales. Algo importante de notar es el elevado número de genes ortólogos a dihidroflavon-4-reductasa (DFR), mientras que para *Arabidopsis* solo se encuentra uno, en cacao se encontraron 18. Los DFR catalizan la reacción que produce los flavan-3, 4-dioles, quienes son los precursores inmediatos de los flavonoides catequina y epicatequina; lo que explicaría la acumulación de estos compuestos fenólicos de entre 8-12% de peso seco de las semillas desgrasadas y que coloca como ya antes se apuntó al cacao como una de las fuentes más ricas conocidas de este fitonutriente (Argout *et al.*, 2011; Tomás-Barberán *et al.*, 2007).

Aroma y sabor del cacao, efecto del procesamiento poscosecha

Además de las características reológicas de la pasta de cacao (obtenida de la molienda de las almendras secas y tostadas), o de las propiedades físicas de la manteca de cacao (extraída por prensado, disolventes u otro método), el aroma y sabor de las almendras y de los derivados de éstas (particularmente el chocolate), son las principales razones de la popularidad del cacao. A su vez, el aroma y sabor está en función de la presencia y/o ausencia de diversos compuestos (Frauendorfer y Schieberle, 2008). Algunos autores (Luna *et al.*, 2002; Rodríguez-Campos *et al.*, 2011) han realizado esfuerzos para relacionar la presencia de algunos compuestos tanto con la calidad sensorial como con la no calidad o con la ausencia de algún proceso importante para exaltar el aroma y/o sabor. En este sentido, se ha indicado que los ácidos isobutírico, isovalérico y propiónico pueden fungir tanto como indicadores de calidad deficiente o aromas no deseables en un cacao (rancio, astringente, acidez extrema, aroma a queso o a jamón) como indicadores de tiempos prolongados en los procesos de fermentación (Rodríguez-Campos *et al.*, 2011). Así mismo, algunos alcoholes producidos durante la fermentación (comentadas más adelante) también son empleados como indicadores de calidad (Oberparleiter y Ziegleder, 1997), o los ácidos orgánicos (principalmente acético) son indicadores de la acidificación de las almendras y consecuentemente de su calidad. En este sentido se pueden revisar reportes más completos para ver la contribución de compuestos volátiles y no volátiles y su implicación en la calidad sensorial.

Otro ejemplo de compuestos implicados en el desarrollo principalmente de sabor final y en menor medida en el aroma del cacao, lo constituyen los polifenoles y los alcaloides. Los polifenoles están implicados tanto en la astringencia de las almendras como en la intensidad

del aroma a cacao. Por su parte, los alcaloides, específicamente la teobromina, la cafeína y la teofilina, son los principales responsables del sabor amargo de las semillas. La concentración de estos compuestos, está influenciada por el procesamiento de las almendras (Luna *et al.*, 2002; Schwan y Wheals, 2004).

De lo anterior, es importante hacer notar que, el sabor y aroma de las almendras producidas por el árbol del cacaotero, o correctamente expresado, sus precursores son el resultado de varios factores. El más determinante es el origen genético, es así que los árboles de la variedad Criollo ofrecen características superiores respecto a las variedades Forastero y Trinitario (Camú *et al.*, 2008). Otro factor determinante es el ambiente donde se desarrollan los árboles, dado por las condiciones edafo-climáticas y la vegetación asociada a los individuos, que se ha comprobado influyen en los aromas y sabores que habrán de desarrollar las almendras. Los demás factores, atribuibles al manejo pre y poscosecha son las prácticas culturales de manejo de la plantación, que en las plantaciones de América reviste de importancia por las plagas y enfermedades que aquejan a este cultivo (moniliasis, mancha negra, hormigas, ardillas), fertilización (sintética u orgánica), momento de cosecha, tiempo transcurrido entre la cosecha y el procesamiento de la mazorca, así como las atribuibles (igualmente importantes) al procesamiento de las mismas (Aikpodpotion, 2010). La fermentación, secado y tostado de las almendras, se han reportado como los principales procesos poscosecha implicados en las características finales de sabor y aroma que habrán de poseer las almendras y consecuentemente los productos de ellas obtenidos como cocoa, chocolate, manteca de cacao (Schwan y Wheals, 2004; Luna *et al.*, 2002).

Por la relevancia que tiene el procesamiento primario de los frutos y las semillas, y dado que existe un consenso en la comunidad científica acerca de que los principales cambios que ocurren a las almendras durante la fermentación, secado y tostado conlleva implicaciones que habrán de reflejarse en el sabor y aroma de las almendras, toda vez que se dice que un cacao que no es procesado de este modo, no puede desarrollar las características sensoriales deseables de un buen cacao, tendiendo a permanecer astringente y con características heterogéneas (Hansen, del Olmo, & Burri, 1998), resulta de interés ver de manera particular como cada proceso poscosecha influye en la producción y/o transformación de biomoléculas.

La fermentación

Las frutos de cacao (mazorcas) son bayas con un espeso y duro exocarpo (cáscara) de entre 7 y 20 mm de grosor, la cual sirve de protección a las semillas dispuestas a lo largo de un eje embrionario, y envueltas por un fluido mucilaginoso, el cual dada su composición química actúa como medio de cultivo para que de manera espontánea se desarrollen una sucesión de microorganismos, que dan origen a la fermentación del mucílago y transformación de las todavía semillas por éste envuelto. Los principales componentes del mucílago o pulpa son agua (82-87%) azúcares (10-15%), pentosanas (2-3%), ácido cítrico

(1-3%), cuyo contenido determina el pH del mucílago (~3,6) y pectina (1-1,5%) (Roelofsen, 1958). Ésta última responsable del aspecto mucilaginoso, aunque contribuyendo de manera importante se han reportado citratos, hemicelulosa y lignina (Pettipher, 1986). También se han reportado como componentes de la pulpa proteínas, aminoácidos, vitaminas, (principalmente vitamina C) y minerales, los cuales complementan los principales requerimientos de levaduras, bacterias ácido lácticas, ácido acéticas y otras que se han reportado participan de la fermentación (Schwan y Wheals, 2004).

La fermentación es un proceso que se lleva a cabo de manera espontánea, en un periodo generalmente de cuatro a seis días, en cajas, cestas, o en plataformas (Baker *et al.*, 1994; Tomlins *et al.*, 1993). Pese a que hay algunos informes del desarrollo e inoculación de cultivos iniciadores (Schwan, 1998; Dzogbefia *et al.*, 1999; Lefeber *et al.*, 2010), el proceso se sigue dando de manera espontánea, al menos entre los pequeños productores. La primera etapa de la fermentación conlleva un aumento moderado de la temperatura a valores de entre 35 y 45 °C (Camú *et al.*, 2008) que provoca la transformación de la pulpa mucilaginoso alrededor de los granos, la muerte del embrión y la aparición de metabolitos sensorialmente deseables (Gotsch, 1997).

Las profundas transformaciones que sufren tanto el mucílago como los granos, son el resultado de la intensa proliferación y actividad microbiana que provocan cambios bioquímicos en el interior de las almendras y que, contribuyen a la reducción de amargor y astringencia; así mismo a la aparición de diferentes compuestos volátiles considerados como indicativos de la calidad del grano (Rodríguez-Campos *et al.*, 2011). A diferencia de otras materias primas fermentables, las enzimas endógenas de las semillas de cacao, al ser activadas juegan un papel fundamental en el desarrollo de moléculas asociadas con el sabor o precursoras del aroma. Estas enzimas, aunadas a la de los microorganismos endófitos (aunque se reportan que son estériles) y a las de los que “contaminan” la pulpa, tienen actividad degradadora importante (De Brito *et al.*, 2000).

Durante la primera fase de la fermentación, el metabolismo de levaduras es favorecido por la acidez del medio (dado por el contenido de ácido cítrico), la riqueza en carbohidratos fermentables y el bajo contenido de oxígeno de la masa; estos microorganismos provocan la despectinización (depolimerización de la pectina) del medio, dando lugar a la aparición de alcoholes simples (principalmente etanol) que junto con el ácido cítrico, al permear a las semillas provocan la muerte del embrión (Quesnel, 1965).

Los metabolitos antes producidos, aunados a azúcares simples (sacarosa, glucosa y fructosa) optimizan el medio para la aparición de bacterias (tolerantes al etanol) ácido lácticas (BAL) microaerofílicas y ácido acéticas (BAA) aerófilas, que promueven la degradación de azúcares y la oxidación (exotérmica) de etanol hasta ácido láctico y acético, respectivamente, además de manitol (Lefeber *et al.*, 2010). Lo anterior conlleva a un ligero aumento en la acidez de la pulpa, la cual, al permear a través de la testa de los

cotiledones disminuye el pH interno (de 6.5 a 4.8), lo que como antes se citó, activa enzimas endógenas implicadas en la transformación de diversas moléculas. Las principales enzimas implicadas son invertasas, glucosidasas, proteasas y polifenol oxidasas (De Brito *et al.*, 2000; Misnawi *et al.*, 2003).

Un aspecto adicional importante en el proceso de fermentación y consecuentemente en el desarrollo de sabor y aroma, son los manejos técnicos o tecnológicos que habrán de tenerse para lograr la correcta sucesión de microorganismos. El tamaño del fermento (pila, caja, etc.) se ha propuesto como un factor importante, ya que puede regular en gran medida la temperatura de la masa y junto con el mezclado y volteado constante (diariamente o hasta dos veces al día) durante los días que dura el proceso resultan medulares para evitar el desarrollo de hongos no deseados (Papalexandratou *et al.*, 2011), o la desaparición abrupta de mesófilos deseables en la aparición de moléculas de interés por la falta de aireación o el aumento de la temperatura (Rohan, 1963). Una revisión más detallada del proceso de fermentación puede ser consultada en Schnaw y Wheals (2004).

En suma, los principales cambios que se asocian con la fermentación pueden ser agrupados en función de la ruta metabólica que habrá de seguir, para estar presentes en las almendras fermentadas:

- 1). Degradación o transformación: proteínas de almacenamiento (mayoritariamente globulinas) por acción de endo – aspartatoproteasas y carboxipeptidasas, sacarosa (por acción de endo-invertasas), glicósidos (antocianinas), epicatequina y antocianidinas (que son oxidadas por acción de las polifenol oxidasas y transformados en quinonas), quinonas que forman complejos con otros polifenoles, proteínas o péptidos.
- 2). Formación, básicamente a expensas de otros metabolitos degradados o acomplexados: glucosa y fructosa (a partir de sacarosa), aminoácidos libres (tanto hidrofílicos como hidrófobos), oligopéptidos, alcoholes (etanol), taninos condensados (a partir de polifenoles), antocianidinas (degradación de antocianinas), galactosa, arabinosa, quinonas, ácidos orgánicos (principalmente acético y láctico).
- 3). Disminución: principalmente por difusión de las células de almacenamiento promovida durante el tiempo de fermentación y que se ha reportado reduce hasta el 30% del total de alcaloides y hasta 20% de polifenoles totales (taninos como las proantocianidinas y flavonoides como la flavan 3-ol).

Estos cambios, generan un amplio número de moléculas agrupadas en familias, de las que sobresalen los esteroides, alcoholes y ácidos (Portillo *et al.*, 2009) que habrán de modificarse o aumentar su contenido con los tratamientos posteriores (Tabla III) pero que desde esta etapa disminuyen la astringencia (disminución de polifenoles totales) y la tonalidad púrpura de las semillas (transformación y degradación de antocianinas), amargor (disminución de alcaloides) de los granos, desarrollo de coloración marrón (presencia de quinonas). Incluso

algunos parámetros que se emplean a nivel industrial para evaluar el grado de fermentación son los contenidos de antocianinas y el color marrón.

El secado

Una vez completada la fermentación, la humedad de los granos que está por el 60% (Rohan, 1964), debe reducirse hasta un nivel que garantice, por un lado que las semillas conservarán su testa, es decir no tan secos que se vuelva quebradiza, pero tampoco con humedad suficiente para permitir la proliferación de hongos que se desarrollan durante el almacenamiento y que pueden generar sabores desagradables o más riesgadamente, producir toxinas. Valores de humedad entre 6 y 8% son altamente deseables (Harrington, 2011). Este proceso se lleva a cabo casi exclusivamente y cuando las condiciones lo permiten con secado solar. Cuando las condiciones ambientales son adversas, es decir lluvias por periodos prolongados, se recurre a secadores mecánicos, los más empleados son del tipo Samoa o estufas de aire forzado (Amoye, 2006). El uso de uno u otro método se ha reportado, puede ejercer efecto sobre las características sensoriales del producto final, en este sentido se ha reportado (Rodríguez-Campos *et al.*, 2011), que los contenidos de alcoholes, ésteres y pirazinas se ven aumentados durante el proceso de secado al sol, mientras que los de ácidos, aldehídos y cetonas disminuidos, lo cual coloca el secado al sol como el mejor método para obtener el máximo sabor (Jinap *et al.*, 1994). Además de los contenidos, las familias de compuestos también se ven modificadas tal como se aprecia en la Tabla III. Es notorio que los grupos que se sugiere disminuyen en concentración durante el secado, no muestran aumento sustancial en el número de familias (ácidos).

Tabla III. Distribución de las familias y cantidad de compuestos volátiles identificados en cacao Criollo (Portillo *et al.*, 2009)

Familias	Cacao fresco	Cacao seco	Cacao tostado
Aldehídos	8	12	11
Alcoholes	15	13	13
Ácidos	14	14	12
Cetonas	9	13	13
Esteres	22	27	26
Hidrocarburos	3	7	3
Pirazinas	3	7	15
Misceláneos	3	5	3
Pirroles	1	4	4
Furanos	6	7	7
Azufres	1	1	2
Terpenos	3	4	3
Fenoles	4	6	6
Oxazoles	0	1	1
Total	92	121	119

Estos cambios, pueden ser explicados debido a la continuidad de la fase oxidativa que se inició en la fermentación, por lo que el secado juega un papel importante en la disminución de la astringencia, amargor y acidez del grano, así como en el desarrollo del color marrón a partir de los compuestos fenólicos mediado como antes se dijo por polifenoloxidasas (Luna *et al.*, 2002; Jinap *et al.*, 1994).

El Tostado

La última etapa del procesamiento primario de las almendras de cacao es el tostado, que cumple varias funciones físico - químico - sensoriales. Esta etapa ayuda a reblandecer y retirar la testa de las almendras, las vuelve más frágiles, como preparación a la molienda o prensado, pero sobretodo, desarrolla a partir de los precursores generados en la fermentación y secado, aún más el sabor del cacao o característico a chocolate.

Este proceso continúa promoviendo la disminución de la acidez, al reducir las concentraciones de ácidos volátiles como el ácido acético (Ramli *et al.*, 2006); aunque no la de otros ácidos no volátiles como el oxálico, cítrico, tartárico, succínico y láctico (Jinap *et al.*, 1998). Sin embargo los principales cambios derivan de las reacciones de Maillard y Strecker, donde participan los aminoácidos libres, péptidos y azúcares reductores (Ziegleder, 1991). La reacción de Maillard es una reacción de oscurecimiento no enzimática que requiere un aminoácido y un azúcar reductor, tal como glucosa o fructosa. Las condiciones ideales para la reacción de Maillard que se produzca son, temperatura elevada y bajo contenido de humedad, condiciones que se dan durante el tostado (Lindsay, 1996). La degradación de Strecker se produce cuando un carbonilo derivado de la reacción de Maillard, reacciona con otros aminoácidos libres del producto. Esto provoca la degradación de los aminoácidos a aldehídos, amoníaco y dióxido de carbono. Los aldehídos que se producen contribuyen al aroma. La degradación de Strecker de cada aminoácido específico, produce un aldehído único con un aroma único (Lindsay, 1996). Se ha reportado que aminoácidos hidrofóbicos como la leucina, alanina, fenilalanina y tirosina, liberados durante la fermentación contribuyen de manera importante al unirse a fructosa y glucosa (reductores) que se producen por la hidrólisis de sacarosa (Voight *et al.*, 1993; Voight *et al.*, 1994; Granvogl *et al.*, 2006).

Conclusión

El aroma y sabor característico de las almendras de cacao, son el producto de la composición química de las mismas y de los procesos que le sobrevienen a la cosecha de los frutos. La composición química a su vez, es la resultante del origen genético principalmente, de los factores ambientales y las prácticas de cultivo. Así las moléculas implicadas en el sabor y aroma son las grasas, proteínas, azúcares, polifenoles y alcaloides. Los polifenoles están asociados con la astringencia y variaciones en la coloración, por lo

que su biosíntesis está regulada por la expresión de enzimas implicadas en la ruta de biosíntesis (vía shikimato), además de responder a inductores ambientales, como la radiación UV-B, lesiones y ataque de patógenos. Los alcaloides, que determinan el gusto amargo de las almendras y otros tejidos donde se acumulan, disminuyen su concentración en las semillas con la maduración del fruto y con el procesamiento poscosecha, y su síntesis es conservada por la síntesis de *novo*, pero difiere incluso entre especies por la vía de salvamento de purinas. Debido a que la enzima cafeína sintasa implicada en la formación de teobromina no posee actividad metiladora en el N1, la teobromina resulta ser el principal alcaloide de las almendras de cacao.

Agradecimientos

El primer autor agradece al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca económica otorgada, así como al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México por la instrucción recibida en los estudios de posgrado.

Referencias

- Aikpokpodion P (2010) Variation in agromorphological characteristics of cacao, *Theobroma cacao* L., in farmers' fields in Nigeria. *New Zeal. J. Crop Hort.* 38: 157-170.
- Álvarez C, Pérez E, Lares M (2007) Caracterización física y química de almendras de cacao fermentadas, secas y tostadas cultivadas en la región de Cuyagua, estado Aragua. *Agron. Trop.* 57: 249-256.
- Amoye S (2006) Cocoa sourcing, world economics and supply. *The Manufact. Confect.* 86: 81e-85e.
- Argout X, Salse J, Aury J, Guiltinan M, Droc G, et al. (2011) The genome of *Theobroma cacao*. *Nat. Genet.* 43: 101-109.
- Ashihara H, Kato M, Crozier A (2011) *Distribution, biosynthesis and catabolism of methylxanthines in plants*. In: B.B. Fredholm (ed.) *Methylxanthines, Handbook of Experimental Pharmacology* 11-26 pp.
- Ashihara H, Sano H, Crozier A (2008) Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochem.* 69: 841-856.
- Ashihara H, Kato M, Ye C (1998) Biosynthesis and metabolism of purine alkaloids in leaves of cocoa tea (*Camellia ptilophylla*). *J. Plant Res.* 111: 599-604.
- Ashihara H, Kubota H (1987) Biosynthesis of purine alkaloids in *Camellia* plants. *Plant Cell Physiol.* 28: 535-539.
- Baker D, Tomlins K, Gray C (1994) Survey of Ghanaian cocoa farmer fermentation practices and their influence on cocoa flavour. *Food Chem.* 51: 425-431.
- Beggs C, Kuhn K, Böcker R, Wellman E (1987) Phytochrome-induced flavonoid biosynthesis in mustard (*Sinapis alba* L.) cotyledons. Enzymic control and differential regulation of anthocyanic and quercetin formation. *Planta* 172: 121-126.

- Bernays A, Oppenheim S, Chapman R, Kwon H, Gould F (2000) Taste sensitivity of insect herbivores to deterrents is greater in specialists than in generalists: a behavioral test of the hypothesis with two closely related caterpillars. *J. Chem. Ecol.* 26: 547-563.
- Bravo L (1998) Polyphenol: chemistry, dietary sources, metabolism and nutritional significance. *Nutr. Rev.* 56: 317-333.
- Camú N, De Winter T, Addo S, Takrama J, Bernaert H, De Vuyst L (2008) Fermentation of cocoa beans: influence of microbial activities and polyphenol concentrations on the flavour of chocolate. *J. Sci. Food Agr.* 88: 2288-2297.
- Crozier S, Preston A, Hurst J, Payne M, Mann J, Hainly L, Miller D (2011) Cacao seeds are a “Super Fruit”: A comparative analysis of various fruit powders and products. *Chem. Central J.* 5: 5.
- Dantas L, Guerra M (2010) Chromatin differentiation between *Theobroma cacao* L. and *T. grandiflorum* Schum. *Genet. Mol. Biol.* 33: 94-98.
- De Brito E, Pezoa-García N, Gallão M, Cortelazzo A, Fevereiro P, Braga M (2000) Structural and chemical changes in cocoa (*Theobroma cacao* L.) during fermentation, drying and roasting. *J. Sci. Food Agric.* 81: 281-288.
- Dillinger T, Barriga P, Escárcega S, Jiménez M, Lowe D, Grivetti L (2000) Food of the Gods: Cure for humanity? A cultural history of the medicinal and ritual use of chocolate. *J. Nutr.* 130: 2057S-2072S.
- Dixon R, Pavia N (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell.* 7: 1085-1097.
- Dzogbefia V, Buamah R, Oldham J (1999) The controlled fermentation of cocoa (*Theobroma cacao* L.) using yeasts: enzymatic process and associated physico-chemical changes in cocoa sweatings. *Food Biotechnol.* 13: 1-12.
- Evans W (2000) *Trease and Evans - Pharmacognosy*. (15th ed). Edinburgh. Ed. Saunders.
- FAO (2014) Organización de las Naciones Unidas para la Agricultura y la Alimentación. *FAOSTAT*. Estadísticas de Producción 2012.
- Fatland B, Ke J, Anderson M, Mentzen W, Cui L, Allred C, Johnston J, Nikolau B, Wurtele E (2004) Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-Coenzyme A in *Arabidopsis*. *Plant Physiol.* 130: 740-756.
- Frauendorfer F, Schieberle P (2008) Changes in key aroma compounds of criollo cocoa beans during roasting. *J. Agric. Food Chem.* 56: 10244-10251.
- Gotsch N (1997). Cocoa biotechnology: status, constraints and future prospects. *Biotechnol. Adv.* 15: 333-352.
- Graham T (1991) Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates. *Plant Physiol.* 95: 594-603.
- Granvogl M, Bagan S, Schieberle P (2006) Formation of amines and aldehydes from parent amino acids during thermal processing of cocoa and model systems: new insights into pathways of the Strecker reaction. *J. Agric. Food Chem.* 54: 1730-1739.
- Grotewold E (2006) *The science of flavonoids*. Ohio, USA. Ed. Springer. 274 p.
- Hammerstone JF, Romanczyk LJ, Aitken WM (1994) Purine alkaloid distribution within *Herrania* and *Theobroma*. *Phytochemistry* 35: 1237-1240.

- Hahlbrock K, Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. *Annu. Rev. Plant Phys.* 40: 347-369.
- Hansen C, del Olmo M, Burri C (1998) Enzyme activities in cocoa beans during fermentation. *J. Sci. Food Agric.* 77: 273-281.
- Harrington W (2011) The effects of roasting time and temperature on the antioxidant capacity of cocoa beans from Dominican Republic, Ecuador, Haiti, Indonesia, and Ivory Coast. Tesis de Maestría, Universidad de Tennessee. E.U.A.
- Herrmann K, Weaver L (1999) The shikimate pathway. Annual Review of Plant Physiology. *Plant Mol. Biol.* 50: 473-503.
- Ishida M, Kitao N, Mizuno K, Tanikawa N, Kato M (2009) Occurrence of theobromine synthase in purine alkaloid-free species of Camellia plants. *Planta* 229: 559-568.
- Jinap S, Siti M, Norsiaty M (1994) Formation of methyl pyrazine during cocoa bean fermentation. *Pertanika J. Trop. Agric. Sci.* 17: 27-32.
- Jinap S, Wan Rosli W, Russly A, Nordin L (1998) Effect of roasting time and temperature on volatile component profiles during nib roasting of cocoa beans (*Theobroma cacao*). *J. Sci. Food Agric.* 77: 441-448.
- Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H (2000) Caffeine synthase gene from tea leaves. *Nature* 406: 956-957.
- Kato M, Mizuno K, Fujimura T, Iwama M, Irie M, Crozier A, Ashihara H (1999) Purification and characterization of caffeine synthase from tea leaves. *Plant Physiol.* 120: 579-586.
- Koyama Y, Tomoda Y, Kato M, Ashihara H (2003) Metabolism of purine bases, nucleosides and alkaloids in theobromine-forming *Theobroma cacao* leaves. *Plant Physiol. Biochem.* 41: 977-984.
- Kubasek W, Shirley B, McKillop A, Goodman H, Briggs W, Ausubel F (1992) Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings. *Plant Cell.* 4: 1229-1236.
- Lefeber T, Janssens M, Camú N, De Vuyst L (2010) Kinetic analysis of strains of lactic acid bacteria and acetic acid bacteria in cocoa pulp simulation media toward development of a starter culture for cocoa bean fermentation. *Appl. Environ. Microb.* 76: 7708-7716.
- Liendo R, Padilla F, Quintana A (1997) Characterization of cocoa butter extracted from Criollo cultivars of *Theobroma cacao* L. *Food Res. Int.* 30: 727-731.
- Lindsay R. (1996) *Food Additives*. In: O. Fennema (Ed). Food Chemistry (3rd ed). New York. Marcel Dekker 768-823 pp.
- Luna F, Cruzillat D, Cirou L, Bucheli P (2002) Chemical composition and flavor of Ecuadorian cocoa liquor. *J. Agric. Food Chem.* 50: 3527-3532.
- Martens S, Mithöfer A (2005) Flavones and flavone synthases. *Phytochemistry* 66: 2399-2407.
- Martínez-Valverde I, Periago M, Ros G (2000) Significado nutricional de los compuestos fenólicos de la dieta. *Arch. Latinoam. Nutr.* 50: 5-18.
- Mazzafera P, Wingsle G, Olsson O, Sandberg G (1994) S-adenosyl-L-methionine:theobromine 1-Nmethyltransferase, an enzyme catalyzing the synthesis of caffeine in coffee. *Phytochemistry* 37: 1577-1584.
- Misnawi S, Jamilah B, Nazamid S (2003) Effects of incubation and polyphenol oxidase enrichment on colour, fermentation index, procyanidins and astringency of unfermented and partly fermented cocoa beans. *Int. J. Food Sci. Technol.* 38: 285-295.

