

**UNIVERSIDAD NACIONAL
TORIBIO RODRÍGUEZ DE MENDOZA DE AMAZONAS**



ESCUELA DE POSGRADO

**TESIS PARA OBTENER EL GRADO ACADÉMICO DE MAESTRO
EN GESTIÓN PARA EL DESARROLLO SUSTENTABLE**

**“Evaluación de la diversidad genética del cacao nativo fino de
aroma (*Theobroma cacao* L.) mediante el uso de marcadores
microsatélites”.**

Autora: Bach. Ingrid Milagros Iliquin Inga

Asesores: Ph.D. Jorge Ronny Díaz Valderrama

Ph.D. Santos Triunfo Leiva Espinoza

Registro:

CHACHAPOYAS – PERÚ

2023

AUTORIZACIÓN DE PUBLICACIÓN DE LA TESIS DE MAESTRÍA EN EL REPOSITORIO INSTITUCIONAL DE LA UNTRM



UNTRM

Reglamento del Proceso de Graduación en la Escuela de Posgrado de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas

ANEXO 6

AUTORIZACIÓN DE PUBLICACIÓN DE LA TESIS EN EL REPOSITORIO INSTITUCIONAL DE LA UNTRM

1. Datos de autor 1

Apellidos y nombres (tener en cuenta las tildes):

Iliguin Inga Ingrid Milagros

DNI N°: 73532254

Correo electrónico: ingrid.iliguin.egg@untrm.edu.pe

Nombre de la Maestría (X)/Doctorado ():

Gestión para el Desarrollo Sustentable

Datos de autor 2

Apellidos y nombres (tener en cuenta las tildes):

DNI N°: _____

Correo electrónico: _____

Nombre de la Maestría ()/Doctorado ():

2. Título de la tesis para obtener el grado académico de Maestro (X) / Doctor ()

Evaluación de la diversidad genética del cacao nativo fino de aroma (Theobroma cacao L.) mediante el uso de marcadores microsatélites.

3. Datos de Asesor

Apellidos y nombres: Díaz Valderrama Jorge Ronny

DNI, Pasaporte, C.E N°: 70205149

ORCID: <https://orcid.org/0000-0002-1015-3751>

Datos de Co-Asesor

Apellidos y nombres: Leiva Espinoza Santos Triunfo

DNI, Pasaporte, C.E N°: 41265287

ORCID: <https://orcid.org/0000-0003-1710-1994>

4. Campo del conocimiento según Organización para la Cooperación y el Desarrollo

Económicos-OCDE 1.00.00 Ciencias naturales → 1.06.00 Biología

5. Originalidad del Trabajo

Con la presentación de esta ficha, el autor o autores señalan expresamente que la obra es original, ya que sus contenidos son producto de su directa contribución intelectual. Se reconoce también que todos los datos y las referencias a materiales ya publicados están debidamente identificados con su respectivo crédito e incluidos en las notas bibliográficas y en las citas que se destacan como tal.

6. Autorización de publicación

El o los titular de los derechos de autor otorga a la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), la autorización para la publicación del documento indicado





en el punto 2, bajo la *Licencia creative commons* de tipo BY-NC: Licencia que permite distribuir, remezclar, retocar, y crear a partir de su obra de forma no comercial por lo que la Universidad deberá publicar la obra poniéndola en acceso libre en el repositorio institucional de la UNTRM y a su vez en el Registro Nacional de Trabajos de Investigación-RENATI, dejando constancia que el archivo digital que se está entregando, contiene la versión final del documento sustentado y aprobado por el Jurado Evaluador.

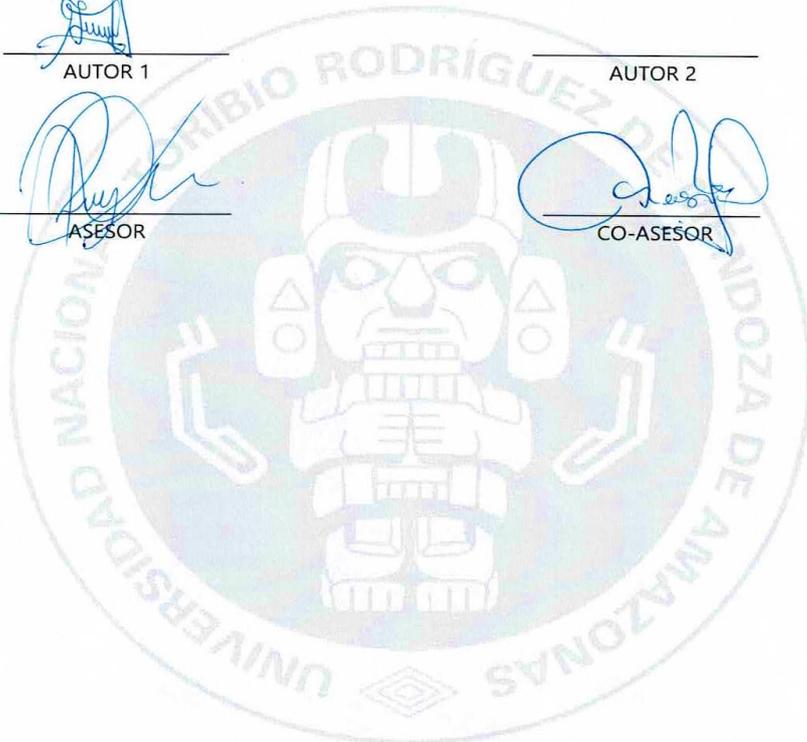
Chachapoyas, 19 de Julio de 2023

AUTOR 1

AUTOR 2

ASESOR

CO-ASESOR



DEDICATORIA

A mi familia, Esperanza, Rosa Manuela y María Luisa, por su apoyo incondicional en cada uno de los momentos de mi vida.

A mi hija Samantha Issabela Tacilla Iliquin, por ser el motivo de mi superación personal y desarrollarme en la parte profesional, y a todos mis amigos que me han brindado su apoyo, su conocimiento y tiempo en el desarrollo de la presente investigación.

Ingrid Milagros

AGRADECIMIENTO

Doy gracias a Dios por haberme otorgado una familia maravillosa, quienes me han apoyado y brindado su amor incondicional a lo largo de mi vida, dándome fortaleza en aquellos momentos difíciles para seguir adelante.

Al Instituto de Investigación para el Desarrollo Sustentable de Ceja de Selva (INDES-CES) de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM-A) por haberme dado la oportunidad de realizar este trabajo de investigación en las instalaciones del Laboratorio de Investigación en Sanidad Vegetal (LABISANV).

Al Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (FONDECYT) por el financiamiento de esta investigación en el marco de la ejecución del proyecto: “Estudio de la diversidad genética de agroecosistema del cacao nativo fino de aroma, con fines de protección del germoplasma y centro de origen, en la zona nor oriental del Perú” – GEN CACAO, con contrato N° 30-2019-FONDECYT-BM-INC.INV, ejecutado por el INDES-CES.

A mis asesores, Ph.D. Jorge Ronny Díaz Valderrama y Ph.D. Santos Triunfo Leiva Espinoza, por sus orientaciones, confianza, paciencia y apoyo incondicional en la ejecución y redacción del presente trabajo de investigación.

A los docentes miembros del jurado, Ph.D. Danilo Edson Bustamante Mostajo, Dr. Segundo Manuel Oliva Cruz, M.Sc. Elí Pariente Mondragón por sus oportunas y valiosas sugerencias.

Al Sr. Marco Antonio Pasapera Alvitres, por su valiosa asistencia en las colectas y proceso de recolección de muestras en campo, por sus enseñanzas y conocimientos compartidos.

A todas las personas que me brindaron ánimos, aliento, sugerencias y apoyo incondicional, inspirándome a seguir adelante.

Ingrid Milagros

**AUTORIDADES DE LA UNIVERSIDAD NACIONAL TORIBIO RODRÍGUEZ
DE MENDOZA DE AMAZONAS**

Ph.D. JORGE LUIS MAICELO QUINTANA

Rector

Dr. OSCAR ANDRÉS GAMARRA TORRES

Vicerrector Académico

Dra. MARÍA NELLY LUJÁN ESPINOZA

Vicerrectora de Investigación

Dr. EFRAÍN MANUELITO CASTRO ALAYO

Director de la Escuela de Posgrado de la UNTRM

JURADO EVALUADOR DE LA TESIS



Ph.D. DANILO EDSON BUSTAMANTE MOSTAJO

Presidente



Dr. SEGUNDO MANUEL OLIVA CRUZ

Secretario



M.Sc. ELÍ PARIENTE MONDRAGÓN

Vocal

CONSTANCIA DE ORIGINALIDAD DE LA TESIS



UNTRM

Reglamento del Proceso de Graduación en la Escuela de Posgrado de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas

ANEXO 3

CONSTANCIA DE ORIGINALIDAD

Los suscritos, miembros del Jurado Evaluador del Proyecto de Tesis ()/Tesis (X) titulado: Evaluación de la diversidad genética del cacao nativo fino de aroma (Theobroma cacao L.) mediante el uso de marcadores microsatélites, presentado por el Aspirante Ingrid Milagros Iliguin Inga para obtener el Grado Académico de Maestro (X)/Doctor () en Gestión para el Desarrollo Sustentable de la Escuela de Posgrado de la UNTRM, hacemos constar que después de revisar la originalidad del Proyecto de Tesis ()/Tesis (X)/Tesis en formato de artículo científico () con el software de prevención de plagio **Turnitin**, verificamos:

a) De acuerdo con el informe de originalidad (adjunto), el Proyecto de Tesis ()/Tesis (X)/ Tesis en formato de artículo científico () tiene 17 % de similitud, que es menor al 25% permitido en la UNTRM.

b) La persona responsable de someter el trabajo al software de prevención de plagio **Turnitin** fue:

Danilo Edson Bustamante Mostajo, y pertenece al área () / oficina () / dependencia (X) de Facultad de Ingeniería Civil y Ambiental.



Chachapoyas, 27 de Junio del 2023

SECRETARIO

PRESIDENTE

VOCAL

OBSERVACIONES:

.....
.....

ACTA DE SUSTENTACIÓN DE LA TESIS



UNTRM

Reglamento del Proceso de Graduación en la
Escuela de Posgrado de la Universidad Nacional
Toribio Rodríguez de Mendoza de Amazonas

ANEXO 5

ACTA DE SUSTENTACIÓN DE TESIS

En el lugar Auditorio EPG. de la ciudad de Chachapoyas, el día 05 de Julio del año 2023, siendo las 16.00 horas, el Aspirante Ingrid Milagros Eliquin Inga, defiende en sesión pública presencial la Tesis titulada: Evaluación de la diversidad genética del Cacao Nativo Fino de Aroma (Theobroma cacao L.) mediante el uso de marcadores microsatélites. para obtener el Grado Académico de Maestro (X)/Doctor () en Gestión para el desarrollo sustentable, a ser otorgado por la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas; ante el Jurado Evaluador, conformado por:

Presidente: Ph. D. Danilo Edson Bustamante Hostajo
Secretario: Dr. Segundo Manuel Oliva Cruz
Vocal: M.Sc. Eli Poriente Hondragón

Procedió el aspirante a hacer la exposición de la Introducción, Material y método, Resultados, Discusión y Conclusiones, haciendo especial mención de sus aportaciones originales. Terminada la defensa de la Tesis, los miembros del Jurado Evaluador pasaron a exponer su opinión sobre la misma, formulando cuantas cuestiones y objeciones consideraron oportunas, las cuales fueron contestadas por el aspirante.

Tras la intervención de los miembros del Jurado Evaluador y las oportunas respuestas del aspirante, el Presidente abre un turno de intervenciones para los presentes en el acto de sustentación, para que formulen las cuestiones u objeciones que consideren pertinentes. Seguidamente, a puerta cerrada, el Jurado Evaluador determinó la calificación global concedida a la sustentación de la Tesis, en términos de:

Aprobada (✓)/Desaprobada () por Unanimidad (✓)/Mayoría ().

Otorgada la calificación, el Secretario del Jurado Evaluador lee la presente Acta en esta misma sesión pública. A continuación se levanta la sesión.

Siendo las 18:20 horas del mismo día y fecha, el Jurado Evaluador concluye el acto de sustentación de la Tesis.




SECRETARIO


VOCAL


PRESIDENTE

OBSERVACIONES:

.....

.....

ÍNDICE O CONTENIDO GENERAL

AUTORIZACIÓN DE PUBLICACIÓN DE LA TESIS DE MAESTRÍA EN EL REPOSITORIO INSTITUCIONAL DE LA UNTRM	ii
DEDICATORIA	iv
AGRADECIMIENTO	v
AUTORIDADES DE LA UNIVERSIDAD NACIONAL TORIBIO RODRÍGUEZ DE MENDOZA DE AMAZONAS	vi
JURADO EVALUADOR DE LA TESIS	vii
CONSTANCIA DE ORIGINALIDAD DE LA TESIS	viii
ACTA DE SUSTENTACIÓN DE LA TESIS	ix
ÍNDICE O CONTENIDO GENERAL	x
ÍNDICE DE TABLAS	xii
ÍNDICE DE FIGURAS	xiii
RESUMEN	xvi
ABSTRACT	xvii
I. INTRODUCCIÓN	18
II. MATERIALES Y MÉTODOS	21
2.1. Área de estudio	21
2.2. Población, muestra y muestreo	23
2.2.1. Población	23
2.2.2. Muestra	23
2.2.3. Muestreo	23
2.3. Metodología	26
2.3.1. Diseño de marcadores microsatélites SSR	26
2.3.1.1. Detección de los loci SSR utilizados para caracterizar los grupos genéticos	26
2.3.1.2. Selección de loci SSR nuevos para caracterizar la diversidad de cacao	27
2.3.2. Caracterización molecular de la diversidad genética de cacao	27
2.3.2.1. Identificación, selección y colecta de muestras	27
2.3.2.2. Preparación de las muestras	30
2.3.2.3. Homogenización de las muestras	31
2.3.2.4. Extracción de ADN	31
2.3.2.5. Cuantificación de ADN	31
2.3.2.6. PCR y análisis de fragmentos	31
2.3.2.7. Electroforesis del ADN producto de la amplificación	32

2.4. Análisis de datos	33
2.4.1. Análisis de fragmentos y determinación de polimorfismo alélico	33
2.4.2. Análisis de diversidad genética	33
2.4.3. Desequilibrio de ligamiento	33
2.4.4. Análisis de la estructura genética	33
2.4.5. Análisis de variancia molecular (AMOVA)	34
2.4.6. Análisis de asociación	34
III. RESULTADOS	35
3.1. Diseño de marcadores microsatélites	35
3.2. Evaluación de la diversidad genética	42
3.2.1. Polimorfismo alélico en marcadores microsatélite SSR	42
3.2.2. Diversidad genética	43
3.2.3. Desequilibrio de ligamiento	45
3.2.4. Estructura genética	49
3.2.5. Análisis de asociación	50
3.2.5.1. Red de extensión mínima	50
3.2.5.2. Análisis de componentes principales (ACP)	51
3.2.5.3. Método de grupo de pares no ponderado con media aritmética (UPGMA)	52
3.3. Evaluación de la diversidad genética de acuerdo a la altitud	54
3.3.1. Diversidad genética	54
3.3.2. Análisis de variancia molecular (AMOVA)	54
3.4. Análisis de componentes principales (ACP) de accesiones de la colección INDES-CES	55
IV. DISCUSIÓN	58
V. CONCLUSIONES	64
VI. RECOMENDACIONES	65
VII. REFERENCIAS BIBLIOGRÁFICAS	66
ANEXOS	76

ÍNDICE DE TABLAS

Tabla 1. Muestras de CNFA de acuerdo a la provincia de muestreo.....	23
Tabla 2. Descriptores morfológicos del fruto y de productividad de las accesiones de cacao.	24
Tabla 3. Información pasaporte de las accesiones de CNFA.....	28
Tabla 4. Condiciones de PCR para la amplificación de muestras de CNFA.	32
Tabla 5. <i>Primers</i> para la amplificación de regiones SSR de <i>Theobroma cacao</i> homólogos a los utilizados por Motomayor <i>et al.</i> (2008).....	36
Tabla 6. <i>Primers</i> diseñados en este estudio para la amplificación de regiones SSR de <i>Theobroma cacao</i>	38
Tabla 7. Alelos observados en marcadores microsatélite SSR <i>Theobroma cacao</i>	43
Tabla 8. Diversidad genética <i>Theobroma cacao</i> basada en el análisis de marcadores SSR.	44
Tabla 9. Diferenciación genética entre provincias en base al índice de fijación G_{st} y el índice de diferenciación D con el conjunto de datos proporcionados por los marcadores SSR	44
Tabla 10. Tabla de diversidad por población con accesiones de cacao de acuerdo con la altitud	54
Tabla 11. Análisis de variancia molecular (AMOVA) generada para accesiones de <i>Theobroma cacao</i>	55
Tabla 12. Distancia genética utilizado en análisis de estructura genética y análisis de varianza molecular (AMOVA).	55
Tabla 13. Análisis de variancia molecular (AMOVA) generada para 20 accesiones de <i>Theobroma cacao</i> de la colección INDES-CES.	57
Tabla 14. Distancia genética utilizado en análisis de estructura genética y análisis de varianza molecular (AMOVA).	57

ÍNDICE DE FIGURAS

Figura 1. Mapa de distribución de accesiones de Cacao Nativo Fino de Aroma (CNFA) procedentes de la región Amazonas.....	22
Figura 2. Ubicación de los primers para amplificar regiones SSR utilizados en este estudio utilizando como fondo la representación gráfica de los cromosomas del genoma de cacao del grupo genético Criollo (accesión B97-61/B2) (Argout <i>et al.</i> , 2017), distribuido bajo la licencia internacional Creative Commons Attribution 4.0 (http://creativecommons.org/licenses/by/4.0/). Las flechas amarillas muestran la ubicación aproximada de los SSR.	41
Figura 3. Productos PCR de marcadores moleculares SSR visualizados en un gel de agarosa al 1%. Se muestra el análisis de cada marcador con 5 accesiones de cacao y un control, corridos de manera paralela a una escalera molecular de 10,000 pb. Los marcadores INDES-mTC12 e INDES-mTC6, indican una correcta amplificación. El marcador INDES-mTC1 indica una PCR nula.	42
Figura 4. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra un locus heterocigoto en la accesión de cacao INDES-6 (colorante 6-FAM TM , azul) para el marcador INDES-mTC12.	42
Figura 5. Curva de genotipos multilocus respecto al número de loci más informativos	45
Figura 6. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.226$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i>	46
Figura 7. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.364$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i> de la provincia de Chachapoyas.	46
Figura 8. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.376$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i> de la provincia de Rodríguez de Mendoza.	47
Figura 9. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.0346$) (línea azul). Este histograma indica el resultado	

de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i> de la provincia de Condorcanqui.....	47
Figura 10. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.172$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i> de la provincia de Bagua y Utcubamba.....	48
Figura 11. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.0213$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de <i>Theobroma cacao</i> de la provincia de Luya.	48
Figura 12. Gráfico de Sedimentación (<i>Scree test</i>) del porcentaje de varianza explicada por cada componente en un PCA de los datos genéticos. El codo del diagrama indica que hay alrededor de $K = 4$ componentes en los datos.....	49
Figura 13. Valores el criterio de entropía cruzada en función del número de factores en las ejecuciones Modelos Mixtos de Factores Latentes (LFMM). Se aprecia que $K = 4$ es el número de poblaciones ancestrales óptimo con el menor valor de entropía cruzada.	49
Figura 14. Estructura poblacional de <i>Theobroma cacao</i> basada en 16 marcadores SSR, utilizando $K = 4$ como número óptimo de hipotéticas poblaciones ancestrales, que tienen su propio color en el gráfico. Cada una de las accesiones de cacao está representada por una barra vertical coloreada según su coeficiente de mezcla genética para cada población ancestral.	50
Figura 15. Red de extensión mínima de <i>Theobroma cacao</i> . Los nodos (círculos) representan genotipos multilocus (MLG); el grosor de línea de conexión entre nodos, la relación basada en el número de alelos diferentes (datos de repetición de secuencia simple [SSR]). El grosor de las líneas se escala proporcionalmente a la inversa de la distancia genética observada, es decir, cuanto más gruesa es la línea, más relacionados están los nodos conectados. Las longitudes de línea son arbitrarias. Cada provincia está indicada por colores.....	51
Figura 16. Análisis de asociación de muestras de 96 accesiones de <i>Theobroma cacao</i> L., basado en 16 marcadores microsatélite SSR. Análisis de componentes principales (PCA) que representa los dos primeros componentes principales acumulando un total del 35.00% de la diversidad. Cada provincia está indicada por colores. Las elipses indican la distribución de las muestras. Las líneas indican el valor del coeficiente de esa variable en ese componente. Cada punto representa una muestra.....	52

Figura 17. Dendograma UPGMA de la diversidad genética de 96 accesiones de *Theobroma cacao*, construido en base a las distancias genéticas de Nei, evaluados con 16 marcadores microsatélite SSR. 53

Figura 18. Análisis de asociación de muestras de 20 accesiones de *Theobroma cacao* L. de la colección INDES-CES, basado en 16 marcadores microsatélite SSR. Análisis de componentes principales (PCA) que representa los dos primeros componentes principales acumulando un total del 36.48% de la diversidad. Cada grupo está indicado por colores. Las elipses indican la distribución de las muestras. Las líneas indican el valor del coeficiente de esa variable en ese componente. Cada punto representa una muestra..... 56

RESUMEN

El cacao (*Theobroma cacao* L.) de Amazonas es genéticamente diverso y posee propiedades sensoriales superiores. Estudios anteriores han sugerido la influencia de la altitud sobre las propiedades sensoriales, pero se desconoce si la genética está influenciada por la altitud. El objetivo de este estudio fue analizar la diversidad genética del cacao de Amazonas, con énfasis en áreas poco exploradas, y evaluar si esta diversidad genética está influenciada por la altitud. Se colectaron y genotiparon 96 accesiones de cacao de seis provincias de Amazonas con 16 marcadores microsatélites. Los índices de diversidad revelaron un perfil genético único para las 96 accesiones, y se encontró que la diversidad genética no estaba influenciada por la altitud. Además, se observó desequilibrio de ligamiento genético, propia de la propagación vegetativa impulsada por el hombre, en todas las provincias excepto en Luya, donde los genotipos encontrados derivaron de eventos de reproducción sexual, probablemente ancestrales. Las poblaciones de Luya y Chachapoyas tuvieron mayores valores de diferenciación genética a las de Condorcanqui, Bagua y Utcubamba, provincias de mayor producción en Amazonas, lo que revela la unicidad del cacao en estas provincias. Las muestras de Luya y Chachapoyas provienen del ecosistema Bosques Tropicales Estacionalmente Secos del Marañón, lo que hace de esta área protegida el hogar de un cacao único. Estos resultados proporcionan la base para la implementación de estrategias de conservación y aprovechamiento máximo del potencial genético del cacao en Amazonas.