- Mizuno K, Okuda A, Kato M, Yoneyama N, Tanaka H, Ashihara H, Fujimura T (2003) Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.) *FEBS Lett.* 534: 75-81.
- Mizuno K, Tanaka H, Kato M, Ashihara H, Fujimura T (2001) *cDNA cloning of caffeine (theobromine) synthase from coffee (Coffea arabica L.)*. In: Proceedings of the International Scientific Colloquium on Coffee. ASIC, Paris, 19, 815-818.
- Oberparleiter S, Ziegleder G (1997) Amyl alcohols as compounds indicative of raw cocoa bean quality. *Eur. Food Res. Technol.* 204: 156-160.
- 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 Microbiol.* 28: 1326-1338.
- Parr A, Bolwell G (2000) Phenols in the plant and in man. The potential for possible nutritional enhancement of the diet by modifying the phenols content or profile. *J. Sci. Food Agric.* 80: 985-1012.
- Pettipher G (1986) Analysis of cocoa pulp and the formulation of a standardised artificial cocoa pulp medium. *J. Sci. Food Agric.* 37: 297-309.
- Ploetz R (2007) Cacao diseases: important threats to chocolate production worldwide. *Phytopathology* 97: 1634-1639.
- Portillo E, Labarca M, Grazziani L, Cros E, Assemat S, Davrieux F, Boulanger R, Marcano M (2009) Formación del aroma del cacao Criollo (*Theobroma cacao* L.) en función del tratamiento poscosecha en Venezuela. *Rev. UDO Agric.* 9: 458-468.
- Powis T, Hurst W, Rodríguez M, Ortiz P, Blake M, Cheetham D, Coe M, Hodgson J (2008) The origins of cacao use in Mesoamerica. *Mexicon* 30: 35-38.
- Powis T, Cyphers A, Gaikwad N, Grivetti L, Cheong K (2011) Cacao use and the San Lorenzo Olmec. *PNAS* 108: 8595-8600.
- Quesnel V (1965) Agents inducing the death of the cacao seeds during fermentation. *J. Sci. Food Agric.* 16: 441-447.
- Ramiro E, Franch A, Castellote C, Andres-Lacueva C (2005) Effect of *Theobroma cacao* flavonoids on immune activation of a lymphoid cell line. *Brit. J. Nutr.* 93: 859-866.
- Ramli N, Hassan O, Said M, Samsudin W, Idris N (2006). Influence of roasting condition on volatile flavour of roasted Malaysian cocoa beans. *J. Food Process. Pres.* 30: 280-298.
- Rice-Evans C, Miller N, Paganga G (1997). Antioxidant properties of phenolic compounds. *Trends Plant Sci.* 2: 2152-2159.
- Rodriguez-Campos J, Escalona-Buendía H, Orozco-Avila I, Lugo-Cervantes E, Jaramillo-Flores M (2011) Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Res. Inter.* 44: 250-258.
- Roelofsen PA (1958) Fermentation, drying, and storage of cocoa beans. *Adv. Food Res.* 8; 225-296.
- Rohan T (1963). The precursors of chocolate aroma: application of gas chromatography in following formation during fermentation of cocoa beans. *J Food Sci.* 32: 402-404.
- Rohan T (1964). The precursors of chocolate aroma: comparative study of fermented and unfermented cocoa beans. *J. Food Sci.* 29: 456-459.

- Rusconi M, Conti A (2010) *Theobroma cacao* L., the food of the gods: A scientific approach beyond myths and claims. *Pharmacol. Res.* 61: 5-13.
- Schwan R (1998) Cocoa fermentations conducted with a defined microbial cocktail inoculum. *Appl. Environ. Microb.* 64: 1477-1483.
- Schwan R, Wheals A (2004) The microbiology of cocoa fermentation and its role in chocolate quality. *Crit. Rev. Food Sci.* 44: 205-221.
- Sotelo A, Álvarez G (1991) Chemical composition of wild *Theobroma* species and their comparison to the cacao bean. *J. Agric. Food Chem.* 39: 1940-1943.
- Smyth D (1992) Effect of methylxanthine treatment on rice seedling growth. *J. Plant Growth Regul.* 11: 125-128.
- Stark T, Bareuther S, Hofmann T (2006) Molecular definition of the taste of roasted cocoa nibs (*Theobroma cacao*) by means of quantitative studies and sensory experiments. *J. Agric. Food Chem.* 54: 5530-5539.
- Stasolla C, Katahira R, Thorpe T, Ashihara H (2003) Purine and pyrimidine nucleotide metabolism in higher plants. *J. Plant Physiol.* 160: 1271-1295.
- Sugiura M, Takeda Y (2000) Nucleic acids. In B.B. Buchanan, W. Gruissen, R.L. Jones (Eds.), *Biochemistry and Molecular Biology of Plants* (260-310 pp). Rockville: American Society of Plant Physiologists.
- Tomas-Barberán F, Cienfuegos-Jovellanos E, Marín A, Muguera B, Gil-Izquierdo A, Cerdá B, Zafrilla P, Morillas J, Mulero J, Ibarra A, Pasamar M, Ramón D, Espín J (2007) A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. *J. Agric. Food Chem.* 55: 3926-3935.
- Tomlins K, Baker D, Daplyn P, Adomako D (1993) Effect of fermentation and drying practices on the chemical and physical profiles of Ghana cocoa. *Food Chem.* 46: 257-263.
- Tsao R (2010) Chemistry and biochemistry of dietary polyphenols. *Nutrients* 2: 1231-1246.
- Uefuji H, Ogita S, Yamaguchi Y, Koizumi N, Sano H (2003) Molecular cloning and functional characterization of three distinct *N*-methyltransferase involved in the caffeine biosynthetic pathway in coffee plants. *Plant Physiol.* 132: 372-380.
- Voigt J, Biehl B, Kamaruddin S (1993) The major seed proteins of *Theobroma cacao* L. *Food Chem.* 47: 145-151.
- Voigt J, Biehl B, Heinrich H, Kamaruddin S, Gaim-Marsoner G, Hugi A (1994) *In vitro* formation of cocoa-specific aroma precursors: aroma related peptides generated from cocoa seed protein by co-operation of an aspartic endoprotease and a carboxypeptidase. *Food Chem.* 49: 173-180.
- Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology. *Plant Physiol.* 126: 485-493.
- 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 Res. Int.* 33: 423-447.
- Ziegler G (1991) Composition of flavor extracts of raw and roasted beans. *Eur. Food Res. Technol.* 192: 521-525.

CAPÍTULO 5

Agrupamiento de muestras de semillas de cacao empleando métodos de estadística multivariada

Con la finalidad de caracterizar las semillas de cacao, se midieron y determinaron características físicas y químicas. Se midió el largo, ancho, circunferencia y peso de las almendras. Del mismo modo se determinó los contenidos de humedad, cenizas, grasa (así como su composición) y proteína. Adicionalmente y por su contribución con estos importantes compuestos, se cuantificó el contenido de compuestos fenólicos y la actividad antioxidante. Se aborda la discusión sobre la composición de las almendras y el bajo efecto que ejerce el ambiente sobre la composición química. Dado ese bajo efecto, se consideró otro criterio de agrupamiento de muestras; con la finalidad de reducir el número de muestras que los jueces evaluadores (en etapa posterior) pueden analizar sin aumentar la probabilidad del error experimental. Así, se incluyó una serie de análisis estadísticos multivariados (Análisis de Componentes Principales, Análisis de k-medias y Análisis Discriminante) que permitieron empleando una muestra de *Theobroma bicolor* (Patate) como control, por un lado encontrar cuales de las características fisicoquímicas explicaron el nuevo agrupamiento (aportan más variación) y por otro lado, permitieron formar 7 grupos de muestras que comporten (dentro de grupos) características similares.

Los resultados de este enfoque de análisis fueron publicados como artículo de investigación en la revista *European Food Research and Technology* (Springer, ISSN versión impresa 1438-2377, versión electrónica 1438-2385). La cita es: Alfredo Vázquez-Ovando, Francisco Molina-Freaner, Juan Nuñez-Farfán, David Betancur-Ancona, Miguel Salvador-Figueroa (2015) Classification of cacao beans (*Theobroma cacao* L.) of Southern Mexico based on chemometric analysis with multivariate approach. *European Food Research and Technology* 240(6): 1117-1128.

Classification of cacao beans (*Theobroma cacao* L.) of southern Mexico based on chemometric analysis with multivariate approach

Alfredo Vázquez-Ovando · Francisco Molina-Freaner ·
Juan Nuñez-Farfán · David Betancur-Ancona ·
Miguel Salvador-Figueroa

Received: 28 August 2014 / Revised: 18 December 2014 / Accepted: 11 January 2015 / Published online: 20 February 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract The aim of this study was to group samples of cacao collected in Southern of Mexico. For this, several physical bean variables (weight, length, width and bean circumference), chemical (moisture, ash, fat, protein, fatty acids) in addition to polyphenol content and antioxidant capacity (ABTS method) were measured. Forty-five cacao samples derived from plants with Criollo phenotype, and a control sample of *T. bicolor* was included. All variables were analyzed by multivariate analysis (principal component analysis—PCA, k-means and discriminant analysis—DA). The four physical bean traits were not useful when incorporated into the analysis altogether with the chemical variables and led to wrong grouping, and the PCA did not separate *T. bicolor* of the cocoa samples, so they were

discarded for further analysis. Among chemical variables, those that contributed most to the variance and explained the total variance (PCA) were palmitoleic, palmitic, oleic and stearic fatty acids, as well as moisture content and polyphenol content. All samples were classified into seven homogeneous groups. The geographical origin of the samples did not correlate with chemometric grouping, which shows little influence of microclimates, the possible dispersion of highly related materials or individuals adapted to the microenvironmental conditions of the region. Multivariate analysis allowed the grouping of individuals of cacao based on six chemical characteristics of the beans.

Keywords Fatty acids · Multivariate analysis · Polyphenol content · Criollo cacao · Moisture · Geographical origin

A. Vázquez-Ovando
Posgrado en Ciencias Biológicas, Universidad Nacional
Autónoma de México (UNAM), México, DF, Mexico
e-mail: jose.vazquez@unach.mx

A. Vázquez-Ovando · M. Salvador-Figueroa (✉)
Centro de Biociencias, Universidad Autónoma de Chiapas,
Boulevard Príncipe Akishino s/n, CP 30890 Tapachula,
Chiapas, Mexico
e-mail: miguel.salvador@unach.mx

F. Molina-Freaner
Departamento de Ecología de la Biodiversidad, Instituto de
Ecología, Unidad Hermosillo, Universidad Nacional Autónoma
de México (UNAM), Hermosillo, Sonora, Mexico

J. Nuñez-Farfán
Laboratorio de Genética Ecológica y Evolución, Departamento de
Ecología Evolutiva, Instituto de Ecología, Universidad Nacional
Autónoma de México (UNAM), México, DF, Mexico

D. Betancur-Ancona
Facultad de Ingeniería Química, Universidad Autónoma de
Yucatán, Mérida, Yucatán, Mexico

Introduction

Cacao (*Theobroma cacao* L.) is the main raw material in the manufacture of chocolate. Currently, its cultivation is mainly widespread in the tropics of America, Africa and Asia. The flavor and aroma are two of the most important features that consumers and manufacturers appreciate in cacao beans. Of the three morphogeographic varieties recognized (Criollo, Forastero and Trinitario), it has been reported that beans from the Criollo variety are of the highest sensory quality [1] and from which the fine cacao is obtained, which currently contributes 6.8 % of world production [2]. In addition to its genetic origin, which can be verified and the individual cataloged in any of the ten genetic groups recently proposed using molecular markers [3], other factors that also influence the quality are the environmental conditions where the

tree grows, and the pre- and post-harvest practices [4]. All factors together affect the composition and concentration of volatile and non-volatile chemical compounds that determine the flavor and aroma of cacao beans [4]. Compositional analysis is a useful tool in characterizing the quality of seeds and consequently of the product obtained from them. In this sense, the traditional approach is to analyze the variances (ANOVA) of the analytes or parameters in particular; however, when variations between samples tend to be very narrow, or factors exist that contribute variability such as geographical origin of samples, the ANOVA is limited. In that case, we can turn to multivariate methods, which provide more useful interpretations [5, 6]. Thus, for example, cluster analysis and canonical discriminant analysis have been used to select indicators of genetic similarity between groups and achieve accurate identification [7], as well as to characterize the genus *Theobroma* individuals using several morphological traits of leaf, flower, fruit and seed [8]. Cluster analysis has also been employed to classify agave fermented drinks by analyzing the combination of volatile compounds [9] or to locate molecules (volatile acids) indicators of off-flavor in cacao by principal component analysis [10]. Using similar multivariate techniques, it has been reported that some morphological and chemical traits taken together are, for example, critical to classify tuber samples by geographical origin [11].

In Mexico, cacao cultivation is concentrated in two regions at southeast of the country. In the most southern region, on the border with Guatemala (in the Chiapas state), 30 % of the national total of cacao is produced [12]. It has been reported empirically, that the beans produced possess aroma and flavor of the highest quality. These features have awakened interest in this cacao from companies in the overseas market that produce fine chocolates; so that, much of the cacao grown in this region is exported [13]. This region, located in what made up pre-colonial Mesoamerica, has a history of Criollo cacao cultivation dating back to 1900 BC [14]. Based on the information from the past in this region and taking into account that in the mid-twentieth century, introduction of genotypes with higher agronomic performance was conducted (presumably Trinitarios and/or Forasteros), the quality of existing plantations may be the result of the presence of cacao with high Criollo ancestry, reported in a study using microsatellite molecular markers [15] or Trinitario genotypes displaying quality characteristics inherited from the parental Criollo. In the latter study, it is concluded that the morphological characteristics of the tree and fruit do not seem to adequately define the variety. Therefore, the objective of this study was to analyze with a chemometric approach the characteristics of cacao seeds using multivariate analysis techniques.

Materials and methods

Collection and processing of cacao beans

Fruits (from two to six per accession) of 45 cacao accessions and one Patate (*T. bicolor*) accession were collected in seven municipalities of the Soconusco region, Chiapas, Mexico (Suchiate, SUED; Frontera Hidalgo, FHSA; Cacahoatán, CAAM; Tapachula, TASG; Mazatán, MAMG; Huehuetán, HUJF, HUTG, Tuzantán, TUCA, TURA and Tuxtla Chico, TCHR) having morphological traits described by Engels [16] of the Criollo variety (sweet pulp, white seeds, elongated pods with pronounced grooves) and desirable sensory characteristics (declared by the farmers). The pods were harvested at full maturity (yellowing of more than 60 % of the pod), were transported to the laboratory in labeled bags and processed the same day as cutoff. The pods were cut transversely using a scalpel, and the seeds were removed from the placenta which was discarded. The seeds were introduced again to the pod and sealed with tape, were placed upright in an acclimation chamber at 28 °C and RH of 80 %, for a period of 6 days, reversing its position every 24 h to ensure homogeneity. After 6 days, the beans were dried at 60 °C × 48 h in a hot air oven (Felisa® Mexico). The dried beans were stored under refrigeration (4 °C) until the time of analysis.

Morphological traits of beans

For ten fermented and dried beans of each accession, the following determinations were made: weight (g) in analytical balance (Ohaus® Pioneer™ Model PA214; USA), length (cm), width (cm) and bean circumference (cm) using digital vernier caliper (Digimatic Caliper Model 500-197-20; Mitutoyo America Corp. Aurora, IL), according to the procedure described by Rangel-Fajardo et al. [17].

Chemical composition of beans

For fermented and dried beans, the following was determined: moisture (method 925.09), ash (method 923.03), fat (method 920.39) and protein (method 954.01) using the AOAC official procedures [18].

Quantification of fatty acids

Following the procedure, 2 g fat of cacao beans was extracted described by Liendo et al. [19]. This fat was used to obtain the fatty acid methyl esters (FAMES) by alkaline transesterification following the method of Ichihara et al. [20] with adaptations in final volume. In a vial, pre-washed with hexane and dried for 2 h to 300 °C, 2 ml of hexane and 40 mg of fat were added. The vial was placed on a dry

block heater (Thermoblock®) at 40 °C for 1 min, and then, 100 µL of sodium methoxide (prepared with NaOH 2 M in absolute methanol) was added, stirred vigorously for 2 min, heated for 5 min (40 °C) and centrifuged (Eppendorf microcentrifuge 5425D) at 1,397g for 10 min. The supernatant containing FAMES was removed and used to determine the fatty acid composition by gas chromatography–mass spectrometry (Thermo Scientific Focus GC chromatograph and quadrupole mass spectrometer, Thermo Scientific DSQ II). One microliter of FAMES was injected, and separation was performed with a Thermo TR-5 ms SQC column (30 m × 0.25 mm ID × 0.25 µm) with the following temperature program: 100 °C, heating 25 °C/min to 200 °C, 2.5 °C/min to 230 °C for 1 min and 10 °C/min to 250 °C. The chromatograph was operated using helium as gas carrier at a pressure of 7.48 psi. The injector was programmed in split mode 60:1 to 280 °C temperature. The mass spectrometer was operated with ionizing voltage of 70 eV, interface temperature of 280 °C in scan mode and mass range of 50–500 m/z. The compounds in each sample were identified by comparing their mass spectra with the compounds of the library of the National Institute of Standards and Technology (NIST). The sum of the abundances of the major fatty acids (myristic, palmitoleic, palmitic, linoleic, oleic, stearic, eicosanoic, behenic) as 100 % fat and the abundance of each fatty acid is reported as a relative percentage (%). From these data, the content of saturated fatty acids (SFA), unsaturated (UFA) and SFA/UFA ratio was calculated.

Total polyphenols

From 0.1 g of previously defatted bean, phenolic compounds were extracted by following the procedures described by Jimenez-Escrig et al. [21] and Cannac et al. [22]. The extraction solution was prepared by methanol and water 1:1 (v/v) mixing and adjusting to pH 2. To each sample macerated with a glass rod, 2 mL of the extract solution was added and was left to agitate for 24 h at room temperature. Then, it was centrifuged at 2,739g for 5 min, and the supernatant removed and stored at 4 °C. With the precipitate, a second extraction was performed under the same conditions. After 24 h, the second supernatant was obtained and added to the previous one, which was used to quantify the total content of polyphenolic compounds, using the Folin-Ciocalteu [23] reagent. The reaction mixture containing 5 µL of sample was prepared, 45 µL of methanol: water (pH 2), 500 µL of Folin reagent and 400 µL of Na₂CO₃ 1 M. This was left 15 min in the dark, and the absorbance was measured at 765 nm using spectrophotometer Thermo Scientific GENESYS 20. A standard curve of gallic acid was prepared, and the results are reported in equivalent terms of gallic acid per gram of sample (mg EAG/g dry weight).

Antioxidant activity

First, an extraction of antioxidant compounds with aqueous methanol solution (80 % v/v) was performed. With the extract, the antioxidant activity was measured by testing discoloration through radical trapping of 2,2'-azino-di-3-ethylbenzothiazoline-6-sulfonic acid (ABTS⁺) [24]. A calibration curve was plotted with sequential dilutions of 4 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (trolox) stock solution to obtain concentrations of 10, 20, 30, 40, 50 µM, and as a blank solution, we used a 80 % methanolic solution. After this, we calculate the antioxidant activity once we obtained the sample value in the calibration curve. Then, the percent inhibition was interpolated vs concentration of trolox, and the results were reported in mmol trolox equivalent per gram of sample (TEAC mmol/g).

Data analysis

Our morphological (weight, length, width, bean circumference) and chemical traits measured in cocoa beans (total 18) were undergoing to principal component analysis (PCA). *T. bicolor* (Pataste) was employed with external control for assessing grouping of samples with similar morphochemometric characteristics. With the traits that explained the new variables (F1 and F2), a cluster analysis was performed and a dissimilarity dendrogram was created. Finally, the membership of each sample to certain group created by the dendrogram was verified using discriminant analysis (DA). The analyses were performed using software XLSTAT™ v 2012 and Infostat™ v 2008.

Results and discussion

Morphological characteristics of seeds

Table 1 shows the values of the four morphological variables evaluated in cacao beans. The average bean weight was 1.25 g. Only some samples showed remarkable deviations from this average, the CAAM01 sample exhibited a weight of 0.63 g (the lowest of all), and MAMG08 recorded the highest weight (1.85 g) in all samples. The latter also exhibited the highest value for the length of the bean (2.60 cm average for all beans). The results of this study are consistent with those reported for Criollo cacao Porcelana regarding bean length (average 1.70 cm) but are slightly higher as the width of bean, since that study reported an average value of 1.5 cm [25]. Overall, the four characteristics analyzed show little variability (small values of variance), so analyzing variances individually (ANOVA) found statistical equality ($P > 0.05$) for all samples (even when the *T. bicolor* sample was introduced to the analysis).

Table 1 Morphological traits assessed in 45 samples of cacao seeds (*Theobroma cacao*) Crolio phenotype and one Patate (*T. bicolor*) collected from cultivated plants

Sample or accession	Weight (g)	Length (cm)	Width (cm)	Bean circumference (cm)
TASG07	1.08 ± 0.09	2.48 ± 0.49	1.82 ± 0.11	3.60 ± 0.14
TASG08	1.31 ± 0.10	2.70 ± 0.30	1.92 ± 0.13	3.92 ± 0.19
TASG09	1.17 ± 0.21	2.36 ± 0.15	1.92 ± 0.11	4.06 ± 0.37
TASG10	1.25 ± 0.04	2.61 ± 0.11	1.82 ± 0.02	3.84 ± 0.15
TASG12	1.39 ± 0.34	2.48 ± 0.24	1.88 ± 0.43	4.06 ± 0.66
TASG13	1.17 ± 0.08	2.58 ± 0.13	1.84 ± 0.19	3.80 ± 0.29
TASG15	1.04 ± 0.07	2.20 ± 0.20	1.60 ± 0.28	3.74 ± 0.18
TASG16	1.52 ± 0.05	2.84 ± 0.27	1.92 ± 0.11	3.98 ± 0.15
TASG17	1.33 ± 0.09	2.84 ± 0.11	2.10 ± 0.34	3.74 ± 0.53
TASG21	1.07 ± 0.15	2.48 ± 0.65	1.62 ± 0.39	4.20 ± 0.30
TASG23	1.35 ± 0.20	2.46 ± 0.21	1.98 ± 0.48	3.98 ± 0.76
CAAM01	0.63 ± 0.18	2.30 ± 0.20	1.52 ± 0.11	2.48 ± 0.58
CAAM06	1.63 ± 0.05	2.54 ± 0.38	1.88 ± 0.21	4.12 ± 0.22
CAAM07	1.12 ± 0.07	2.64 ± 0.11	1.70 ± 0.16	3.74 ± 0.18
CAAM08	1.46 ± 0.03	3.10 ± 0.12	1.78 ± 0.08	3.64 ± 0.13
CAAM09	1.49 ± 0.09	2.80 ± 0.58	2.18 ± 0.13	4.44 ± 0.33
CAAM11	1.32 ± 0.17	2.60 ± 0.19	1.72 ± 0.16	4.08 ± 0.22
CAAM14	1.48 ± 0.13	2.50 ± 0.25	2.14 ± 0.27	4.48 ± 0.29
PHSA03	0.89 ± 0.06	2.60 ± 0.29	1.52 ± 0.04	3.16 ± 0.15
PHSA04	1.25 ± 0.05	2.61 ± 0.16	1.82 ± 0.01	3.84 ± 0.21
PHSA05	1.22 ± 0.08	2.68 ± 0.13	1.82 ± 0.08	3.70 ± 0.16
PHSA06	1.41 ± 0.16	2.96 ± 0.21	1.96 ± 0.15	3.84 ± 0.18
SUED01	1.63 ± 0.12	3.08 ± 0.28	2.36 ± 0.13	4.48 ± 0.11
SUED03	1.14 ± 0.11	2.60 ± 0.20	1.90 ± 0.14	3.98 ± 0.15
SUED05	1.12 ± 0.09	2.26 ± 0.18	1.68 ± 0.38	3.92 ± 0.18
SUED06	1.18 ± 0.07	2.78 ± 0.13	1.84 ± 0.11	3.62 ± 0.13
MAMG01	1.04 ± 0.23	1.98 ± 0.41	1.36 ± 0.40	3.40 ± 0.46
MAMG02	1.14 ± 0.10	2.54 ± 0.18	2.00 ± 0.53	3.66 ± 0.18
MAMG03	1.25 ± 0.12	2.61 ± 0.23	1.82 ± 0.09	3.84 ± 0.07
MAMG04	0.90 ± 0.12	2.66 ± 0.42	1.68 ± 0.08	3.52 ± 0.15
MAMG05	1.25 ± 0.09	2.61 ± 0.23	1.82 ± 0.11	3.84 ± 0.22
MAMG06	1.17 ± 0.05	2.38 ± 0.15	1.58 ± 0.41	4.06 ± 0.13
MAMG08	1.85 ± 0.20	3.16 ± 0.29	2.02 ± 0.24	4.08 ± 0.40
MAMG10	1.25 ± 0.09	2.61 ± 0.19	1.82 ± 0.05	3.84 ± 0.19
MAMG11	1.14 ± 0.08	2.64 ± 0.21	1.80 ± 0.23	3.78 ± 0.15
MAMG13	1.07 ± 0.14	2.50 ± 0.22	1.88 ± 0.11	3.72 ± 0.28
MAMG14	1.23 ± 0.08	2.58 ± 0.13	1.90 ± 0.07	3.88 ± 0.08
HUTG03	1.00 ± 0.06	2.38 ± 0.19	1.62 ± 0.04	3.42 ± 0.58
HUTG04	1.25 ± 0.12	2.61 ± 0.12	1.82 ± 0.23	3.84 ± 0.26
TUCA02	1.30 ± 0.06	2.78 ± 0.08	1.88 ± 0.08	3.84 ± 0.15
TURA01	1.25 ± 0.05	2.88 ± 0.16	1.64 ± 0.21	3.68 ± 0.11
TCHR05	1.25 ± 0.04	2.61 ± 0.09	1.82 ± 0.19	3.84 ± 0.09
HUFR04	1.02 ± 0.25	2.38 ± 0.20	1.82 ± 0.11	3.72 ± 0.30
HUFR05	1.45 ± 0.11	2.42 ± 0.38	1.66 ± 0.27	4.12 ± 0.36
HUFR06	1.35 ± 0.04	2.72 ± 0.04	1.74 ± 0.11	3.98 ± 0.08
Average	1.25	2.60	1.82	3.84
Variance	0.06	0.06	0.04	0.13
<i>T. bicolor</i>	1.75 ± 0.18	2.72 ± 0.19	1.90 ± 0.16	4.18 ± 0.22

Values are the mean ± standard deviation of ten measurements

This coincides with previous studies [8], which show that cacao and Patate beans are similar in length but different in width, reporting an average value of 1.40 cm width for *T. cacao*. In this study, only the seeds of MAMG01 (1.36 cm) sample were near those reported. The cited study also reported average seed weight of 6.25 g for wild *T. bicolor* in South America, while in this study, the value for this sample was much lower (1.75 g). The latter may be due to the adaptation of *T. bicolor* sample to the collection region (southern Mexico), possible genetic differentiation or simply a consequence of the domestication process, as the seeds in the study used for comparison come from wild accessions, and it is documented that the process of domestication tends to reduce both the size of the seed and the fruit [26].

Proximal composition, polyphenols and antioxidant activity

The fat content in 45 samples studied was consistent for the Criollo variety cacao reported in other latitudes. Contents (dry basis) have been reported from 46/100 to 56/100 g [17]. In this study, we found minimum values of 35.53/100 g (HUTG04 sample) and maximum of 59.02/100 g (TCHR05 sample) with mean value of 47.10/100 g (Table 2). It has been suggested [27] that the cacaos with highest commercial value, because they contain fat, are those with values greater than 56/100 g. In this sense, only three samples (TASG23, CAAM06 and TCHR05) of those analyzed here present content greater than said value. It has been suggested that the lipid content of the beans, in addition to its association with genetic origin, can be influenced by environmental conditions [19], and although cacao is grown in the intertropical world zone, where environmental conditions are roughly similar, the microclimates in the region of Mexico where the samples studied are located may have influenced the accumulation of lipids.

Regarding the content of protein (dry basis), a fairly wide range of values was found from 10.27/100 to 18.54/100 g. For most samples, the protein content was low compared with Forastero, where values are reported as 16 % [28]. The values found are similar to samples obtained from the Dominican Republic (11.8/100 g) and Ecuador (12/100 g) [29]. This may indicate that the most individuals analyzed have little correspondence with Forastero genotype and greater connection with Criollo or Trinitario varieties or at least that varieties in the Americas share this characteristic.