Palabras clave: Cacao, genotipos, microsatélites, polimorfismo alélico.

ABSTRACT

Cocoa (*Theobroma cacao* L.) from Amazonas is genetically diverse and possesses superior sensory properties. Previous studies have suggested the influence of altitude on sensory properties, but it is unknown whether genetics are influenced by altitude. The objective of this study was to analyze the genetic diversity of Amazonian cacao, with emphasis on poorly explored areas, and to evaluate whether this genetic diversity is influenced by altitude. Ninety-six cacao accessions from six provinces of Amazonas were collected and genotyped with 16 microsatellite markers. Diversity indices revealed a unique genetic profile for the 96 accessions, and genetic diversity was found not to be influenced by altitude. In addition, genetic linkage disequilibrium, typical of human-driven vegetative propagation, was observed in all provinces except Luya, where the genotypes found were derived from sexual, probably ancestral, reproductive events. The populations of Luya and Chachapoyas had higher values of genetic differentiation than those of Condorcanqui, Bagua and Utcubamba, provinces of higher production in Amazonas, which reveals the uniqueness of cacao in these provinces. The samples from Luya and Chachapoyas come from the Marañon Seasonally Dry Tropical Forest ecosystem, making this protected area home to a unique cacao. These results provide the basis for the implementation of strategies to conserve and maximize the genetic potential of cacao in Amazonas.

Key words: Cacao, genotypes, microsatellites, allelic polymorphism.

I. INTRODUCCIÓN

El cacao (*Theobroma cacao* L.) es un cultivo originario de Sudamérica de gran importancia económica para muchos países de zonas tropicales (Argout *et al.*, 2017). La evidencia más antigua de su uso y domesticación datan de hace aproximadamente 5,300 años en las ruinas de la cultura Mayo-Chinchipec, al sureste del Ecuador en el Alto Amazonas (Zarrillo *et al.*, 2018). Por su parte, Olivera-Núñez (2018) menciona que posiblemente el templo de la Huaca Montegrande, ruinas también de la cultura Mayo-Chinchipec, ubicado en la provincia de Jaén, Perú, guarde vestigios de la existencia del cacao más antiguo del mundo, lo cual indica que el cacao también fue cultivado y consumido en el norte del país. Por lo tanto, además de económica, el cacao tiene una importancia histórica para el país (Díaz-Valderrama *et al.*, 2020).

La Amazonía peruana alberga una amplia diversidad y variabilidad genética de cacao (García-Carrión, 2008; Thomas *et al.*, 2012). El 60% de la biodiversidad de cacao existente en el mundo (material genético) se encuentra en Perú (Ministerio de Desarrollo Agrario y Riego [MIDAGRI], 2020). Además, el Perú es considerado uno de los principales productores y proveedores de cacao fino, siendo el segundo país productor mundial de cacao orgánico después de Ecuador (Camacho, 2019). La producción nacional de cacao en grano viene incrementándose desde hace diez años y crece a una tasa anual de 12.6% (MIDAGRI, 2022). Sin embargo, a pesar de su crecimiento y alta diversidad del cultivo, se reportan pocos estudios sobre los ecotipos nativos de cacao existentes y su diversidad. Esto origina que la base genética del cultivo esté basado principalmente en la introducción de clones o variedades mejoradas, cuyos atributos sensoriales tienden a ser inferiores (Bartley, 2005).

En el Perú, el caso más conocido es el híbrido Colección Castro Naranjal #51 (CCN-51), introducido desde Ecuador, por sus características de alta productividad, y resistencia a plagas y enfermedades (Díaz-Valderrama *et al.*, 2020). Existen reportes, que, en Perú, sólo el 44% del área total del cacao cultivado es nativo y el 56% se encuentran cultivadas con CCN-51 (Ceccarelli *et al.*, 2022). Por lo tanto, el cacao CCN-51 representa una fuerte amenaza de erosión genética para el material genético nativo debido al reemplazo de plantaciones y a cruces no controladas (Bartley, 2005; Solórzano *et al.*, 2012; Lehmann, 2015; Viena-Vásquez, 2020). Asimismo, el cacao CCN-51 generalmente es cultivado bajo sistemas de producción intensiva, siendo poco amigable con el medio ambiente y sin

un enfoque sostenible. En este contexto, debido a la amplia diversidad genética del cacao en Amazonas, Perú (Bustamante *et al.*, 2022), urge la necesidad de desarrollar un sistema estratégico de producción, tomando un modelo de agricultura sostenible, planteando estrategias de inserción en mercados diferenciados, que permitan obtener mejoras económicas en todos los eslabones de la cadena de valor, principalmente en los pequeños productores (García-Carrión, 2008).

La región Amazonas en 2018 registró una superficie agrícola de 520,995.42 ha, donde el cultivo de cacao abarcó una superficie cosechada de 7,856.50 ha, mostrando un notable incremento de 7,873.0 ha para el 2019; involucrando a 11,644 productores (SIEA, 2022). Por lo tanto, en Amazonas, el cultivo de cacao es de mucha importancia para la economía familiar del ámbito rural. Los pequeños productores, cultivan este producto como parte de una unidad de producción diversificada (Goñas *et al.*, 2022). El Cacao Nativo Fino de Aroma (CNFA) de la región Amazonas es de alta demanda en el mercado internacional, debido a sus características de aroma y sabor, siendo uno de los diez productos peruanos con denominación de origen, protegido bajo el nombre de “Cacao Amazonas Perú” (INDECOPI, 2016; Oliva-Cruz, 2020). Esto otorga una ventaja a los productores de Amazonas para posicionarse en el mercado nacional y mundial como producto diferenciado. Por lo tanto, la conservación de los recursos genéticos y utilización de la diversidad genética existente, son esenciales para el cultivo sostenible del cacao (Laliberté *et al.*, 2018).

Muchas investigaciones han caracterizado la diversidad genética del cacao utilizando principalmente marcadores microsatélites o “*Simple Sequence Repeats*” (SSRs), o de “*Single Nucleotide Polymorphism*” (SNPs), con fines de protección, conservación y mejoramiento genético (Avenidaño-Arrazate *et al.*, 2018; Arevalo-Gardini *et al.*, 2019; Ricaño-Rodríguez *et al.*, 2019; Nieves-Orduña *et al.*, 2021; Bustamante *et al.*, 2022). Motamayor *et al.* (2008), en su investigación sobre la diferenciación poblacional geográfica y genética del cacao, genotiparon 1,241 accesiones de cacao con 106 marcadores microsatélites, encontrando 10 grupos genéticos: Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañon, Nacional, Nanay y Purús. Por su parte, Zhang *et al.* (2012) genotiparon 164 accesiones del Cacao Nacional Boliviano “CNB” con 15 marcadores SSR, para el cual incluyeron accesiones cultivadas y silvestres de las regiones amazónicas de La Paz y Beni. Estos investigadores realizaron un análisis de varianza molecular y encontraron que el CNB tiene un perfil genético único que es

significativamente diferente a los 10 grupos de germoplasma de cacao conocidos en Sudamérica (Motamayor *et al.*, 2008), y que las poblaciones cultivadas de CNB y las silvestres de CNB en el río Beni comparten un perfil genético similar, sugiriendo que el CNB cultivado es originario de Bolivia.

Ecotipos nativos de cacao del Perú también han sido analizados molecularmente. Zhang *et al.* (2006) caracterizaron 105 accesiones de cacao de los valles del Huallaga y Ucayali con 15 marcadores microsatélites, quienes encontraron que las accesiones de Ucayali presentaban una diversidad genética más amplia que las de Huallaga, corroborando que la diversidad genética del cacao tiene una estructura espacial, donde la adopción de medidas de conservación es importante. Por otra parte, Arevalo-Gardini, *et al.* (2019) estudiaron la identidad genética y origen del cacao nativo Piura Porcelana. Para ello compararon, a través del análisis de 96 marcadores SNP, la identidad genética del cacao Piura Porcelana con 11 grupos de germoplasma de cacao conocidos existentes en los bancos de genes de cacao *ex situ* (Turnbull y Hadley, 2016). Arevalo-Gardini, *et al.* (2019) encontraron que el cacao Piura Porcelana comparte la misma pertenencia poblacional general con el cacao Nacional de Ecuador a pesar de tener diferencias genéticas significativas, lo que le sugiere un estatus genético único.

A pesar de todos los esfuerzos por caracterizar la diversidad genética de cacao en Perú, la caracterización molecular de la base genética del CNFA recién está siendo tomada en cuenta. Bustamante *et al.* (2022) evaluaron la diversidad genética y estructura poblacional del CNFA (Bagua, Condorcanqui, Jaén, Mariscal Cáceres y Utcubamba; principalmente la Región Amazonas). Estos investigadores caracterizaron 143 árboles de cacao con 192 marcadores SNP. Los 143 árboles de cacao se compararon con 1,838 perfiles de cacao de referencia del International Cocoa Genebank Trinidad. Se encontraron que 134 árboles de cacao eran muestras únicas y 104 muestras tenían al menos un 30% de ascendencia del grupo genético Nacional, siendo este el bagaje genético predominante en el CNFA con algo de contribución de Criollo (Bustamante *et al.*, 2022).

El CNFA también ha sido caracterizado morfológicamente. Oliva-Cruz *et al.* (2022), utilizando descriptores morfológicos y sensoriales, identificaron y seleccionaron 146 accesiones de CNFA de tres regiones del Perú (Cajamarca, Amazonas y San Martín). Las accesiones seleccionadas cumplieron exigencias de vigorosidad, tolerancia a plagas y enfermedades, historial de rendimiento y calidad. Estos investigadores establecieron 5

grupos morfológicos: los Toribianos, los INDES, los Bagüinos, los Utkus y los Cajas. Dentro de estos grupos, los INDES y los Bagüinos muestreados a una altitud mayor a los 500 m.s.n.m son los que reportaron mejores características sensoriales y rendimiento (Oliva-Cruz *et al.*, 2022). Por lo tanto, la altitud puede influir en las características morfológicas y los atributos sensoriales de las accesiones de cacao en Amazonas. Sin embargo, hasta ahora no se ha estudiado la influencia de la altitud en la diversidad genética del cacao. Por otro lado, aún existen centros de producción de cacao en Amazonas poco explorados (como la provincia de Condorcanqui) y otras zonas como los distritos de Providencia, Ocalli, Pisuquia y Camporredondo pertenecientes a la provincia de Luya, que no han sido considerados en estudios a pesar de su producción cacaotera (SIEA, 2022). En este contexto, el objetivo general de este estudio fue evaluar la diversidad genética del CNFA en la región Amazonas muestreadas en dos altitudes (menor a los 500 m.s.n.m y mayor a los 500 m.s.n.m). Los objetivos específicos comprendieron el diseño de marcadores moleculares tipo microsatélite (SSR) y la evaluación de la diversidad genética de cacao analizando 16 marcadores moleculares presentes en el ADN genómico y del cloroplasto.

II. MATERIALES Y MÉTODOS

2.1. Área de estudio

Se utilizaron muestras de tejido vegetal de *T. cacao* procedentes de diferentes puntos geográficos de la región Amazonas, principalmente en áreas poco exploradas para la diversidad del cacao. Amazonas se encuentra en el nororiente del Perú entre la cordillera andina y la llanura amazónica. Las colectas del tejido vegetal fueron realizadas durante el mes de abril de 2021 y abril de 2022. Las áreas muestreadas fueron lugares de producción de cacao a una altitud mayor a los 500 m.s.n.m. en las provincias de Chachapoyas (distrito de Balsas), Luya (Providencia, Ocalli, Pisuquia, Camporredondo), Rodríguez de Mendoza (Chirimoto), Bagua (Copallín, El Parco) y Utcubamba (Cumba, Cajaruro, Lonya Grande, Bagua Grande). También se colectaron muestras en áreas ubicadas a menos a los 500 m.s.n.m en las provincias de Condorcanqui (Nieva, El Cenepa y Río Santiago) y Utcubamba (Bagua Grande) (Figura 1, Tabla 1). Asimismo, para comparar las accesiones de CNFA se incluyeron dos muestras de CCN-51 provenientes de la provincia de Bagua y Condorcanqui (Tabla 1).

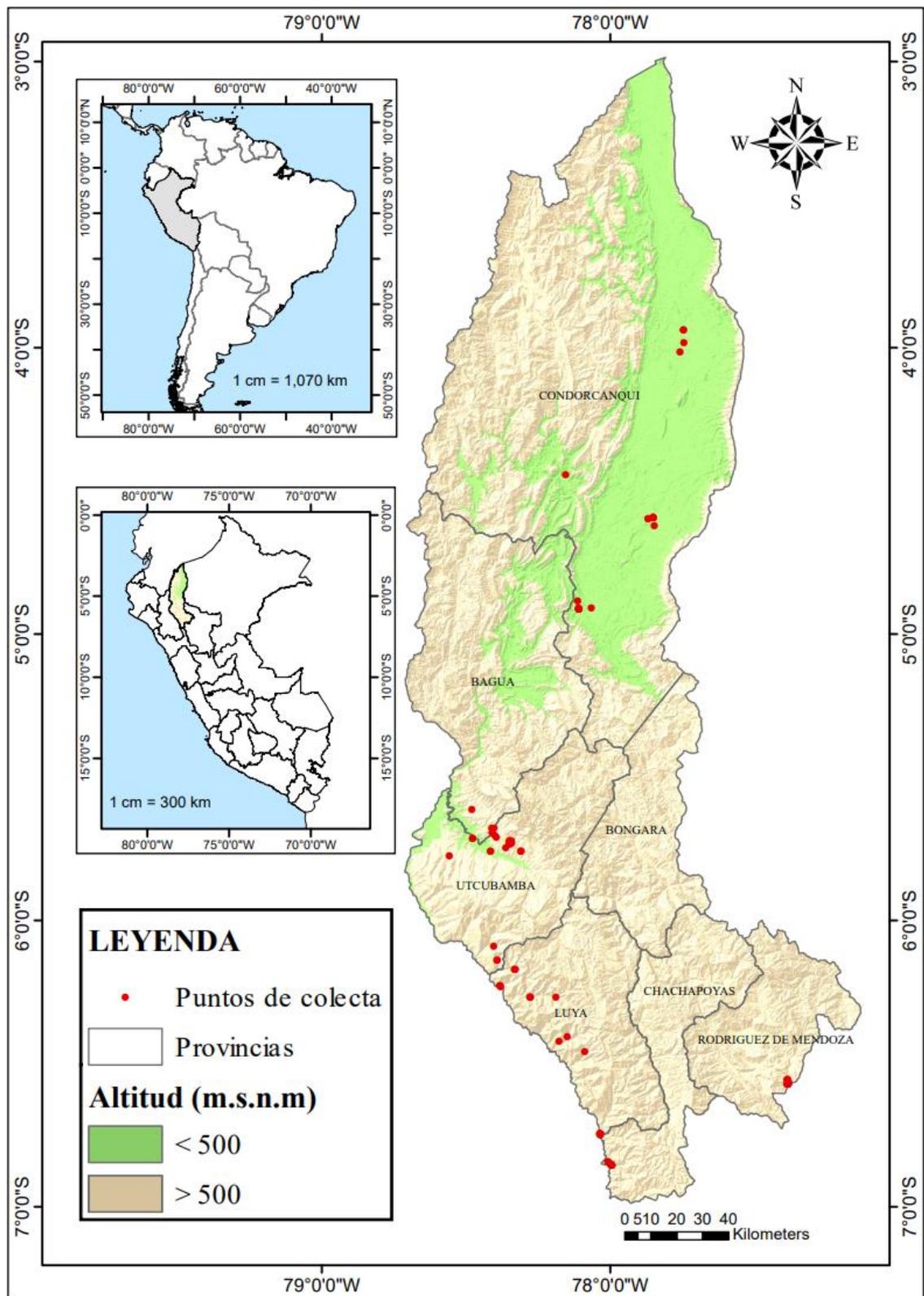


Figura 1. Mapa de distribución de accesiones de Cacao Nativo Fino de Aroma (CNFA) procedentes de la región Amazonas.

2.2. Población, muestra y muestreo

2.2.1. Población

La población estuvo constituida por todas las accesiones de CNFA de las provincias cacaoteras de la región Amazonas.

2.2.2. Muestra

Se trabajó con una muestra de 96 accesiones de CNFA (Figura 1), de las cuales 20 fueron de la colección INDES-CES procedentes de la provincia de Bagua y Utcubamba (Oliva-Cruz *et al.*, 2022), 19 de la provincia de Chachapoyas (distrito de Balsas), 21 de la provincia de Condorcanqui (Nieva, El Cenepa, Río Santiago), 12 de Luya (distrito de Providencia, Ocalli, Pisuquia, Camporredondo), 16 de Rodríguez de Mendoza (Chirimoto), 3 de Utcubamba que no pertenecen a la colección INDES-CES (distrito de Lonya Grande), 3 muestras de cacao identificados como Cacao Amazonas Perú “CAP” (INDECOPI, 2016) y 2 muestras de CCN-51 (uno procedente de la provincia de Bagua “CCN-51” y otro de la provincia de Condorcanqui “CCN-51-C”). Las muestras fueron agrupadas a nivel de provincias:

Tabla 1. Muestras de CNFA de acuerdo a la provincia de muestreo.