For the content of phenolic compounds (dry basis), values were from 0.67 g EAG/100 g for the MAMG03 sample to 6.85 g EAG/100 g for the FHSA03 sample (Table 2). It has been reported that the fermentation process of the

seeds is critical to the content of phenolic compounds since unfermented samples of Forastero variety reach values up to 8 g EAG/100 g, while for fermented samples content is reported from 2 to 4 g EAG/100 g [30]. *T. bicolor* together with six accessions of cacao had values below 2 mg EAG/g, which could lead to the taste possessed by the beans and products derived from them. The correlation between polyphenol content and antioxidant capacity of the beans in the studied samples, although positive, was low ($R^2 = 0.0034$ excluding *T. bicolor*), contrary to the strong correlation previously reported [31, 32]. This discrepancy could be attributed to the fact that some proteins, protein fractions or peptides may contribute to the total antioxidant capacity [33]. In this sense, the values found for total antioxidant capacity ranged from 7.86 to 15.38 mmol TEAC/g (Table 2). These results agree with reports for cacao powder [32] and are superior to tea and wine samples reported by the same authors, and are even higher when compared with fruit, classified by their potent antioxidant activity as "superfruits" among European consumers [34].

Content and fatty acid composition

The cumulative content of saturated fatty acids of different cacao accessions studied was between 59 and 72 % (Table 3). It has been reported that the soft texture of cacao butter (a highly desirable characteristic) is associated with a high proportion of saturated fatty acids [19]. In this sense, all accessions (including *T. bicolor*) have this feature and values above those reported for other cultivars of cacao in South America (62 % [35]). This also demonstrates little or no effect of mean daily temperatures on the content of saturated fatty acids, which can be increased with increasing the temperature of the fruit [36].

Classification of samples using multivariate analysis

In order to reduce the dimensionality of the data presented in Tables 1, 2 and 3, a series of multivariate analyzes were conducted. Prior to this analysis, it is important to note that the variables palmitic acid, stearic acid (Table 3) and polyphenol content (Table 2) when analyzed together, including Patate (*T. bicolor*) as control exhibited high variance values (data not shown) which, as shown later, are useful as discriminatory variables.

For the PCA, 18 morphological and chemical traits (Tables 1, 2 and 3), obtained from the characterization of the beans and their components, were entered. As a result of the PCA, the new F1 and F2 factors were obtained, contributing the greatest variability in the observations, 18.67 and 13.69 %, respectively. From the preceding, the original variables that contribute more than 10 % of the components 1 and 2 were localized. Those variables were palmitoleic

Table 2 Chemical parameters determined for samples of cacao beans (*Theobroma cacao*) Criollo phenotype ($n = 45$) and Patate (*T. bicolor*) ($n = 1$) obtained from cultivated plants

Sample	Moisture (%)	Ash (g/100 g db)	Fat (g/100 g db)	Protein (g 100/g db)	Polyphenols (g GAE/100 g db)	Antioxidant capacity (mmol TEAC/g db)
TASG07	4.11	4.74	49.88	14.42	2.71	7.86
TASG08	4.57	4.27	43.80	11.65	2.58	10.30
TASG09	2.56	3.43	51.84	10.27	4.34	10.47
TASG10	1.24	4.55	48.10	13.94	2.75	15.25
TASG12	3.83	3.31	46.41	13.00	3.88	9.48
TASG13	4.18	3.46	51.67	13.09	3.09	9.29
TASG15	2.19	5.71	44.47	13.53	4.26	9.62
TASG16	2.18	4.87	40.79	11.32	3.78	10.13
TASG17	2.97	4.69	51.33	14.43	4.19	10.34
TASG21	2.17	4.24	37.28	11.49	4.36	11.84
TASG23	4.09	6.03	56.25	12.48	2.25	9.12
CAAM01	4.06	7.03	40.65	13.47	3.39	11.51
CAAM06	2.08	2.94	57.13	12.42	3.11	9.03
CAAM07	3.07	5.82	47.34	12.51	5.67	12.13
CAAM08	4.28	3.02	36.90	11.52	3.54	12.29
CAAM09	2.29	4.08	48.36	13.23	2.82	13.47
CAAM11	4.87	3.55	38.29	13.78	3.15	10.07
CAAM14	4.44	4.34	49.71	11.93	6.25	12.07
FHSA03	4.27	4.26	41.80	12.84	6.85	10.98
FHSA04	2.30	3.58	50.14	15.49	5.50	15.30
FHSA05	4.10	3.20	55.94	14.55	3.28	8.04
FHSA06	4.70	4.49	41.81	15.95	5.23	15.18
SUED01	6.53	6.56	52.78	13.94	5.66	11.32
SUED03	3.46	5.33	49.49	13.69	1.54	8.40
SUED05	5.61	5.21	50.60	11.84	3.16	9.24
SUED06	4.63	6.40	45.09	12.14	1.97	11.74
MAMG01	6.22	4.94	52.53	12.04	3.70	10.54
MAMG02	5.82	4.92	53.10	11.61	1.33	7.94
MAMG03	3.80	7.48	44.70	18.54	0.67	14.85
MAMG04	5.95	5.58	52.89	12.96	2.90	10.02
MAMG05	2.30	6.75	49.15	10.66	1.68	14.98
MAMG06	6.14	4.64	55.38	12.88	4.19	11.83
MAMG08	3.01	5.41	38.50	13.52	1.79	11.25
MAMG10	4.43	5.71	43.19	13.48	2.09	8.27
MAMG11	5.24	6.26	47.80	11.32	2.25	9.02
MAMG13	4.39	4.58	39.69	13.11	2.60	10.11
MAMG14	4.22	3.55	48.53	14.94	5.43	12.04
HUTG03	5.64	4.99	43.84	13.92	2.49	10.84
HUTG04	2.20	4.19	35.53	15.96	1.25	15.38
TUCA02	2.69	5.73	43.43	12.68	5.40	10.56
TURA01	5.11	3.48	45.79	14.87	2.88	9.44
TCHR05	5.11	4.73	59.02	14.20	4.27	14.96
HUJF04	7.06	4.12	41.28	14.41	4.87	10.97
HUJF05	4.97	3.80	50.26	13.31	3.42	12.20
HUJF06	4.93	3.81	46.80	14.63	3.62	9.84
<i>T. bicolor</i>	4.40	3.45	41.31	22.06	0.68	8.11

GAE gallic acid equivalent, TEAC trolox equivalent antioxidant capacity

Table 3 Contents (relative abundance, %) of the principal fatty acids of *T. cacao* beans ($n = 45$) with Criollo phenotype and *T. bicolor* ($n = 1$) of cultivated plants

Sample/accession	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C20:0	C22:0	SFA	UFA	SFA/UFA
TASG07	0.10	26.24	0.18	38.10	31.31	2.23	0.94	0.91	66.28	33.72	1.97
TASG08	0.12	20.76	0.15	49.44	26.68	1.13	1.30	0.43	72.04	27.96	2.58
TASG09	0.28	23.44	0.14	44.85	27.62	1.55	1.19	0.93	70.68	29.32	2.41
TASG10	0.06	24.06	0.10	39.17	34.21	1.14	0.99	0.27	64.55	35.45	1.82
TASG12	0.18	18.06	0.20	51.08	27.48	1.39	1.25	0.36	70.93	29.07	2.44
TASG13	0.13	14.99	0.10	49.47	32.37	1.24	0.95	0.75	66.29	33.71	1.97
TASG15	0.14	22.70	0.12	45.37	28.92	1.75	0.71	0.29	69.21	30.79	2.25
TASG16	0.06	24.81	0.17	44.41	27.12	2.16	0.95	0.33	70.56	29.44	2.40
TASG17	0.10	22.21	0.11	39.71	33.46	3.07	1.08	0.27	63.37	36.63	1.73
TASG21	0.62	18.89	0.09	49.20	27.78	1.65	1.58	0.20	70.49	29.51	2.39
TASG23	0.06	20.30	0.11	46.16	29.96	1.91	1.30	0.20	68.02	31.98	2.13
CAAM01	0.05	24.41	0.13	42.54	29.48	1.91	1.30	0.17	68.48	31.52	2.17
CAAM06	0.07	19.23	0.18	46.82	30.00	2.21	1.17	0.30	67.60	32.40	2.09
CAAM07	0.06	22.57	0.18	41.43	32.24	2.26	0.97	0.29	65.32	34.68	1.88
CAAM08	0.07	22.21	0.16	43.47	30.52	2.52	0.87	0.17	66.79	33.21	2.01
CAAM09	0.08	27.17	0.23	40.87	27.95	2.50	0.89	0.31	69.31	30.69	2.26
CAAM11	0.07	21.83	0.14	42.66	31.76	2.21	1.05	0.30	65.90	34.10	1.93
CAAM14	0.08	21.43	0.17	41.04	33.90	1.97	1.08	0.33	63.96	36.04	1.77
FHSA03	0.18	25.27	0.24	39.73	31.07	1.82	1.25	0.44	66.87	33.13	2.02
FHSA04	0.06	24.71	0.20	39.38	32.55	1.73	1.08	0.28	65.52	34.48	1.90
FHSA05	0.50	30.47	0.17	34.97	31.05	1.72	0.98	0.14	67.06	32.94	2.04
FHSA06	0.11	25.00	0.25	39.39	32.38	1.65	0.92	0.30	65.72	34.28	1.92
SUED01	0.11	26.50	0.16	39.68	30.80	1.36	1.21	0.18	67.68	32.32	2.09
SUED03	0.07	25.42	0.15	38.35	32.63	2.38	0.86	0.14	64.84	35.16	1.84
SUED05	0.09	25.47	0.16	38.76	32.67	1.58	1.10	0.18	65.59	34.41	1.91
SUED06	0.08	23.39	0.20	42.52	30.82	1.32	1.47	0.21	67.66	32.34	2.09
MAMG01	0.13	24.52	0.25	33.53	37.83	2.31	1.20	0.22	59.61	40.39	1.48
MAMG02	0.14	21.78	0.16	41.31	34.77	0.70	0.99	0.15	64.37	35.63	1.81
MAMG03	0.08	28.97	0.31	36.50	31.36	1.52	1.11	0.16	66.80	33.20	2.01
MAMG04	0.12	27.51	0.25	36.26	33.76	0.76	1.12	0.21	65.22	34.78	1.88
MAMG05	0.07	23.33	0.20	41.94	31.86	1.16	1.26	0.19	66.78	33.22	2.01
MAMG06	0.05	23.93	0.17	41.54	31.96	0.97	1.19	0.20	66.91	33.09	2.02
MAMG08	0.07	25.54	0.16	38.69	32.60	1.32	1.27	0.34	65.91	34.09	1.93
MAMG10	0.10	22.90	0.11	41.51	33.08	0.48	1.31	0.50	66.33	33.67	1.97
MAMG11	0.09	23.70	0.13	39.90	32.26	1.40	1.59	0.93	66.21	33.79	1.96
MAMG13	0.16	29.36	0.17	34.72	31.48	2.50	0.69	0.93	65.85	34.15	1.93
MAMG14	0.84	20.79	0.19	42.82	32.54	1.26	0.90	0.66	66.02	33.98	1.94
HUTG03	0.07	24.70	0.22	36.83	33.69	1.75	1.33	1.42	64.34	35.66	1.80
HUTG04	0.09	27.99	0.31	33.89	34.16	1.57	1.50	0.49	63.96	36.04	1.77
TUCA02	0.09	26.17	0.20	37.38	32.78	1.93	0.85	0.59	65.08	34.92	1.86
TURA01	0.11	23.04	0.14	36.67	35.50	2.46	0.78	1.30	61.90	38.10	1.62
TCHR05	0.06	21.08	0.24	42.91	32.41	1.75	0.96	0.59	65.60	34.40	1.91
HUJF04	0.09	25.70	0.26	38.34	31.79	2.13	0.89	0.81	65.82	34.18	1.93
HUJF05	0.07	25.28	0.28	39.64	31.50	1.70	1.20	0.32	66.51	33.49	1.99
HUJF06	0.33	24.35	0.25	38.68	32.54	2.41	1.00	0.44	64.80	35.20	1.84
<i>T. bicolor</i>	0.03	10.88	0.06	49.52	36.55	1.16	1.57	0.23	62.23	37.77	1.65

acid (F2, 11.52 %), palmitic acid (F1, 10.70 % and F2 13.61 %), stearic acid (F1, 14.85 % and F2, 13.80 %), antioxidant activity (F2, 11.15 %), weight (F1, 18.58 %), length (F2, 13.54 %), width (F1, 12.65 %) and bean circumference (F1, 15.88 %) of the bean. It is noteworthy how the four morphological traits contribute substantially to the composition of the new factors (weight, width and bean circumference of the bean for F1 and length of the bean for F2). However, this contribution is associated with the strong homogeneity of the data (low values in the variances, Table 1), which proved of little use as grouping criteria. Similar to this, Niemenak et al. [37] reported little usefulness of morphological traits like shell weight and number of seeds to group individuals through cacao comparative study using PCA, assuming the lack of significant correlation with the variables that scantily explain the total variability may derive from developmental differences based on environmental factors. In that study, it was found that the variables analyzed as a whole accounted for 93.76 % of the total variability, having as dominant features for the first principal component (58.60 % of total variation) catechin, epicatechin, anthocyanin content and total phenols attributing to the latter the separation among the groups of cacao clones, and for the second principal component (21.60 % of the total variability) epicatechin content. In the present study, by integrating the morphological traits to PCA, it was not possible to visualize the expected segregation of Pataste (Fig. 1a), an accession not corresponding to cacao. For this reason, a new PCA was performed incorporating only the 14 chemical traits. Derived from this analysis, Fig. 1b provides a representation of the observations regarding the new factors, F1 and F2. Both factors explained 32.8 % of the total variability. In this representation, a consistent separation of Pataste (*T. bicolor*) sample is shown.

After excluding the four morphological traits, PCA showed that six of the original variables are contributing largely to the new factors F1 and F2. Such traits are fatty acids (palmitoleic, palmitic, oleic and stearic acid) as well as moisture and the polyphenol content. These variables were used thereafter. With data from these six chemical variables, the dendrogram shown in Fig. 2 was constructed. The formation of seven groups shows differentiation among cacao accessions themselves (cluster) and the accession of Pataste (dissimilarity value > 100), which is left alone in a separate cluster. The cluster that grouped more accessions was cluster 1 (C1, Fig. 2), whereas cluster 6 (C6, Fig. 2) included only two accessions. Accessions corresponding to clusters 5 and 6 were the most dissimilar (>5,000) from the other accessions.

Verifying the usefulness of the six chemical characters in the classification, a new cluster analysis was performed, excluding the Pataste, so that the final grouping

was obtained. To verify the actual membership of each of the accessions in the group assigned with the cluster analysis (dendrogram), a discriminant analysis (DA) was performed. The results of this test revealed 8.89 % overall error in assigning a priori. Cluster 1 (33.3 %) and cluster 5 (25 %) were the only one in which the errors occurred. Two accessions of each cluster belonged to different clusters. The four misclassified individuals were grouped into their respective new cluster, and with this matrix, the DA was run again, finding 0 % error in the assignment. The 45 accessions of cacao plantations from Soconusco, Chiapas, Mexico, after statistical analysis based on the chemical composition of its beans, remain grouped as shown in Table 4.

With the exception of cluster 3 (Table 4) which grouped six accessions of the same geographical population, little correlation was found between geographical origins of individuals with DA obtained from the chemical characteristics grouping, as the accessions collected in the same site (marked with the same four-letter code) were placed in different groups. Since the chemical composition is associated with quality characteristics exhibited by cacao, the hypothesis that a microclimate can enhance the unique characteristics of flavor quality and aroma, supported by other authors [38], turns out to have a minimum contribution. This phenomenon may have several explanations: (1) either the characteristics of the microclimate in our study are not sufficiently decisive regarding the chemical composition, (2) the accessions have a strong genetic association (highly related individuals) as demonstrated in genetic studies using molecular markers [15] or (3) or trees in adaptive process to the environmental conditions of the Soconusco, Mexico, which could be promoting the particular flavor and other characteristics as demonstrated occurs in other places. It has suggested using the term "terroir" to differentiate the distinctive characteristic of regional cocoas [39]. Even with the above, the classification proposed here is undoubtedly critical for the sensory quality which accessions must exhibit, given that chemical variables are related directly with the sensorial character, and several studies have even proposed the use of chemometric approach with a large number of tools to try to predict the sensory quality [40].

Conclusions

The chemometric multivariate analysis of 18 physical-chemical traits of cacao beans allowed select six traits that were decisive for clustering of samples of similar macrogeographical origin. These characteristics were the content of palmitoleic acid, palmitic acid, oleic acid, stearic acid, polyphenols and moisture of the cocoa beans.

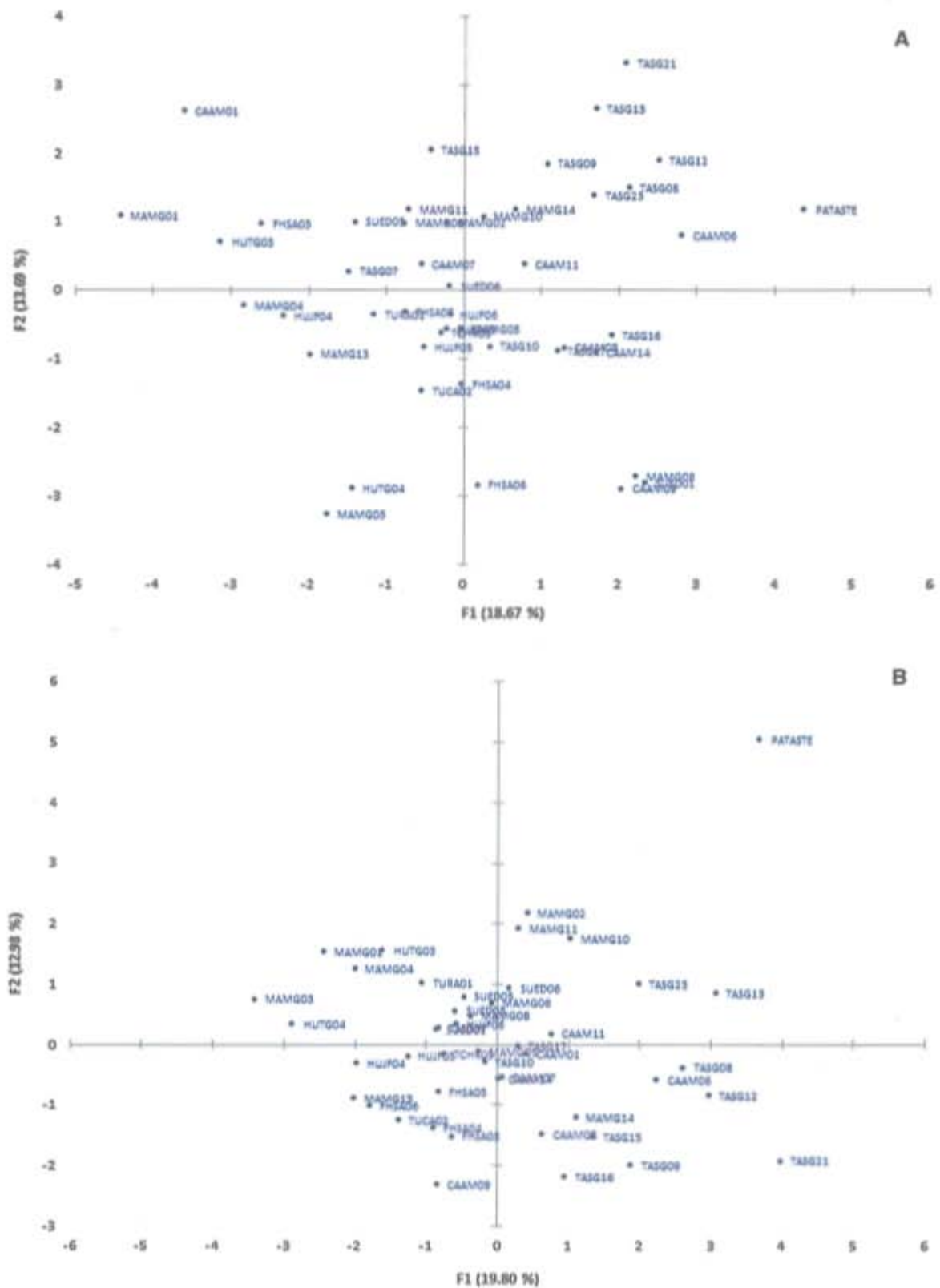


Fig. 1 Representation of observations (samples) on the factors (F1 and F2) obtained from principal component analysis. **a** With 18 morpho-chemical traits (Tables 1, 2 and 3). **b** With 14 chemical traits (Tables 2 and 3)

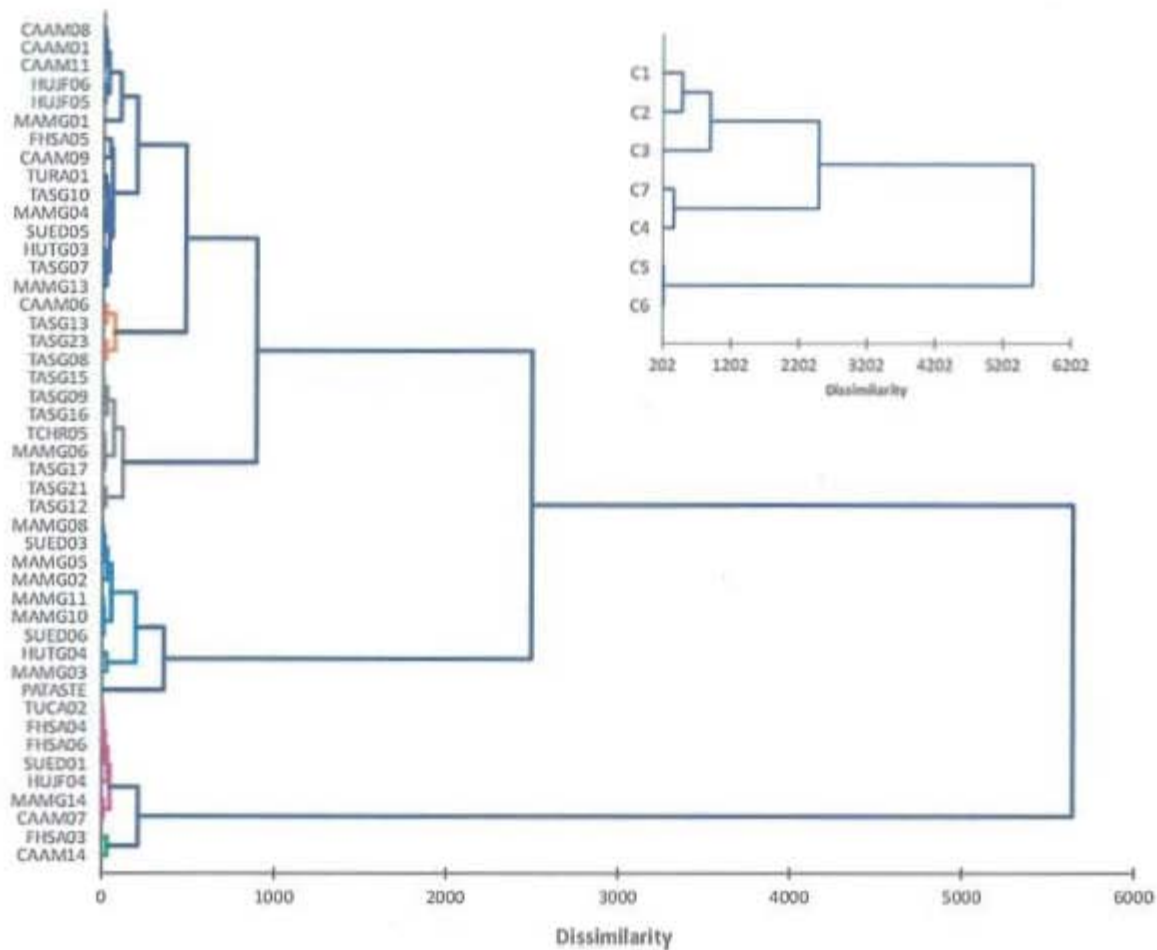


Fig. 2 Dendrogram based on Euclidean distance constructed with data from the six chemical variables evaluated in cacao beans ($n = 45$) and Pataste ($n = 1$) that best explain the variability of the F1 and F2 factors in PCA

Table 4 Grouping using discriminant analysis of cacao seeds from 45 accessions with Criollo phenotype based on the chemical composition

1	2	3	4	5	6	7
TASG07	TASG08	TASG09	CAAM07	FHSA05	TASG10	MAMG03
CAAM09	TASG13	TASG12	CAAM14	SUE005	SUE003	HUTG04
MAMG11	TASG23	TASG15	FHSA03	MAMG01	SUE006	
HUTG03	CAAM01	TASG16	FHSA04	MAMG04	MAMG02	
TURA01	CAAM06	TASG17	FHSA06	HUJF05	MAMG05	
MAMG13	CAAM08	TASG21	SUE001	HUJF06	MAMG08	
	CAAM11	MAMG06	TUCA02		MAMG10	
		TCHR05	HUJF01			
			MAMG14			

Morphological traits (length, width, weight and bean circumference) of the cacao beans are not useful to segregate individuals even of different species. For samples of cocoa beans from the southern region of Mexico, mismatch between the microsite cultivation of plants and the chemical composition of beans was found.

Acknowledgments Thanks are due to the Posgrado en Ciencias Biológicas of the Universidad Nacional Autónoma de México (UNAM), for the academic instruction in doctoral studies for the first author, and to the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexico) for a financial grant awarded to the first author. Special thanks to Samuel Guillifa for contact with farmers and to Lisbeth Chacón and Andy Villarreal for technical support.

Conflict of interest None.

Compliance with Ethics Requirements This article does not contain any studies with human or animal subjects.

References

- Camú N, De Winter T, Addo S, Takrama J, Bernaert H, De Vuyst L (2008) Fermentation of cocoa beans: influence of microbial activities and polyphenol concentrations on the flavour of chocolate. *J Sci Food Agric* 88:2288–2297
- ICCO (2012) International Cocoa Organization. *Quart Bull Cocoa Stat 6 XXXVIII—No. 2. Cocoa year 2010/11* (May 2013)
- Motamayor JC, Lachenaud P, da Silva e Mota JW, Looer R, Kuhn DN, Brown JS, Schnell RJ (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS One* 3:e3311
- Afoakwa EO (2012) Chocolate and cocoa, flavor and quality. *Kirk-Othmer Enc Chem Technol* 1–19. doi:10.1002/0471238961.chocafaa.a01
- Adamska E, Cegielska-Taras T, Kaczmarek Z, Szala L (2004) Multivariate approach to evaluating the fatty acid composition of seed oil in a doubled haploid population of winter oilseed rape (*Brassica napus* L.). *J Appl Genet* 45:419–425
- Harrison JM, Howard D, Malven M, Halls SC, Culler AH, Harrigan GG, Wolfinger RD (2013) Principal variance component analysis of crop composition data: a case study on herbicide-tolerant cotton. *J Agric Food Chem* 61:6412–6422
- Yeater KM, Bollero GA, Bullock DG, Rayburn AL, Rodriguez-Zas S (2004) Assessment of genetic variation in hairy vetch using canonical discriminant analysis. *Crop Sci* 44:185–189
- Santos RC, Pires JL, Correa RX (2012) Morphological characterization of leaf, flower, fruit and seed traits among Brazilian *Theobroma* L. species. *Genet Resour Crop Evol* 59:327–345
- De León-Rodríguez A, Escalante-Minakata P, Jiménez-García MI, Ordoñez-Acevedo LG, Flores Flores JL, Barba de la Rosa AP (2008) Characterization of volatile compounds from ethnic Agave alcoholic beverages by gas chromatography-mass spectrometry. *Food Technol Biotechnol* 46:448–455
- Rodríguez-Campos J, Escalona-Buendía H, Orozco-Avila I, Lugo-Cervantes E, Jaramillo-Flores M (2011) Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Res Int* 44:250–258
- López-Cortés I, Salazar-García DC, Malheiro R, Guardiola V, Pereira JA (2013) Chemometrics as a tool to discriminate geographical origin of *Cyperus esculentus* L. based on chemical composition. *Ind Crop Prod* 51:19–25
- SIAP (2012) Servicio de Información Agroalimentaria y Pesquera. SAGARPA México. Producción agrícola por cultivo, reporte 2012. http://www.siap.gob.mx/index.php?option=com_wrapper&view=wrapper&Itemid=350. Accessed 30 July 2013
- Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J, Salvador-Figueroa M (2012) Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad. *BioTecnología* 16:36–56
- Powis T, Hurst W, Rodríguez M, Ortiz P, Blake M, Cheetham D, Coe M, Hodgson J (2008) The origins of cacao use in Mesoamerica. *Mexicon* 30:35–38
- Vázquez-Ovando JA, Molina-Freaner F, Nuñez-Farfán J, Ovando-Medina I, Salvador-Figueroa M (2014) Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico. *Genet Mol Res* 13:10404–10414
- Engels JM (1983) A systematic description of cacao clones. I. The discriminative value of quantitative characteristics. *Euphytica* 32:377–385
- Rangel-Fajardo MA, Zavaleta-Mancera HA, Cordova-Tellez L, Lopez-Andrade AP, Delgado-Alvarado A, Vidales-Fernandez I, Villegas-Monter A (2012) Anatomía e histoquímica de la semilla del cacao (*Theobroma cacao* L.) criollo mexicano. *Rev Fitotec Mex* 35:189–197
- AOAC (1997) Official methods of analysis. Association of Official Analytical Chemists, 17th ed. In: Horwitz W (ed). Washington, DC
- Liendo R, Padilla F, Quintana A (1997) Characterization of cocoa butter extracted from Criollo cultivars of *Theobroma cacao* L. *Food Res Int* 30:727–731
- Ichihara K, Shibajara A, Yamamoto K, Nakayama T (1996) An improved method for rapid analysis of the fatty acids of glycerolipids. *Lipids* 31:535–539
- Jiménez-Escrig A, Dragsted L, Daneshvar B, Pulido R, Saura-Calixto F (2003) *In vitro* antioxidant activities of edible artichoke (*Cynara scolymus* L.) and effect on biomarkers of antioxidants in rats. *J Agric Food Chem* 51:5540–5545
- Cannac M, Pasqualini V, Greff S, Fernandez C, Ferrat L (2007) Characterization of phenolic compounds in *Pinus laricio* needles and their responses to prescribed burnings. *Molecules* 12:1614–1622
- Pourmorad F, Hosseini-mehr SJ, Shahabimajd N (2006) Antioxidant activity, phenol and flavonoid contents of some selected Iranian medicinal plants. *Afr J Biotechnol* 5:1142–1145
- Vázquez-Ovando A, Rosado-Rubio G, Chel-Guerrero L, Betancur-Ancona D (2009) Physicochemical properties of a fibrous fraction from chia (*Salvia hispanica* L.). *LWT- Food Sci Technol* 42:168–173
- Chacón RI, Ramis C, Gómez C (2011) Morphological description of fruits and seeds of the Criollo Porcelana cocoa (*Theobroma cacao* L.) in south of Maracaibo Lake. *Rev Fac Agron* 28(Suppl. 1):1–13
- Orsi CH, Tanksley SD (2009) Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. *PLoS Genet* 5:e1000347
- Pires JL, Mattos-Cascardo JC, Lambert SV, Figueira A (1998) Increasing cocoa butter yield through genetic improvement of *Theobroma cacao* L.: seed fat content variability, inheritance, and association with seed yield. *Euphytica* 103:115–121
- Álvarez C, Pérez E, Lares M (2007) Physical-chemical characterization of fermented, dried and roasted cocoa beans cultivated in the region of Cuyagua, Aragua state. *Agron Trop* 57:249–256
- Bertazzo A, Comai S, Brunato I, Zancato M, Costa CVL (2011) The content of protein and non-protein (free and protein-bound) tryptophan in *Theobroma cacao* beans. *Food Chem* 124:93–96
- Tomas-Barberán F, Cienfuegos-Jovellanos E, Marín A, Muguerra B, Gil-Izquierdo A, Cerdá B, Zafrilla P, Morillas J, Mulero J, Ibarra A, Pasamar M, Ramón D, Espín J (2007) A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. *J Agric Food Chem* 55:3926–3935
- Abbe-Maleyki MJ, Ismail A (2010) Antioxidant properties of cocoa powder. *J Food Biochem* 34:111–128
- Lee KW, Kim YJ, Lee HJ, Lee CY (2003) Cocoa has more phenolic phytochemicals and a higher antioxidant capacity than teas and red wine. *J Agric Food Chem* 51:7292–7295
- Preza AM, Jaramillo ME, Puebla AM, Mateos JC, Hernández R, Lugo E (2010) Antitumor activity against murine lymphoma L5178Y model of proteins from cacao (*Theobroma cacao* L.) seeds in relation with *in vitro* antioxidant activity. *BMC Complement Altern Med* 10:61

34. Crozier S, Preston A, Hurst J, Payne M, Mann J, Hainly L, Miller D (2011) Cacao seeds are a “Super Fruit”: a comparative analysis of various fruit powders and products. *Chem Cent J* 5:5
35. Padilla FC, Liendo R, Quintana A (2000) Characterization of cocoa butter extracted from hybrid cultivars of *Theobroma cacao* L. *Arch Lat Nut* 50:200–205
36. Lehrian DW, Keeney PG, Butler DR (1980) Triglyceride characteristics of cocoa butter from cacao fruit matured in a microclimate of elevated temperature. *J Am Oil Chem Soc* 57:66–69
37. Niemenak N, Rohsius C, Elwers S, Ndoumou DO, Lieberei R (2006) Comparative study of different cocoa (*Theobroma cacao* L.) clones in terms of their phenolics and anthocyanins contents. *J Food Compos Anal* 19:612–619
38. Aikpokpodion P (2010) Variation in agro-morphological characteristics of cacao, *Theobroma cacao* L., in farmers’ fields in Nigeria. *N Zeal J Crop Hort* 38:157–170
39. Sukha DA, Butler DR, Comissiong EA, Umaharan P (2014) The impact of processing location and growing environment on flavor in cocoa (*Theobroma cacao* L.)—implications for “Terroir” and certification. *Acta Hort (ISHS)* 1047:255–262
40. Wood JE, Allaway D, Boulton E, Scott IM (2010) Operationally realistic validation for prediction of cocoa sensory qualities by high-throughput mass spectrometry. *Anal Chem* 82:6048–6055

CAPÍTULO 6

Descriptores de olor y sabor en semillas de cacao de Soconusco, Chiapas, México

Con el agrupamiento de muestras planteado en el Capítulo 5, se realizó la evaluación sensorial. Para lo anterior se estableció un panel de jueces entrenados. El presente Capítulo aborda desde el proceso de entrenamiento hasta la obtención de los perfiles de sabor y olor de las muestras. Los resultados de este análisis revelaron diferentes grados de calidad. Las muestras de mayor calidad sensorial se encontraron asociadas con la presencia de notas deseables (dulce, almendras) y la ausencia de notas no deseables. Mediante análisis multivariado fue posible establecer las relaciones estadísticas entre los descriptores (osabor y olor) y las muestras. Se reporta la presencia del descriptor “picante o pungente” en las muestras y se discute sobre la ausencia de notas florales o frutales detectadas en muestras de cacao de otras regiones del mundo.

Estos resultados fueron publicados en la revista Food Science and Technology (Sociedade Brasileira de Food Science and Technology, ISSN versión impresa 0101-2061, versión electrónica 1678-457X). La cita es: Alfredo Vázquez-Ovando, Lisbeth Chacón-Martínez, David Betancur-Ancona, Héctor Escalona-Buendía, Miguel Salvador-Figueroa (2015) Sensory descriptors of cocoa beans from cultivated trees of Soconusco, Chiapas, Mexico. Food Science and Technology DOI: 10.1590/1678-457X.6552.

Sensory descriptors of cocoa beans from cultivated trees of Soconusco, Chiapas, Mexico

Alfredo VÁZQUEZ-OVANDO¹, Lisbeth CHACÓN-MARTÍNEZ¹, David BETANCUR-ANCONA²,
Héctor ESCALONA-BUENDÍA³, Miguel SALVADOR-FIGUEROA^{1*}

Abstract

The odor and taste profile of cocoa bean samples obtained from trees cultivated in southern Mexico were evaluated by trained panelists. Seven representative samples (groups) of a total of 45 were analyzed. Four attributes of taste (sweetness, bitterness, acidity and astringency), and nine of odor (chocolate, nutty, hazelnut, sweet, acidity, roasted, spicy, musty and off-odor) were evaluated. A sample (G7) with higher scores in sweet taste and sweet and nutty odors was detected, as well as a high association between these descriptors and the sample, analyzed through principal component analysis (PCA). Similarly, samples that showed high scores for non-desired odors in cocoas such as off-odor and musty were identified and related by PCA to roasted odor and astringent taste (G2 and G4). Based on this scores, the samples were listed in descending order by their sensory quality as G7> G5> G6> G3> G1> G4> G2.

Keywords: quantitative descriptive analysis; Criollo cocoa; chocolate odor; spicy odor.

Practical Application: Descriptors generated are useful for characterizing the sensory quality of cocoa grown in southern Mexico.

1 Introduction

From its origin as a spicy frothy drink (Coe & Coe, 2007) or fermented beverage prepared by several peoples that lived in Mesoamerica nearly 4,000 years ago (Crown, 2013), chocolate has evolved into one of the non-basic food products of highest popularity and acceptance especially in Western countries. Its texture, the sensory stimuli of pleasure and mood produced by its consumption (Macht & Dettmer, 2006), in addition to the amount and composition of its phenolic compounds which provide health benefits, are the main reasons for its acceptance (Djoussé et al., 2011; d'El-Rei et al., 2013).

The main raw materials for the production of chocolate are dried cocoa beans, which are derived from the cocoa tree (*Theobroma cacao* L.). In 2012, > 5 million tons of dry beans were produced in tropical regions of the world (Food and Agriculture Organization, 2013). About 65% of the beans are obtained from trees grown in Africa, but only those obtained from trees grown in some parts of the Americas are considered of the highest quality due to the sensory characteristics they possess and conserve from Criollo parents (Smith, 1999). These characteristics which have a direct impact on the quality of chocolate are due to genetic, environmental and post harvest processing factors (Voight, 2013). The first determining factor associated with the beans obtained from the Criollo variety is genetic. This variety was domesticated by the peoples of Mesoamerica (Cruz et al., 1995) and there are reports of the existence of trees with high ancestry in several countries which once formed that region (Motilal et al., 2010; Trognitz et al., 2011; Vázquez-Ovando et al., 2014). Similarly, the environmental factor depends on the farm region and it has been reported to exert a strong effect on the chemical characteristics

and consequently on the sensory properties exhibited by beans (Afoakwa et al., 2008). Post harvest processing also influences the final quality of beans. An inefficient management (untimely cutting of fruits, non-fermentation, prolonged fermentation, improper drying, among others) may reduce quality (Afoakwa et al., 2008). If post harvest practices remain constant, it can be assumed that genetic and edapho-climatic conditions where the trees grow determines the quality of cocoa. As a result of these processes, it has been shown that cocoa grown in certain regions of the world possess very specific sensory characteristics (Sukha et al., 2008).

In the southeast region of Mexico, the history of cocoa farming dates back to pre-colonial times, when it was used even as currency by the Aztec people and their tributaries. This fact shows the importance that cocoa had (Crown, 2013). The Criollo variety was cultivated by the Olmecs, Mayas, Mokayas and others peoples as recorded in several codices. In the mid-twentieth century, genotypes of greater production were introduced in the region, which when recombined with native plantations gave rise to the trees cultivated at present. The differentiated genetic and chemical characteristics of these trees suggest that sensory characteristics may also be particular. Some specialized industrialists from various European countries have shown specific interest in cocoa obtained from plantations in the Soconusco region, in southern Mexico. However, there are no sensory descriptors to characterize such quality. Therefore, the aim of this work was to obtain sensory descriptors for taste and odor of cocoa beans from trees cultivated in Soconusco, Mexico, with the participation of trained panelists.

Received 30 Oct., 2014

Accepted 08 Mar., 2015

¹Centro de Biociencias, Universidad Autónoma de Chiapas, Tapachula, Chiapas, Mexico

²Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, Mérida, Yucatán, Mexico

³Departamento de Biotecnología, Universidad Autónoma Metropolitana Iztapalapa, México, DF, Mexico

*Corresponding author: jose.vazquez@unach.mx

2 Materials and methods

2.1 Collection and processing of samples

Fruits (2-6 per tree) were collected from 45 cocoa trees in seven municipalities (Suchiate, SUEJ; Frontera Hidalgo, FHSA; Cacahoatán, CAAM; Tapachula, TASG; Mazatán, MAMG; Huehuetán, HUJF, HUTG and Tuzantán, TUCA, TURA, TCHR) in the Soconusco region of Chiapas State in Mexico. The selection was based on phenotypic traits described by Engels (1983) for the Criollo variety (sweet pulp, white beans, elongated pods with pronounced grooves) and high sensory quality (declared by cocoa farmers). The fruits were harvested at consumption ripening (yellowing of more than 60% of the pod), labeled, transported to the laboratory in polyethylene bags and processed the same day. In the laboratory, the pods were washed in clean water and sanitized surface (70% ethanol) and placed in a laminar flow hood. Using a sanitized scalpel, each pod was hand dissected from the middle (equatorial), the two sections were separated carefully and the seed were freed of the surrounding. The connecting thread (placenta) was removed; the seed and pulp were reintroduced in the pod: The re-armed pod was sealed with tape Millipore® (Vázquez-Ovando et al., 2015). Then the pods were placed vertically in an acclimation chamber and maintained at 28 °C and RH of 80% for a period of six days (to ensure the full fermentation in the fruit with more pulp content), reversing their position every 24 h. Once micro-fermentation was completed, the beans were dried at 60 °C for 48 hours in a forced air oven (Felisa® Mexico). The dried beans were stored under refrigeration (4 °C) in sealed glass jars until the day of sensory evaluation.

2.2 Grouping of samples prior to sensory evaluation

In order to obtain enough samples that would allow panelists to carry out the sensory evaluation in just one session, the 45 samples were grouped using multivariate analysis (principal component analysis [PCA] and discriminant analysis [DA]) of morphological and chemical characteristics of the beans (Vázquez-Ovando et al., 2015). Derived from this analysis, seven groups were formed with a minimum of two and up to nine samples per group. The groups were previously reported by Vázquez-Ovando et al. (2015). The average chemical compositions of the groups are shown in Table 1. Composite samples (groups) were formed after mixing all samples belonging to the same group. A representative sample from each group was taken to be evaluated by the trained panelists.

2.3 Selection and training of panelists

Based on procedures of descriptive analysis (QDA method, Lawless & Heymann, 1999), personal invitations were issued to 40 undergraduate students (between the ages of 20 and 23 y) in order to pre-select panelists. The pre-selection questionnaire was administered to assess health status, eating habits and the ability of panelists to express the taste sensations caused by tasting a derivative of cocoa. The screening criteria included good health (self report), non-smoker, ability to work well on a panel, interest in participating, a consumer of cocoa or its derivatives or at least did not dislike its consumption, no allergic reactions to consumption of cocoa and its derivatives, and most importantly the skills to describe and express the sensations perceived.

Table 1. Classification of 45 samples into seven groups based on the multivariate analysis of chemical parameters of cocoa beans. The value of parameters is a mean and standard deviation of number (n = 2 to 9) of samples belonging to this group. Made with data reported by Vázquez-Ovando et al. (2015).

Group	G1	G2	G3	G4	G5	G6	G7
Parameter	TASG07 CAAM09 MAMG11 HUTG03 TURA01 MAMG13	TASG08 TASG13 TASG23 CAAM01 CAAM06 CAAM08 CAAM11	TASG09 TASG12 TASG15 TASG16 TASG17 TASG21 MAMG06 TCHR05	CAAM07 CAAM14 FHSA03 FHSA04 FHSA06 SUED01 TUCA02 HUJF04 MAMG14	FHSA05 SUED05 MAMG01 MAMG04 HUJF05 HUJF06	TASG10 SUED03 SUED06 MAMG02 MAMG05 MAMG08 MAMG10	MAMG03 HUTG04
Moisture (g.100g ⁻¹)	4.46 ± 1.20	4.02 ± 0.90	3.39 ± 1.51	4.36 ± 1.61	5.30 ± 0.78	3.56 ± 1.54	3.00 ± 1.13
Ash (g.100g ⁻¹ db)	4.69 ± 0.94	4.33 ± 1.59	4.45 ± 0.79	4.72 ± 1.06	4.42 ± 0.95	5.58 ± 0.78	5.84 ± 2.33
Fat (g.100g ⁻¹ db)	45.89 ± 3.70	46.38 ± 8.52	48.32 ± 7.40	46.31 ± 4.31	51.50 ± 3.07	46.66 ± 4.81	40.12 ± 6.48
Protein (g.100g ⁻¹ db)	13.48 ± 1.25	12.63 ± 0.87	12.64 ± 1.48	13.85 ± 1.44	13.22 ± 1.19	12.72 ± 1.26	17.25 ± 1.82
Polyphenols (g GAE.100 g ⁻¹ db)	2.63 ± 0.23	3.02 ± 0.45	4.16 ± 0.21	5.65 ± 0.58	3.35 ± 0.30	1.88 ± 0.46	0.96 ± 0.41
Fatty acids (% db of total fatty acid)							
C14:0	0.10 ± 0.03	0.08 ± 0.03	0.19 ± 0.19	0.18 ± 0.25	0.21 ± 0.17	0.08 ± 0.02	0.09 ± 0.00
C16:1	0.18 ± 0.04	0.14 ± 0.03	0.15 ± 0.05	0.21 ± 0.04	0.23 ± 0.05	0.15 ± 0.04	0.31 ± 0.00
C16:0	25.70 ± 2.36	20.53 ± 2.94	21.89 ± 2.39	24.24 ± 2.10	26.27 ± 2.35	23.77 ± 1.35	28.48 ± 0.69
C18:2	2.14 ± 0.46	1.88 ± 0.52	1.79 ± 0.62	1.40 ± 0.81	1.35 ± 0.82	1.22 ± 0.60	1.54 ± 0.03
C18:1	32.03 ± 2.54	30.11 ± 1.83	29.59 ± 2.58	32.23 ± 0.93	33.23 ± 2.45	32.85 ± 1.34	32.76 ± 1.98
C18:0	37.85 ± 2.26	45.79 ± 3.00	44.88 ± 3.77	39.91 ± 1.64	36.97 ± 2.43	40.50 ± 1.71	35.19 ± 1.84
C20:0	1.04 ± 0.35	1.14 ± 0.18	1.11 ± 0.25	1.02 ± 0.15	1.10 ± 0.10	1.16 ± 0.22	1.30 ± 0.28
C22:0	0.96 ± 0.39	0.33 ± 0.21	0.40 ± 0.25	0.43 ± 0.21	0.25 ± 0.11	0.26 ± 0.13	0.33 ± 0.24

A meeting was held with the screening panelists (25 people) to explain the aim of the study, its importance, scope and level of commitment. Subsequently, detection tests were carried out to familiarize panelists with characteristic descriptors of taste, using solutions with concentrations of 12 g.L⁻¹ sucrose (sweet), 0.27 g.L⁻¹ of caffeine (bitter), 0.6 g.L⁻¹ citric acid (acidity), 2 g.L⁻¹ sodium chloride (salty) and 0.5 g.L⁻¹ tannic acid for the astringent attribute (International Organization for Standards, 1993). Stimulus perception tests were performed by each of the panelists, and a candidate was considered acceptable when correct in 100% of the tests. Subsequently, ordering tests were performed using the concentrations shown in Table 2. Panelists were trained in sessions of 2 h per week for a period of 22 weeks by ordering tests. For variables of taste, it was verified that the panelists had the ability to detect differences and order concentrations properly. Panelists were presented with a series of four concentrations for each stimulus in a circular order and requested to order them from lowest to highest concentration. During the first sessions standard solutions (Table 2) were used, and later cocoa bean samples were presented to familiarize the panelists with their taste.

Similarly, the selected panelists were trained to associate descriptions of specific odors for cocoa beans, either by previous experience or references provided (cocoa powder, cocoa paste, dried and roasted cocoa beans) so that all panelists agreed on the same sensory language. Nine odor descriptors were considered: chocolate, nutty, hazelnut, acidity, sweet, spicy, roasted, and the auxiliary descriptors associated with a defective sample as musty odor and off-odor. The data generated by panelist and attribute (or stimulus) were subjected to analysis of variance and eight panelists (5 women and 3 men) were chosen on the basis of their discriminative capacity ($p < 0.30$) and repeatability ($p > 0.05$) as suggested by Melo et al. (2009).

2.4 Sample preparation and sensory evaluation

Prior to the evaluation, 300 g of beans from each group (composite sample) were roasted on an electric hot plate (150 °C for 15 min) and the rind was removed at the end of the process.

Table 2. Materials and concentrations of aqueous solutions used in training tests for panelists (International Organization for Standards, 1993).

Taste	Material	Concentration (g.L ⁻¹)
Sweet	Sucrose	0.025, 0.25, 1.25, 2.5
Astringency	Tannic acid	0.05, 0.08, 0.125, 0.25
Acidity	Tartaric acid	0.025, 0.037, 0.055, 0.085
Bitter	Caffeine	0.037, 0.055, 0.085, 0.127

Table 3. Means value of descriptors of taste obtained by trained panelists from seven cocoa samples (representing 45 samples).

Attribute	G1	G2	G3	G4	G5	G6	G7
Sweet	1.19 ^{abcd}	1.43 ^{bc}	1.70 ^b	0.75 ^{cd}	0.39 ^d	1.08 ^{bcd}	2.83 ^a
Bitter	4.19 ^{ab}	3.34 ^{bc}	3.07 ^c	4.66 ^a	3.82 ^{abc}	3.76 ^{abc}	3.23 ^{bc}
Acidity	2.32 ^{ab}	3.09 ^a	2.90 ^a	2.87 ^a	1.59 ^{bc}	1.37 ^c	1.23 ^c
Astringent	0.96 ^b	1.75 ^{ab}	1.21 ^b	1.93 ^{ab}	0.98 ^b	2.28 ^a	1.42 ^{ab}

Values with the same letter by row do not differ significantly ($p > 0.05$). Values were obtained in 9 cm unstructured line.

0.5 g of roasted beans were placed in plastic containers of 15 cm³. A total of 168 containers for odor testing and 168 for taste were prepared. The containers were sealed and left for 6 h at a temperature of 28 °C to promote the accumulation of volatiles. The final evaluation was conducted in a single session in a spacious, enclosed area with adequate lighting and a temperature of 25 °C. Containers coded with three digits were presented to panelists. First, odor testing was performed, for which a series of seven containers arranged in a circle were presented to the panelist. Each of the panelists was asked to remove the lid of the sealed container, inhale the volatiles and cast their judgment; to this end each panelist was given seven pages (one for each sample) in which they were asked to mark on an unstructured linear scale of 90 mm the stimulus perceived for each of the attributes (one line per attribute). This procedure was conducted two more times (in triplicate) using codes of three different digits (21 different combinations) for each sample and repetition. Similarly, panelists were asked to evaluate the samples for the four predefined attributes of taste (Lawless & Heymann, 1999). For this evaluation, the panelists were asked to rinse their mouth with water between samples. When required, the panelists used tea bags to desaturate their olfactory system or flavourless biscuits to desaturate taste.

2.5 Data analysis

Distances were measured at the origin of the mark where the panelist evaluated the attribute and the corresponding values were recorded for each of the descriptors (cm). Considering the data as continuous variables, an analysis of variance was performed. The means were classed by the Newman-Keuls test at 5% limit (Assemat et al., 2005).

With data from all descriptors (a total of 13), a principal component analysis (PCA) was carried out, and with the graphic representation of the components that most explained the variation, the correlation between descriptors and samples was established (Thamke et al., 2009). All statistical analyses were carried out using XLSTAT Software (version 2013.2.07 Addinsoft® SARL).

3 Results and discussion

For the four attributes of taste, a significant difference was found in at least one sample (Table 3). It may be noted that the attributes of acidity and bitterness presented higher values than sweetness. The lowest value of sweetness (0.39) was found for sample G5. It has been reported that the compounds present in cocoa beans contributing to sweet taste (monosaccharides, disaccharides, oligosaccharides and some L-amino acids) are in a concentration lower than the perception threshold of a panelist (Stark et al., 2006); this may indicate that the samples analyzed

could have had a concentration above the threshold mainly of sucrose, which is reported to be the sugar contributing the most to this attribute. Regarding the bitter taste, the G4 sample had the highest score and was significantly different ($p < 0.05$) from other samples. The bitter taste of cocoa beans is mainly due to the alkaloids present, although not exclusively, since molecules such as diketopiperazines, free L-amino acids or peptides also contribute to the perception of bitterness, as proven by Stark et al. (2006). Other molecules that may also contribute to the perception of bitterness, not directly but rather as a "bitter-astringent" sensation, are low molecular weight tannins such as epicatechins, catechins and procyanidins, and that together as total polyphenols, are in high concentrations in G4 (Table 1). However, the PCA (Figure 1) indicated that the astringent and bitter taste were not correlated and conversely a stronger association was observed between the acid and astringent tastes. This may explain why the results of this study do not coincide with the order reported by Stark et al. (2006) in samples of cocoa (bitter > astringent > sour), or perhaps something similar to that proposed happens with sweet taste, i.e. the contents of organic acids (lactic, citric, succinic, acetic, malic) are found in higher amounts than in cocoa at other latitudes and have a determining impact on this attribute. This has already been reported for cocoa in several regions of the world, where panelists have given scores for acidity even higher than those of bitterness (Jinap et al., 1995).

Figure 2 shows the odor profile of each sample evaluated by the panel. Significant differences among samples ($p < 0.05$) were found for the nutty, sweet, musty and off-odor attributes. The highest score of chocolate odor (2.48) was found in the G1 sample, but since there was no difference between samples ($p > 0.05$) only in the G3, G5 and G6 samples was a correlation between chocolate and hazelnut odors found by the PCA analysis (Figure 1). Some aldehydes, ketones and pyrazines which are formed during fermentation and roasting partially contribute to chocolate odor (Rodríguez-Campos et al., 2011).

In general terms, the highest scores were found for the sweet and nutty odors. Particularly for the sweet odor, the G7 sample had a score of 3.39. According to Afoakwa et al. (2009) this odor is associated with molecules such as alcohols, ketones, aldehydes, esters and pyrazines, which are highly desirable in cocoa beans, as they contribute to the high quality of chocolates. These compounds are formed from precursor molecules during fermentation and are further enhanced in roasting (Frauendorfer & Schieberle, 2008). The cocoas from the Mazatán location (code MAMG, Table 1) were found to have the highest values of sweet odor which may be related to the fruit pulp content, since upon opening the pods it was noted that the samples from this region contained a remarkably higher content of pulp surrounding the beans. The high content of pulp probably increased the sugar content that can be transformed to molecules which contribute to the sweet odor or serve as a cultivation means that may enhance the multiplication of microorganisms responsible for making those changes (Quesnel, 1965).