Nº	Provincia	Lugar de procedencia /Distrito	Número de muestras
1	Chachapoyas	Balsas	19
2	Luya	Providencia, Ocalli, Pisuquia y Camporredondo	12
3	Rodríguez de Mendoza	Chirimoto	16
4	Condorcanqui	Nieva, El Cenepa y Río Santiago	22*
5	Bagua y Utcubamba	<i>Bagua:</i> Copallín, El Parco <i>Utcubamba:</i> Cumba, Bagua Grande, Cajaruro, Lonya Grande	27**
Total			96

* incluido la muestra de cacao CCN-51-C

**incluido la muestra de cacao CCN-51

2.2.3. Muestreo

Se colectaron muestras de hojas de cacao. El muestreo fue de tipo no probabilístico. Las muestras fueron tomadas de acuerdo a las características morfológicas del fruto y productivas recomendadas en Oliva-Cruz (2020) (Tabla 2; Anexo 1). Por otro lado, las muestras que no tenían frutos se eligieron de acuerdo a las sugerencias de diferenciación morfológica descritas por los cacaoteros de cada localidad muestreada.

Tabla 2. Descriptores morfológicos del fruto y de productividad de las accesiones de cacao.

Nº	Código	Clasificación de Oliva-Cruz (2020)	Características del fruto											Características de la productividad			
			Color del fruto inmaduro				Forma				Rugosidad de la superficie				Tamaño del fruto		
			1: verde	2: verde pigmentado	3: rojo	4: rojo pigmentado	1: oblongo	2: elíptico	4: esférico	5: oblato	0: ausente	1: ligero	5: intermedio	7: fuerte	1: Grande: (20-24 cm)	2: Mediano: (15-19 cm)	3: Pequeño: (10-14 cm)
1	INDES-4	Los Toribianos	X					X					X				X
2	INDES-32		X					X					X				X
3	INDES-35		X					X					X				X
4	INDES-80		X					X					X				X
5	INDES-14	Los INDES	X					X					X				X
6	INDES-27		X					X					X				X
7	INDES-50		X				X						X				X
8	INDES-63		X					X					X				X
9	INDES-6	Los Bagüinos	X					X					X				X
10	INDES-31		X					X					X				X
11	INDES-53		X					X					X				X
12	INDES-70		X					X					X				X
13	INDES-64	Los Utkus				X			X				X				X
14	INDES-98					X			X				X				X
15	INDES-99					X		X		X			X				X
16	INDES-100					X		X		X			X				X
17	INDES-12	Los Cajas			X				X				X				X
18	INDES-38		X						X				X				X
19	INDES-78		X						X				X				X
20	INDES-81					X			X				X				X
21	TCCH-2	SI	X						X				X				X
22	TCCH-3		X						X				X				X
23	TCCH-6		X						X				X				X
24	TCCH-8		X						X				X				X

2.3. Metodología

El presente trabajo de investigación inició con actividades de campo, realizando la recolección de muestras de tejidos foliares de *T. cacao* de accesiones de diferentes zonas cacaoteras de la región Amazonas (Figura 1; Anexo 2). Las demás actividades como el diseño de marcadores SSR, procedimientos de biología molecular, y el análisis bioinformático de datos se realizó en el Laboratorio de Investigación en Sanidad Vegetal (LABISANV) del Instituto de Investigación para el Desarrollo Sustentable de Ceja de Selva (INDES-CES) de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM-A), Perú (Anexos 3–5).

2.3.1. Diseño de marcadores microsatélites SSR

El diseño de marcadores SSR se basó en dos genomas de cacao públicos: uno de la accesión B97-61/B2 proveniente de Belice del grupo genético Criollo y otro del cultivar Matina 1-6 de Costa Rica del grupo genético Amelonado (Argout *et al.*, 2017; Motamayor *et al.*, 2013). Primero, el genoma de B96-61/B2 se analizó con el programa QDD 3.1 (Megléczy *et al.*, 2014), que sirvió para detectar *in silico* todas las regiones SSR presentes en dicho genoma. Luego, se utilizó el mismo programa QDD para diseñar *primers* o cebadores para amplificar, mediante PCR, las regiones SSR identificadas. Los *primers* diseñados tuvieron la misma temperatura de fusión, lo que facilitó la estandarización del termo ciclado de PCR para cada región SSR.

2.3.1.1. Detección de los loci SSR utilizados para caracterizar los grupos genéticos

Se utilizó la información de secuencia de los *primers* de amplificación PCR de cada uno de los 106 marcadores SSR utilizados en la caracterización de cacao donde se establecieron los grupos genéticos (Motamayor *et al.*, 2008). Para saber en qué cromosoma (de los 10 que tiene el cacao) se ubica cada región SSR en el genoma del cultivar criollo B96-61/B2, se utilizó el programa BLASTN 2.8.1+ (Camacho *et al.*, 2009) de manera local, con la opción de búsqueda optimizada para secuencias cortas ('blastn-short'), debido a que las secuencias de *primers* tienen alrededor de 18 a 20 nucleótidos (Anexo 4). Una vez que se obtuvieron las ubicaciones de los loci SSR, se identificaron los nuevos *primers* diseñados anteriormente por QDD para esa misma región SSR del genoma de cacao. Luego, usando la misma búsqueda 'blastn-short', se evaluaron si las secuencias de los *primers* diseñados por QDD homólogos a los *primers* originales (Motamayor *et al.*, 2008), también estaban presentes en el otro genoma público de cacao, el del cultivar amelonado Matina 1-6 (Motamayor *et al.*, 2013) (Anexo 5), con

el fin de verificar que los nuevos *primers* tengan el potencial de amplificar en ambos genomas. Esto nos permitió cubrir la diversidad del cacao de la mejor manera posible, y de tener mayor probabilidad de éxito a la hora de realizar el análisis en laboratorio con las muestras colectadas. Seguidamente, se seleccionaron 20 loci SSR, de tal manera que haya dos por cada uno de los diez cromosomas que posee el cacao (Argout *et al.*, 2017; Motamayor *et al.*, 2013), distanciados por lo menos 5 Mb (millones de bases o nucleótidos) entre sí.

2.3.1.2. Selección de loci SSR nuevos para caracterizar la diversidad de cacao

A partir de las regiones SSR identificadas en el genoma del cultivar criollo B96-61/B2 y de los *primers* diseñados por QDD, se seleccionaron otros 21 loci SSR adicionales, nuevos y diferentes a los de Motamayor *et al.* (2008). Además de considerar los mismos criterios anteriormente descritos (dos por cromosoma con una distancia mínima de 5 Mb), se seleccionaron los loci con mayor número de repeticiones de ADN repetitivo. Además, se seleccionó un locus SSR adicional proveniente del genoma del cloroplasto. La presencia de las secuencias de los *primers* diseñados para estos nuevos loci en el genoma del cultivar amelonado Matina 1-6 (Motamayor *et al.*, 2013) se corroboró como se describió anteriormente.

2.3.2. Caracterización molecular de la diversidad genética de cacao

2.3.2.1. Identificación, selección y colecta de muestras

En total se identificaron y seleccionaron 96 accesiones de cacao, 94 de CNFA y 2 de CCN-51 (Tabla 3). La característica principal que se consideró para la identificación y selección de muestras de accesiones de cacao nativo fue la siguiente: en cada uno de los lugares visitados, se realizó una interacción agricultor-investigador para identificar las accesiones de CNFA diferenciadas principalmente por características morfológicas y productivas (Tabla 2; Anexo 1). En la provincia de Rodríguez de Mendoza sólo se realizaron muestreos en Zarumilla, distrito de Chirimoto, y en Chachapoyas específicamente en Tuen, distrito de Balsas (Tabla 3), por ser los únicos distritos donde se cultiva cacao en dichas provincias. Una vez identificado, el árbol fue georreferenciado y etiquetado (Anexo 2), como lo detalla Oliva-Cruz *et al.* (2022). Para los cacaos nativos denominados INDES como ya se encontraban identificados por Oliva-Cruz *et al.* (2022) se procedió directamente a la colecta de muestras en el jardín clonal de la institución. Para la colecta de muestras de tejido foliar se consideraron tres hojas distales por cada accesión de cacao, con tamaño mediano y en buen estado fitosanitario (Anexo 2), como lo

describen Lins *et al.* (2016) y Suazo-Ubieta (2020). Las muestras de hojas fueron colectadas y colocadas en sobres manila A4 hasta su traslado al LABISANV.

Tabla 3. Información pasaporte de las accesiones de CNFA.

Nº	Código	Lugar	Provincia	Fecha de colecta	Altitud	Coordenadas		
						Este	Norte	Zona UTM
1	TCCH-1	Hornopampa	Chachapoyas	15/08/2021	1067	169179.3	9241546.6	18
2	TCCH-2	Hornopampa	Chachapoyas	15/08/2021	1079	169142.5	9241531.0	18
3	TCCH-3	Hornopampa	Chachapoyas	15/08/2021	1047	169100.2	9241413.9	18
4	TCCH-4	Hornopampa	Chachapoyas	15/08/2021	1051	168880.8	9241603.2	18
5	TCCH-5	Hornopampa	Chachapoyas	15/08/2021	1057	168920.7	9241609.6	18
6	TCCH-6	Hornopampa	Chachapoyas	15/08/2021	1048	168732.3	9241767.4	18
7	TCCH-7	Hornopampa	Chachapoyas	15/08/2021	1079	168744.6	9241756.1	18
8	TCCH-8	Lumbay	Chachapoyas	4/10/2021	1022	830855.8	9242736.4	17
9	TCCH-9	Tuen	Chachapoyas	5/10/2021	861	830809.5	9242705.9	17
10	TCCH-10	Tuen	Chachapoyas	5/10/2021	869	827911.8	9253301.4	17
11	TCCH-11	Tuen	Chachapoyas	5/10/2021	878	827899.4	9253276.9	17
12	TCCH-12	Tuen	Chachapoyas	5/10/2021	878	827844.3	9253311.0	17
13	TCCH-13	Tuen	Chachapoyas	5/10/2021	882	827844.2	9253304.9	17
14	TCCH-14	Tuen	Chachapoyas	5/10/2021	869	827844.2	9253295.7	17
15	TCCH-15	Tuen	Chachapoyas	5/10/2021	876	827844.2	9253298.7	17
16	TCCH-16	Tuen	Chachapoyas	5/10/2021	877	827819.6	9253292.7	17
17	TCCH-17	Tuen	Chachapoyas	5/10/2021	881	827789.1	9253336.0	17
18	TCCH-18	Tuen	Chachapoyas	5/10/2021	885	827776.7	9253317.6	17
19	TCCH-19	Tuen	Chachapoyas	5/10/2021	888	827913.7	9253104.6	17
20	TCL-1	Delo	Luya	2/09/2021	1389	811219.7	9306446.2	17
21	TCL-2	Playa Jumeth	Luya	2/09/2021	1001	801304.8	9306406.3	17
22	TCL-3	Playa Jumeth	Luya	2/09/2021	1002	801268.1	9306452.6	17
23	TCL-4	Playa Jumeth	Luya	2/09/2021	1014	801258.8	9306437.3	17
24	TCL-5	La Playa Grande	Luya	2/09/2021	592	789969.1	9310580.0	17
25	TCL-6	La Playa Grande	Luya	2/09/2021	590	789693.3	9310796.5	17
26	TCL-7	La Playa Grande	Luya	3/09/2021	589	789677.9	9310802.7	17
27	TCL-8	Santa Rosa	Luya	3/09/2021	1298	795338.6	9317030.0	17
28	TCL-9	Santa Rosa	Luya	3/09/2021	1235	795788.8	9317227.5	17
29	TCL-10	Balsapucro	Luya	6/10/2021	1876	822092.5	9285242.5	17
30	TCL-11	Guayal	Luya	6/10/2021	1748	812332.5	9289364.9	17
31	TCL-12	Lloque	Luya	6/10/2021	1768	815359.3	9291106.8	17
32	TCRM-1	Zarumilla	Rodríguez de Mendoza	4/08/2021	1198	236581.5	9273358.7	18
33	TCRM-2	Zarumilla	Rodríguez de Mendoza	4/08/2021	1190	236664.2	9273417.5	18
34	TCRM-3	Zarumilla	Rodríguez de Mendoza	5/08/2021	1186	236667.2	9273426.7	18
35	TCRM-4	Zarumilla	Rodríguez de Mendoza	5/08/2021	1230	236195.3	9273123.3	18
36	TCRM-5	Zarumilla	Rodríguez de Mendoza	5/08/2021	1239	236216.5	9273184.8	18
37	TCRM-6	Zarumilla	Rodríguez de Mendoza	5/08/2021	1236	236219.6	9273184.8	18

38	TCRM-7	Zarumilla	Rodríguez de Mendoza	5/08/2021	1237	236185.9	9273172.4	18
39	TCRM-8	Zarumilla	Rodríguez de Mendoza	5/08/2021	1240	236187.9	9273175.5	18
40	TCRM-9	Zarumilla	Rodríguez de Mendoza	5/08/2021	1199	236575.7	9273275.7	18
41	TCRM-10	Zarumilla	Rodríguez de Mendoza	5/08/2021	1199	236575.6	9273297.2	18
42	TCRM-11	Zarumilla	Rodríguez de Mendoza	5/08/2021	1201	236566.4	9273300.2	18
43	TCRM-12	Zarumilla	Rodríguez de Mendoza	5/08/2021	1152	236495.0	9274089.7	18
44	TCRM-13	Zarumilla	Rodríguez de Mendoza	5/08/2021	1166	236343.6	9274255.0	18
45	TCRM-14	Zarumilla	Rodríguez de Mendoza	5/08/2021	1156	236227.0	9274856.8	18
46	TCRM-15	Zarumilla	Rodríguez de Mendoza	6/08/2021	1159	236220.9	9274847.5	18
47	TCRM-16	Zarumilla	Rodríguez de Mendoza	6/08/2021	1160	236221.0	9274832.2	18
48	TCC-1	Pampa Hermosa	Condorcanqui	6/07/2021	243	183698.2	9491906.3	18
49	TCC-2	Pampa Hermosa	Condorcanqui	6/07/2021	196	183695.2	9491881.7	18
50	TCC-3	Puerto Galilea	Condorcanqui	9/07/2021	183	193746.4	9555645.9	18
51	TCC-4	Villa Gonzalo	Condorcanqui	9/07/2021	182	195097.3	9564368.3	18
52	TCC-5	Villa Gonzalo	Condorcanqui	9/07/2021	182	195134.2	9564408.4	18
53	TCC-6	Huabal	Condorcanqui	9/07/2021	196	195250.3	9559198.4	18
54	TCC-7	Nieva	Condorcanqui	11/07/2021	200	181738.8	9491243.6	18
55	TCC-8	Nieva	Condorcanqui	11/07/2021	202	181717.2	9491234.3	18
56	TCC-9	Huampami	Condorcanqui	11/07/2021	233	815860.3	9508358.6	17
57	TCC-10	Kuzumatak	Condorcanqui	25/04/2022	359	820645.7	9456526.8	17
58	TCC-11	Kuzumatak	Condorcanqui	25/04/2022	345	820679.5	9456517.4	17
59	TCC-12	Kuzumatak	Condorcanqui	25/04/2022	312	820750.2	9456452.5	17
60	TCC-13	Kuzumatak	Condorcanqui	25/04/2022	277	820801.9	9456280.2	17
61	TCC-14	Tokio	Condorcanqui	26/04/2022	254	820254.1	9459406.1	17
62	TCC-15	Tsuan	Condorcanqui	26/04/2022	288	820597.1	9456714.5	17
63	TCC-16	Najaim Paraiso	Condorcanqui	27/04/2022	220	825569.2	9456975.7	17
64	TCC-17	Pampa Hermosa	Condorcanqui	28/04/2022	208	183704.4	9491894.0	18
65	TCC-18	Pampa Hermosa	Condorcanqui	28/04/2022	208	183689.0	9491890.9	18
66	TCC-19	Pampa Hermosa	Condorcanqui	28/04/2022	195	183692.1	9491884.7	18
67	TCC-20	Pampa Hermosa	Condorcanqui	28/04/2022	195	183729.2	9491866.4	18
68	TCC-21	Ebron	Condorcanqui	28/04/2022	208	184162.4	9488443.4	18
69	INDES-4	Sector La Concordia	Utcubamba	24/05/2021	753	788671.0	9368556.0	17
70	INDES-32	Sector Quebrada Seca	Utcubamba	24/05/2021	421	779564.0	9367833.0	17
71	INDES-35	Sector Quebrada Seca	Utcubamba	24/05/2021	423	779584.0	9367821.0	17

72	INDES-80	Diamante Bajo	Utcubamba	26/05/2021	840	794303.0	9367139.0	17
73	INDES-14	Sector El Limoncito	Utcubamba	26/05/2021	817	793728.0	9366961.0	17
74	INDES-27	Sector La Cruz	Utcubamba	14/04/2021	761	786945.0	9369993.0	17
75	INDES-50	Diamante Bajo	Utcubamba	14/04/2021	730	794447.0	9366031.0	17
76	INDES-63	Naranjos Altos	Utcubamba	27/05/2021	727	793806.0	9365734.0	17
77	INDES-6	Sector El Chalan	Utcubamba	24/05/2021	754	787894.0	9369168.0	17
78	INDES-31	Sector Jahuanga	Utcubamba	14/04/2021	629	770538.0	9361358.0	17
79	INDES-53	Diamante Bajo	Utcubamba	14/04/2021	727	794453.0	9366666.0	17
80	INDES-70	Lluhuana	Utcubamba	24/05/2021	970	787756.0	9371938.0	17
81	INDES-64	Naranjos Altos	Utcubamba	25/05/2021	665	792251.0	9364133.0	17
82	INDES-98	El Tigre	Utcubamba	26/05/2021	793	798067.0	9362827.0	17
83	INDES-99	El Tigre	Utcubamba	26/05/2021	793	798067.0	9362827.0	17
84	INDES-100	Pan de Azucar	Utcubamba	27/05/2021	793	798067.0	9362827.0	17
85	INDES-12	Sector El Chalan	Utcubamba	28/05/2021	812	793633.0	9366965.0	17
86	INDES-38	Sector Copallín	Bagua	28/05/2021	947	786789.0	9372113.0	17
87	INDES-78	Tañuspe	Bagua	27/05/2021	575	779260.0	9378861.0	17
88	INDES-81	El Limoncito	Utcubamba	27/05/2021	837	794270.0	9367134.0	17
89	TCU-1	Cocahuayco	Utcubamba	3/09/2021	1187	788718.4	9320874.3	17
90	TCU-2	La Fila Larga	Utcubamba	3/09/2021	1119	788696.9	9320889.8	17
91	TCU-3	La Fila Larga	Utcubamba	3/09/2021	1139	787391.6	9326432.1	17
92	CAP-6	Bagua Grande	Utcubamba	14/02/2022	464	786433.0	9363054.4	17
93	CAP-10	Bagua Grande	Utcubamba	14/02/2022	464	786436.2	9363063.6	17
94	CAP-126	Bagua Grande	Utcubamba	14/02/2022	754	793550.0	9366144.8	17
95	CCN-51	Cajaruro	Utcubamba	14/02/2022	751	793516.0	9366120.4	17
96	CCN-51-C	Kuzumatak	Condorcanqui	25/04/2022	281	820789.8	9456341.7	17

2.3.2.2. Preparación de las muestras

Las muestras de hojas fueron limpiadas con agua destilada estéril y secadas con papel toalla. Luego, con ayuda de un bisturí y pinzas estériles, se cortaron secciones de aproximadamente 15 x 0.2 mm. Los cortes se realizaron entre la nervadura central y el borde de la hoja, considerando solo el área de la hoja que no contenía nervaduras secundarias. Después se pesaron 100 mg de tejido foliar (peso fresco) y se colocaron en tubos Eppendorf de 2 ml de capacidad. Las tiras estériles de hojas no utilizadas se

colocaron en bolsas Ziploc de 80 x 85 cm con geles de sílica para su conservación y futuros análisis moleculares.

2.3.2.3. Homogenización de las muestras

Para la homogenización de muestras del tejido vegetal, se utilizó un triturador Benchmark Scientific BeadBug D1030-E. Se colocaron en tubos Eppendorf de 2 ml de capacidad microesferas de acero inoxidable de 2.8 mm y muestras de tejido foliar. Luego, los microtubos fueron enfriados en nitrógeno líquido, colocados en el triturador y agitados a 2,800 rpm durante 45 segundos, hasta que el material vegetal quedara homogenizado o pulverizado. Durante todo el proceso, se evitó que las muestras se descongelaran, puesto que las hojas de cacao presentan un alto índice de oxidación.

2.3.2.4. Extracción de ADN

Para la extracción de ADN se utilizó el kit NucleoSpin® Plant II (Macherey-Nagel, Düren, Alemania), siguiendo el protocolo del fabricante indicado para la obtención de ADN genómico de plantas (Anexo 3). Finalmente, se etiquetó a cada una de las muestras y el ADN extraído se almacenó a -20° C.

2.3.2.5. Cuantificación de ADN

Cada muestra de ADN genómico extraído fue cuantificada utilizando el espectrofotómetro Eppendorf BioSpectrometer® Basic (Eppendorf, Hamburgo, Alemania) a 260 nm de longitud de onda (Anexo 3). Con los datos obtenidos, se realizaron diluciones con agua Mili-Q a una concentración de 1.0 ng/μl para ser utilizado en las reacciones de PCR.

2.3.2.6. PCR y análisis de fragmentos

Para analizar la longitud de los productos de la PCR se utilizó la técnica de electroforesis capilar. Este método requiere el uso de *primers* fluorescentes para marcar los productos de PCR. Se siguió el protocolo implementado por Schuelke (2000) para marcar con fluorescencia los productos PCR de manera económica. Este protocolo, a diferencia de un PCR regular donde se utiliza dos *primers* (uno *forward* y otro *reverse*), requiere de un tercer *primer*, el *primer* universal M13 (-21) (5'-TGTAACGACGGCCAGT- 3') [a partir de ahora sólo lo llamaremos como *primer* M13] como *primer* fluorescente para todos los marcadores SSR analizados. Se utilizó los colores estándar del conjunto de colorantes el DS-33 de Applied Biosystems® (6-FAM™, azul; VIC®, verde; NED™, amarillo; and PET®, rojo). El uso de los cuatro colores permitió combinar cuatro

reacciones de PCR diferentes en una sola muestra, lo que reduce el número de muestras a analizar y a la vez los costos cuando se trabaja con múltiples muestras y marcadores.

El cóctel de PCR estuvo conformado por los siguientes tres *primers*: *primer* M13 fluorescente, *forward primer* con la secuencia M13 en el extremo 5', y el *reverse primer*. Por ejemplo: VIC® 5'-TGTAACGACGGCCAGT-3', el *primer* M13 fluorescente; 5'-TGTAACGACGGCCAGT XXXXXXXXXXXXXXXXXXXX-3', *forward primer* con la secuencia M13 en el extremo 5'; y 5'-YYYYYYYYYYYYYYYYYYYY-3', el *reverse primer*. Las reacciones de PCR se llevaron a cabo en un volumen total de 12.5 µL, que contenía 6.25 µL de GoTaq® G2 Green Master Mix, 0.469 µL del *primer* M13 fluorescente, 0.156 µL M13 *forward primer*, 0.625 *reverse primer* y 5 µL de ADN genómico a una concentración de 1.0 ng/µl. Las condiciones de termo ciclado (Tabla 4) siguieron las especificaciones de Schuelke (2000). La replicabilidad de las amplificaciones se evaluó mediante la repetición de PCR de 20 muestras seleccionadas al azar.

Tabla 4. Condiciones de PCR para la amplificación de muestras de CNFA.

Paso inicial	Pasos de amplificación			Nº de ciclos	Paso final
94°C/ 5min	94°C/ 30s	57°C/ 45s	72°C/ 45s	30	72°C/ 10 min
	94°C/ 30s	53°C/ 45s	72°C/ 45s	8	

2.3.2.7. Electroforesis del ADN producto de la amplificación

Los productos amplificados se confirmaron visualmente mediante electroforesis en un tanque Cleaver Scientific Ltd (Rugby, Reino Unido). Se cargó 1.5 µL de la muestra en los pocillos de gel de agarosa al 1% (1.5 g de agarosa; 150 ml de Tris-acetate-EDTA (TAE Buffer) 1X; 6 µL de SYBR® Safe DNA Gel Stain), sumergido en TAE Buffer 1X. En la cámara de electroforesis se aplicaron las siguientes condiciones: 100 voltios (V) y 100 amperios (A) durante 30 minutos tiempo, en el cual las regiones de ADN amplificadas en la PCR migraron hacia el polo positivo del tanque. Las bandas resultantes en los geles se dimensionaron con la escalera de ADN “Bench Top 1 Kb DNA Ladder” (250 pb hasta 10,000 pb) (Promega, Madison, Estados Unidos), y se visualizaron con luz azul emitida por el equipo de electroforesis utilizado (Anexo 7). Una vez verificados los productos de la PCR, las muestras fueron enviadas a la empresa MacroGen (Seúl, Corea del Sur), para el análisis de fragmentos a través de electroforesis capilar con el analizador ABI 3730XL.

2.4. Análisis de datos

2.4.1. Análisis de fragmentos y determinación de polimorfismo alélico

Los datos del análisis de fragmentos fueron obtenidos a través de electroferogramas. El tamaño y polimorfismo de alelos se obtuvo utilizando GeneScan 500 LIZ como regla molecular estándar y el Software Geneious Prime Versión 2023.0.4.

2.4.2. Análisis de diversidad genética

El análisis de diversidad genética se realizó con el paquete “*poppr*” versión 2.9.3 (Kamvar *et al.*, 2014, 2015) en RStudio (R version 4.2.0). Se calcularon los índices de diversidad genotípica, mediante el índice de Shannon - Wiener (H), Simpson (lambda) y Stoddard y Taylor (G). Asimismo, se calculó la diversidad alélica mediante el índice de heterocigosidad de Nei (Hexp), número de Genotipos Multilocus (MLG), polimorfismos de los marcadores microsatélites (número de alelos por locus y alelos efectivos). La distribución de la diversidad de genotipos se calculó con el índice de Evenness (*E5*) (Grünwald *et al.*, 2003). Todos estos análisis se realizaron utilizando el paquete “*poppr*”, a menos que se especifique algún otro paquete o software utilizado. Las diferencias genéticas entre provincias se evaluaron utilizando el paquete DEMEtics, este análisis nos permitió calcular el índice de fijación *Gst* (Nei, 1973) y el índice de diferenciación *D* (Jost, 2008) a un intervalo de confianza de $p < 0.05$ (Gerlach *et al.*, 2010).

2.4.3. Desequilibrio de ligamiento

El desequilibrio de ligamiento (*linkage disequilibrium*) permite evaluar la asociación no aleatoria de alelos de diferentes loci (Slatkin, 2008). Se calculó mediante el Índice de asociación (*I_A*) (Brown *et al.*, 1998) y el Índice de asociación estandarizado ($\bar{r}d$) (Agapow & Burt, 2001). Este análisis se realizó con la función “*ia*” con 1,000 permutaciones para calcular la probabilidad de significancia. Esto permitió conocer si los alelos obtenidos de las muestras son significativamente iguales o diferentes a lo que se esperaría dentro una población con reproducción sexual, i.e., si las muestras se encuentran en equilibrio.

2.4.4. Análisis de la estructura genética

La estructura genética de las muestras de cacao se analizó con el paquete LEA de R (Frichot & François, 2015). LEA permite realizar un análisis de estructura genética dedicado al estudio de genética de poblaciones, que a diferencia de programas ampliamente utilizados como STRUCTURE, no asume que las muestras se encuentren en equilibrio o desequilibrio de ligamiento. Para la valoración del número óptimo de

poblaciones ancestrales hipotéticas (K) se utilizó el gráfico de sedimentación o *Scree test* (Cattell, 1996) y el criterio de entropía cruzada (Frichot & François, 2015). Se calculó el *Scree test* mediante la función “*pca*”. Para obtener el valor de entropía cruzada, se utilizó la función “*snmf*” considerando un rango de valores de K desde 1 hasta 15, que representan las estimaciones de grupos genéticos de cacao conocidos. Se realizaron 100 repeticiones para cada valor de K, utilizando un parámetro de regularización del algoritmo de 100 y 200 interacciones en el algoritmo. Se seleccionó el valor más bajo de entropía cruzada para su posterior análisis y visualización en STRUCTURE PLOT v. 2.0 (Ramasamy *et al.*, 2014).

2.4.5. Análisis de variancia molecular (AMOVA)

El AMOVA permitió determinar la variación molecular presentes en las muestras analizadas basado en un modelo jerárquico, por grupos (altitud, provincia y muestra). Este análisis atribuye un porcentaje de variación entre altitud, entre provincia dentro de altitud, entre muestras dentro de provincia y de manera individual. Se utilizó la función “*poppr.amova*”.

2.4.6. Análisis de asociación

El análisis de asociación se realizó en base a una red de extensión mínima utilizando la función “*poppr.msn*” (Csardi & Nepusz, 2006), un Análisis de Componentes principales (ACP) usando las herramientas del paquete *ade4* (Dray & Dufour, 2007), y mediante un dendograma proporcionado por el Método de Grupo de Pares No Ponderado con Media Aritmética (UPGMA) construido en base a las distancias genéticas (distancias de Nei), con 100 *bootstraps*. Estos análisis ayudan a explorar las asociaciones de las accesiones de cacao con respecto a los alelos encontrados por los marcadores SSR, así como evaluar las contribuciones a la diversidad genética de cada alelo encontrado. En paralelo, se realizó el análisis de 20 accesiones de cacao de la colección del INDES-CES, las cuales fueron agrupadas según descriptores morfológicos y sensoriales: los Toribianos, los Indes, los Bagüinos, los Utkus y los Cajas (Oliva-Cruz *et al.*, 2022). Se evaluaron 4 accesiones de cacao por cada grupo identificado. Se investigó la relación genética y su influencia en las características morfológicas y sensoriales, considerando los grupos reportados en el estudio de Oliva-Cruz *et al.* (2022).

Todos estos análisis nos ayudaron a obtener conclusiones respecto a la variabilidad genética del CNFA con respecto a la altitud de las zonas muestreadas. Además, nos ayudó

a identificar las regiones de mayor diversidad y empezar a plantear estrategias de conservación del CNFA.

III. RESULTADOS

3.1. Diseño de marcadores microsatélites

Se lograron identificar 41 loci SSR y diseñar 41 pares de *primers* para cada locus, de los cuales 20 pares de *primers* son homólogos a los SSR trabajados en Motamayor *et al.* (2008), es decir, amplifican la misma región SSR. Estos 20 SSR se encuentran distribuidos en los diez cromosomas del cacao (dos por cromosoma y distancia mínima de 5 Mb) (Tabla 5). Para los 21 SSR restantes, los *primers* fueron diseñados considerando los mismos criterios anteriormente descritos, seleccionando los loci con mayor número de repeticiones de ADN repetitivo, 20 SSR ubicados en el ADN genómico nuclear, y uno ubicado en el cloroplasto (Tabla 6).

Tabla 5. *Primers* para la amplificación de regiones SSR de *Theobroma cacao* homólogos a los utilizados por Motomayor *et al.* (2008)

Nº	Primer	Secuencia de primer (5'-3')	Primer en Motomayor <i>et al.</i> (2008)	Cromosoma Número	Posición	Tamaño Aprox. PCR	Patrón repetitivo de ADN
1	INDES_mTC59_F* INDES_mTC59_R*	TGTA AACGACGGCCAGT TGACACTACCATCATGAGATTTC AGAAACAACACACCAACAACCC	mTcCIR184	1	871383	163	(AG)8
2	INDES_mTC35_F INDES_mTC35_R	TGTA AACGACGGCCAGT CAGAGTAATCTACAGGCACAGGG CGAAAGATTAAGAACTGGTCCTG	mTcCIR106	1	2459505	123	(AG)7
3	INDES_mTC22_F* INDES_mTC22_R*	TGTA AACGACGGCCAGT TCAGAAGATGCTCAGCCTTG CGGCGTGTGTCTATAACCATT	mTcCIR60	2	19251737	215	(AC)12
4	INDES_mTC78_F INDES_mTC78_R	TGTA AACGACGGCCAGT CCAAGAGGCATTGCGATTAC TCTCCCTTCATACACCAGGG	mTcCIR261	2	40322008	172	(AG)13
5	INDES_mTC84_F* INDES_mTC84_R*	TGTA AACGACGGCCAGT TCACTATACATAACAGCAGCCATT GTTACGTTTCTTCCGCCACT	mTcCIR289	3	29547629	134	(AG)8
6	INDES_mTC87_F INDES_mTC87_R	TGTA AACGACGGCCAGT AGAAGAACTAACGGCGGACA ATTGGAGGCTCATCCAGAT	SHRSTc7	3	35296078	239	(AG)13
7	INDES_mTC6_F* INDES_mTC6_R*	TGTA AACGACGGCCAGT TTGCCTCATATTTCTGACCC TCTGCCATCCTTCTTGATCTC	mTcCIR12	4	24533034	284	(AC)6
8	INDES_mTC17_F INDES_mTC17_R	TGTA AACGACGGCCAGT TTGTTCTGTTATCTTGTGCTGA AGCATCACATCTGATCCACAC	mTcCIR33	4	28292949	217	(ACGC)7
9	INDES_mTC86_F INDES_mTC86_R	TGTA AACGACGGCCAGT AAGCATGTAAATGCCATACGA TGTGTGGATGAGTCACATGAT	SHRSTc4	5	32592266	122	(AG)8
10	INDES_mTC38_F INDES_mTC38_R	TGTA AACGACGGCCAGT TGTTTGATTGACCTCTGTTCTCT TTGTTCTGGAAAGTGTAGGAAAG	mTcCIR109	5	39040794	134	(AG)14
11	INDES_mTC3_F INDES_mTC3_R	TGTA AACGACGGCCAGT TCCCTCTAAACTACCCTAAATATTCTT TCTCTAGTCCAATCAGCAATTC	mTcCIR6	6	219614	192	(AG)15
12	INDES_mTC77_F* INDES_mTC77_R*	TGTA AACGACGGCCAGT ATCTGCAGGACTTGGACCAC TCAAGCCTTACAGCATTCCC	mTcCIR255	6	23520517	204	(AC)11

13	INDES_mTC49_F*	TGTA AACGACGGCCAGT TTTGCAATTCTTCCTCTCTCTTT	mTcCIR147	7	9135080	204	(AG)6
	INDES_mTC49_R*	TATGGAGCTCACATGGACCC					
14	INDES_mTC83_F	TGTA AACGACGGCCAGT TGGAAATCAGATATGGGTAAAGG	mTcCIR277	7	12802611	377	(AG)12
	INDES_mTC83_R	AACGATATCCAGATTTGACACA					
15	INDES_mTC1_F	TGTA AACGACGGCCAGT GTCTGCTTTGTGCAATTGGG	mTcCIR1	8	271705	150	(AG)8
	INDES_mTC1_R	GGCTCAGTGAAGCAAAGGAA					
16	INDES_mTC14_F	TGTA AACGACGGCCAGT TCATCAGTTACATACCAATAGAATCAT	mTcCIR26	8	5600841	256	(AG)16
	INDES_mTC14_R	ATTTGTCAATGAGAATCAGATATAGTG					
17	INDES_mTC12_F*	TGTA AACGACGGCCAGT CCCTTATGGTAAGATTCTCGCA	mTcCIR24	9	6355004	191	(AG)14
	INDES_mTC12_R*	TGCATGTGAGGCATTTCTGT					
18	INDES_mTC4_F*	TGTA AACGACGGCCAGT GTTGATTTGCCAGCACAAGA	mTcCIR8	9	22510859	217	(AG)8
	INDES_mTC4_R*	TTAATCCATGCCTGGACCTT					
19	INDES_mTC23_F*	TGTA AACGACGGCCAGT AAGGCGAAGTGCCACATATT	mTcCIR61	10	14256196	223	(AC)12
	INDES_mTC23_R*	TTTGCTCATTGCCACTCAA					
20	INDES_mTC19_F	TGTA AACGACGGCCAGT TTCGGACAATGCTTAACGTG	mTcCIR37	10	20886312	167	(AT)5
	INDES_mTC19_R	TTCCAGTAAAGGCAGTTGGG					

Nota: Los *primers forward* tienen la secuencia TGTA AACGACGGCCAGT “M13” en el extremo 5', técnica propuesta por Schuelke (2000).

* *Primers* utilizados para la amplificación de las regiones microsatélite de *Theobroma cacao*.

Tabla 6. *Primers* diseñados en este estudio para la amplificación de regiones SSR de *Theobroma cacao*.

N°	Primer	Secuencia de primer (5'–3')	Cromosoma		Tamaño Aprox. PCR	Patrón repetitivo de ADN***
			Número	Posición		
1	INDES_mTC92_F INDES_mTC92_R	TGTA AACGACGGCCAGT CTTCTGGCTCATAGTTAAGTACTCTTT GGTCTCACTAATAAACTAAGGCATGT	1	2988765	292	(AT)32
2	INDES_mTC93_F* INDES_mTC93_R*	TGTA AACGACGGCCAGT AGAGGGTTCAGGTTTCTCCC TTTGGTAACGCTTTGGCTCT	1	35571464	283	(AG)27
3	INDES_mTC95_F INDES_mTC95_R	TGTA AACGACGGCCAGT AACCACATGGCTCTGCTTCT TGATACACAGCTTTGGTGGG	2	2529764	298	(AT)38
4	INDES_mTC97_F INDES_mTC97_R	TGTA AACGACGGCCAGT GCCTAGTTGCAACCACTCTTATC GCTTTAAGAGAGAGGGCAAGG	2	40955264	293	(AT)28
5	INDES_mTC98_F* INDES_mTC98_R*	TGTA AACGACGGCCAGT TCATCCAACTATGACCAAGTGA TTGGGTGCCAGCTTACTCTT	3	789419	298	(AG)104
6	INDES_mTC100_F INDES_mTC100_R	TGTA AACGACGGCCAGT TACATGTGGACTTATGGGCAA CATAAGAGGTCTATATGGAATTTATCC	3	17395445	290	(AT)27
7	INDES_mTC103_F INDES_mTC103_R	TGTA AACGACGGCCAGT TGGGATTTAGGGATCGAACTT CTCTGCCAAATAAAGCTGCC	4	19287236	300	(AT)27
8	INDES_mTC101_F INDES_mTC101_R	TGTA AACGACGGCCAGT TTGGAGCATTTGGTTGAGC AGGGAACACAAGGAACGACT	4	24309005	197	(AT)47
9	INDES_mTC129_F INDES_mTC129_R	TGTA AACGACGGCCAGT CGATGGAAGTTTCAGGTTCGT AGAGGCAATCCAAGCACAAC	5	22081335	257	(AT)26
10	INDES_mTC104_F* INDES_mTC104_R*	TGTA AACGACGGCCAGT TGTTGGTTGGTCTTGCTGAG AAAGGACGGCACATAGATCG	5	33872115	263	(AG)30
11	INDES_mTC107_F*	TGTA AACGACGGCCAGT TGTGTCTATTTACCTCGCAG	6	6686311	250	(AT)24

	INDES_mTC107_R*	TGGGAATTGCATACCCAATAA					
12	INDES_mTC108_F	TGTA AACGACGGCCAGT CGAAGGGAAGGACCTCTCA	6	19330677	300	(AT)27	
	INDES_mTC108_R	ACATACATCAGCCACTGTTAATACTAC					
13	INDES_mTC112_F	TGTA AACGACGGCCAGT TGCTTCCTGAGGTGAGCATA	7	606548	220	(AT)26	
	INDES_mTC112_R	TCAGCAAGGTCATCTTCAACC					
14	INDES_mTC111_F	TGTA AACGACGGCCAGT GCCAATCTCAATAAGCCAGG	7	19431893	229	(AT)24	
	INDES_mTC111_R	CGTAACATTAGCTTTGGGAGC					
15	INDES_mTC113_F*	TGTA AACGACGGCCAGT TTGATCTTGTGGAATTCCTG	8	411586	219	(AT)33	
	INDES_mTC113_R*	TGTACATACCTTGCTCCGAAT					
16	INDES_mTC125_F*	TGTA AACGACGGCCAGT TATCATGACAAGCAGCGGTC	8	9830566	248	(AT)24	
	INDES_mTC125_R*	CCCTCAACATTTGAAAGTTCC					
17	INDES_mTC127_F	TGTA AACGACGGCCAGT GCCAGGAAGTTGTAGGTGGT	9	4289610	239	(AT)28	
	INDES_mTC127_R	CCTGGATACA ACTAATTCCACAG					
18	INDES_mTC119_F	TGTA AACGACGGCCAGT GAGAAGGAAAGGTGAGAGCG	9	34358901	292	(AG)107	
	INDES_mTC119_R	TCCCAATATTTGAGTGAATGAA					
19	INDES_mTC116_F	TGTA AACGACGGCCAGT AATAGTATTTCCCTCCGTCTTTCTTT	10	549721	283	(AT)34	
	INDES_mTC116_R	TCCACATTGATTATGAGTTTATGACTT					
20	INDES_mTC117_F	TGTA AACGACGGCCAGT TGTGTAGGAACATACTCCTTTCTCT	10	18984488	244	(AT)25	
	INDES_mTC117_R	CGGGTAAATATTCCATCATCG					
21	INDES_mTC90_F*	TGTA AACGACGGCCAGT TCGATATTTAGTCATTGGAGGG		Cloroplasto**	44393	128 (AT)8	
	INDES_mTC90_R*	CGTTCTATTTTCATCTGATATTTTCG					

Nota: Los *primers forward* tienen la secuencia TGTA AACGACGGCCAGT “M13” en el extremo 5’, técnica propuesta por Schuelke (2000).