For attributes where a significant difference was found, the G2 sample showed the highest values for musty (1.89) and off-odor (2.48). The same sample was correlated with the acidity, acidity

odor, astringency, roasted odor and off odor attributes (Figure 1). The presence of undesirable odors such as musty (1.89) and off-odor are indicators of inappropriate post-harvest handling, in this sense the amount of mucilage in pods may have determined

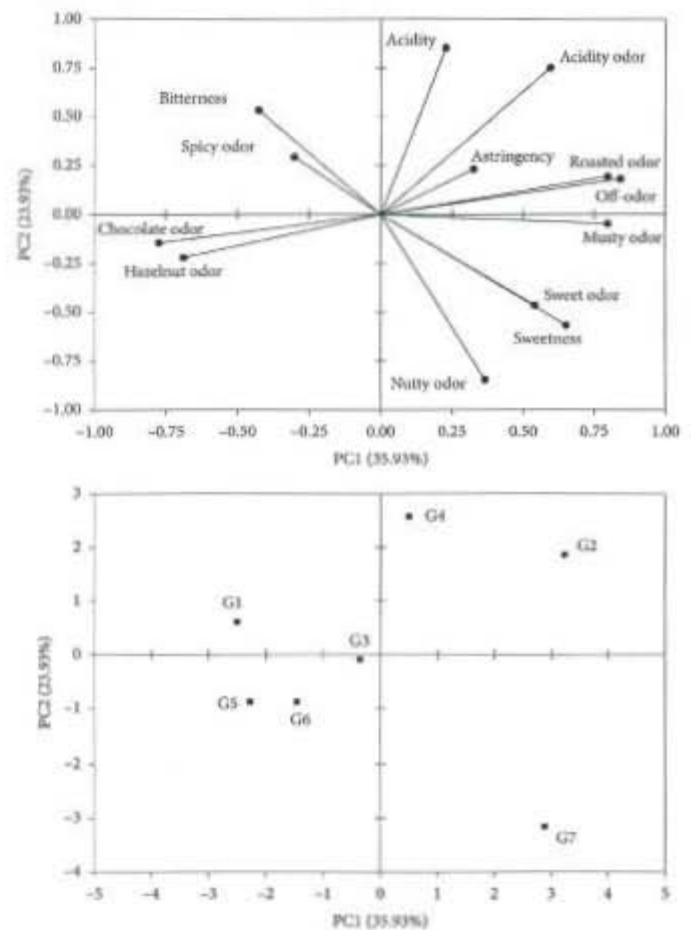


Figure 1. Principal component analysis (PCA) loadings for descriptors (up) and cocoa samples (down).

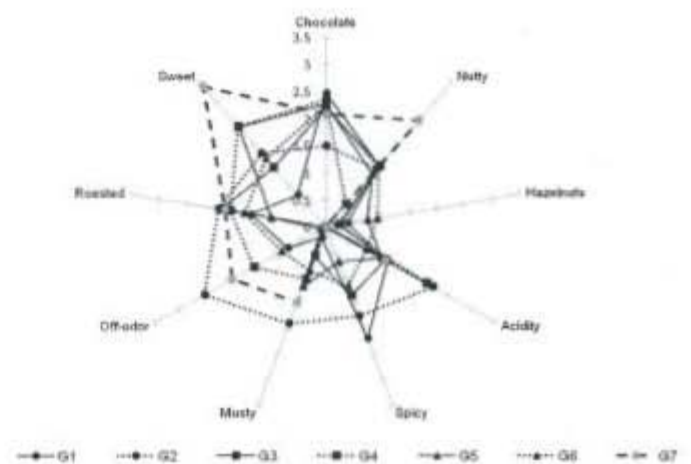


Figure 2. Attributes of odor described for seven samples of fermented cocoa obtained from plants cultivated in the southern region of Mexico (representing 45 samples).

the presence of these odors. García-Alamilla et al. (2007) have reported that upon the exhaustion of the sugars and water from the fermentation medium (basically the mucilage content in the sample), the formation of acetic, propionic, isobutyric and isovaleric acids begin. These acids are associated with the occurrence of undesired odors such as rancidity, fat, and cheese in cocoa, which were detected and described as off-odor

The presence of the spicy descriptor detected by panelists may be related to the presence of molecules as dihydrocapsaicin, norhydrocapsaicin, homocapsaicin, carotenoids, flavonols, all of phenolic nature (Serra-Bonvehí, 2005). The content of phenolic compounds was significant in the analyzed samples (Table 1). The strong correlation found between bitterness and spicy odors also supports this hypothesis since the contribution of polyphenols to bitterness was inferred before (Figure 1). The use of the spicy descriptor may also be associated with local terminology, since the region has chili of the genus *Capsicum* sp. called "chile chocolate" in Spanish and panelists accustomed to this smell created the descriptor and associated it with the powder of this chili.

Unlike other studies (Sukha & Butler, 2005) which set out descriptors of flower and fruit aromas, in this study the panelists did not perceive, and therefore did not propose, such descriptors. Conversely, predominant in all samples were the chocolate, spicy, hazelnut, acidity and roasting odors, which may be used as the characteristics that identify cocoas from this part of Mexico (Soconusco). Only for some samples was a positive correlation found between the content of polyphenols and the astringent taste, i.e. samples where the polyphenol content and the scores for the astringent taste were higher. The same happened with bitter taste. This is contrary to what has been reported (Crozier et al., 2011), since positive correlations have been found between polyphenol content and astringent taste. This lack of total correlation may be explained by the fact that in some samples, roasting promoted greater oxidation of phenolic compounds (quantified in unfermented seeds), consequently greater formation of quinones derived from some polyphenols such as catechins, epicatechins, and condensation of proteins, resulting in reduced astringency and bitter taste (Forsyth & Quesnel, 1957).

Based on the results, the samples considered with higher sensory quality are G7 with lower acidity (1.23), higher sweet taste (2.83) and sweet odor (3.39) and less bitter taste (3.23); the G3 sample where undesirable odors were absent and which had low astringent taste (1.21); and the G1 sample showed higher chocolate (2.48), chili (2.18) and roasted (1.35) odors. However, only the first two samples (G7 and G3) were found to be correlated (Figure 1) with desirable descriptors in quality samples (sweetness, chocolate, nuts, hazelnuts), together with the G5 and G6 samples. At the other extreme, the samples found with low sensory quality were G2 and G4, which despite having some desirable odors, in general they had high values of acidity, off-odor and musty. From the above it can be established that the sensory quality of cocoas may be defined as G7 > G5 > G6 > G3 > G1 > G4 > G2.

The results indicate a low contribution from the environment, since samples with similar codes came from the same collection site but were sensorially different. The results from the chemical

composition analyses of these beans performed in our group support this hypothesis. Since the post-harvest processing was the same for all samples, the only variable that explains the differentiation is allelic composition. It has been demonstrated that cocoas included in this study have little Criollo ancestry, i.e. a group not genetically corresponding to the old Criollo cultivated by the Mayas in Mesoamerica (Vázquez-Ovando et al., 2014). Possibly we are dealing with a genetic group in a process of adapting to the conditions of Soconusco that has boosted its own odor and taste attributes, which as noted earlier, are of interest for international markets.

4 Conclusions

The cocoas of the region of Soconusco, Mexico have sensory characteristics unique to and different from cocoa cultivated in other world regions. No descriptors were identified that associated with flower and fruit odor as in other cacao types. We achieved differentiated samples (groups) of high and low sensory quality. The highest sensory quality cocoas were associated with descriptors of sweet taste, less bitterness and chocolate and hazelnut odors. Conversely, the low quality cocoas were associated with bitter taste and off-odors. The spicy-odor descriptor is reported for the first time for samples of cocoa beans.

Acknowledgments

The first author extends thanks to the Posgrado en Ciencias Biológicas of the Universidad Nacional Autónoma de México (UNAM), for the academic instruction in doctoral studies. To the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexico) for a financial grant awarded to the first author. The authors also thank the efforts of the sensory panelists involved in this study.

References

- Afoakwa, E. O., Paterson, A., Fowler, M., & Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: a critical review. *Critical Reviews in Food Science and Nutrition*, 48(9), 840-857. <http://dx.doi.org/10.1080/10408390701719272>. PMID:18788009
- Afoakwa, E., Paterson, A., Fowler, M., & Ryan, A. (2009). Matrix effects on flavour volatiles release in dark chocolates varying in particle size distribution and fat content using GC-mass spectrometry and GC-olfactometry. *Food Chemistry*, 113(1), 208-215. <http://dx.doi.org/10.1016/j.foodchem.2008.07.088>.
- Assemat, S., Lachenaud, P., Ribeyre, E., Davrieux, F., Pradon, J. L., & Cros, E. (2005). Bean quality traits and sensory evaluation of wild Guianan cocoa populations (*Theobroma cacao* L.). *Genetic Resources and Crop Evolution*, 52(7), 911-917. <http://dx.doi.org/10.1007/s10722-003-6117-2>.
- Coe, S. D., & Coe, M. D. (2007). *The true history of chocolate* (2nd. ed.). New York: Thames and Hudson.
- Crown, P. L. (2013). Prehispanic use of cacao. In R. Watson, V. R. Preedy & S. Zibadi (Eds.), *Chocolate in Health and Nutrition* (cap. 1, pp. 3-9). New York: Humana Press. http://dx.doi.org/10.1007/978-1-61779-803-0_1
- Crozier, S. J., Preston, A. G., Hurst, J. W., Payne, M. J., Mann, J., Hainly, L., & Miller, D. L. (2011). Cacao seeds are a "Super Fruit": A comparative analysis of various fruit powders and products.

- Chemistry Central Journal*, 5(1), 5. <http://dx.doi.org/10.1186/1752-153X-5-5>. PMID:21299842
- Cruz, M., Whitkus, R., Gómez-Pompa, A., & Mota-Bravo, L. (1995). Origins of cacao cultivation. *Nature*, 375(6532), 542-543. <http://dx.doi.org/10.1038/375542a0>.
- d'El-Rei, J., Cunha, A. R., Burlá, A., Burlá, M., Olgman, W., Neves, M. F., Viridis, A., & Medeiros, F. (2013). Characterisation of hypertensive patients with improved endothelial function after dark chocolate consumption. *International Journal of Hypertension*, 2013, 985087. <http://dx.doi.org/10.1155/2013/985087>. PMID:23533716.
- Djoussé, L., Hopkins, P. N., North, K. E., Pankow, J. S., Arnett, D. K., & Ellison, R. C. (2011). Chocolate consumption is inversely associated with prevalent coronary heart disease: the National Heart, Lung, and Blood Institute Family Heart Study. *Clinical Nutrition (Edinburgh, Lothian)*, 30(2), 182-187. <http://dx.doi.org/10.1016/j.clnu.2010.08.005>. PMID:20858571
- Engels, J. M. (1983). A systematic description of cacao clones. I. The discriminative value of quantitative characteristics. *Euphytica*, 32(2), 377-385. <http://dx.doi.org/10.1007/BF00021446>.
- Food and Agriculture Organization - FAO. (2013). *FAOSTAT Online Statistical Service*. Retrieved from <http://faostat.fao.org/>.
- Forsyth, W. G. C., & Quesnel, V. C. (1957). Cacao glycosidase and colour changes during fermentation. *Journal of the Science of Food and Agriculture*, 8(9), 505-509. <http://dx.doi.org/10.1002/jsfa.2740080902>.
- Frauentorfer, F., & Schieberle, P. (2008). Changes in key aroma compounds of Criollo cocoa beans during roasting. *Journal of Agricultural and Food Chemistry*, 56(21), 10244-10251. <http://dx.doi.org/10.1021/jf802098f>. PMID:18925740
- García-Alamilla, P., Salgado-Cervantes, M. A., Barel, M., Berthomieu, G., Rodríguez-Jiménes, G. C., & García-Alvarado, M. A. (2007). Moisture, acidity and temperature evolution during cacao drying. *Journal of Food Engineering*, 79(4), 1159-1165. <http://dx.doi.org/10.1016/j.jfoodeng.2006.04.005>.
- International Organization for Standards - ISO. (1993). *International Standard 8586-1: sensory analysis: general guidance for the selection, training and monitoring of assessors*. Geneva.
- Jinap, S., Dimick, P. S., & Hollender, R. (1995). Flavour evaluation of chocolate formulated from cocoa beans from different countries. *Food Control*, 6(2), 105-110. [http://dx.doi.org/10.1016/0956-7135\(95\)98914-M](http://dx.doi.org/10.1016/0956-7135(95)98914-M).
- Lawless, H. T., & Heymann, H. (1999). *Sensory evaluation of food-principles and practices*. Maryland: Aspen Publisher.
- Macht, M., & Dettmer, D. (2006). Everyday mood and emotions after eating a chocolate bar or an apple. *Appetite*, 46(3), 332-336. <http://dx.doi.org/10.1016/j.appet.2006.01.014>. PMID:16546294
- Melo, L. L. M. M., Bolini, H. M. A., & Efraim, P. (2009). Sensory profile, acceptability, and their relationship for diabetic/reduced calorie chocolates. *Food Quality and Preference*, 20(2), 138-143. <http://dx.doi.org/10.1016/j.foodqual.2008.09.001>.
- Motilal, L., Zhang, D., Umaharan, P., Mischke, B. S., Mooledhar, V., & Meinhardt, L. W. (2010). The relic Criollo cacao in Belize genetic diversity and relationship with Trinitario and other cacao clones held in the International cocoa Genebank, Trinidad. *Plant Genetic Resources*, 8(02), 106-115. <http://dx.doi.org/10.1017/S1479262109990232>.
- Quesnel, V. (1965). Agents inducing the death of the cacao seeds during fermentation. *Journal of the Science of Food and Agriculture*, 16(8), 441-447. <http://dx.doi.org/10.1002/jsfa.2740160804>.
- Rodríguez-Campos, J., Escalona-Buendía, H., Orozco-Avila, L., Lugo-Cervantes, E., & Jaramillo-Flores, M. (2011). Dynamics of volatile and non-volatile compounds in cacao (*Theobroma cacao* L.) during fermentation and drying processes using principal components analysis. *Food Research International*, 44(1), 250-258. <http://dx.doi.org/10.1016/j.foodres.2010.10.028>.
- Serra-Bonvehí, J. (2005). Investigation of aromatic compounds in roasted cocoa powder. *European Food Research and Technology*, 221(1-2), 19-29. <http://dx.doi.org/10.1007/s00217-005-1147-y>.
- Smith, N. (1999). *The Amazon River Forest: a natural history of plants, animals, and people*. New York: Oxford University Press.
- Stark, T., Bareuther, S., & Hofmann, T. (2006). Molecular definition of the taste of roasted cocoa nibs (*Theobroma cacao*) by means of quantitative studies and sensory experiments. *Journal of Agricultural and Food Chemistry*, 54(15), 5530-5539. <http://dx.doi.org/10.1021/jf0608726>. PMID:16848542
- Sukha, D. A., Butler, D. R., Umaharan, P., & Boulton, E. (2008). The use of an optimised organoleptic assessment protocol to describe and quantify different flavour attributes of cocoa liquors made from Ghana and Trinitario beans. *European Food Research and Technology*, 226(3), 405-413. <http://dx.doi.org/10.1007/s00217-006-0551-2>.
- Sukha, D., & Butler, D. (2005). The CFC/ICCO/INIAP Cocoa Flavour Project - Investigating the spectrum of fine flavour within genotypes and between origins. *INGENIC Newsletter*, 10, 22-25.
- Thamke, I., Dürschmid, K., & Rohma, H. (2009). Sensory description of dark chocolates by consumers. *LWT - Food Science and Technology*, 42, 534-539. <http://dx.doi.org/10.1016/j.lwt.2008.07.006>.
- Trognitz, B., Scheldeman, X., Hansel-Hohl, K., Kuant, A., Grebe, H., & Hermann, M. (2011). Genetic population structure of cacao plantings within a young production area in Nicaragua. *PLoS ONE*, 6(1), e16056. <http://dx.doi.org/10.1371/journal.pone.0016056>. PMID:21264251
- Vázquez-Ovando, J. A., Molina-Freaner, F., Nuñez-Farfán, J., Betancur-Ancona, D., & Salvador-Figueroa, M. (2015). Classification of cacao beans (*Theobroma cacao* L.) of southern Mexico based on chemometric analysis with multivariate approach. *European Food Research and Technology*. In press. <http://dx.doi.org/10.1007/s00217-015-2415-0>.
- Vázquez-Ovando, J. A., Molina-Freaner, F., Nuñez-Farfán, J., Ovando-Medina, I., & Salvador-Figueroa, M. (2014). Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico. *Genetics and Molecular Research*, 13(4), 10404-10414. <http://dx.doi.org/10.4238/2014.December.12.2>. PMID:25511024.
- Voight, J. (2013). Chocolate and cocoa aroma. In R. Watson, V. R. Preedy & S. Zibadi (Eds.), *Chocolate in health and nutrition* (cap. 7, pp. 89-101). New York: Humana Press. http://dx.doi.org/10.1007/978-1-61779-803-0_7.

CAPÍTULO 7

De las relaciones entre los marcadores moleculares SSR y características químicas y sensoriales de semillas de cacao

En este capítulo se describe el procedimiento metodológico empleado para correlacionar estadísticamente los fragmentos obtenidos del análisis con los microsatélites (SSR) y las características químicas y sensoriales obtenidas de las semillas de cacao de plantas cultivadas en el Soconusco, Chiapas. Se propone el uso del análisis de regresión lineal múltiple para lograr este objetivo y se discute sobre los fragmentos que después del análisis podrían ser considerado verdaderos marcadores de la calidad química y sensorial de cacao.

Con los resultados de este apartado se preparó una nota científica que será sometida para ser considerada a publicación. El título tentativo en español es “Asociación entre marcadores moleculares SSR y características químicas y sensoriales de muestras de cacao cultivado en Soconusco, Chiapas; México”.

Asociación entre marcadores moleculares SSR y características químicas y sensoriales de muestras de cacao cultivado en Soconusco, Chiapas; México

Introducción

Theobroma cacao, conocida mundialmente como cacao, es una planta domesticada y cultivada en Mesoamérica desde hace 4000 años (De la Cruz et al., 1995; Henderson et al., 2007). Las semillas que se obtienen de los frutos se emplean en alimentación como fuente de lípidos y por sus propiedades estimulantes y terapéuticas (Dillinger et al., 2000). Después de la llegada de los españoles al continente Americano, el chocolate, principal producto obtenido del cacao, se popularizó primero en Europa y luego en el resto del mundo (Monteiro et al., 2009). Actualmente esta industria genera divisas por más de 73,000 millones de dólares al año, lo que convierte a *T. cacao* en la especie de mayor importancia comercial dentro del género *Theobroma* (ICCO, 2012).

Aunque durante todo el tiempo que se ha cultivado cacao se ha realizado mejoramiento genético, fue hasta la segunda mitad del siglo veinte cuando dio inicio los mayores esfuerzos por obtener características deseables en los árboles, los frutos y consecuentemente en las semillas. Árboles de mayor porte, productividad, tolerancia a plagas y enfermedades y la calidad de frutos y semillas fueron las principales características objetivo (Monteiro et al., 2009). Las técnicas de mejoramiento tradicional que fueron muy útiles, poco a poco se han ido sustituyendo por otras apoyadas en marcadores moleculares. Así, los análisis basados en mapeo enfocados en la detección de *loci* ligados a caracteres cuantitativos (QTL) son una herramienta que se ha popularizado para muchas especies vegetales. Sin embargo son altamente dependientes de costos además de laboriosas. En especies vegetales de rápido desarrollo son una opción muy valiosa, sin embargo para especies arbóreas y de desarrollo más lento suponen varios años de trabajo continuado. Para *T. cacao*, varios grupos de trabajo han realizado mapeos y localizado varios candidatos QTL's. El enfoque fundamental ha estado orientado hacia la localización de regiones asociados con caracteres de interés agronómico (Cervantes-Martínez et al., 2006; Clement et al., 2003; Feltus et al., 2011) o con la resistencia a enfermedades (Brown et al., 2007; Lanaud et al., 2009; Pugh et al., 2004). La búsqueda de marcadores asociados con características fenotípicas implicadas

en la calidad sensorial ha sido limitada (Vázquez-Ovando et al., 2012). La complejidad de estos caracteres (aroma, sabor, textura) puede ser la principal razón.

Se ha demostrado que las características sensoriales del cacao guardan relación o son parcialmente explicadas por la composición química de las semillas, por lo cual puede ser útil localizar marcadores moleculares asociados con las características químicas o sensoriales. En este sentido, los estudios de asociación entre marcadores moleculares y caracteres fenotípicos empleando herramientas estadísticas como el análisis de regresión múltiple (Virk et al. 1996; Khadivi-Khub, 2014; Khadivi-Khub et al., 2014), puede resultar un modelo valioso de análisis de datos como criterio previo para la selección de regiones genómicas candidatas a ser evaluados en la búsqueda de QTL's; para describir modelos matemáticos con fines predictivos de la composición o calidad; o como verdaderos marcadores. Por lo anterior el objetivo del presente estudio fue buscar asociación de bandas generadas con marcadores moleculares SSR y algunas características químicas y sensoriales de semillas de cacao mediante el empleo de análisis de regresión múltiple.

Materiales y métodos

Obtención de datos

Se emplearon los datos de fragmentos generadas de la caracterización molecular de árboles de cacao cultivados en la región Soconusco con 12 marcadores microsatélites (mTcCIR1, mTcCIR3, mTcCIR6, mTcCIR7, mTcCIR8, mTcCIR11, mTcCIR12, mTcCIR15, mTcCIR19, mTcCIR21, mTcCIR25, mTcCIR28) obtenidos por Vázquez-Ovando et al. (2014). Se obtuvieron los datos de la caracterización química de semillas de cacao pertenecientes a las mismas muestras (Vázquez-Ovando et al., 2015a); para este análisis se consideraron los contenidos de humedad, cenizas, grasa, proteína, ácidos grasos (C14:0, C16:0, C16:1, C18:0, C18:1, C18:2, C20:0, C22:0), compuestos polifenólicos y capacidad antioxidante. Finalmente se obtuvieron los datos numéricos de caracteres sensoriales reportados por Vázquez-Ovando et al. (2015b) obtenidos por panelistas entrenados, que incluye olores a chocolate, nueces, avellanas, ácido, dulce, picante, tostado, humedad y mal-olor así como sabores amargo, ácido, astringente y dulce. Para el análisis de asociación marcador molecular-característica química se incluyeron datos de

44 muestras. Así mismo, para los caracteres sensoriales se incluyó la caracterización de siete muestras.

Análisis de asociación

Basado en el procedimiento sugerido por Virk et al. (1996), se construyó una matriz binaria de presencia (1), ausencia (0) de cada tamaño de alelo revelado por todos los *loci* SSR para 44 muestras. La matriz fue complementada con los datos de cada una de las características químicas o sensoriales. Los datos moleculares fueron considerados como variables independientes (cuantitativos) y los caracteres químicos o sensoriales como variables dependientes (cuantitativos). Se realizó un análisis de regresión múltiple (ARM) empleando los métodos de “análisis de regresión lineal” con la opción “paso a paso” con el software XLStat Pro v 2013. Este análisis está basado en el modelo $Y = a + b_1m_1 + b_2m_2 + \dots + b_jm_j + \dots + b_nm_n + d + e$, que relaciona la variación en la variable dependiente (Y o carácter cuantitativo) como una función lineal del conjunto de variables (m_j) independientes, representados por los marcadores SSR. Los términos b_j son los coeficientes de regresión parcial que especifican las relaciones empíricas entre Y y m_j , d representa los valores residuales entre accesiones después de la regresión y e es el error aleatorio de Y que incluye la variación ambiental (Khadivi-Khub, 2014; Virk et al., 1996). Para seleccionar las variables independientes de la ecuación de regresión, valores de F con 0.045 y 0.099 de probabilidades se utilizaron como criterios de “entrada” y “retirada”, respectivamente (Roy y Bargmann 1957; Affifi y Clark 1984). Los valores de R^2 denotan los coeficientes de correlación múltiple. Adicionalmente, los marcadores seleccionados fueron probados de forma independiente mediante su ajuste a una curva lineal usando modelos lineales para confirmar la significancia de los estadísticos β para cada banda identificada por el análisis. Beta se define como el coeficiente de regresión estandarizado $= BS_x/S_y$, donde B es el coeficiente de regresión o pendiente y S_x y S_y son las desviaciones estándar de las variables independientes (x) y dependientes (y) (Affifi y Clark, 1984; Kar et al., 2008). Se realizó la prueba t de Student para evaluar la significancia entre los valores estimados de caracteres fenotípicos de cada accesión donde los marcadores estuvieron presentes y ausentes. Solo aquellos marcadores que mostraron valores de regresión significativos ($p < 0.05$) fueron considerados como asociados con el carácter en cuestión.

Resultados y Discusión

Para todos los caracteres químicos evaluados se encontró asociación significativa con al menos dos bandas siendo todos los modelos significativos (valores $p < 0.05$, análisis de la varianza). Así mismo, se encontraron valores de R^2 (proporción de variación fenotípica) acumulados en el intervalo de 0.204 (ácido graso linoleico) a 0.849 (ácido graso eicosanoico). En el Cuadro 1 se muestran solo los caracteres fenotípicos cuyo valor de R^2 acumulado fue > 0.5 . En total, 38 bandas se asociaron significativamente (positiva o negativamente) con los 13 caracteres fenotípicos y 14 bandas tuvieron asociación con más de un carácter (Cuadro 1). La banda con mayor frecuencia fue mTcCIR12₁₈₈ la cual se encontró asociada con los contenidos de ácido eicosanoico, humedad, grasa y polifenoles. Tres de estos caracteres, también se encontraron asociados con el marcador mTcCIR15₂₄₈. Lo anterior podría ser indicio de que estos rasgos se encuentran también asociados entre ellos. Se ha sugerido que los marcadores que encuentran asociación con más de un carácter pueden resultar altamente útiles en programas de mejoramiento (Khadivi-Khub, 2014).

Como se precisó antes, otra utilidad de estos datos así como de todos los marcadores asociados con los rasgos fenotípicos evaluados en cacao, es con enfoque predictivo. En ese sentido, todos los caracteres fenotípicos evaluados presentaron ecuaciones con valores de ajuste $R^2 > 0.5$, (Cuadro 2) denotando ajustes adecuados al modelo. En general, entre más marcadores estuvieron asociados con cada uno de los caracteres el coeficiente de determinación (R^2) fue mayor. A pesar de que todos los caracteres encontraron un modelo que resultó significativo ($p < 0.05$), solo algunos caracteres muestran una distribución de datos predichos frente a datos observados congruente con el ajuste (Figura 1) y coincide con los valores R^2 más altos. Así los ácidos palmítico, esteárico y eicosanoico resultan los caracteres predictores más adecuados.

Cuadro 1. Marcadores microsatélites asociados con caracteres de la composición química y descriptores sensoriales de semillas de cacao y sus respectivos coeficientes

Carácter	Marcador SSR (alelo)	<i>r</i>	<i>R</i> ²	Coefficiente estandarizado beta	Valor <i>t</i>	Valor <i>p</i>
Ácido mirístico	mTcCIR11 ₃₇₀	0.684	0.468	0.640	6.547	<0.0001
	mTcCIR15 ₂₂₈	0.749	0.561	0.256	2.533	0.015
	mTcCIR06 ₃₃₅	0.780	0.609	0.253	2.583	0.014
Ácido palmitoleico	mTcCIR06 ₂₀₉	0.812	0.659	0.224	2.372	0.023
	mTcCIR28 ₃₃₀	0.417	0.174	-0.414	-3.490	0.001
	mTcCIR06 ₂₆₁	0.557	0.310	-0.265	-2.215	0.033
	mTcCIR06 ₂₅₁	0.636	0.404	0.342	2.966	0.005
Ácido palmítico	mTcCIR19 ₃₂₈	0.709	0.502	0.323	2.774	0.008
	mTcCIR19 ₂₁₄	0.479	0.229	-0.534	-5.056	<0.0001
	mTcCIR06 ₂₃₇	0.568	0.323	-0.389	-3.744	0.001
	mTcCIR19 ₃₁₀	0.653	0.426	0.249	2.514	0.017
	mTcCIR11 ₃₇₀	0.727	0.528	-0.393	-4.005	0.000
	mTcCIR15 ₂₄₈	0.766	0.586	-0.334	-3.203	0.003
	mTcCIR06 ₂₄₇	0.796	0.633	0.350	3.007	0.005
Ácido esteárico	mTcCIR28 ₃₃₈	0.823	0.677	-0.244	-2.198	0.034
	mTcCIR19 ₂₁₄	0.480	0.230	0.475	4.824	<0.0001
	mTcCIR11 ₃₇₀	0.584	0.341	0.597	5.556	<0.0001
	mTcCIR06 ₂₃₇	0.690	0.476	0.486	4.852	<0.0001
	mTcCIR15 ₂₂₈	0.756	0.572	-0.445	-4.125	0.000
	mTcCIR07 ₁₅₀	0.786	0.618	-0.257	-2.495	0.017
	mTcCIR28 ₃₅₁	0.817	0.668	-0.236	-2.348	0.024
Ácido eicosanoico	mTcCIR11 ₃₇₀	0.583	0.340	0.273	3.461	0.002
	mTcCIR15 ₂₄₈	0.658	0.433	0.340	4.289	0.000
	mTcCIR25 ₁₇₅	0.733	0.537	-0.587	-7.655	<0.0001
	mTcCIR07 ₁₇₂	0.771	0.594	-0.544	-6.536	<0.0001
	mTcCIR28 ₃₅₁	0.810	0.656	-0.357	-4.564	<0.0001
	mTcCIR19 ₂₆₈	0.838	0.702	0.220	2.669	0.12
	mTcCIR12 ₁₈₈	0.862	0.743	-0.442	-5.163	<0.0001
	mTcCIR08 ₃₁₉	0.887	0.787	-0.434	-5.082	<0.0001
	mTcCIR03 ₂₂₇	0.908	0.825	-0.223	-2.949	0.006
	mTcCIR07 ₁₄₂	0.921	0.849	-0.207	-2.707	0.011
Contenido de grasa	mTcCIR03 ₂₀₉	0.395	0.156	0.495	4.359	<0.0001
	mTcCIR12 ₁₈₈	0.544	0.296	0.443	3.979	0.000
	mTcCIR19 ₂₈₆	0.638	0.407	-0.431	-3.712	0.001
	mTcCIR07 ₁₅₀	0.694	0.482	-0.268	-2.314	0.026
	mTcCIR19 ₃₁₀	0.737	0.543	-0.253	-2.260	0.030
Contenido de proteína	mTcCIR19 ₂₆₈	0.401	0.161	-0.436	-3.835	0.000
	mTcCIR06 ₂₅₁	0.550	0.302	0.449	3.879	0.000
	mTcCIR15 ₂₆₆	0.630	0.397	-0.337	-2.931	0.006
	mTcCIR15 ₂₄₈	0.678	0.460	-0.267	-2.345	0.024
	mTcCIR11 ₃₀₄	0.718	0.516	0.240	2.086	0.044

Cuadro 1. Continuación

Carácter	Marcador SSR (alelo)	<i>r</i>	<i>R</i> ²	Coefficiente estandarizado <i>beta</i>	Valor <i>t</i>	Valor <i>p</i>
Amargor	mTcCIR28 ₃₆₂	0.844	0.713	-0.651	---	---
	mTcCIR12 ₁₈₈	0.961	0.924	-0.494	---	---
	mTcCIR06 ₂₃₇	0.992	0.985	0.181	---	---
	mTcCIR15 ₂₄₈	0.999	0.999	-0.151	---	---
Acidez	mTcCIR19 ₂₆₈	1.000	1.000	0.057	---	---
	mTcCIR11 ₂₉₂	0.946	0.895	-1.015	---	---
	mTcCIR11 ₃₀₄	0.987	0.975	-0.354	---	---
	mTcCIR19 ₂₆₈	0.998	0.997	0.129	---	---
	mTcCIR06 ₂₅₁	1.000	1.000	-0.081	---	---
Olor a nuez	mTcCIR28 ₃₅₁	1.000	1.000	-0.020	---	---
	mTcCIR01 ₁₄₇	0.831	0.690	0.716	29.198	<0.0001
	mTcCIR12 ₁₈₈	0.993	0.987	0.0575	23.437	0.000
Olor ácido	mTcCIR15 ₂₂₈	0.999	0.998	-0.112	-4.583	0.020
	mTcCIR12 ₁₆₈	0.673	0.453	0.731	16.175	0.001
	mTcCIR12 ₁₈₈	0.958	0.918	-0.635	-14.040	0.001
Olor a moho	mTcCIR19 ₃₀₀	0.997	0.994	-0.286	-6.322	0.008
	mTcCIR12 ₁₆₈	0.682	0.465	0.870	7.510	0.005
	mTcCIR25 ₁₃₇	0.922	0.851	-0.807	-7.129	0.006
Olor a tostado	mTcCIR28 ₃₆₂	0.986	0.972	0.407	3.624	0.036
	mTcCIR15 ₂₂₈	0.780	0.608	-0.522	-17.904	0.003
	mTcCIR28 ₃₆₂	0.943	0.889	0.702	24.108	0.002
	mTcCIR12 ₁₈₈	0.990	0.981	-0.401	-13.777	0.005
	mTcCIR15 ₂₄₈	0.999	0.999	0.161	5.534	0.031

Es importante hacer notar que el contenido de grasa es un factor determinante tanto de la variedad como de la calidad del cacao, y en este análisis se encontró una asociación significativa entre este carácter y varios marcadores moleculares. En este sentido los marcadores microsatélites asociados con dicha característica pueden resultar candidatos a ser empleados en programas para localizar cacaos con mayor contenido de grasa o indicadores de la calidad de la manteca, puesto que el contenido de ácido esteárico, determinante de las características químicas de la manteca de cacao, presentó fuerte asociación con seis marcadores moleculares. Este conjunto de seis marcadores puede resultar útil en la predicción de la calidad de los cacaos.

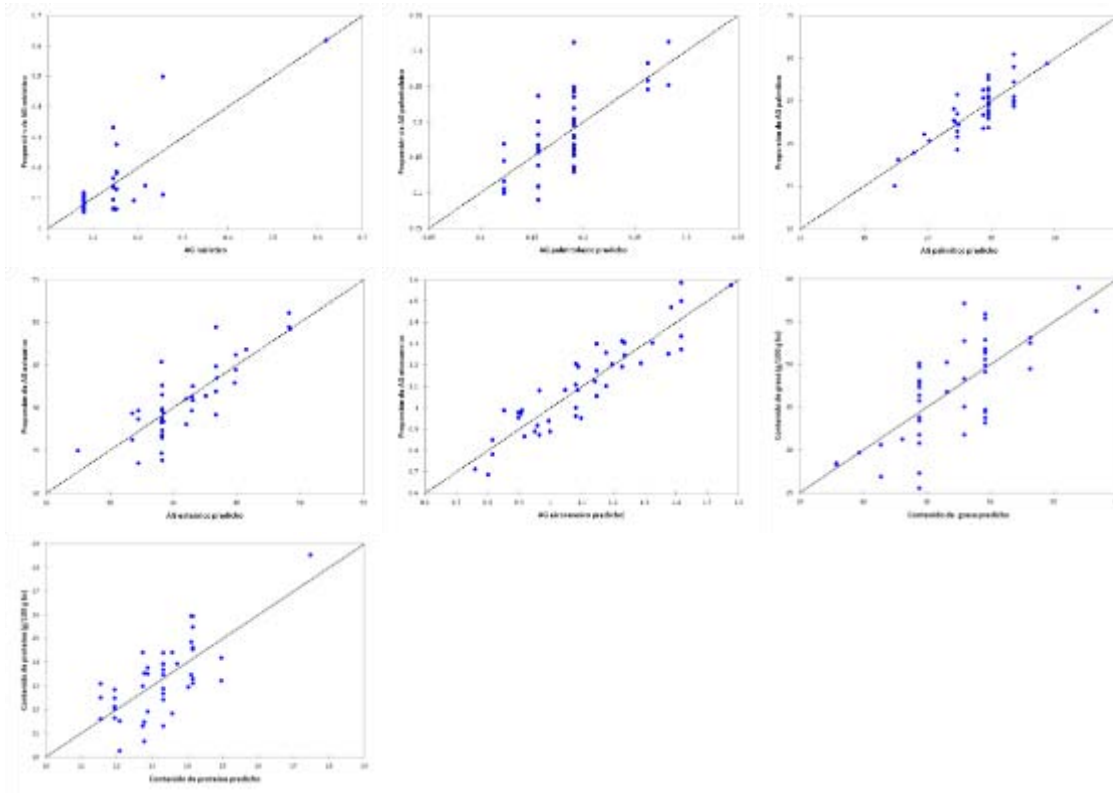


Figura 1. Diagramas de correlación entre los valores observados frente a los predichos para caracteres de la composición química de semillas de cacao. Obtenidos de la correlación-regresión múltiple de datos de 37 fragmentos provenientes de la amplificación de 11 marcadores microsatélites.

Los descriptores sensoriales amargor y acidez estuvieron asociados con cinco fragmentos (cada uno de los descriptores) de ADN amplificados por microsatélites con valores R^2 de 1 (Cuadro 2). Otros cuatro descriptores lo estuvieron con una significancia superior a 0.99. Resulta de interés anotar que en todos los descriptores, con solamente una banda se logra un porcentaje de predicción sobre el carácter sensorial de arriba del 50 % (valores R^2 de 0.5). El fragmento mTcCIR28₃₆₂ se encontró asociado con los caracteres sensoriales amargor, olor a moho y olor a tostado. Los dos primeros caracteres a su vez son indicadores de cacaos más amargos o con notas no deseables en cacaos finos. Dado que todas las muestras fueron fermentadas, secadas y tostadas del mismo modo, en la asociación del fragmento mTcCIR28₃₆₂ con los caracteres podría revelar utilidad de este fragmento como un verdadero marcador molecular de caracteres sensoriales no deseables en cacaos finos. Las representaciones graficas de valores obtenidos frente a

los valores predichos estimados por el modelo se muestran en la Figura 2. A pesar de que el número de muestras analizadas para el perfil sensorial fue de siete (representativas de 45 muestras de cacao que fueron agrupadas en *clusters* en función de su composición química), los modelos predictivos de ajuste resultan en dos descriptores con ajustes “exactos” (amargor y acidez).

La posición física de varios de los marcadores moleculares empleados en este estudio que se asociaron con caracteres químicos o sensoriales puede dar claridad sobre las regiones del cromosoma de cacao que guardan relación intrínseca con caracteres fenotípicos. Si bien es cierto que los SSR amplifican fragmentos muy pequeños, y todos los incluidos en este estudio están básicamente asociados con regiones no codificantes (di y tetranucleotidos), no puede dejarse de lado la posibilidad de que puedan estar cercanos a regiones codificantes. El número de bandas asociadas con ácidos grasos puede reforzar este supuesto. Argout et al. (2011) reportan en cacao la presencia de un número de genes que codifican a enzimas que tienen relación directa con la biosíntesis de ácidos grasos superior al de otras dicotiledóneas oleaginosas. Un análisis *in silico* dirigido puede ayudar a esclarecer estas posibles asociaciones físicas.

Pocos estudios han analizado la asociación de marcadores moleculares con características diferentes al rendimiento agronómico o la tolerancia a enfermedades Araújo et al., 2009 reportan algunos marcadores relacionados con el contenido de grasa y dureza de la manteca de cacao. Dos fragmentos obtenidos con el marcador mTcCIR19 que se localiza en el cromosoma 7 asociados con el contenido de grasa coinciden con la región reportada como marcador QTL en el mismo grupo de ligamiento (Araújo et al., 2009), sin embargo dos fragmentos adicionales revelados por los *loci* mTcCIR03 y mTcCIR07 se localizan en el cromosoma 2 y podrían resultar útiles para explicar este carácter fenotípico. Otros reportes analizan la presencia de fragmentos del ADN relacionados con el color de la mazorca del cacao (Brown et al., 2007; Motamayor et al., 2013) como un indicio de la calidad potencial del producto final. Tales marcadores se localizan en el cromosoma 4 del genoma de cacao. En el presente análisis no se localizó marcadores moleculares asociados con moléculas de naturaleza polifenólica quienes participan activamente en el color de la mazorca, sin embargo en este mismo cromosoma se localizaron fragmentos revelados por el microsatélite mTcCIR12 (mTcCIR12₁₆₈, mTcCIR12₁₈₈) asociados con varias de las características sensoriales. Esta asociación parece tener el mismo sentido en la correlación (Cuadro 2), es decir la presencia de los

fragmentos se relaciona con aumento en los olores deseables (a nuez) y descenso de los no deseables (olores a moho y ácido).

Cuadro 2. Ecuación del modelo predictivo para los rasgos de la composición química y descriptores sensoriales de semillas de cacao

Carácter	Ecuación del modelo	R^2	F	Pr>F
Ácido mirístico	$Y = 0.08 + (0.07 * mTcCIR06_{209}) + (0.11 * mTcCIR06_{335}) + (0.47 * mTcCIR11_{370}) + (0.07 * mTcCIR15_{228})$	0.659	18.817	<0.0001
Ácido palmitoleico	$Y = 0.19 + (0.09 * mTcCIR06_{251}) - (0.03 * mTcCIR06_{261}) + (0.07 * mTcCIR19_{328}) - (0.07 * mTcCIR28_{330})$	0.502	9.827	<0.0001
Ácido palmítico	$Y = 26.77 - (2.42 * mTcCIR06_{237}) + (7.01 * mTcCIR06_{247}) - (7.88 * mTcCIR11_{370}) - (2.03 * mTcCIR15_{248}) - (4.65 * mTcCIR19_{214}) + (5.00 * mTcCIR19_{310}) - (2.30 * mTcCIR28_{338})$	0.677	10.768	<0.0001
Ácido esteárico	$Y = 41.54 + (4.21 * mTcCIR06_{237}) - (2.41 * mTcCIR07_{150}) + (16.71 * mTcCIR11_{370}) - (4.29 * mTcCIR15_{228}) + (5.78 * mTcCIR19_{214}) - (2.35 * mTcCIR15_{228})$	0.668	12.405	<0.0001
Ácido eicosanoico	$Y = 1.27 - (0.14 * mTcCIR03_{227}) - (0.13 * mTcCIR07_{142}) - (0.28 * mTcCIR07_{172}) - (0.27 * mTcCIR08_{319}) + 0.39 * mTcCIR11_{370} - (0.19 * mTcCIR12_{188}) + (0.15 * mTcCIR15_{248}) + (0.10 * mTcCIR19_{268}) - (0.60 * mTcCIR25_{175}) - (0.18 * mTcCIR28_{351})$	0.849	17.386	<0.0001
Grasa	$Y = 47.92 + (13.90 * mTcCIR03_{209}) - (3.52 * mTcCIR07_{150}) + (5.19 * mTcCIR12_{188}) - (6.54 * mTcCIR19_{268}) - (9.92 * mTcCIR19_{310})$	0.543	9.029	<0.0001
Proteína	$Y = 14.15 + (3.32 * mTcCIR06_{251}) + (0.80 * mTcCIR11_{304}) - (0.84 * mTcCIR15_{248}) - (1.24 * mTcCIR15_{266}) - (1.38 * mTcCIR19_{268})$	0.516	8.092	<0.0001
Amargor	$Y = 4.66 + (0.21 * mTcCIR06_{237}) - (0.74 * mTcCIR12_{188}) - (0.16 * mTcCIR15_{248}) + 0.06 * mTcCIR19_{268} - (0.69 * mTcCIR28_{362})$	1	--	--
Acidez	$Y = 2.9 - (0.17 * mTcCIR06_{251}) - (1.5 * mTcCIR11_{292}) - (0.74 * mTcCIR11_{304}) + (0.19 * mTcCIR19_{268}) - (0.03 * mTcCIR28_{351})$	1	--	--
Olor a nuez	$Y = 0.56 + (1.115 * mTcCIR01_{147}) + (0.90 * mTcCIR12_{188}) - (0.18 * mTcCIR15_{228})$	0.998	592.74	0.000
Olor acido	$Y = 2.07 + (0.985 * mTcCIR12_{168}) - (0.86 * mTcCIR12_{188}) - (0.39 * mTcCIR19_{300})$	0.994	173.83	0.001
Olor a moho	$Y = 1.02 + (1.29 * mTcCIR12_{168}) - (0.84 * mTcCIR25_{137}) + (0.43 * mTcCIR28_{362})$	0.972	35.183	0.008
Olor a tostado	$Y = 1.72 - (0.37 * mTcCIR12_{188}) - (0.48 * mTcCIR15_{228}) + (0.10 * mTcCIR15_{248}) + (0.45 * mTcCIR28_{362})$	0.999	432.47	0.002

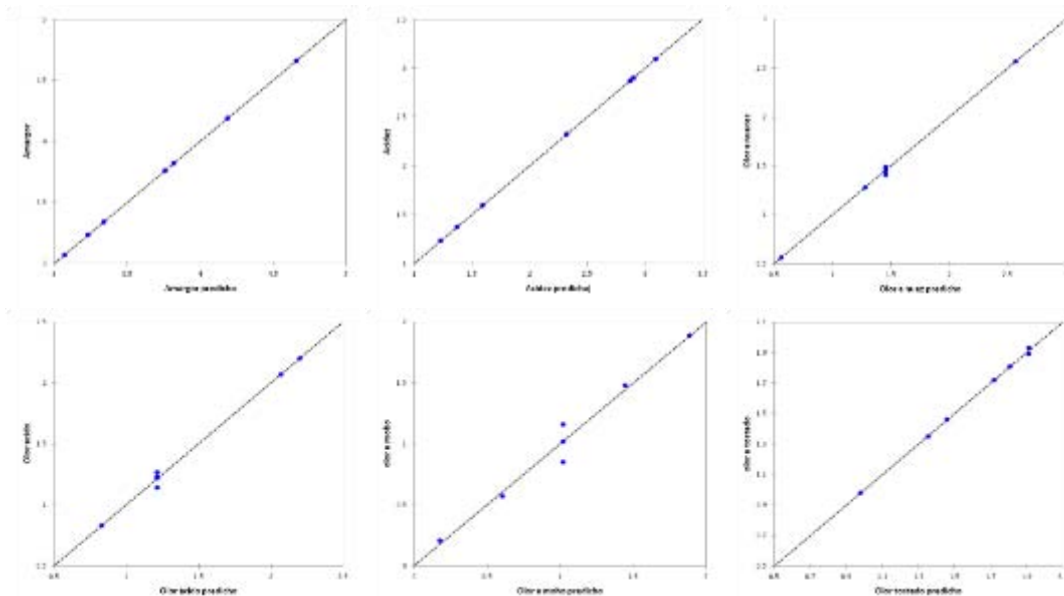


Figura 2. Diagramas de correlación entre los valores observados frente a los predichos para caracteres de sabor y olor evaluados en muestras de semillas de cacao. De izquierda a derecha, arriba: amargor, acidez y olor a nueces; abajo: olor a ácido, mal olor y olor a tostado. Obtenidos de la correlación-regresión múltiple de datos de 37 fragmentos provenientes de la amplificación de 11 marcadores microsatélites.

Los ajustes encontrados en el estudio de asociación revelan la utilidad del análisis de regresión múltiple para correlacionar los fragmentos generados por marcadores moleculares microsatélites y las características fenotípicas ácido mirístico, ácido palmitoleico, ácido palmítico, ácido esteárico, ácido eicosanoico, contenido de grasa, contenido de proteína, amargor, acidez, olor a nuez, olor ácido, olor a moho, olor a tostado. Finalmente, la validación empírica de los modelos propuestos sería la etapa adicional para convertir a los fragmentos en marcadores de calidad fenotípica.

Referencias

- Affifi AA, Clark V (1984) Computer-aided multivariate analysis. Van Nostrand Reinhold, New York.
- Brown J, Phillips-Mora W, Power E, Krol C, Cervantes-Martinez C, Motamayor J, Schnell J (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. *Crop Science* 47:1851-1858.

Araújo IS, de Souza Filho GA, Pereira MG, Faleiro FG, de Queiroz VT, Guimarães CT, Moreira MA, de Barros EG, Machado RCR, Pires JL, Schenell R, Lopes UV (2009) Mapping of Quantitative Trait Loci for butter content and hardness in cocoa beans (*Theobroma cacao* L.) Plant Molecular Biology Reporter 27:177-183.

Argout X, Salse J, Aury J, Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahiri M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C (2011) The genome of *Theobroma cacao*. Nature Genetics 43(2): 101-109.

Brown J, Phillips-Mora W, Power E, Krol C, Cervantes-Martinez C, Motamayor J, Schnell J (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. Crop Science. 47: 1851-1858.

Cervantes-Martinez C, Brown J, Schnell R (2006) Combining ability for disease resistance, yield, and horticultural traits of cacao (*Theobroma cacao* L.) clones. Journal of the American Society for Horticultural Science 131(2):231-241.

Clement D, Risterucci A, Motamayor J, N'Goran J, Lanaud C (2003) Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. Genome 46:103-111.

De la Cruz M, Whitkus R, Gómez-Pompa A, Mota-Bravo L (1995) Origins of cacao cultivation. Nature 375:542-543.

Dillinger T, Barriga P, Escárcega S, Jiménez M, Lowe D, Grivetti L (2000) Food of the Gods: Cure for humanity? A cultural history of the medicinal and ritual use of chocolate. Journal of Nutrition 130: 2057S-2072S.

Feltus F, Saski C, Mockaitis K, Haiminen N, Parida L, Smith Z, Ford J, Staton M, Ficklin S, Blackmon B, Cheng C, Schnell R, Kuhn D, Motamayor J (2011) Sequencing of a QTL-rich region of the *Theobroma cacao* genome using pooled BACs and the identification of trait specific candidate genes. BMC Genomics 12:379.

Henderson JS, Joyce RA, Hall GR, Hurst WJ, McGovern PE (2007). Chemical and archaeological evidence for the earliest cacao beverages. *Proceedings of the National Academy of Sciences of the United States of America* 104:18937-18940.

ICCO (International Cocoa Organization) (2012). *Quarterly Bulletin of Cocoa Statistics*. Vol. XXXVIII-No.2. Cocoa Year 2010/11 London (May 2012).

Kar PK, Srivastava PP, Awasthi AK, Urs SR (2008) Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India. *Tree Genetics & Genomes* 4:75–83.

Khadivi-Khub A (2014). Regression association analysis of fruit traits with molecular markers in cherries. *Plant Systematic and Evolution* 300:1163-1173.

Khadivi-Khub A, Karimi E, Hadian J (2014) Population genetic structure and trait associations in forest savory using molecular, morphological and phytochemical markers. *Gene* 546:297-308.

Lanaud C, Fouet O, Clément D, Boccara M, Risterucci A, Surujdeo-Maharaj S, Legavre T & Argout X (2009) A meta-QTL analysis of disease resistance traits of *Theobroma cacao* L. *Molecular Breeding* 24:361-374.

Monteiro WR, Lopes UV, Clement D (2009) Genetic Improvement in Cocoa, in: Jain, S. M. Priyadarshan, P. M. (eds). *Breeding Plantation Tree Crops: Tropical Species*. Springer New York pp 589-626.

Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo O, Findley SD, Zheng P, Utro P, Royaert S, Saski C, Jenkins J, Podicheti R, Zhao M, Scheffler BE, Stack JC, Feltus FA, Mustiga GM, Amores F, Phillips W, Marelli JP, May GD, Shapiro H, Ma J, Bustamante CD, Schnell RJ, Main D, Gilbert D, Parida L, Kuhn DN (2013) The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology* 14:r53.

Pugh T, Fouet O, Risterucci A, Brottier P, Abouladze M, Deletrez C, Courtois B, Clement D, Larmande P, N’Goran J & Lanaud C (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theoretical and Applied Genetics* 108(6):1151-1161.

Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J and Salvador-Figueroa M (2012). Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad. *BioTecnología* 16:36-56.

Vázquez-Ovando JA, Molina-Freaner F, Nuñez-Farfán J, Ovando-Medina I, Salvador-Figueroa M (2014) Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico, *Genetic and Molecular Research* 13:10404-10414.

Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J, Betancur-Ancona D, Salvador-Figueroa M (2015a) Classification of cacao beans (*Theobroma cacao* L.) of Southern Mexico based on chemometric analysis with multivariate approach. *European Food Research and Technology*. Aceptado

Vázquez-Ovando A, Chacón-Martínez L, Betancur-Ancona D, Escalona-Buendía H, Salvador-Figueroa M (2015b) Sensory descriptors of cocoa beans from cultivated trees of Soconusco, Chiapas, Mexico. *Food Science and Technology*. Sometido

Virk PS, Ford-Lloyd BF, Jackson MT, Pooni HS, Clemenot TP, Newbury HJ (1996). Predicting quantitative variation within rice germplasm using molecular markers. *Heredity* 76:296-304.

CAPÍTULO 8

DISCUSIÓN GENERAL Y CONCLUSIONES

Los datos generados durante el estudio de la caracterización molecular resultaron útiles en dos aspectos. Primeramente, permitieron realizar análisis de asociación entre varios de los fragmentos generados por marcadores moleculares microsatélites (SSR's) y las características químicas y sensoriales de muestras de semillas de cacao cultivadas en la región Soconusco al sureste de México; por lo que se proponen como marcadores predictivos útiles de caracteres fenotípicos asociados con la calidad de las semillas de cacao (Capítulo 7). Además, este fue el objetivo general de esta investigación. De manera paralela, los datos de microsatélites permitieron resolver el parentesco genético Criollo de 27 árboles muestreados en las parcelas de Soconusco. Si bien es cierto que esto no coincide con lo esperado, pues al intencionar el muestreo hacia plantas que comparten características fenotípicas con la variedad Criollo, era de esperarse mayor correspondencia de los 108 árboles muestreados con los árboles empleados como controles Criollo, cultivados por los pueblos precolombinos y colectados de regiones donde no hay cultivo intensivo en la actualidad pero si se reporta de alta producción en la época pre y Colonial. La falta de correspondencia fue discutida en su momento, pero las conclusiones revisten de importancia práctica porque están revelando, que la diversidad fenotípica que se encuentra en las parcelas de Soconusco no corresponde precisamente con la diversidad genética, que según los resultados de nuestro estudio fue moderada. También estaría denotando que los caracteres fenotípicos Criollo pueden ser fácilmente confundidos con la variedad posiblemente Trinitario.