* *Primers* utilizados para la amplificación de las regiones microsatélite de *Theobroma cacao*.

** *Primer* ubicado en el cloroplasto.

*** Nótese el mayor número de repeticiones del patrón de ADN en estos marcadores, en comparación con las regiones SSR de la Tabla 5.

De los 41 marcadores moleculares SSR, se seleccionaron 16 para este estudio. Los marcadores INDES_mTC4, INDES_mTC6, INDES_mTC12, INDES_mTC22, INDES_mTC23, INDES_mTC49, INDES_mTC59, INDES_mTC77 e INDES_mTC84, son homólogos a los marcadores SSR utilizados por Motomayor *et al.* (2008) (Tabla 5), y los marcadores INDES_mTC90, INDES_mTC93, INDES_mTC98, INDES_mTC104, INDES_mTC107, INDES_mTC113 e INDES_mTC125, pertenecieron a los diseñados en este estudio (Tabla 6). Los *primers* para las regiones SSR homólogos a los de Motomayor *et al.* (2008) se encuentran ubicados en el mismo cromosoma de origen y región SSR del genoma del cacao; sin embargo, no tienen la misma secuencia original, ya que los *primers* nuevos se diseñaron para que tengan la misma temperatura de fusión y facilitar el trabajo de laboratorio, y para que los productos PCR sean analizados mediante electroforesis capilar a través de la técnica de análisis económica propuesta por Schuelke (2000). De los 16 marcadores SSR descritos, 15 se encuentran distribuidos en los 10 cromosomas del cacao y 1 SSR (INDES_mTC90) está ubicado en el cloroplasto (Figura 2).

Los 16 marcadores SSR seleccionados en este estudio (Figura 2) fueron los que presentaron amplificaciones menos ambiguas, menor número de alelos nulos, y los que mostraron mayor nitidez en las bandas amplificadas (fragmentos de PCR) (Figura 3). La replicabilidad de estos PCR fue corroborada analizando los electroferogramas como resultado del análisis de fragmentos (Figura 4).

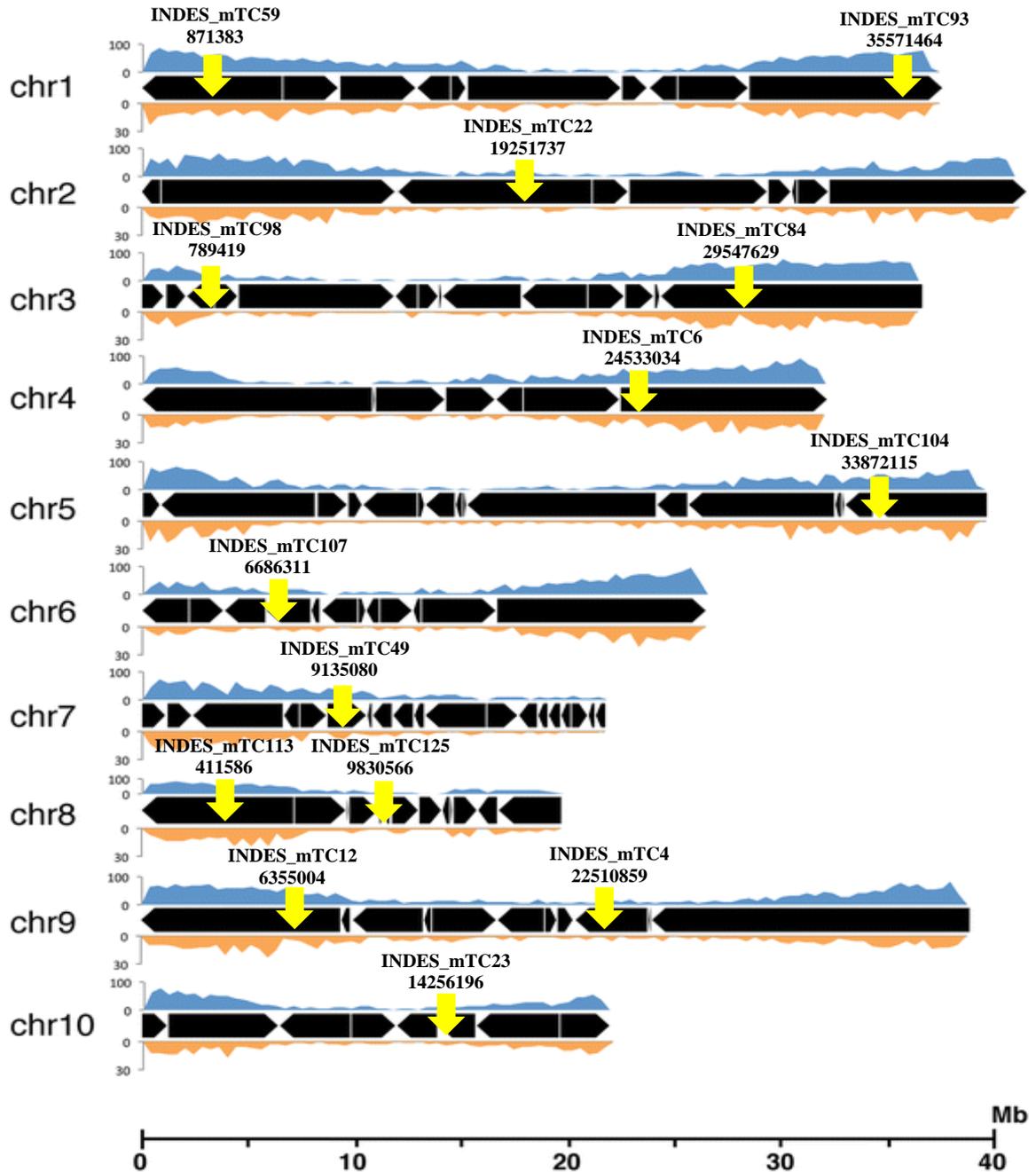


Figura 2. Ubicación de los *primers* para amplificar regiones SSR utilizados en este estudio utilizando como fondo la representación gráfica de los cromosomas del genoma de cacao del grupo genético Criollo (accesión B97-61/B2) (Argout *et al.*, 2017), distribuido bajo la licencia internacional Creative Commons Attribution 4.0 (<http://creativecommons.org/licenses/by/4.0/>). Las flechas amarillas muestran la ubicación aproximada de los SSR.

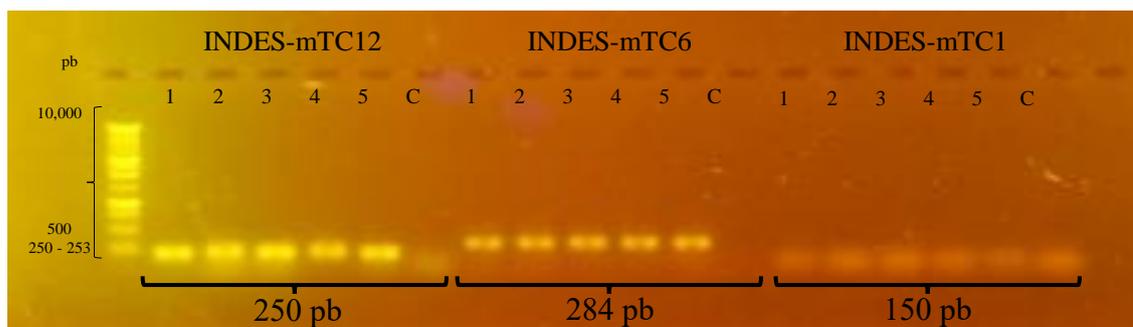


Figura 3. Productos PCR de marcadores moleculares SSR visualizados en un gel de agarosa al 1%. Se muestra el análisis de cada marcador con 5 accesiones de cacao y un control, corridos de manera paralela a una escalera molecular de 10,000 pb. Los marcadores INDES-mTC12 e INDES-mTC6, indican una correcta amplificación. El marcador INDES-mTC1 indica una PCR nula.

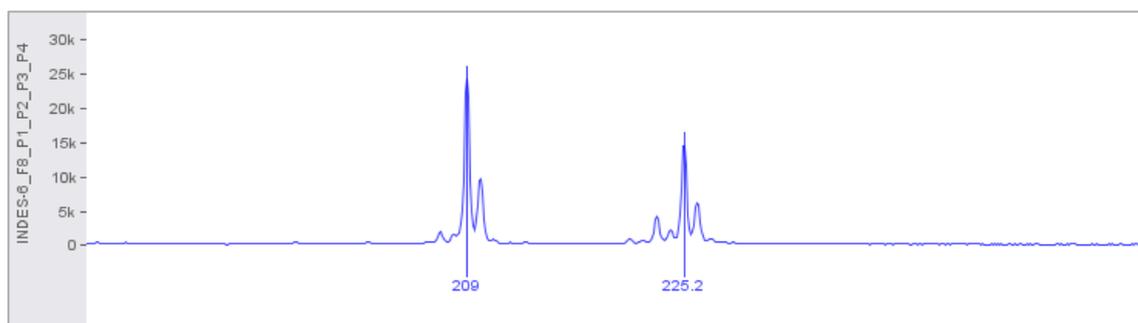


Figura 4. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra un locus heterocigoto en la accesión de cacao INDES-6 (colorante 6-FAMTM, azul) para el marcador INDES-mTC12.

3.2. Evaluación de la diversidad genética

3.2.1. Polimorfismo alélico en marcadores microsatélite SSR

Los 16 marcadores microsatélites SSR evaluados mostraron diferentes niveles de polimorfismo alélico en las muestras de *T. cacao*. Quince (15) loci (INDES_mTC4, INDES_mTC6, INDES_mTC12, INDES_mTC22, INDES_mTC23, INDES_mTC49, INDES_mTC59, INDES_mTC77, INDES_mTC84, INDES_mTC93, INDES_mTC98, INDES_mTC104, INDES_mTC107, INDES_mTC113 e INDES_mTC125), fueron polimórficos con una variación de 2 a 17 alelos (Tabla 7). El INDES_mTC90, loci que se encontraba en el cloroplasto fue el único locus monomórfico (1 alelo), este marcador no se incluye en este análisis al no presentar polimorfismo.

Los 9 marcadores microsatélites SSR homólogos a los utilizados por Motomayor *et al.* (2008) obtuvieron una media de 6 alelos (Tabla 7). Los 6 marcadores SSR nuevos, identificados en este estudio, presentaron una media de 12.33 alelos, sin incluir el cloroplastidial (Tabla 7). Los índices de Simpson (1-D) y Diversidad de Nei (Hexp) mostraron que los marcadores SSR nuevos obtuvieron una media total de 0.69 en ambos casos, superior a la media de los marcadores SSR utilizados por Motomayor *et al.* (2008) que obtuvieron valores totales de 0.51 en los dos parámetros evaluados (Tabla 7).

Tabla 7. Alelos observados en marcadores microsatélite SSR *Theobroma cacao*.

SSR	Allele	1-D	Hexp	Descripción
INDES_mTC4	2	0.22	0.22	
INDES_mTC6	9	0.73	0.73	
INDES_mTC12	6	0.55	0.55	
INDES_mTC22	7	0.45	0.45	
INDES_mTC23	13	0.79	0.80	
INDES_mTC49	2	0.43	0.43	
INDES_mTC59	5	0.45	0.46	
INDES_mTC77	4	0.50	0.51	
INDES_mTC84	6	0.44	0.44	
Mean	6	0.51	0.51	
INDES_mTC93	9	0.54	0.54	
INDES_mTC98	12	0.83	0.83	
INDES_mTC104	13	0.84	0.84	
INDES_mTC107	13	0.76	0.76	
INDES_mTC113	17	0.77	0.78	
INDES_mTC125	10	0.37	0.37	
Mean	12.33	0.69	0.69	

Nota: Allele = alelo; 1-D = Índice de Simpson; Hexp = Diversidad de genes Nei (Heterocigosidad esperada)

3.2.2. Diversidad genética

Se evaluaron 96 accesiones de cacao y se identificó 96 genotipos multilocus (MLG), es decir, cada accesión fue genéticamente única. Los resultados mostraron altos valores de diversidad de genes Nei (Hexp = 0.545) para el total de accesiones evaluadas. A nivel de provincias, el valor más alto lo obtuvo Condorcanqui (0.667), seguido de Bagua y Utcubamba (0.529) (Tabla 8). La diversidad genotípica según el índice de Shannon-Weiner (H), mostró un valor total de 4.56. El índice total de diversidad de Simpson (lambda) fue de 0.990. El índice de diversidad de Stoddard y Taylor (G) mostró un valor total de 96. En los tres últimos análisis de diversidad descritos, Bagua y Utcubamba unidos obtuvieron los valores más altos (H = 3.30, lambda = 0.963 y G = 27), seguido

por Condorcanqui ($H = 3.09$, $\lambda = 0.955$ y $G = 22$). Luya obtuvo el valor más bajo en todos los índices de diversidad evaluados ($Hexp = 0.217$, $H = 2.48$, $\lambda = 0.917$ y $G = 12$). El índice de Evenness ($E5$) obtuvo un valor de 1, esto indica que los genotipos se encuentran distribuidos de manera uniforme en las provincias evaluadas (Tabla 8).

Por otro lado, el índice de fijación Gst y el índice de diferenciación D indicó que todas las provincias eran significativamente diferentes entre sí ($p < 0.05$) (Tabla 9). Sin embargo, se observó que Luya y Chachapoyas obtuvieron mayores valores de diferenciación genética a las de Condorcanqui, Bagua y Utcubamba (Tabla 9).

Tabla 8. Diversidad genética *Theobroma cacao* basada en el análisis de marcadores SSR.

Provincia	N	MLG	Hexp	H	lambda	G	E5
Chachapoyas	19	19	0.478	2.94	0.947	19	1
Luya	12	12	0.217	2.48	0.917	12	1
Rodríguez de Mendoza	16	16	0.383	2.77	0.938	16	1
Condorcanqui	22	22	0.667	3.09	0.955	22	1
Bagua y Utcubamba	27	27	0.529	3.30	0.963	27	1
Total	96	96	0.545	4.56	0.990	96	1

Nota: N = Número de muestras, MLG = Número de genotipos multilocus o linajes multilocus, Hexp = Diversidad de genes Nei (Heterocigosidad esperada), H = Índice de Shannon-Weiner, lambda = Índice de Simpson, G = Índice de diversidad de Stoddard y Taylor, E5 = Índice de Evenness.

Tabla 9. Diferenciación genética entre provincias en base al índice de fijación Gst y el índice de diferenciación D con el conjunto de datos proporcionados por los marcadores SSR

Provincia	Chachapoyas	Luya	Rodríguez de Mendoza	Condorcanqui	Bagua y Utcubamba
Chachapoyas					
Luya	0.061*				
Rodríguez de Mendoza	0.113*	0.106*			
Condorcanqui	0.250*	0.383*	0.300*		
Bagua y Utcubamba	0.153*	0.180*	0.100*	0.149*	

Nota: D = Índice de diferenciación (* valores altamente significativos); $p < 0.05$ es significativo

Por otro lado, de acuerdo a la curva de genotipos multilocus (MLG) se muestra que los 15 marcadores SSR utilizados son suficientes para estudiar la diversidad en cacao, ya que al añadir un marcador adicional no incrementaría el número de genotipos encontrados en las muestras evaluadas (Figura 5).

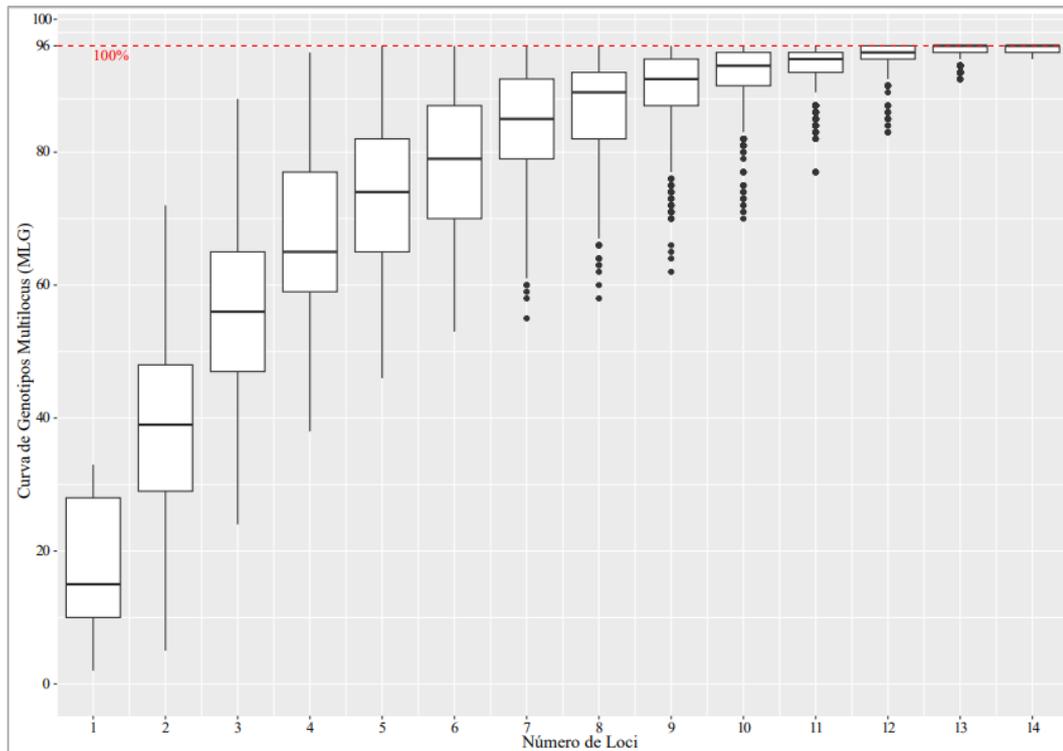


Figura 5. Curva de genotipos multilocus respecto al número de loci más informativos

3.2.3. Desequilibrio de ligamiento

El análisis de *linkage disequilibrium* mostró un índice de asociación (I_A) de 3.099 e índice de asociación estandarizado ($\bar{r}d$) de 0.226, altamente significativos ($p = 0.000999$) (Figura 6). Estos resultados fueron corroborados con un análisis a nivel de provincias, evidenciándose que la provincia de Chachapoyas (Figura 7), Rodríguez de Mendoza (Figura 8), Condorcanqui (Figura 9), Bagua y Utcubamba (Figura 10), obtuvieron valores altamente significativos ($p = 0.000999$), indicando que las muestras evaluadas de *T. cacao* presentan un desequilibrio de ligamiento, por lo tanto, presentan reproducción clonal. Sin embargo, este mismo análisis en la provincia de Luya mostró un valor no significativo (0.228) (Figura 11), lo cual indica que los genotipos obtenidos de esta provincia han sido el resultado de una diseminación basada en la reproducción sexual.