Aunque actualmente existe una propuesta de clasificar a los cacaos de América en al menos 10 grupos genéticos, el genotipo Criollo sigue siendo un grupo aparte y marcadamente diferente de los demás. Varios trabajos argumentan que el origen de esta planta esté en la cuenca del Amazonas, con base en la diversidad genética que allí se localiza y que en algún momento evolutivo se trasladaron accesiones a Mesoamérica donde se inició la diferenciación genética así como posteriormente su domesticación y cultivo extensivo. La moderada diversidad genética en los cacaos de Soconusco puede ser un apoyo a esta hipótesis. Sin embargo, esta diversidad, puede también deberse a la enorme presión de selección a la que estuvieron sometidas las plantaciones durante la domesticación, que al estar solo dirigida a caracteres de forma y sabor, reducen la

diversidad del material cultivado. Aun con esto, nuestros resultados revelan que hoy es posible localizar plantas cultivadas en las parcelas de Soconusco que comparten similitud genética con los Criollos ancestrales.

Sin embargo, al no poder elucidar el origen genético del otro grupo genético (K=2) hallado mediante análisis estadístico Bayesiano (posteriormente dividido en dos subgrupos, ver Capítulo 3), por ausencia de individuos control, se hipotetiza solamente basado en la historia del cultivo de cacao en esta región del estado de Chiapas, México. Los árboles pueden corresponder a genotipos producto de la recombinación de los Criollos nativos y las variedades introducidas después de la llegada de los españoles. Resolver este cuestionamiento incluyendo otros grupos genéticos controles, para comprender cabalmente la composición genética de las plantas que hoy se cultivan al sur de México, proporcionará información necesaria de la historia del cacao en el sitio donde se sabe se domesticó esta planta por pueblos precolombinos, además servirá como base para estudios aplicados en la búsqueda de cacaos con mejores características sobre todo sensoriales.

Al ser el aspecto sensorial una de las características más importantes en el comercio de cacao, el enfoque de calidad abordado en esta investigación se centró en esta característica, ligada a la composición química de las semillas. La caracterización química permitió establecer un conjunto de seis caracteres que permiten agrupar a los cacaos de Soconusco, Chiapas con base en su composición química. Cuatro de los seis caracteres están relacionados con la composición de ácidos grasos de la manteca de cacao, lo que deja ver de manera clara la influencia de esta característica que ya en varios estudios ha sido propuesta como marcador fenotípico de calidad de cacao. Al ser el cacao una planta de gran interés comercial, lo que demuestran los aproximadamente 5 millones de toneladas que se comercializan al año, no resulta extraño que, a la fecha del planteamiento de esta investigación, muchos caracteres químicos como el contenido y composición de la grasa, el contenido de teobromina, de polifenoles y otros han sido propuestos como caracteres fenotípicos de segregación. Sin embargo la gran mayoría de estos trabajos se centran en el agrupamiento de muestras de origen geográfico o genético marcadamente diferenciado. El aporte de nuestra investigación consiste en plantear caracteres que permiten agrupar muestras que fenotípicamente corresponden a un grupo de cacao y que genéticamente (según se demostró) corresponden a dos grupos diferentes. Otro aporte de la investigación es la propuesta en concordancia con algunos autores, del término de origen francés “terroir”, tal como ocurre en otros sectores (quesos,

café, vinos, te) donde se asigna nombres particulares a muestra de determinada región y que corresponde a la respuesta del genotipo al ambiente donde se desarrolla la planta. Cacao Real de Soconusco es una propuesta empírica que se ha argumentado entre los involucrados en la cadena de producción de cacao de esta región y que puede en el futuro ser alentada con resultados como los obtenidos en esta investigación.

Los caracteres o descriptores sensoriales obtenidos con la ayuda de un panel entrenado que se generaron para cacaos de Soconusco apoyan la propuesta anterior. Como resultado de la evaluación sensorial de las muestras de cacao se proponen descriptores muy particulares para las muestras analizadas. Estos descriptores, que están asociados con grupos de moléculas específicas, son el resultado de la composición genética de los árboles y la interacción con las condiciones edafo-climáticas de esta región de México. Aunque la región donde se muestrearon los cacaos incluidos en este estudio comprende una franja de unos 200 km paralela a la costa del estado de Chiapas, se pueden localizar algunos microclimas. Sin embargo, en nuestro estudio no se encontró asociación entre el origen geográfico de las muestras (parcelas o municipios) y la composición genética, o composición química e incluso sensorial. Lo anterior puede ser el resultado de la mutación de genes específicos como resultado de los intensos procesos de fijación de caracteres fenotípicos muy específicos y de interés humano, que fueron empleados durante la domesticación de los cacaos y a las posibles recombinaciones que han ocurrido desde la introducción de los genotipos Forastero y Trinitario. Los productos de tales recombinaciones han sido llamados por algunos autores como cacaos Criollos Modernos. En nuestro estudio, no se pudo incluir el total de árboles muestreadas en el estudio genético (108) para el estudio de la composición química y sensorial, debido a la presencia de la enfermedad sumamente agresiva conocida como moniliasis que ataca básicamente a los frutos y no permite su desarrollo. Un estudio químico y sensorial de las accesiones no incluidas en este estudio o con un universo muestral mayor podría apoyar la inclusión de los descriptores aquí propuestos y dar origen como antes se propuso a la denominación de origen.

Otras pruebas de asociación estadística, como el Análisis de Componentes Principales, de Correspondencia Canónica o de Discriminantes, podrían aportar información de la asociación entre los marcadores generados y las características fenotípicas desde diferentes enfoques. Éstos resultados, podrían ser comparados con la desarrollada en el Capítulo 7 que exhibe suficiente robustez.

REFERENCIAS

- AOAC (1997). Official methods of analysis. Association of Official Analytical Chemists, 17th ed. In: Horwitz W (ed). Washington, DC.
- Abbe--Maleyki MJ, Ismail A (2010). Antioxidant properties of cocoa powder. *Journal of Food Biochemistry* 34: 111–128.
- Adamska E, Cegielska Taras T, Kaczmarek Z, Szała L (2004). Multivariate approach to evaluating the fatty acid composition of seed oil in a doubled haploid population of winter oilseed rape (*Brassica napus* L.). *Journal of Applied Genetics* 45: 419–425.
- Affifi AA, Clark V (1984). Computer-aided multivariate analysis. Van Nostrand Reinhold, New York.
- Afoakwa E, Paterson A, Fowler M, Ryan A (2009). Matrix effects on flavour volatiles release in dark chocolates varying in particle size distribution and fat content using GC–mass spectrometry and GC-olfactometry. *Food Chemistry* 113: 208-215.
- Afoakwa E, Paterson A, Fowler M, Ryan A (2008). Flavor formation and character in cocoa and chocolate: a critical review. *Critical Reviews in Food Science* 48: 840-857.
- Afoakwa EO (2012). Chocolate and cocoa, flavor and quality. Kirk Othmer Encyclopedia of Chemical Technology 1-19.
- Aikpokpodion P (2010). Variation in agro-morphological characteristics of cacao, *Theobroma cacao* L., in farmers' fields in Nigeria. *New Zealand Journal of Crop and Horticultural Science* 38: 157-170.
- Álvarez C, Pérez E, Lares M (2007). Caracterización física y química de almendras de cacao fermentadas, secas y tostadas cultivadas en la región de Cuyagua, estado Aragua. *Agronomía Tropical* 57: 249-256.
- Amoye S (2006). Cocoa sourcing, world economics and supply. *Manufacture Confectionery* 86: 81e-85e.
- Araújo IS, de Souza Filho GA, Pereira MG, Faleiro FG, de Queiroz VT, Guimarães CT, Moreira MA, de Barros EG, Machado RCR, Pires JL, Schenell R, Lopes UV (2009). Mapping of Quantitative Trait Loci for butter content and hardness in cocoa beans (*Theobroma cacao* L.). *Plant Molecular Biology Reporter* 27: 177-183.
- Argout X, Salse J, Aury J, Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahí M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C (2011). The genome of *Theobroma cacao*. *Nature Genetics* 43: 101-109.

- Ashihara H, Kato M, Crozier A (2011). Distribution, biosynthesis and catabolism of methylxanthines in plants. In: B.B. Fredholm (ed.) Methylxanthines, Handbook of Experimental Pharmacology 11-26 pp.
- Ashihara H, Sano H, Crozier A (2008). Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* 69: 841-856.
- Ashihara H, Kato M, Ye C (1998). Biosynthesis and metabolism of purine alkaloids in leaves of cocoa tea (*Camellia ptilophylla*). *Journal of Plant Research* 111: 599-604.
- Ashihara H, Kubota H (1987). Biosynthesis of purine alkaloids in *Camellia* plants. *Plant Cell Physiology* 28: 535-539.
- Assemat S, Lachenaud P, Ribeyre F, Davrieux F, Pradon JL, Cros E (2005). Bean quality traits and sensory evaluation of wild Guianan cocoa populations (*Theobroma cacao* L.). *Genetic Resources and Crop Evolution* 52: 911-917.
- Azofeita-Delgado A (2006). Uso de marcadores moleculares en plantas; Aplicaciones en frutales del trópico. *Agronomía Mesoamericana* 17: 221-242.
- Baker D, Tomlins K, Gray C (1994). Survey of Ghanaian cocoa farmer fermentation practices and their influence on cocoa flavour. *Food Chemistry* 51: 425-431.
- Bartley B, Cope F (1973). Practical aspects of self-incompatibility in *Theobroma cacao* L. In: *Agricultural Genetics-Selected Topics*. Moav R (ed) John Wiley & Sons, Inc. New York. pp 109-134.
- Beggs C, Kuhn K, Böcker R, Wellman E (1987). Phytochrome-induced flavonoid biosynthesis in mustard (*Sinapis alba* L.) cotyledons. Enzymic control and differential regulation of anthocyanic and quercetin formation. *Planta* 172: 121-126.
- Bernays A, Oppenheim S, Chapman R, Kwon H, Gould F (2000). Taste sensitivity of insect herbivores to deterrents is greater in specialists than in generalists: a behavioral test of the hypothesis with two closely related caterpillars. *Journal of Chemical Ecology* 26: 547-563.
- Bertazzo A, Comai S, Brunato I, Zancato M, Costa CVL (2011). The content of protein and non-protein (free and protein-bound) tryptophan in *Theobroma cacao* beans. *Food Chemistry* 124:93-96.
- Bhattacharjee R, Aikpokpodion P, Kolesnikova-Allen M, Badaru K, Schnell R (2004). West African Cocoa: A pilot study on DNA fingerprinting the germplasm from Cross River State of Nigeria. *INGENIC Newsletter* 9: 15-20.
- Bhattacharjee R, Kolesnikova-Allen M, Aikpokpodion P, Taiwo S, Ingelbrecht I (2004). An improved semiautomated rapid method of extracting genomic DNA for molecular marker analysis in cocoa, *Theobroma cacao* L. *Plant Molecular Biology Reporter* 22: 435-436.
- Borrone J, Kuhn D, Schnell R (2004). Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theoretical and Applied Genetics* 109: 495-507.

- Brown J, Phillips-Mora W, Power E, Krol C, Cervantes-Martinez C, Motamayor J, Schnell J (2007). Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. *Crop Science* 47: 1851-1858.
- Brown J, Kuhn D, Lopez U, Schnell R (2005). Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an F₂ population using SSR markers and candidate genes. *Journal of the American Society for Horticulture Science* 130: 366-373.
- Bravo L (1998). Polyphenol: chemistry, dietary sources, metabolism and nutritional significance. *Nutrition Reviews* 56: 317-333.
- Camú N, De Winter T, Addo S, Takrama J, Bernaert H, De Vuyst L (2008). Fermentation of cocoa beans: influence of microbial activities and polyphenol concentrations on the flavour of chocolate. *Journal of the Science of Food and Agriculture* 88: 2288-2297.
- Cannac M, Pasqualini V, Greff S, Fernandez C, Ferrat L (2007) Characterization of phenolic compounds in *Pinus laricio* needles and their responses to prescribed burnings. *Molecules* 12:1614–1622.
- Cervantes-Martinez C, Brown J, Schnell R (2006). Combining ability for disease resistance, yield, and horticultural traits of cacao (*Theobroma cacao* L.) clones. *Journal of the American Society for Horticulture Science* 131: 231-241.
- Chacón RI, Ramis C, Gómez C (2011) Morphological description of fruits and seeds of the Criollo Porcelana cocoa (*Theobroma cacao* L.) in south of Maracaibo Lake. *Revista de la Facultad de Agronomía LUZ* 28(Supl. 1):1–13.
- Cheesman E (1944). Notes on the nomenclature, classification and possible relationships of cocoa populations. *Tropical Agriculture* 21: 144-159.
- Chiapas.gob (2010) Cacao del Soconusco, de los mejores del mundo: empresarios franceses. Instituto de Comunicación Social. Diciembre de 2010. Consultado junio de 2011. Disponible: <http://www.cocoso.chiapas.gob.mx/documento.php?id=20101212124448>.
- Ciferri R, Ciferri F (1957). The evolution of cultivated cacao. *Evolution* 11: 381-397.
- Clapperton J (1994). A review of research to identify the origins of cocoa flavour characteristics. *Cocoa Growers' Bulletin* 48: 7-16.
- Clement D, Risterucci A, Motamayor J, N'Goran J, Lanaud C (2003) Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. *Genome* 46: 204-212.
- Clement D, Risterucci A, Motamayor J, N'Goran J, Lanaud C (2003) Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. *Genome* 46: 103-111.
- Coe SD, Coe MD (2007). *The true history of chocolate* (2. Ed.). Thames and Hudson Ltd., New York.
- Coe SD, Coe MD (1996). *The True History of Chocolate*. 1st edn. Thames and Hudson Ltd., New York.

- Couch J, Zintel H, Fritz P (1993) The genome of the tropical *Theobroma cacao* L. *Molecular and General Genetics* 237: 123-128.
- Crouzillat D, Menard B, Mora A, Phillips W, Petiard V (2000). Quantitative trait analysis in *Theobroma cacao* using molecular markers. *Euphytica* 114: 13-23.
- Crouzillat D, Lerceteau E, Petiard V, Morera J, Rodriguez H, Walker D, Phillips W, Ronning C, Schnell R, Osei J, Fritz P (1996). *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. *Theoretical and Applied Genetics* 93: 205-214.
- Crozier S, Preston A, Hurst J, Payne M, Mann J, Hainly L, Miller D (2011). Cacao seeds are a "Super Fruit": A comparative analysis of various fruit powders and products. *Chemistry Central Journal* 5: 5.
- Crown PL (2013). Prehispanic use of cacao. In R. Watson, V. R. Preedy & S. Zibadi (Eds.), *Chocolate in Health and Nutrition* (chap. 1; pp 3-9). New York: Humana Press.
- Crown PL, Hurst WJ (2009). Evidence of cacao use in the Prehispanic American Southwest. *Proceedings of the National Academy of Sciences of the United States of America* 106: 2110-2113.
- Crozier S, Preston A, Hurst J, Payne M, Mann J, Hainly L, Miller D (2011). Cacao seeds are a "Super Fruit": A comparative analysis of various fruit powders and products. *Chemical Central Journal* 5: 5.
- Cuatrecasas J (1964). Cacao and its allies: a taxonomic revision of the genus *Theobroma*. *Contributions from the United States National Herbarium* 35: 379-614.
- Dantas L, Guerra M (2010). Chromatin differentiation between *Theobroma cacao* L. and *T. grandiflorum* Schum. *Genetics and Molecular Biology* 33: 94-98.
- De Brito E, Pezoa-García N, Gallão M, Cortelazzo A, Fevereiro P, Braga M (2000). Structural and chemical changes in cocoa (*Theobroma cacao* L) during fermentation, drying and roasting. *Journal of the Science of Food and Agriculture* 81: 281-288.
- De la Cruz M, Whitkus R, Gomez-Pompa A, Mota-Bravo L (1995). Origins of cacao cultivation. *Nature* 375: 542-543.
- d'El-Rei J, Cunha AR, Burlá A, Burlá M, Oigman W, Neves MF, Viridis A, Medeiros F (2013). Characterisation of hypertensive patients with improved endothelial function after dark chocolate consumption. *International Journal of Hypertension* Article ID 985087.
- De León-Rodríguez A, Escalante-Minakata P, Jiménez-García MI, Ordoñez-Acevedo LG, Flores-Flores JL, Barba-de la Rosa AP (2008). Characterization of volatile compounds from ethnic *Agave* alcoholic beverages by gas chromatography--mass spectrometry. *Food Technology and Biotechnology* 46: 448-455.
- Dillinger T, Barriga P, Escárcega S, Jiménez M, Lowe D, Grivetti L (2000). Food of the Gods: Cure for humanity? A cultural history of the medicinal and ritual use of chocolate. *Journal of Nutrition* 130: 2057S-2072S.
- Dixon R, Pavia N (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell* 7: 1085-1097.

- Djoussé L, Hopkins PN, North KE, Pankow JS, Arnett DK, Ellison RC (2011). Chocolate consumption is inversely associated with prevalent coronary heart disease: the national heart, lung, and blood Institute Family Heart Study. *Clinical Nutrition* 30(2): 182-187.
- Dongo L, Aigbekaen E, Jayeola C, Emaku L, Orisajo S (2009). Influence of farmers practices on cocoa bean quality: Nigeria field experience. *African Crop Science Conference Proceedings* 9: 299-302.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Dzogbefia V, Buamah R, Oldham J (1999). The controlled fermentation of cocoa (*Theobroma cacao* L.) using yeasts: enzymatic process and associated physico-chemical changes in cocoa sweatings. *Food Biotechnology* 13: 1-12.
- Earl DA, von Holdt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361.
- Efombagn M, Sounigo O, Eskes A, Motamayor J, Manzanares-Dauleux M, Schnell R, Nyassé S (2009). Parentage analysis and outcrossing patterns in cacao (*Theobroma cacao* L.) farms in Cameroon. *Heredity* 103: 46-53.
- Efombagn M, Nyassé S, Sounigo O, Kolesnikova-Allen M, Eskes A (2007). Participatory cocoa (*Theobroma cacao*) selection in Cameroon: *Phytophthora* pod rot resistant accessions identified in farmers' fields. *Crop Protection* 26: 1467-1473.
- Engels JM (1983) A systematic description of cacao clones. I. The discriminative value of quantitative characteristics. *Euphytica* 32:377-385.
- Eskes A, Guarda D, García L, García P (2007). Is genetic variation for sensory traits of cocoa pulp related to fine flavor cocoa traits? *INGENIC Newsletter* 11: 22-28.
- Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.
- Evans W (2000) Trease and Evans - Pharmacognosy. (15th ed). Edinburgh. Ed. Saunders.
- Faleiro F, Queiroz V, Lopes U, Guimarães C, Pires J, Yamada M, Araújo I, Pereira M, Schnell R, De Souza G (2006). Mapping QTLs for witches' broom (*Crinipellis pernicioso*) resistance in cacao (*Theobroma cacao* L.). *Euphytica* 149: 227-235.
- Faleiro F, Yamada M, Lopes U, Faleiro A, Siqueira R, Costa L, Santos R, Dos Santos R (2002). Genetic similarity of *Theobroma cacao* L. accessions maintained in duplicates in the Cacao Research Center germplasm collection, based on RAPD markers. *Crop Breeding and Applied Biotechnology* 2: 439-444.
- FAO (2014). Organización de las Naciones Unidas para la Agricultura y la Alimentación. FAOSTAT. Estadísticas de Producción 2012.
- FAO (Food and Agriculture Organization). (2013). FAOSTAT Online Statistical Service. Retrieved from <http://faostat.fao.org/>. Accessed August 2014.

- FAO (Food and Agriculture Organization) (2012). FAOSTAT Online Statistical Service. Available at [http://faostat.fao.org/]. Accessed September 12, 2013.
- FAO (2011). Organización de las Naciones Unidas para la agricultura y la Alimentación. Anuario Estadístico 2010. Tabla B7.
- Fatland B, Ke J, Anderson M, Mentzen W, Cui L, Allred C, Johnston J, Nikolau B, Wurtele E (2004). Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-Coenzyme A in Arabidopsis. *Plant Physiology* 130: 740-756.
- Feltus F, Saski C, Mockaitis K, Haiminen N, Parida L, Smith Z, Ford J, Staton M, Ficklin S, Blackmon B, Cheng C, Schnell R, Kuhn D, Motamayor J (2011). Sequencing of a QTL-rich region of the *Theobroma cacao* genome using pooled BACs and the identification of trait specific candidate genes. *BMC Genomics* 12: 379.
- Figueira A, Albuquerque S, Leal G (2006). Genetic mapping and differential gene expression of Brazilian alternative resistance sources to witches' broom (causal agent *Crinipellis perniciosa*). 15th International Cocoa Research Conference p. 166
- Figueira A, Janick J, Levy M, Goldsbrough P (1994). Reexamining the classification of *Theobroma cacao* L. using molecular markers. *Journal of the American Society for Horticultural Science* 119: 1073-1082.
- Flament M, Kebe I, Clement D, Pieretti I, Risterucci A, N'Goran J, Cilas C, Despréaux D, Lanaud C (2001). Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. *Genome* 44: 79-85.
- Forsyth WGC, Quesnel VC. (1957). Cacao glycosidase and colour changes during fermentation. *Journal of the Science of Food and Agriculture* 8: 505-509.
- Frauendorfer F, Schieberle P (2008). Changes in key aroma compounds of criollo cocoa beans during roasting. *Journal of Agricultural and Food Chemistry* 56: 10244-10251.
- García-Alamilla P, Salgado-Cervantes MA, Barel M, Berthomieu G, Rodríguez-Jiménez GC, García-Alvarado MA. (2007). Moisture, acidity and temperature evolution during cacao drying. *Journal of Food Engineering* 79: 1159-1165.
- Gómez-Pompa A, Flores JS, Fernández MA (1990). The sacred cacao groves of the Maya. *Latin American Antiquity* 1: 247-257.
- Gotsch N (1997). Cocoa biotechnology: status, constraints and future prospects. *Biotechnology Advances* 15: 333-352.
- Graham T (1991). Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates. *Plant Physiology* 95: 594-603.
- Granvogl M, Bugan S, Schieberle P (2006). Formation of amines and aldehydes from parent amino acids during thermal processing of cocoa and model systems: new insights into pathways of the Strecker reaction. *Journal of Agriculture and Food Chemistry* 54: 1730-1739.
- Grotewold E (2006). *The science of flavonoids*. Ohio, USA. Ed. Springer. 274 p.

- Guiltinan M, Verica J, Zhang D, Figueira A (2008). Genomics of the *Theobroma cacao*, "The Food of the Gods". In: Genomics of Tropical Crop Plants. Moore P, Mings R (eds). Springer Science+Business Media. NY USA. pp. 145-170.
- Haddon AV (1961). Variety trials of seedling cacao in Malaya. Malayan Agriculture. Journal 43: 169-232.
- Hammerstone JF, Romanczyk LJ, Aitken WM (1994). Purine alkaloid distribution within *Herrania* and *Theobroma*. Phytochemistry 35: 1237-1240.
- Hahlbrock K, Scheel D (1989). Physiology and molecular biology of phenylpropanoid metabolism. Annual Review of Plant Physiology 40: 347-369.
- Hamelin C, Sempere G, Jouffe V, Ruiz M (2013). TropGeneDB, the multi-tropical crop information system updated and extended. Nucleic Acids Research 41: D1172-D1175.
- Hansen C, del Olmo M, Burri C (1998). Enzyme activities in cocoa beans during fermentation. Journal of the Science of Food and Agriculture 77: 273-281.
- Harrington W (2011). The effects of roasting time and temperature on the antioxidant capacity of cocoa beans from Dominican Republic, Ecuador, Haiti, Indonesia, and Ivory Coast. Tesis de Maestría, Universidad de Tennessee. E.U.A.
- Harrison JM, Howard D, Malven M, Halls SC, Culler AH, Harrigan GG, Wolfinger RD (2013) Principal variance component analysis of crop composition data: a case study on herbicide tolerant cotton. Journal of Agricultural and Food Chemistry 61:6412-6422.
- Henderson JS, Joyce RA, Hall GR, Hurst WJ, McGovern PE (2007). Chemical and archaeological evidence for the earliest cacao beverages. Proceedings of the National Academy of Sciences of the United States of America 104: 18937-18940.
- Herrmann K, Weaver L (1999). The shikimate pathway. Annual Review of Plant Physiology. Plant Molecular Biology 50: 473-503.
- Hurst WJ, Tarka SM Jr, Powis TG, Valdez F Jr, Hester TR (2002). Cacao usage by the earliest Maya civilization. Nature 418: 289-290.
- ICCO (International Cocoa Organization) (2012). Quarterly Bulletin of Cocoa Statistics. Vol. XXXVIII-No.2. Cocoa Year 2010/11 London (May 2012).
- Ichihara K, Shibajara A, Yamamoto K, Nakayama T (1996) An improved method for rapid analysis of the fatty acids of glycerolipids. Lipids 31:535-539.
- INAH (2007). Instituto Nacional de Antropología e Historia. Comisión de defensa del patrimonio cultural. Considerations and alternative proposal for the encroachment on the archaeological site by the Tapachula-Talismán highway. Disponible: www.mesoweb.com/reports/lzapa.pdf.
- Ishida M, Kitao N, Mizuno K, Tanikawa N, Kato M (2009). Occurrence of theobromine synthase in purine alkaloid-free species of Camellia plants. Planta 229: 559-568.
- ISO. (1993). International standard 8586-1: Sensory analysis—General guidance for the selection, training and monitoring of assessors. (March 15, 1993).

- Irizarry H, Goenaga R (2000). Clonal selection in cacao base on early yield performance of grafted trees. *Journal of Agriculture- University of Puerto Rico* 84: 153-163.
- Jiménez-Escrig A, Dragsted L, Daneshvar B, Pulido R, Saura-Calixto F (2003) *In vitro* antioxidant activities of edible artichoke (*Cynara scolymus* L.) and effect on biomarkers of antioxidants in rats. *Journal of Agricultural and Food Chemistry* 51:5540-5545.
- Jinap S, Siti M, Norsiaty M (1994). Formation of methyl pyrazine during cocoa bean fermentation. *Pertanika Journal of Tropical Agriculture Science* 17: 27-32.
- Jinap S, Dimick PS, Hollender R. (1995). Flavour evaluation of chocolate formulated from cocoa beans from different countries. *Food Control* 6(2): 105-110.
- Jinap S, Wan Rosli W, Russly A, Nordin L (1998). Effect of roasting time and temperature on volatile component profiles during nib roasting of cocoa beans (*Theobroma cacao*). *Journal of the Science of Food and Agriculture* 77: 441-448.
- Johnson E, Bekele F, Brown S, Song Q, Zhang D, Meinhardt L, Schnell R (2009). Population structure and genetic diversity of the Trinitario cacao (*Theobroma cacao* L.) from Trinidad and Tobago. *Crop Science* 49: 564-572.
- Kar PK, Srivastava PP, Awasthi AK, Urs SR (2008). Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India. *Tree Genetics and Genomes* 4: 75-83.
- Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H (2000). Caffeine synthase gene from tea leaves. *Nature* 406: 956-957.
- Kato M, Mizuno K, Fujimura T, Iwama M, Irie M, Crozier A, Ashihara H (1999). Purification and characterization of caffeine synthase from tea leaves. *Plant Physiology* 120: 579-586.
- Khadivi-Khub A (2014). Regression association analysis of fruit traits with molecular markers in cherries. *Plant Systematic and Evolution* 300: 1163-1173.
- Khadivi-Khub A, Karimi E, Hadian J (2014). Population genetic structure and trait associations in forest savory using molecular, morphological and phytochemical markers. *Gene* 546: 297-308.
- Koyama Y, Tomoda Y, Kato M, Ashihara H (2003). Metabolism of purine bases, nucleosides and alkaloids in theobromine-forming *Theobroma cacao* leaves. *Plant Physiology and Biochemistry* 41: 977-984.
- Kubasek W, Shirley B, McKillop A, Goodman H, Briggs W, Ausubel F (1992). Regulation of flavonoid biosynthetic genes in germinating *Arabidopsis* seedlings. *Plant Cell* 4: 1229-1236.
- Lachenaud P, Zhang D (2008). Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana. *Annals of Forest Science* 65: 310.
- Lachenaud P, Sounigo O, Oliver G (2004). Genetic structure of Guianan wild cocoa (*Theobroma cacao* L.) described using isozyme electrophoresis. *Plant Genetic Resources Newsletter* 139: 24-30.