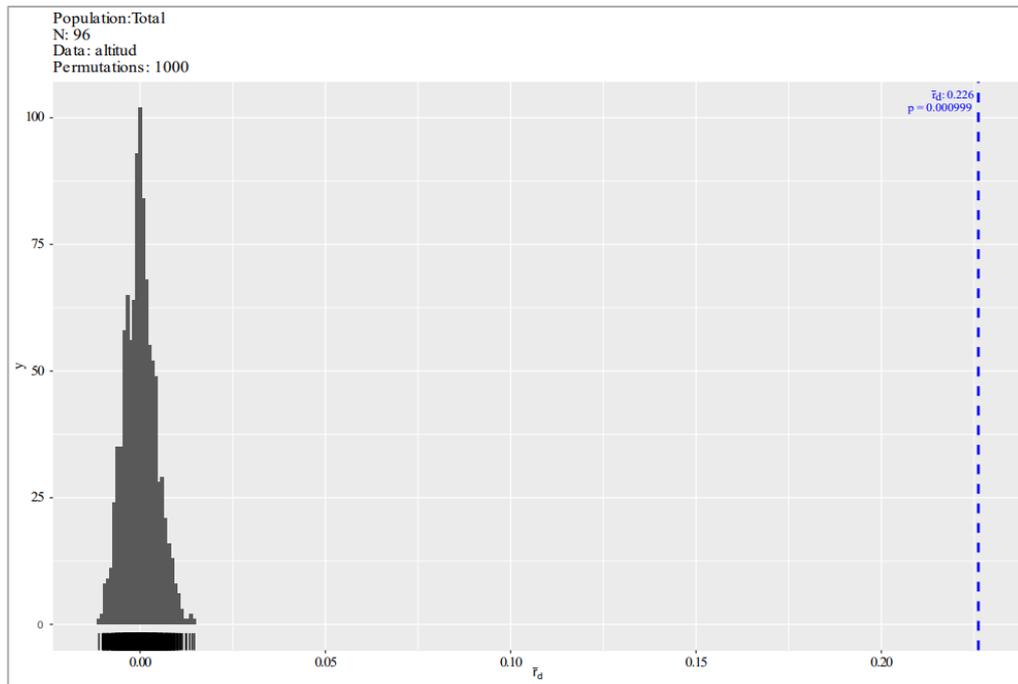


Figura 6. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}_d = 0.226$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao*.

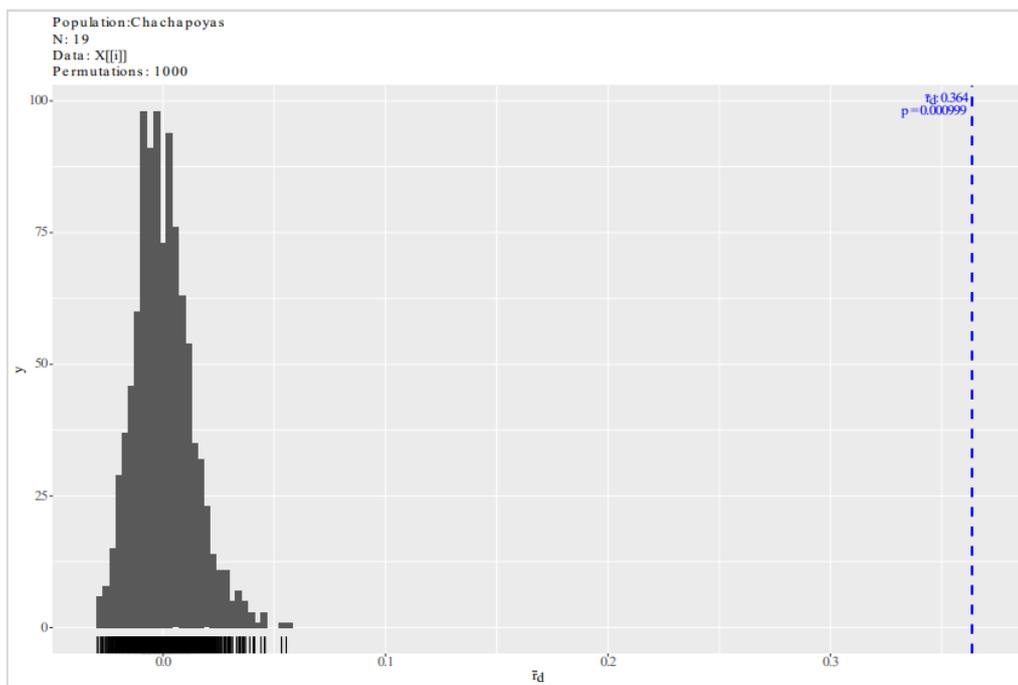


Figura 7. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}_d = 0.364$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao* de la provincia de Chachapoyas.

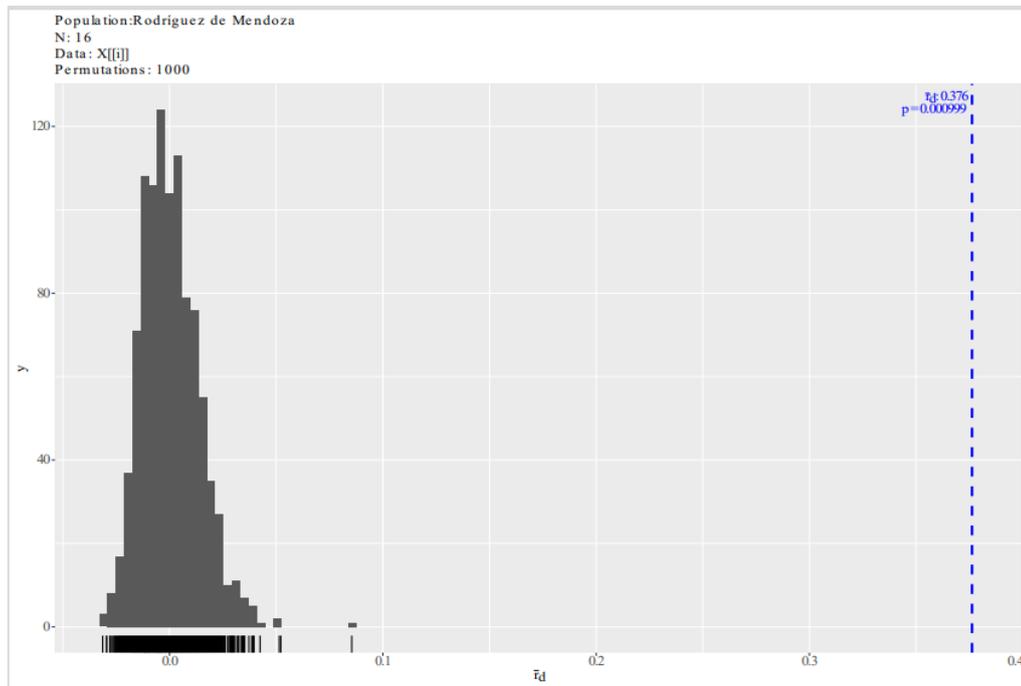


Figura 8. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}_d = 0.376$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao* de la provincia de Rodríguez de Mendoza.

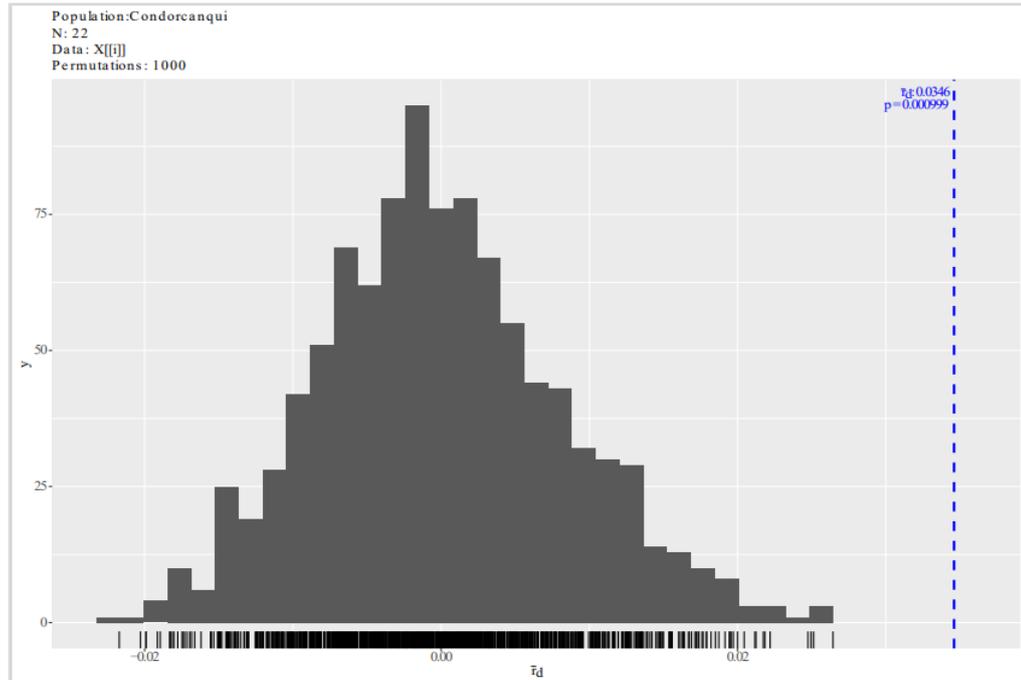


Figura 9. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}_d = 0.0346$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao* de la provincia de Condorcanqui.

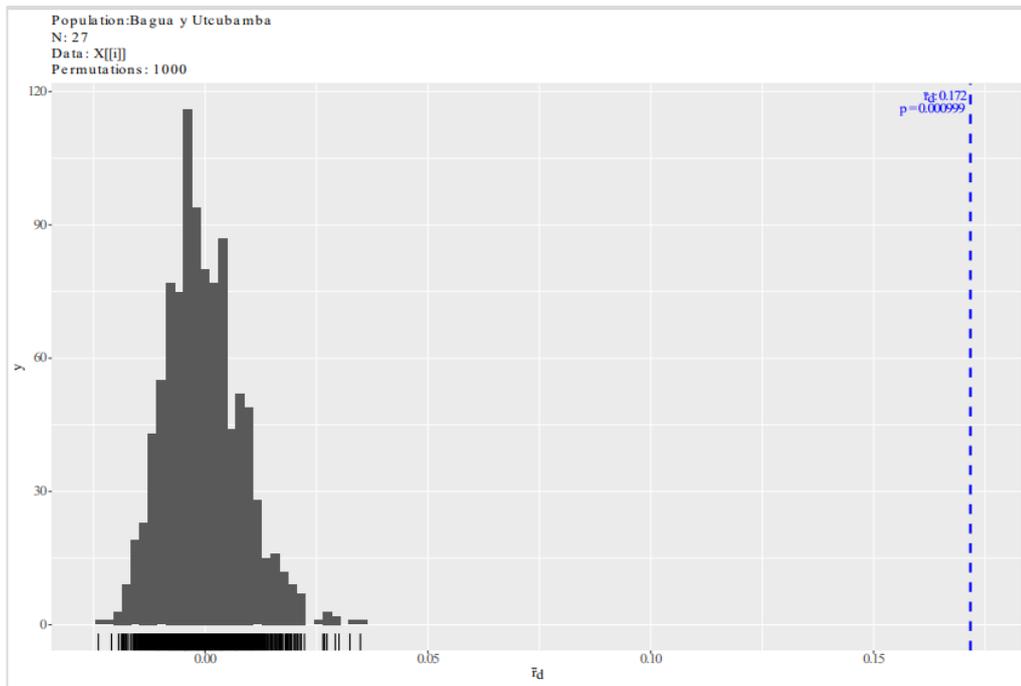


Figura 10. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.172$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao* de la provincia de Bagua y Utcubamba.

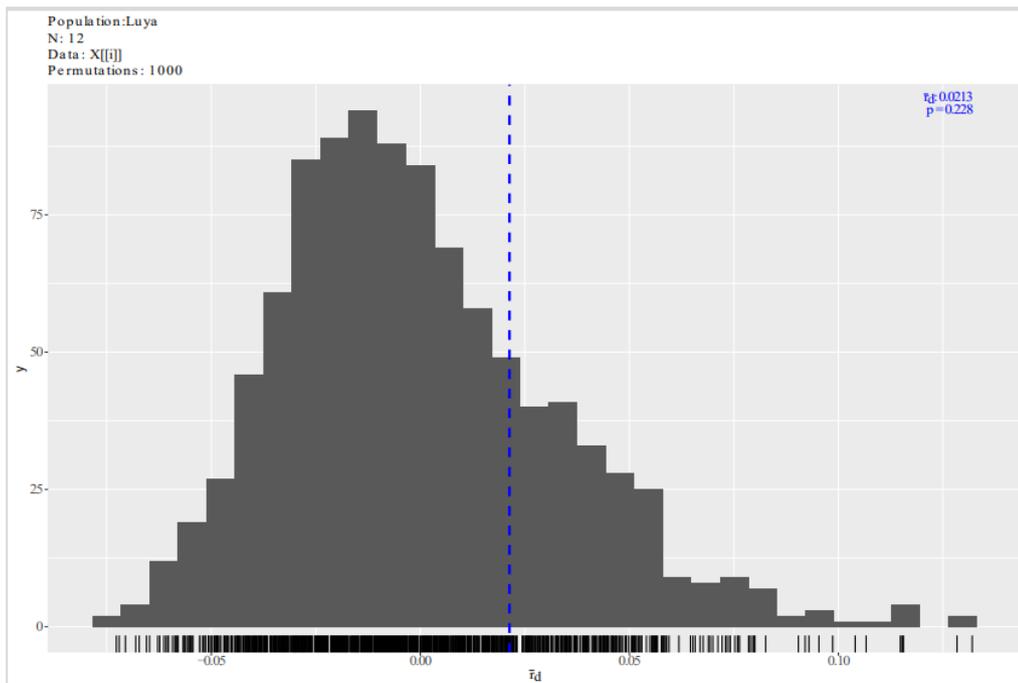


Figura 11. Gráfica de desequilibrio de ligamiento. Se muestra el valor del índice de asociación estandarizado ($\bar{r}d = 0.0213$) (línea azul). Este histograma indica el resultado de 1,000 permutaciones aleatorias en muestras de *Theobroma cacao* de la provincia de Luya.

3.2.4. Estructura genética

La estructura genética entre genotipos realizada en LEA mostró que las accesiones de cacao derivaban de cuatro hipotéticas poblaciones ancestrales ($K = 4$) basado en el gráfico *Scree test* (Figura 12) y el criterio de entropía cruzada (Figura 13).

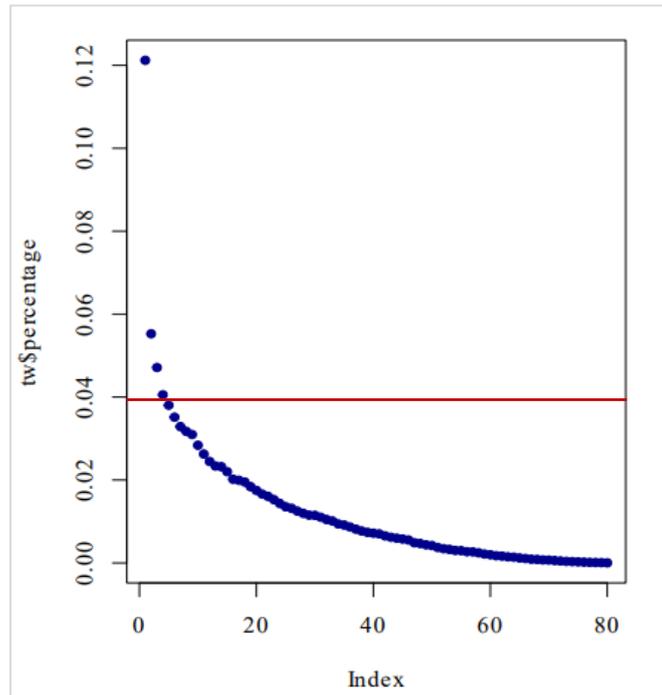


Figura 12. Gráfico de Sedimentación (*Scree test*) del porcentaje de varianza explicada por cada componente en un PCA de los datos genéticos. El codo del diagrama indica que hay alrededor de $K = 4$ componentes en los datos.

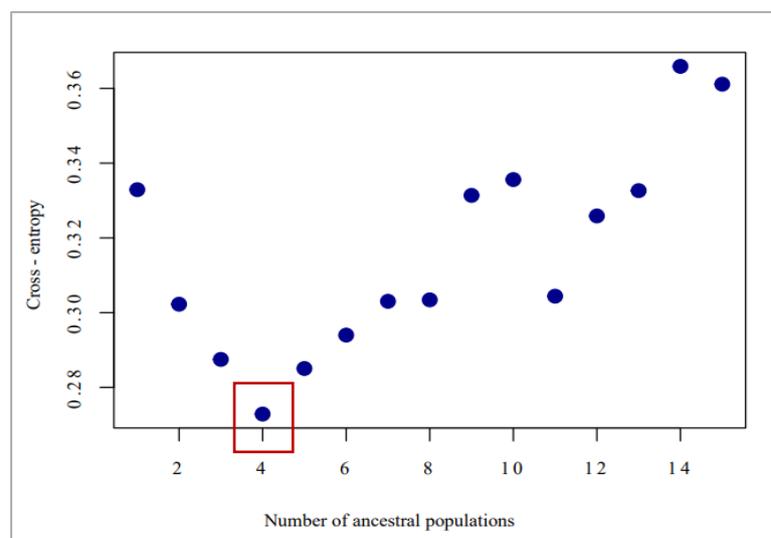


Figura 13. Valores el criterio de entropía cruzada en función del número de factores en las ejecuciones Modelos Mixtos de Factores Latentes (LFMM). Se aprecia que $K = 4$ es el número de poblaciones ancestrales óptimo con el menor valor de entropía cruzada.

La provincia con accesiones derivadas del mayor número de poblaciones ancestrales fue Condorcanqui con $K = 4$, seguida por las accesiones de Chachapoyas, Bagua y Utcubamba con $K = 3$. Las muestras de la provincia de Luya derivan de una sólo población ancestral ($K = 1$), la cual se encuentra distribuida en todas las demás provincias (Figura 14). El coeficiente de mezcla genética (*admixture*) indica que la mayoría de accesiones evaluadas derivan de una sola población ancestral (*admixture* > 0.99970, y se ven graficadas por un solo color, ya que las otras aportan ínfimamente en el coeficiente de mezcla; Anexo 14). Los valores de mezcla genética indican que en la provincia de Condorcanqui se evidencia una mayor mezcla genética (mezcla genética de la población ancestral 1, $V_{K1} = 0.22737$, $V_{K2} = 0.32026$, $V_{K3} = 0.16900$, $V_{K4} = 0.28337$), seguido por Bagua y Utcubamba ($V_{K1} = 0.14819$, $V_{K2} = 0.53073$, $V_{K3} = 0.32098$, $V_{K4} = 0.00010$). Luya fue la provincia donde se evidenció menores valores de mezcla genética ($V_{K1} = 0.00010$, $V_{K2} = 0.99970$, $V_{K3} = 0.00010$, $V_{K4} = 0.00010$). Se muestra que la población K_2 es la que contribuye casi en su totalidad al bagaje genético de todas las accesiones de cacao evaluadas de Luya (Figura 14; Anexo 14).

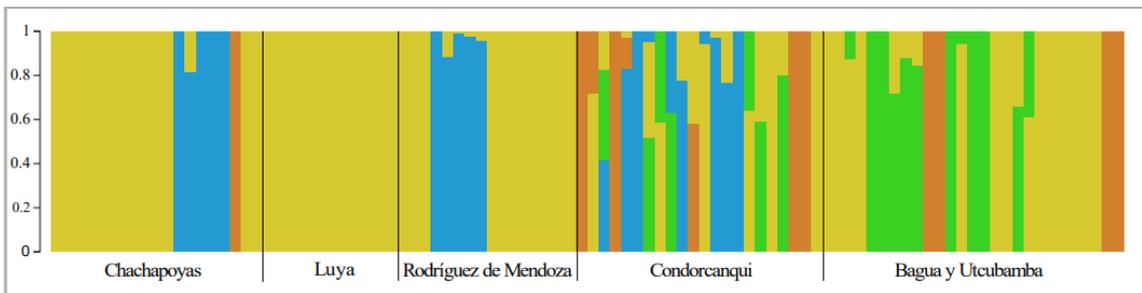


Figura 14. Estructura poblacional de *Theobroma cacao* basada en 16 marcadores SSR, utilizando $K = 4$ como número óptimo de hipotéticas poblaciones ancestrales, que tienen su propio color en el gráfico. Cada una de las accesiones de cacao está representada por una barra vertical coloreada según su coeficiente de mezcla genética para cada población ancestral.

3.2.5. Análisis de asociación

3.2.5.1. Red de extensión mínima

La red de extensión mínima de genotipos multilocus mostró que las 96 accesiones de cacao evaluadas pertenecieron a genotipos multilocus únicos, incluyendo las dos muestras de cacao CCN-51 (Figura 15). Se encontró que las accesiones de cacao de las provincias de Chachapoyas y Luya se agruparon en la parte inferior de la red, mostrando una mayor similitud genética en la mayoría de accesiones ubicada en la esquina inferior

derecha. En contraste, las accesiones de la provincia de Rodríguez de Mendoza se ubicaron en la parte superior. Por otro lado, las accesiones de cacao de las provincias de Condorcanqui, Bagua y Utcubamba se distribuyeron a lo largo de toda la red mínima de extensión (Figura 15).

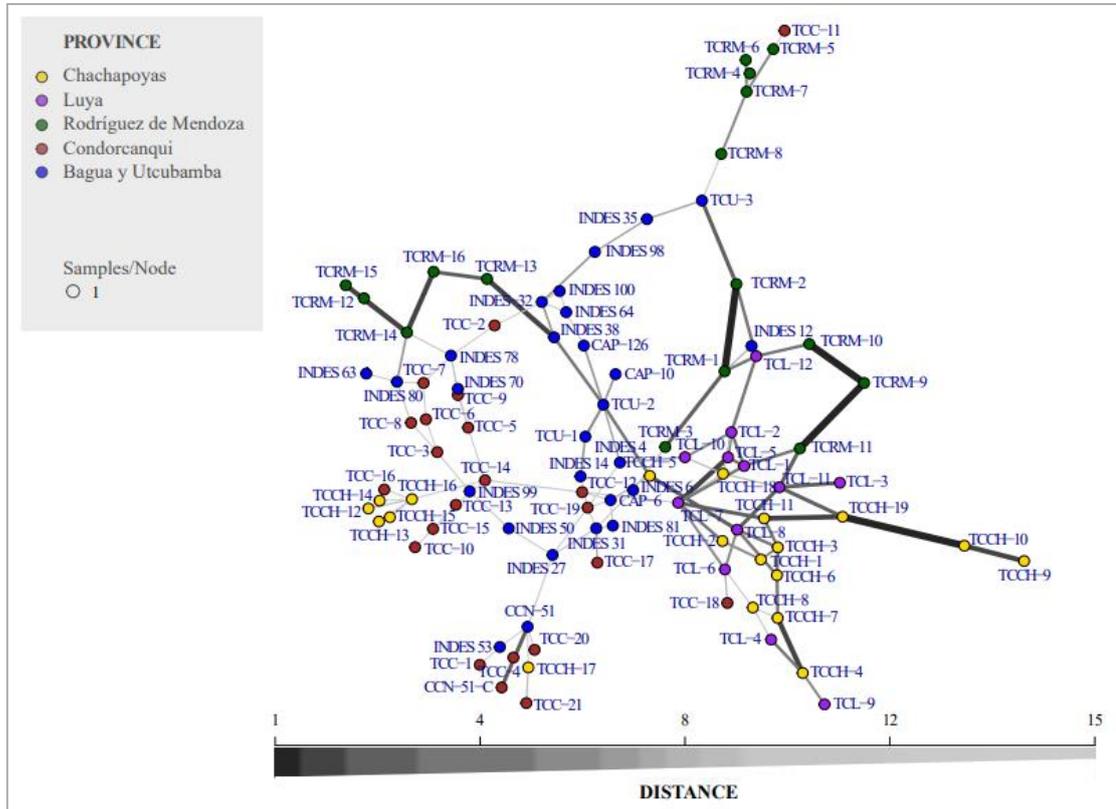


Figura 15. Red de extensión mínima de *Theobroma cacao*. Los nodos (círculos) representan genotipos multilocus (MLG); el grosor de línea de conexión entre nodos, la relación basada en el número de alelos diferentes (datos de repetición de secuencia simple [SSR]). El grosor de las líneas se escala proporcionalmente a la inversa de la distancia genética observada, es decir, cuanto más gruesa es la línea, más relacionados están los nodos conectados. Las longitudes de línea son arbitrarias. Cada provincia está indicada por colores.

3.2.5.2. Análisis de componentes principales (ACP)

En el gráfico de ACP, los dos primeros componentes principales indican una varianza de 35.00%, donde el PC1 abarca el 24.55% y el PC2, el 10.45%. Este análisis mostró relación entre las accesiones que conforman la provincia de Chachapoyas, Rodríguez de Mendoza, Condorcanqui, Bagua y Utcubamba. Sin embargo, se aprecia que las muestras de Luya no presentan relación con las accesiones pertenecientes a la provincia de Condorcanqui, Bagua y Utcubamba (Figura 16).

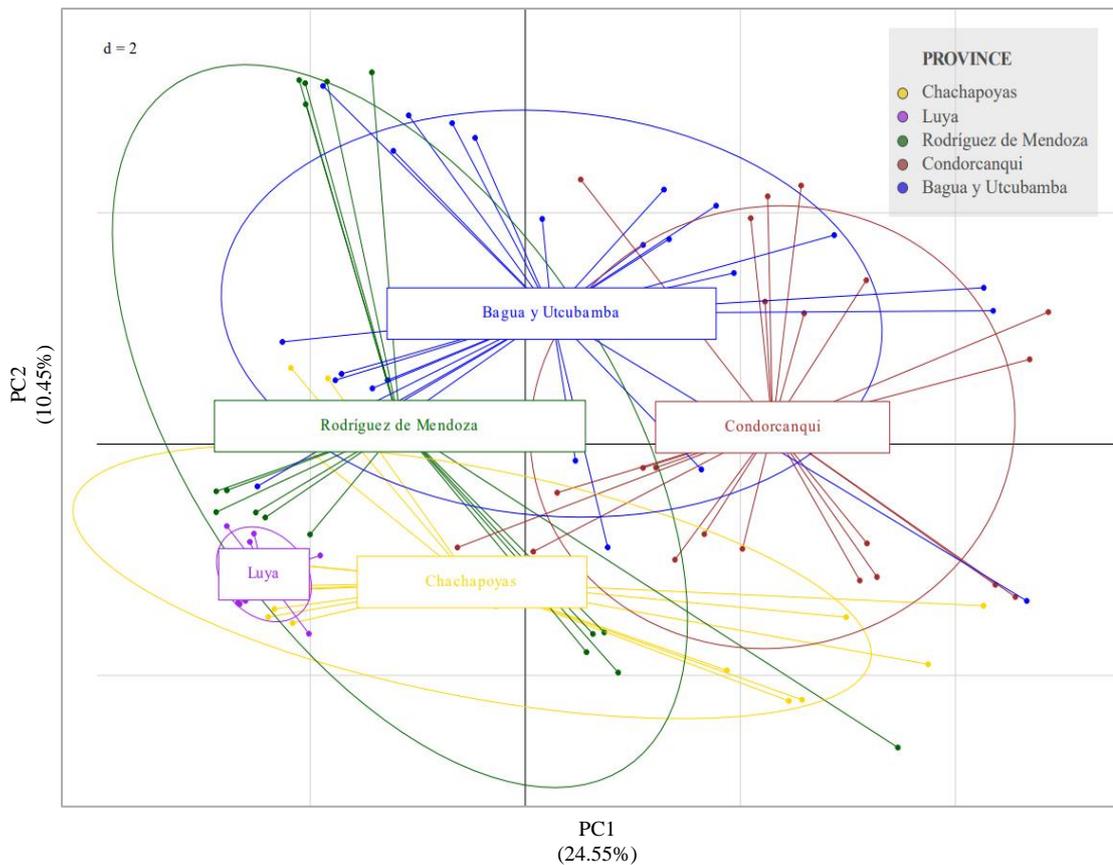


Figura 16. Análisis de asociación de muestras de 96 accesiones de *Theobroma cacao* L., basado en 16 marcadores microsatélite SSR. Análisis de componentes principales (PCA) que representa los dos primeros componentes principales acumulando un total del 35.00% de la diversidad. Cada provincia está indicada por colores. Las elipses indican la distribución de las muestras. Las líneas indican el valor del coeficiente de esa variable en ese componente. Cada punto representa una muestra.

3.2.5.3. Método de grupo de pares no ponderado con media aritmética (UPGMA)

El dendograma de UPGMA evidencia la presencia de cuatro grupos diferenciados, aunque con bajo soporte. En el primer grupo se encuentran un total de 12 accesiones de CNFA pertenecientes a accesiones de la provincia de Condorcanqui, Bagua y Utcubamba. En el segundo grupo se encuentran 14 accesiones de CNFA de la provincia de Condorcanqui, Rodríguez de Mendoza y Chachapoyas. En el tercer grupo se encuentran 64 accesiones de las cinco provincias evaluadas. En el cuarto grupo sólo se encuentran a 6 accesiones de cacao, de las cuales 2 son los CCN-51 (CCN-51 y CCN-51-C), una muestra de la provincia de Chachapoyas (TCCH-17) y las 3 restantes pertenecen a la provincia de Condorcanqui (TCC-4, TCC-20 y TCC-21) (Figura 17).

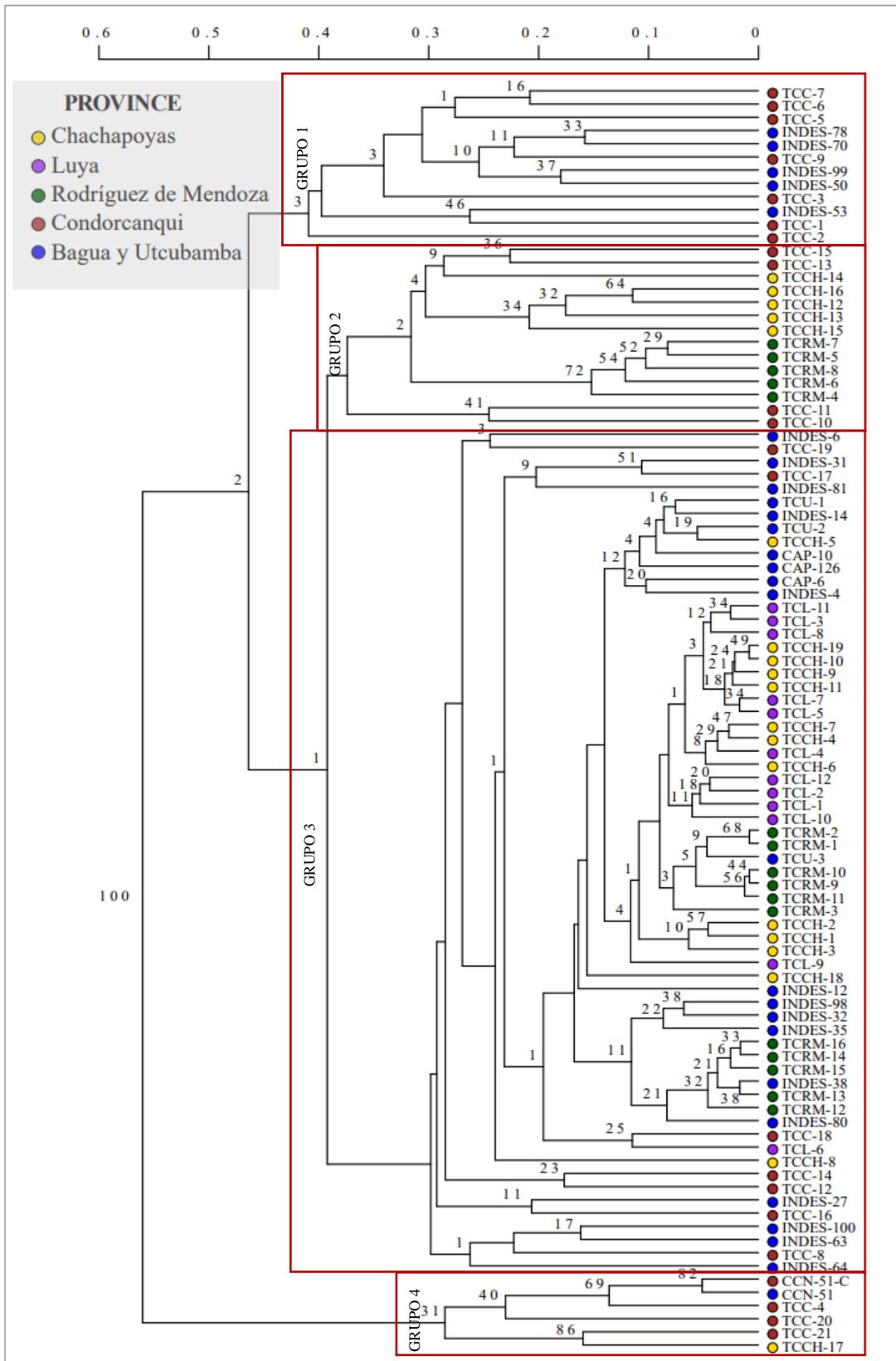


Figura 17. Dendrograma UPGMA de la diversidad genética de 96 accesiones de *Theobroma cacao*, construido en base a las distancias genéticas de Nei, evaluados con 16 marcadores microsatélite SSR.

3.3. Evaluación de la diversidad genética de acuerdo a la altitud

3.3.1. Diversidad genética

De las 96 accesiones evaluadas, 70 accesiones se encuentran a una altitud mayor a los 500 m.s.n.m y 26 accesiones se encuentran a menos de 500 m.s.n.m. Los resultados para el análisis de datos mostraron un mayor índice de diversidad de genes Nei (Hexp) en las accesiones que se encuentran a menos de los 500 m.s.n.m de altitud, alcanzando un valor de 0.650, frente a 0.478 para las muestras de arriba de 500 m.s.n.m. (Tabla 10). Sin embargo, en el índice de Shannon-Weiner (H) y Simpson (lambda) la población de cacao a una altitud mayor de los 500 m.s.n.m alcanza valores superiores con 4.25 y 0.986, respectivamente, versus valores de 3.26 y 0.962, respectivamente, para las muestras a menos de 500 m.s.n.m. (Tabla 10). Los resultados de Hexp se encuentran influenciados por la heterocigosidad presente en las muestras de Condorcanqui, evidenciándose una mayor variabilidad genética en esta provincia y menor grado de *inbreeding*. Por otro lado, el índice de Evenness (*E5*) indica que los genotipos se encuentran distribuidos de manera uniforme en todas las provincias evaluadas (*E5* = 1) (Tabla 10).

Tabla 10. Tabla de diversidad por población con accesiones de cacao de acuerdo con la altitud

Provincia	N	MLG	Hexp	H	lambda	G	E5	Altitud
Condorcanqui	22	22	0.667	3.09	0.955	22	1	
Bagua y Utcubamba	4	4	0.359	1.39	0.750	4	1	
Total	26	26	0.650	3.26	0.962	26	1	< 500 m.s.n.m
Chachapoyas	19	19	0.478	2.94	0.947	19	1	
Luya	12	12	0.217	2.48	0.917	12	1	
Rodríguez de Mendoza	16	16	0.383	2.77	0.938	16	1	> 500 m.s.n.m
Bagua y Utcubamba	23	23	0.552	3.14	0.957	23	1	
Total	70	70	0.478	4.25	0.986	70	1	

Nota: N = Número de muestras, MLG = Número de genotipos multilocus o linajes multilocus, Hexp = Diversidad de genes Nei (Heterocigosidad esperada), H = Índice de Shannon-Weiner, lambda = Índice de Simpson, G = Índice de diversidad de Stoddard y Taylor, E5 = Índice de Evenness.

3.3.2. Análisis de variancia molecular (AMOVA)

El AMOVA mostró que un gran componente de la variación total, equivalente al de 63.08%, es atribuible a las variaciones dentro de las muestras, seguido por variaciones entre muestras dentro de provincia con un valor total de 23.17%. Por el contrario, se evidencia una baja variación entre altitud, alcanzando un porcentaje total de 3.96%

acumulado (Tabla 11). El valor de $\Phi = 0.37$, calculado entre el total de muestras, alcanzó el mayor valor, lo cual evidencia una diferenciación genética entre ellas. La altitud total alcanzó el menor valor ($\Phi = 0.04$; Tabla 12). En cuanto a los tipos de variación analizados, se muestra que la variación entre provincia dentro de altitud, entre muestras dentro de provincia y dentro de las muestras es significativo ($p = 0.001$). Sin embargo, en la variación entre altitud los resultados indican que no es significativo ($p = 0.337$; Tabla 12). Esto indicaría que la altitud no influye en la diversidad genética del cacao en Amazonas. Por el contrario, se encontró que la procedencia de la muestra tiene influencia en la estructura genética.

Tabla 11. Análisis de variancia molecular (AMOVA) generada para accesiones de *Theobroma cacao*.

Tipo de variación	Df	Sum Sq	Mean Sq	Sigma	%	<i>p</i>
Entre altitud	1	71.28	71.28	0.36	3.96	0.337
Entre provincia dentro de altitud	4	143.92	35.98	0.89	9.78	0.001
Entre muestras dentro de provincia	90	897.72	9.97	2.11	23.17	0.001
Dentro de las muestras	96	552.00	5.75	5.75	63.08	0.001
Total	191	1664.91	8.72	9.12	100.00	

Nota: Df = grados de libertad, Sum Sq = Suma de cuadrados observada, Mean Sq = media de cuadros observada, Sigma = varianza estimada, % = variación, $p < 0.05$ es significativo

Tabla 12. Distancia genética utilizado en análisis de estructura genética y análisis de varianza molecular (AMOVA).

	Phi
Phi-muestras-total	0.37
Phi-muestras-provincia	0.27
Phi-provincia-altitud	0.10
Phi-altitud-total	0.04

Nota: El valor de Phi oscila entre 0 y 1. Cuanto mayor sea el valor, mayor será la estructura genética entre los grupos.

3.4. Análisis de componentes principales (ACP) de accesiones de la colección INDES-CES

En este análisis se examinaron 20 accesiones de cacao de la colección INDES-CES (Oliva-Cruz *et al.*, 2022) para verificar la influencia de la genética en las características morfológicas y organolépticas. Se observó, que los dos primeros componentes principales explicaron el 36.48% de la varianza, donde el PC1 representó el 24.12% y el PC2, el 12.36%. Las accesiones de cacao caracterizados como los Bagüinos difieren totalmente

de los Toribianos y los Cajas (Figura 18). Estas variaciones pueden estar relacionadas con características morfológicas como la forma y rugosidad de la superficie, así como con características productivas como el tamaño de los frutos. Los Bagüinos presentaban frutos oblongos, mientras que los Toribianos y los Cajas tenían frutos de forma elíptica. Sin embargo, el AMOVA indicó que la mayor variación total (80.20%) se encontraba dentro de las muestras, independientemente a los 5 grupos evaluados, y no estaba relacionada con las características morfológicas. La variación entre grupos fue no significativo ($p = 0.088$) con un nivel de significancia establecido en $p < 0.05$ (Tabla 13). Estos hallazgos se corroboraron mediante el análisis de distancia genética, donde valor Phi para el grupo total fue el más bajo en comparación con otros parámetros evaluados, con un valor de 0.04 (Tabla 14).