- Lanaud C, Fouet O, Clément D, Boccara M, Risterucci A, Surujdeo-Maharaj S, Legavre T, Argout X (2009). A meta-QTL analysis of disease resistance traits of *Theobroma cacao* L. *Molecular Breeding* 24: 361-374.
- Lanaud C, Risterucci A, Pieretti I, Falque M, Bouet A, Lagoda P (1999). Isolation and characterization of microsatellites in *Theobroma cacao* L. *Molecular Ecology* 8: 2141-2152.
- Lanaud C, Risterucci A, N'Goran A, Clement D, Flament M, Laurent V, Falque M (1995) A genetic linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* 91: 987-993.
- Lanaud C (1986). Genetic studies of *Theobroma cacao* L. with the help of enzymatic markers. I: Genetic control and linkage of nine enzymatic markers. *Cafe Cacao The* 30: 259-270.
- Lawless HT, Heymann H. (1999). *Sensory evaluation of food-principles and practices*. Aspen Publisher Inc. Maryland, USA.
- Lee KW, Kim YJ, Lee HJ, Lee CY (2003) Cocoa has more phenolic phytochemicals and a higher antioxidant capacity than teas and red wine. *Journal of Agricultural and Food Chemistry* 51: 7292–7295.
- Lefeber T, Janssens M, Camú N, De Vuyst L (2010). Kinetic analysis of strains of lactic acid bacteria and acetic acid bacteria in cocoa pulp simulation media toward development of a starter culture for cocoa bean fermentation. *Applied and Environmental Microbiology* 76: 7708-7716.
- Lehrian DW, Keeney PG, Butler DR (1980) Triglyceride characteristics of cocoa butter from cacao fruit matured in a microclimate of elevated temperature. *Journal of the American Oil Chemists' Society* 57: 66–69.
- Lerceteau E, Robert T, Pétiard V, Crouzillat D (1997). Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theoretical and Applied Genetics* 95: 10-19.
- Liendo R, Padilla F, Quintana A (1997). Characterization of cocoa butter extracted from Criollo cultivars of *Theobroma cacao* L. *Food Research International* 30: 727-731.
- Lindsay R (1996). *Food Additives*. In: O. Fennema (Ed). *Food Chemistry* (3rd ed). New York. Marcel Dekker 768-823 pp.
- Loor-Solorzano RG, Fouet O, Lemainque A, Pavek S, Boccara M, Argout X, Amores F, Courtois B, Risterucci AM, Lanaud C (2012). Insight into the wild origin, migration and domestication history of the fine flavour Nacional *Theobroma cacao* L. variety from Ecuador. *PLoS One* 7: e48438.
- Lopéz-Cortés I, Salazar-García DC, Malheiro R, Guardiola V, Pereira JA (2013) Chemometrics as a tool to discriminate geographical origin of *Cyperus esculentus* L. based on chemical composition. *Industrial Crops and Products* 51: 19-25.
- López-Mendoza R (1987). *El cacao en Tabasco*. Colección Cuadernos Universitarios. Serie Agronómica no. 13. Universidad Autónoma de Chapingo, Texcoco.
- Luna F, Crouzillat D, Cirou L, Bucheli P (2002). Chemical composition and flavor of Ecuadorian cocoa liquor. *Journal of the Agriculture and Food Chemistry* 50: 3527-3532.

- Macht M, Dettmer D. (2006). Everyday mood and emotions after eating a chocolate bar or an apple. *Appetite* 46: 332-336.
- Martens S, Mithöfer A (2005). Flavones and flavone synthases. *Phytochemistry* 66: 2399-2407.
- Martínez-Valverde I, Periago M, Ros G (2000). Significado nutricional de los compuestos fenólicos de la dieta. *Archivos Latinoamericanos de Nutrición* 50: 5-18.
- Mazzafera P, Wingsle G, Olsson O, Sandberg G (1994). S-adenosyl-L-methionine:theobromine 1-Nmethyltransferase, an enzyme catalyzing the synthesis of caffeine in coffee. *Phytochemistry* 37: 1577-1584.
- Melo LLMM, Bolini HMA, Efraim P. (2009). Sensory profile, acceptability, and their relationship for diabetic/reduced calorie chocolates. *Food Quality and Preference* 20: 138-143
- Misnawi S, Jamilah B, Nazamid S (2003). Effects of incubation and polyphenol oxidase enrichment on colour, fermentation index, procyanidins and astringency of unfermented and partly fermented cocoa beans. *International Journal of Food Science and Technology* 38: 285-295.
- Mizuno K, Okuda A, Kato M, Yoneyama N, Tanaka H, Ashihara H, Fujimura T (2003) Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.) *FEBS Letters* 534: 75-81.
- Mizuno K, Tanaka H, Kato M, Ashihara H, Fujimura T (2001) *cDNA cloning of caffeine (theobromine) synthase from coffee (Coffea arabica L.)*. In: Proceedings of the International Scientific Colloquium on Coffee. ASIC, Paris, 19, 815-818.
- Molina J, Córdova L (2006) Recursos fitogenéticos de México para la alimentación y la agricultura: Informe Nacional 2006. Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación y Sociedad Mexicana de Fitogenética, A.C. Chapingo, México. pp. 19-20.
- Monteiro WR, Lopes UV, Clement D (2009). Genetic Improvement in Cocoa, in: Jain, S. M. Priyadarshan, P. M. (eds). *Breeding Plantation Tree Crops: Tropical Species*. Springer New York pp 589-626.
- Moreno Y, Melgarejo L, Hernández M, Quintero L, Vargas G (2004). Caracterización molecular de un banco de germoplasma del género *Theobroma* mediante la técnica de RAPD. *Revista Colombiana de Biotecnología* 6: 15-24.
- Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo O, Findley SD, Zheng P, Utro P, Royaert S, Saski C, Jenkins J, Podicheti R, Zhao M, Scheffler BE, Stack JC, Feltus FA, Mustiga GM, Amores F, Phillips W, Marelli JP, May GD, Shapiro H, Ma J, Bustamante CD, Schnell RJ, Main D, Gilbert D, Parida L, Kuhn DN (2013). The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology* 14: r53.
- Motamayor J, Lachneaud P, da Silva e Mota J, Loo R, Kuhn D, Brown J, Schnell R (2008). Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS One* 3: e3311.

- Motamayor J, Risterucci A, Heath M, Lanaud C (2003). Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity* 91: 322-330.
- Motamayor J, Risterucci A, Lopez P, Ortiz C, Moreno A, Lanaud C (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89: 380-386.
- Motilal L, Zhang D, Pathmanathan U, Mischke S, Pinney S, Meinhardt L (2011). Microsatellite fingerprinting in the International Cocoa Genebank, Trinidad: accession and plot homogeneity information for germplasm management. *Plant Genetic Resources* 9: 430-438.
- Motilal L, Zhang D, Umaharan P, Mischke BS, Mooleedhar V, Meinhardt LW (2010). The relic Criollo cacao in Belize genetic diversity and relationship with Trinitario and other cacao clones held in the International cocoa Genebank, Trinidad. *Plant Genetic Resources* 8: 106-110.
- Motilal LA, Zhang D, Umaharan P, Mischke S, Boccara M, Pinney A (2009). Increasing accuracy and throughput in large-scale microsatellite fingerprinting of cacao field germplasm collections. *Tropical Plant Biology* 2: 23-37.
- Muñoz D, García del Valle Y, Franco B, Estrada E, Magaña M (2002). Estudio del patrón de actividad general de monos aulladores (*Alouatta palliata*) en el parque Yumká, Tabasco, México. *Neotropical Primates* 10: 11-16.
- Niemenak N, Rohsius C, Elwers S, Ndoumou DO, 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: 612–619.
- N'Goran J, Laurent V, Risterucci A, Lanaud C (1994). Comparative genetic diversity studies of *Theobroma cacao* using RFLP and RAPD markers. *Heredity* 73: 589-597.
- Oberparleiter S, Ziegler G (1997). Amyl acohols as compounds indicative of raw cocoa bean quality. *European Food Research and Technology* 204: 156-160.
- Ogata N (2003). Domestication and Distribution of the Chocolate Tree (*Theobroma cacao* L.) in Mesoamerica. In: *The Lowland Maya Area. Three Millennia at the Human-Wildland Interface* (Gómez-Pompa A, Allen MF, Fedick SL, Jiménez-Osorio JJ, eds.). The Haworth Press Inc., New York, 415-438.
- Orsi CH, Tanksley SD (2009) Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. *PLoS Genetics* 5:e1000347.
- Padilla FC, Liendo R, Quintana A (2000) Characterization of cocoa butter extracted from hybrid cultivars of *Theobroma cacao* L. *Archivos Latinoamericanos de Nutrición* 50: 200–205.
- Paetkau D, Slade R, Burden M and Estoup A (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* 13: 55-65.
- Papalexandratou Z, Camu N, Falony G, De Vuyst L (2011). Comparison of the bacterial species diversity of spontaneous cocoa bean fermentations carried out at selected farms in Ivory Coast and Brazil. *Food Microbiology* 28: 964-973.

- 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: 1326-1338.
- Parr A, Bolwell G (2000). Phenols in the plant and in man. The potential for possible nutritional enhancement of the diet by modifying the phenols content or profile. *Journal of the Science of Food and Agriculture* 80: 985-1012.
- Peakall R and Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.
- Perry M, Davey M, Power J, Lowe K, Bligh H, Roach P, Jones C (1998). DNA isolation and AFLP genetic fingerprinting of *Theobroma cacao* (L.). *Plant Molecular Biology Reporter* 16: 49-59.
- Pettipher G (1986). Analysis of cocoa pulp and the formulation of a standardised artificial cocoa pulp medium. *Journal of the Science of Food and Agriculture* 37: 297-309.
- Phillips-Mora W, Castillo J, Krauss U, Rodríguez E, Wilkinson M (2005). Evaluation of cacao (*Theobroma cacao*) clones against seven Colombian isolates of *Moniliophthora roreri* from four pathogen genetic groups. *Plant Pathology* 54: 483-490.
- Phillips-Mora W, Coutiño A, Ortiz C, López A, Hernández J, Aime M (2006). First report of *Moniliophthora roreri* causing frosty pod rot (moniliasis disease) of cocoa in México. *Plant Pathology* 55: 584.
- Pires JL, Mattos-Cascardo JC, Lambert SV, Figueira A (1998) Increasing cocoa butter yield through genetic improvement of *Theobroma cacao* L.: seed fat content variability, inheritance, and association with seed yield. *Euphytica* 103: 115–121.
- Ploetz R (2007). Cacao diseases: important threats to chocolate production worldwide. *Phytopathology* 97: 1634-1639.
- Portillo E, Labarca M, Grazziani L, Cros E, Assemat S, Davrieux F, Boulanger R, Marcano M (2009). Formación del aroma del cacao Criollo (*Theobroma cacao* L.) en función del tratamiento poscosecha en Venezuela. *Revista UDO Agrícola* 9: 458-468.
- Pourmorad F, Hosseinimehr SJ, Shahabimajd N (2006) Antioxidant activity, phenol and flavonoid contents of some selected Iranian medicinal plants. *African Journal of Biotechnology* 5: 1142–1145.
- Powis T, Cyphers A, Gaikwad N, Grivetti L, Cheong K (2011). Cacao use and the San Lorenzo Olmec. *Proceedings of the National Academy of Sciences of the United States of America* 108: 8595-8600.
- Powis T, Hurst W, Rodríguez M, Ortiz P, Blake M, Cheetham D, Coe M, Hodgson J (2008). The origins of cacao use in Mesoamerica. *Mexicon* 30: 35-38.
- Powis T, Hurst J, Rodríguez MC, Ortiz P, Blake M, Cheetham D, Coe M, Hodgson J (2007). Oldest chocolate in the New World. *Antiquity* 81: 314.

- Preza AM, Jaramillo ME, Puebla AM, Mateos JC, Hernández R, Lugo E (2010) Antitumor activity against murine lymphoma L5178Y model of proteins from cacao (*Theobroma cacao* L.) seeds in relation with in vitro antioxidant activity. *BMC Complementary and Alternative Medicine* 10:61.
- Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Pugh T, Fouet O, Risterucci A, Brottier P, Abouladze M, Deletrez C, Courtois B, Clement D, Larmande P, N'Goran J, Lanaud C (2004). A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theoretical and Applied Genetics* 108: 1151-1161.
- Queiroz V, Guimarães C, Anher T, Schuster I, Daher T, Pereira M, Miranda V, Loguercio L, Barros E, Moreira M (2003). Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. *Plant Breeding* 122: 268-272.
- Quesnel V (1965). Agents inducing the death of the cacao seeds during fermentation. *Journal of the Science of Food and Agriculture* 16: 441-447.
- Ramiro E, Franch A, Castellote C, Andres-Lacueva C (2005). Effect of *Theobroma cacao* flavonoids on immune activation of a lymphoid cell line. *British Journal of Nutrition* 93: 859-866.
- Ramli N, Hassan O, Said M, Samsudin W, Idris N (2006). Influence of roasting condition on volatile flavour of roasted Malaysian cocoa beans. *Journal of Food Processing and Preservation* 30: 280-298.
- Rangel M (2009) Anatomía y tolerancia a la desecación de semillas de cacao (*Theobroma cacao* L.) Tesis de doctorado en Ciencias. Colegio de Posgraduados. Texcoco, Estado de México. pp.1-49.
- Rangel-Fajardo MA, Zavaleta-Mancera HA, Cordova-Tellez L, Lopez-Andrade AP, Delgado-Alvarado A, Vidales-Fernandez I, Villegas-Monter A (2012) Anatomía e histoquímica de la semilla del cacao (*Theobroma cacao* L.) criollo mexicano. *Revista Fitotecnia Mexicana* 35: 189-197.
- Rice-Evans C, Miller N, Paganga G (1997). Antioxidant properties of phenolic compounds. *Trends in Plant Science* 2: 2152-2159.
- Riju A, Rajesh MK, Sherin PT, Chandrasekar A, Apsara SE, Arunachalam V(2009). Mining of expressed sequence tag libraries of cacao for microsatellite markers using five computational tools. *Journal of Genetics*. 88: 217-225.
- Risterucci A, Paulin D, Ducamp M, N'Goran J, Lanaud C (2003). Identification of QTLs related to cocoa resistance to three species of Phytophthora. *Theoretical and Applied Genetics* 108: 168-174.
- Risterucci A, Grivet L, N'Goran J, Pieretti I, Flament M, Lanaud C (2000). A high-density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* 101: 948-955.
- Rodriguez-Campos J, Escalona-Buendía H, Orozco-Avila I, Lugo-Cervantes E, Jaramillo-Flores M (2011). Dynamics of volatile and non-volatile compounds in cocoa (*Theobroma cacao* L.) during

- fermentation and drying processes using principal components analysis. *Food Research International* 44: 250-258.
- Roelofsens PA (1958). Fermentation, drying, and storage of cocoa beans. *Advances in Food Research* 8: 225-296.
- Rohan T (1963). The precursors of chocolate aroma: application of gas chromatography in following formation during fermentation of cocoa beans. *Journal of Food Science* 32: 402-404.
- Rohan T (1964). The precursors of chocolate aroma: comparative study of fermented and unfermented cocoa beans. *Journal of Food Science* 29: 456-459.
- Ronning C, Schnell R (1994). Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *Journal of Heredity* 85: 291-295.
- Ruiz J, Roa O, Marin I (2011). Molecular ecology of genetic diversity of cacao cultivated in the south-east region of Nicaragua. *International Research Journal of Agricultural Science and Soil Science* 1: 6-13.
- Rusconi M, Conti A (2010). *Theobroma cacao* L., the food of the gods: A scientific approach beyond myths and claims. *Pharmacology Research* 61: 5-13.
- Russell J, Hosein F, Johnson E, Waugh R, Powell W (1993). Genetic differentiation of cocoa (*Theobroma cacao* L.) populations revealed by RAPD analysis. *Molecular Ecology* 2: 89-97.
- Santos RC, Pires JL, Correa RX (2012) Morphological characterization of leaf, flower, fruit and seed traits among Brazilian *Theobroma* L. species. *Genetic Resources and Crop Evolution* 59: 327-345.
- Saunders JA, Mischke S, Leamy EA, Hemeida AA (2004). Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theoretical and Applied Genetics* 110: 41-47.
- Schnell R, Olano C, Brown J, Meerow A, Cervantes-Martínez C , Nagai C, Motamayor JC(2005). Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. *Journal of the American Society for Horticultural Science* 130: 181-190.
- Schultes R (1984). Amazonian cultigens and their northward and westward migrations in pre-Columbian times. *In: Pre-Columbian plant migration. Papers of the Peabody Museum of Archaeology and Ethnology. Vol 76: Mass. Stone D (ed). Harvard University Press. Cambridge. pp 69-83.*
- Schwan R (1998). Cocoa fermentations conducted with a defined microbial cocktail inoculum. *Applied and Environmental Microbiology* 64: 1477-1483.
- Schwan R, Wheals A (2004). The microbiology of cocoa fermentation and its role in chocolate quality. *Critical Reviews in Food Science and Nutrition* 44: 205-221.
- Sereno M, Albuquerque P, Vencovsky R, Figueira A (2006). Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conservation Genetics* 7: 13-24.

- SIAP (2012). Servicio de Información Agroalimentaria y Pesquera. SAGARPA México. Produccion agrícola por cultivo, reporte 2012. http://www.siap.gob.mx/index.php?option=com_wrapper&view=wrapper&Itemid=350. Accessed 30 July 2013.
- SIAP (Servicio de Información Agroalimentaria y Pesquera, SAGARPA México) (2013). Produccion Agrícola por Cultivo, Reporte 2012. Available at [http://www.siap.gob.mx/index.php?option=com_wrapper&view=wrapper&Itemid=350]. Accessed July 30, 2013.
- Smulders M, Esselink D, Amores F, Ramos G, Sukha D, Butler D, Vosman B, Van Loo E (2010). Identification of cocoa (*Theobroma cacao* L.) varieties with different quality attributes and parentage analysis of their beans. *INGENIC Newsletter* 12: 1-13.
- Smyth D (1992). Effect of methylxanthine treatment on rice seedling growth. *Journal of Plant Growth Regulation* 11: 125-128.
- Smith N. (1999). *The Amazon River Forest: a natural history of plants, animals, and people*. Oxford University Press, New York.
- Sotelo A, Álvarez G (1991). Chemical composition of wild *Theobroma* species and their comparison to the cacao bean. *Journal of Agricultural and Food Chemistry* 39: 1940-1943.
- Sounigo O, Umaharan R, Christopher Y, Sankar A, Ramdahin S (2005). Assessing the genetic diversity in the International Cocoa Genebank, Trinidad (ICG,T) using isozyme electrophoresis and RAPD. *Genetic Resources and Crop Evolution* 52: 1111-1120.
- Stark T, Bareuther S, Hofmann T (2006). Molecular definition of the taste of roasted cocoa nibs (*Theobroma cacao*) by means of quantitative studies and sensory experiments. *Journal of Agricultural and Food Chemistry* 54: 5530-5539.
- Stasolla C, Katahira R, Thorpe T, Ashihara H (2003). Purine and pyrimidine nucleotide metabolism in higher plants. *Journal of Plant Physiology* 160: 1271-1295.
- Sugiura M, Takeda Y (2000). Nucleic acids. In B.B. Buchanan, W. Gruissen, R.L. Jones (Eds.), *Biochemistry and Molecular Biology of Plants* (260-310 pp). Rockville: American Society of Plant Physiologists.
- Sukha DA, Butler DR, Comissiong EA, Umaharan P (2014) .The impact of processing location and growing environment on flavor in cocoa (*Theobroma cacao* L.)—implications for “Terroir” and certification. *Acta Horticulturae (International Society for Horticultural Science)* 1047: 255–262.
- Sukha DA, Butler DR, Umaharan P, Boulton E. (2008). The use of an optimised organoleptic assessment protocol to describe and quantify different flavour attributes of cocoa liquors made from Ghana and Trinitario beans. *European Food Research and Technology* 226: 405-413.
- Sukha D, Butler D (2005). The CFC/ICCO/INIAP Cocoa Flavour Project - Investigating the spectrum of fine flavour within genotypes and between origins. *INGENIC Newsletter* 10: 22-25.

- Takrama J, Cervantes-Martínez C, Philips-Mora W, Brown J, Motamayor J, Schnell R (2005). Determination of off-types in a cacao breeding program using microsatellites. *INGENIC Newsletter* 10: 2-7.
- Thamke I, Dürrschmid K, Rohma H. (2009). Sensory description of dark chocolates by consumers. *LWT-Food Science and Technology* 42: 534-539.
- Tomas-Barberán F, Cienfuegos-Jovellanos E, Marín A, Muguerza B, Gil-Izquierdo A, Cerdá B, Zafrilla P, Morillas J, Mulero J, Ibarra A, Pasamar M, Ramón D, Espín J (2007). A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. *Journal of Agricultural and Food Chemistry* 55: 3926-3935.
- Tomlins K, Baker D, Daplyn P, Adomako D (1993). Effect of fermentation and drying practices on the chemical and physical profiles of Ghana cocoa. *Food Chemistry* 46: 257-263.
- Trognitz B, Scheldeman X, Hansel-Hohl K, Kuant A, Grebe H, Hermann M (2011). Genetic population structure of cacao plantings within a young production area in Nicaragua. *PLoS One* 6: e16056.
- Tsao R (2010). Chemistry and biochemistry of dietary polyphenols. *Nutrients* 2: 1231-1246.
- Turnbull CJ, Wadsworth RM, Hadley P (2003). International Cocoa Germplasm Database. ICGD 2003, Version 5.2. Reading, UK: University of Reading.
- Uefuji H, Ogita S, Yamaguchi Y, Koizumi N, Sano H (2003). Molecular cloning and functional characterization of three distinct *N*-methyltransferase involved in the caffeine biosynthetic pathway in coffee plants. *Plant Physiology*. 132: 372-380.
- Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538.
- Vázquez-Ovando JA, Molina-Freaner F, Nuñez-Farfán J, Ovando-Medina I, Salvador-Figueroa M (2014). Genetic identification of *Theobroma cacao* L. trees with high Criollo ancestry in Soconusco, Chiapas, Mexico. *Genetics and Molecular Research* 13: 10404-10414.
- Vázquez-Ovando A, Molina-Freaner F, Nuñez-Farfán J, Salvador-Figueroa M (2012). Potencial de los marcadores moleculares para el rescate de individuos de *Theobroma cacao* L. de alta calidad. *BioTecnología* 16: 36-56.
- Vázquez-Ovando A, Rosado-Rubio G, Chel-Guerrero L, Betancur-Ancona D (2009). Physicochemical properties of a fibrous fraction from chia (*Salvia hispanica* L.). *LWT-Food Science and Technology* 42: 168-173.
- Virk PS, Ford-Lloyd BF, Jackson MT, Pooni HS, Clemenot TP, Newbury HJ (1996). Predicting quantitative variation within rice germplasm using molecular markers. *Heredity* 76: 296-304.
- Voight J (2013). Chocolate and cocoa aroma. In R Watson, V R Preedy, S Zibadi (Eds.), *Chocolate in Health and Nutrition* (chap. 7; pp 89-101). New York: Humana Press.

- Voigt J, Biehl B, Heinrich H, Kamaruddin S, Gaim-Marsoner G, Hugi A (1994) *In vitro* formation of cocoa-specific aroma precursors: aroma related peptides generated from cocoa seed protein by co-operation of an aspartic endoprotease and a carboxypeptidase. *Food Chemistry*. 49: 173-180.
- Voigt J, Biehl B, Kamaruddin S (1993) The major seed proteins of *Theobroma cacao* L. *Food Chemistry*. 47: 145-151.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.
- Warren J (1994) Isozyme variation in a number of populations of *Theobroma cacao* obtained through various sampling regimes. *Euphytica* 72: 121-126.
- Whitkus R, de la Cruz M, Mota-Bravo L, Gómez-Pompa A (1998). Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theoretical and Applied Genetics* 96: 621-627.
- Wilde J, Waugh R, Powell W (1992). Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. *Theoretical and Applied Genetics* 83: 871-877.
- Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology. *Plant Physiology*. 126: 485-493.
- 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: 423-447.
- Wood JE, Allaway D, Boulton E, Scott IM (2010). Operationally realistic validation for prediction of cocoa sensory qualities by high-throughput mass spectrometry. *Analytical Chemistry* 82: 6048-6055.
- Wood GAR, Lass RA (1985). *Cocoa*. Tropical Agriculture Series. 4th edn. Blackwell Science Logman Group Ltd., New York.
- Yeater KM, Bollero GA, Bullock DG, Rayburn AL, Rodriguez Zas S (2004). Assessment of genetic variation in hairy vetch using canonical discriminant analysis. *Crop Science* 44: 185–189.
- Zhang D, Gardini EA, Motilal LA, Baligar V, Bailey B, Zuñiga-Cernades L, Arevalo-Arevalo CE, Meinhardt L (2011). Dissecting genetic structure in farmer selections of *Theobroma cacao* in the Peruvian Amazon: implications for on farm conservation and rehabilitation. *Tropical Plant Biology* 4: 106-116.
- Zhang D, Boccara M, Motilal L, Mischke S, Johnson E, Butler D, Bailey B, Meinhardt L (2009). Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers. *Tree Genetics & Genomes* 5: 595-607.
- Zhang D, Boccara M, Motilal L, Butler D, Umaharan P, Mischke S, Meinhardt L (2008). Microsatellite variation and population structure in the 'Refractario' cacao of Ecuador. *Conservation Genetics* 9: 327-337.

- Zhang D, Arevalo-Gardini E, Mischke S, Zuñiga-Cernades L, Barreto-Chavez A, Adriaola J (2006). Genetic diversity and structure of managed and semi-natural populations of cocoa (*Theobroma cacao*) in the Huallaga and Ucayali Valleys of Peru. *Annals of Botany* 98: 647-655.
- Zhang D, Mischke S, Goenaga R, Hemeida A, Saunders J (2006). Accuracy and reliability of high-throughput microsatellite genotyping for cacao clone identification. *Crop Science* 46: 2084-2092.
- Ziegler G (1991). Composition of flavor extracts of raw and roasted beans. *European Food Research and Technology* 192: 521-525.