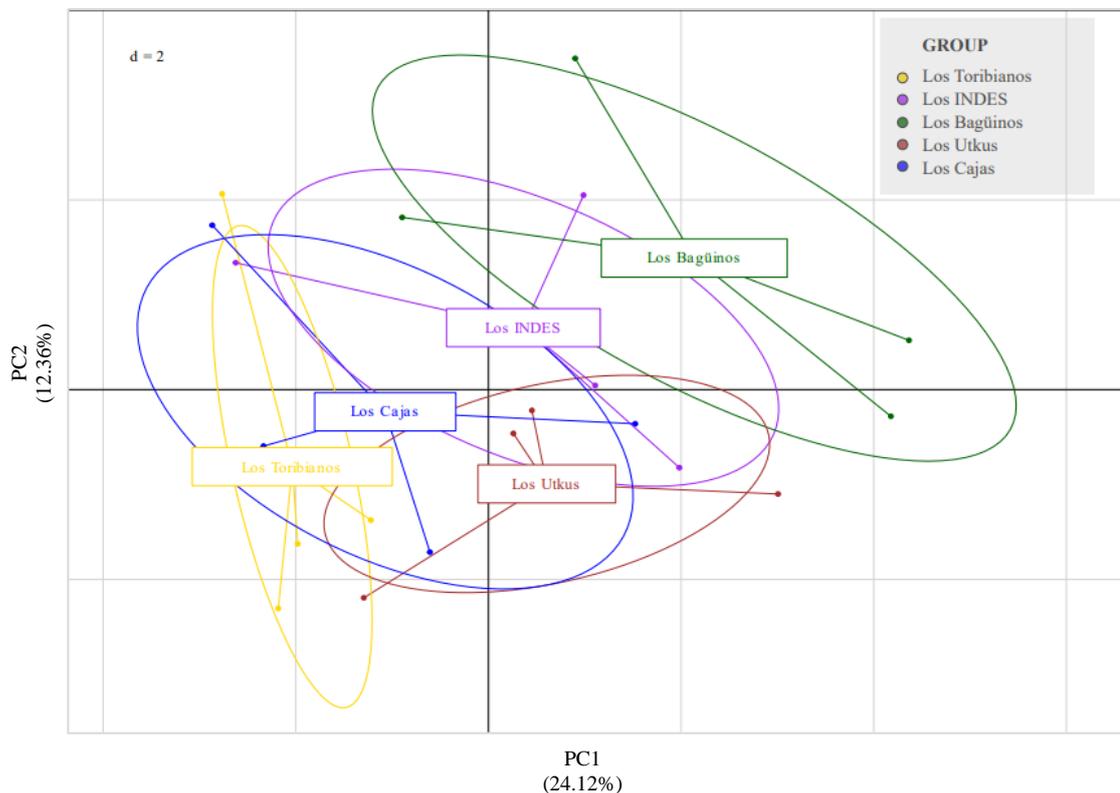


Figura 18. Análisis de asociación de muestras de 20 accesiones de *Theobroma cacao* L. de la colección INDES-CES, basado en 16 marcadores microsatélite SSR. Análisis de componentes principales (PCA) que representa los dos primeros componentes principales acumulando un total del 36.48% de la diversidad. Cada grupo está indicado por colores. Las elipses indican la distribución de las muestras. Las líneas indican el valor del coeficiente de esa variable en ese componente. Cada punto representa una muestra.

Tabla 13. Análisis de variancia molecular (AMOVA) generada para 20 accesiones de *Theobroma cacao* de la colección INDES-CES.

Tipo de variación	Df	Sum Sq	Mean Sq	Sigma	%	<i>p</i>
Entre grupo	4	49.90	12.48	0.31	3.52	0.088
Entre muestras dentro de grupo	15	149.75	9.98	1.44	16.28	0.004
Dentro de las muestras	20	142.00	7.10	7.10	80.20	0.002
Total	39	341.65	8.76	8.85	100.00	

Nota: Df = grados de libertad, **Sum Sq** = Suma de cuadrados observada, **Mean Sq** = media de cuadros observada, **Sigma** = varianza estimada, % = variación, **p < 0.05** es significativo

Tabla 14. Distancia genética utilizado en análisis de estructura genética y análisis de varianza molecular (AMOVA).

	Phi
Phi-muestras-total	0.20
Phi-muestras-grupo	0.17
Phi-grupo-total	0.04

Nota: El valor de Phi oscila entre 0 y 1. Cuanto mayor sea el valor, mayor será la estructura genética entre los grupos.

IV. DISCUSIÓN

Los grupos genéticos del cacao fueron identificados a partir del uso de marcadores SSR nucleares ubicados dentro de los diez cromosomas de cacao (Motamayor *et al.*, 2008; Irish *et al.*, 2010; Everaert *et al.*, 2017; López *et al.*, 2021). El estudio de mayor ámbito registrado con marcadores SSR fue realizado por Motamayor *et al.* (2008), en la que proporcionó la clasificación del cacao agrupada en 10 grupos genéticos (Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañón, Nacional, Nanay y Purús). Asimismo, los marcadores microsatélites SSR no sólo tienen el potencial para ser utilizados en estudios de diversidad genética, sino también son utilizados en el análisis de parentesco de poblaciones (Azpeitia-Morales *et al.*, 2017; Izzah *et al.*, 2022) e identificación de genes de resistencia a enfermedades como escoba de bruja (Brown *et al.*, 2005) y muerte regresiva por rayado vascular (Chandrakant *et al.*, 2019). En este estudio se logró diseñar 41 marcadores microsatélites SSR.

De los 41 marcadores SSR diseñados, se utilizaron 16 marcadores microsatélites SSR (15 genómicos y uno del cloroplasto), por presentar amplificaciones menos ambiguas y menor número de alelos nulos, para la caracterización de genotipos del cacao de Amazonas. Esto indica que no todos los marcadores diseñados presentaron amplificaciones de los fragmentos PCR, pese a las mismas condiciones de fusión de los *primers* al momento del diseño. Resultados similares fueron encontrados por Herrera-García *et al.* (2015), donde el marcador MTcCIR12 fue el locus más polimórfico e informativo, comportándose mejor que los demás marcadores utilizados; mientras que los marcadores MTcCIR22 y MTcCIR15 obtuvieron menor número de alelos y muestras amplificadas. Por su parte, Barlaan *et al.* (2018), en el desarrollo de marcadores genéticos asistidos por genómica funcional y bioinformática, encontraron en 215 muestras de biblioteca, un total de 578 loci SSR, con 271 loci perfectos y 307 ambiguos. Por otro lado, el único marcador monomórfico obtenido, fue el que provenía del cloroplasto (INDES_mTC90). Esto difiere con otros estudios como el de Yang *et al.* (2011) y Nieves-Orduña *et al.* (2021), quienes encontraron niveles importantes de diversidad de cacao con marcadores microsatélite de cloroplasto. A pesar de todo esto, los 16 marcadores SSR utilizados en este estudio fueron capaces de alcanzar la planicie en la curva de acumulación de genotipos en todas las accesiones evaluadas. Esto nos indica que, para el número de muestras trabajadas, la adición de marcadores extras no incrementaría el

número de genotipos observados. En otras palabras, no había la necesidad de genotipar las accesiones de cacao con más marcadores, lo que a su vez redujo los costos de análisis.

La diversidad genética del cacao ha sido objeto de numerosos estudios que han utilizado marcadores moleculares SSR. Estas investigaciones han abarcado desde estudios locales a nivel de país como Bolivia (Zhang *et al.*, 2011), Honduras (López *et al.*, 2021) y Ecuador (Fouet *et al.*, 2022) hasta estudios de alcance internacional como el de Motamayor *et al.* (2008) (con muestras de Belice, Brasil, Colombia, Costa Rica, Ecuador, Guayana Francesa, Ghana, México, Nicaragua, Panama, Perú y Venezuela), y el de Nieves-Orduña *et al.* (2021) (con muestras de Brasil, Colombia, Ecuador, Guayana Francesa, Perú). Para nuestro estudio, se analizaron 96 accesiones de cacao que fueron recolectadas en distritos cacaoteros de Amazonas que han sido escasamente investigados. Estos distritos incluyen Providencia, Ocalli, Pisuquia y Camporedondo, que se encuentran en la provincia de Luya.

De los 16 marcadores moleculares SSR utilizados, nueve fueron homólogos a los utilizados por Motamayor *et al.* (2008) y siete fueron diseñados *de novo* para este estudio. El número de alelos de cada marcador SSR influye en el nivel de polimorfismo en las accesiones de cacao. En este estudio, los niveles de polimorfismo variaron de 2 a 17 alelos por cada locus (sin contar con el marcador monomórfico cloroplastidial), identificando un total de 129 alelos, con una media de 6 para los marcadores homólogos a los de Motamayor *et al.* (2008) y de 12.33 para los diseñados *de novo* en este estudio. Everaert *et al.* (2017) amplificaron 123 alelos en 14 marcadores SSR, obteniendo una variación de 4 a 13 alelos con una media de 9 por marcador, y encontrando que los marcadores mTcCIR15, mTcCIR33 y mTcCIR37 fueron los más polimórficos. Sin embargo, para nuestro estudio el marcador INDES-mTC17, homólogo a mTcCIR33, y INDES-mTC1, homólogo a mTcCIR37, presentaron amplificaciones nulas, por lo que no fueron seleccionados dentro de los 16 marcadores SSR utilizados. Por su parte, Fouet *et al.* (2022) reportaron un número de alelos observados por locus entre 2 y 19, con un promedio de 7.69 en toda la población. López *et al.* (2021) reportó una media de 5.5 alelos con una variación de 3 a 11, con un total de 88 alelos, donde el MTcCIR33 fue el más polimórfico. El marcador INDES_mTC113, diseñado y seleccionado en este estudio debido a su alta cantidad de repeticiones en su patrón de ADN repetitivo, demostró ser el marcador con mayor polimorfismo en las accesiones evaluadas, presentando un total de 17 alelos. Este marcador no ha sido utilizado anteriormente en ningún estudio de

diversidad con marcadores SSR, por lo que se debe tener en cuenta en futuros estudios de diversidad.

Los 16 marcadores moleculares SSR mostraron una alta diversidad genética ($H_{exp} = 0.545$) respecto a lo reportado por otros investigadores que también caracterizaron accesiones de cacao con SSR (López *et al.* (2021), Honduras: 0.540; Quiroz-Vera *et al.* (2023), Ecuador: 0.447). Sin embargo, Fouet *et al.* (2022) utilizando 48 marcadores SSR caracterizaron 283 accesiones de cacao procedentes de las provincias amazónicas ecuatorianas Zamora-Chinchipec, Morona-Santiago y Pastaza, encontrando una diversidad genética superior al de este estudio ($H_{exp} = 0.603$). Por su parte, en Perú, principalmente de la región Amazonas, Bustamante *et al.* (2022) caracterizaron 146 árboles de cacao con 192 marcadores SNP registrando que las muestras tenían una diversidad genética moderada ($H_{exp} = 0.336$) y ascendencia compartida con los grupos Nacional, Amelonado, Iquitos y Criollo. La alta heterocigosidad encontrada en las muestras analizadas puede explicarse por el origen del cacao (Saltini *et al.*, 2013; Ceccarelli *et al.*, 2022). Nieves-Orduña *et al.* (2021) encontraron una alta diversidad genética en la Amazonia occidental, donde Perú y Ecuador fueron los que presentaron mayor diversidad de haplotipos. Esto sugiere que la distribución de la diversidad de haplotipos tiene mucha relación con el centro de origen del cacao, donde la región occidental de la Amazonía es el principal centro de diversificación del cacao (Tschardt *et al.*, 2022).

Los marcadores moleculares son herramientas esenciales en el análisis genético del cacao (Barlaan *et al.*, 2018). Los SSR son flexibles a la hora de analizar un pequeño número de individuos y tienen una gran capacidad para diferenciarse entre ellos (Van-Inghelandt *et al.*, 2010). En este estudio, los nuevos marcadores SSR diseñados *de novo* proporcionaron niveles más altos de diversidad genética que los históricamente utilizados. Estos resultados evidencian las ventajas que ofrecen los marcadores microsatélites SSR en la caracterización de la diversidad genética del cacao.

Oliva-Cruz *et al.* (2022) reportaron que ecotipos de CNFA en altitudes medias y altas, por encima de los 500 m.s.n.m, les permite expresar de una mejor manera sus atributos sensoriales y productivos. Sin embargo, en este estudio no se encontró influencia de la altitud en la diversidad genética. El análisis de la diversidad a nivel altitudinal mostró índices de diversidad genotípica más altos en las accesiones que se encontraban ubicadas mayor a los 500 m.s.n.m; pero se evidenció lo contrario en el índice de heterocigosidad.

Estos resultados, podrían explicarse debido a la alta heterocigosidad encontrada en las muestras, todas a menos de 500 m.s.n.m., de la provincia de Condorcanqui (Hexp = 0.667). Según el análisis de AMOVA realizado, sugirió que la mayor variación de diversidad genética (63.08%) se encuentra dentro de las muestras, seguido por entre muestras dentro de provincia (23.17%), y con un aporte minúsculo y no significativo, la altitud (3.96%).

El dendograma UPGMA y la red de extensión mínima, mostraron que cada una de las accesiones evaluadas presentaban un perfil genético único. Estos resultados, son similares a lo reportado por Bustamante *et al.* (2022) donde indicaron que las muestras de cacao recolectadas en el norte de Perú eran generalmente distintas y únicas. Por lo tanto, estos estudios evidencian la alta diversidad genética nativa existente en el Perú. Por su parte, Avendaño-Arrazate *et al.* (2018) realizaron el análisis de 99 accesiones de cacao provenientes del Banco de Germoplasma del Campo Experimental Rosario Izapa, México; encontrando en el dendograma UPGMA que el genotipo CRIM-88, procedente de Chiapas, fue el único que presentó diferencias en comparación con las demás accesiones. Estos resultados revelaron que en México existen índices bajos de diversidad genética en el cacao, lo cual contrasta significativamente con los hallazgos de este estudio.

En cuanto al PCA se encontraron que las 96 accesiones evaluadas obtuvieron una variación total de 35.00%. Bidot-Martínez *et al.* (2016) en el análisis genético de 537 accesiones de cacao tradicional cubano, reportaron que los dos primeros componentes principales obtuvieron un valor de 33.95%. Por su parte, Boza *et al.* (2012) en el análisis de diversidad genética de 955 accesiones de cacao procedentes de República Dominicana, encontraron que los ejes de los dos primeros componentes principales representaron 12.8 % de la variación total. Por lo tanto, los valores encontrados en este PCA no difieren y en algunos casos explican mejor la variación que en estudios anteriores.

El número de poblaciones ancestrales encontrados en este estudio ($K = 4$) son similares a lo estimado por Fouet *et al.* (2022) donde encontraron que 283 accesiones de cacao del Ecuador derivaban de cuatro poblaciones ancestrales. López *et al.* (2021) en Honduras y Everaert *et al.* (2017) en Vietnam, reportaron que sus muestras procedían de dos poblaciones ancestrales ($K = 2$). Los resultados obtenidos corroboran a los que se han reportado en Bustamante *et al.* (2022). De acuerdo con dicho estudio, se analizaron 134 muestras de cacao y se encontró que estas muestras tenían aportes genéticos de 2 a 7

grupos distintos, que habían sido previamente identificados en el trabajo de Motomayor *et al.* (2008). Además, se observó que la combinación de tres grupos genéticos fue la más común entre las muestras analizadas. Esto significa que, en la mayoría de las muestras de cacao estudiadas, se encontró una mezcla de material genético proveniente de tres grupos específicos (Bustamante *et al.*, 2022). Estos resultados son importantes para comprender la diversidad genética presente en el cacao y brindan información sobre la composición genética de las muestras analizadas. También sugieren que existe variabilidad genética en estas muestras y que hay una mezcla de diferentes linajes genéticos en la población estudiada. Asimismo, las muestras provenientes de Condorcanqui tenían un mayor número de poblaciones ancestrales ($K = 4$). Se infiere que los altos niveles de heterocigosidad encontrados en las muestras de Condorcanqui podrían estar relacionadas con su estado silvestre. Esto se debe a que, en ecosistemas naturales y poco intervenidos por la actividad humana, como los bosques silvestres, hay una mayor disponibilidad de diversidad genética. Por lo tanto, la heterocigosidad del cacao es influenciada por factores como la región geográfica y las prácticas de agricultura, que tienden a reducir la heterocigosidad.

Un caso interesante y particular lo integran las accesiones de Luya. Los análisis de asociación mostraron que las accesiones de la provincia de Luya son diferentes a las accesiones de Condorcanqui, Bagua y Utcubamba, principales provincias cacaoteras de la región Amazonas. Las muestras de la provincia de Luya provienen de un área donde el cacao crece en las playas del río Marañón dentro del ecosistema “Bosques Tropicales Estacionalmente Secos del Marañón”, en climas secos y semicálidos. El cacao de esta zona prácticamente nunca ha sido estudiado, siendo este estudio el primero en analizar la diversidad genética del cacao en esta zona. Dada la ubicación en el PCA, y en base a las conversaciones anecdóticas de los agricultores dueños de las parcelas donde se colectó, se puede inferir que el cacao de Luya ha sido llevado a través de propagación clonal a otras zonas como Balsas y Rodríguez de Mendoza. Esto habla de su potencial de buenas características organolépticas y sensoriales. Asimismo, todas las muestras de Luya derivaron de una sola población ancestral y sus genotipos fueron el resultado de la reproducción y diseminación sexual, inferido por el análisis *linkage disequilibrium*, que tuvo que haber ocurrido hace mucho tiempo, probablemente hace miles de años, debido a que actualmente el cacao se propaga vegetativamente. Claramente, el ecosistema Bosques Tropicales Estacionalmente Secos del Marañón albergaron uno de las

poblaciones ancestrales del cacao de Amazonas. Faltan hacer estudios para corroborar si la genética de este cacao corresponde a alguno de los grupos genéticos previamente identificados (Motomayor *et al.*, 2008). Aquí se evidencian la importancia de la conservación *in situ* de estos cacaos que se encuentran expuestas a ser reemplazadas por material genético foráneo.

Asimismo, la influencia de la genética y las características morfológicas del cacao han sido previamente estudiadas. Bidot-Martínez *et al.* (2021) evaluaron la diversidad morfológica y genética del cacao cubano, encontrando que las plantas evaluadas indicaron una alta variabilidad morfológica y una baja variabilidad genética. Bekele *et al.* (2022) en su estudio sobre asociación de todo el genoma del cacao y los ensayos de selección genómica en accesiones de cacao, identificaron asociaciones significativas entre los marcadores y rasgos fenotípicos relacionados con el rendimiento. Resultados similares fueron reportados por Vázquez-Ovando *et al.* (2018) donde encontraron asociaciones entre marcadores moleculares SSR y características químicas y sensoriales en muestras de cacao de México. Lo anterior refleja que las características fenotípicas y productivas tienen relación con la genética del cacao. Oliva-Cruz *et al.* (2022) clasificó al cacao de Amazonas en cinco grupos basados en descriptores morfológicos y sensoriales: los Toribianos, los INDES, los Bagüinos, los Utkus y los Cajas. En este estudio, se logró analizar 5 muestras representativas de cada uno de estos grupos, y se las analizaron para ver si se lograban encontrar influencias del genotipo en las características de los grupos morfológicos y sensoriales de Oliva-Cruz *et al.* (2022), encontrándose que la mayor genética es dentro de las muestras (80.20%).

Todos estos hallazgos, confirman que la región Amazonas presenta una alta diversidad y variabilidad genética de cacao, puesto que cada una de las accesiones evaluadas presentaron un perfil genético único. Mantener esta diversidad genética en su centro de origen es de importancia mundial, especialmente para la mejora del cultivo y la selección de nuevas variedades mejoradas en el futuro (Ficiciyan *et al.*, 2018). La conservación de las variedades nativas del cacao en Perú es un gran desafío y tienen que ser abordadas debido al alto valor sociológico (Tschardtke *et al.*, 2022), y la región Amazonas posee un gran potencial genético autóctono y único que pueden posicionarse en el mercado que actualmente se encuentra en crecimiento. Asociado a ello, la selección de material genético con características promisorias para programas de mejoramiento, resistencia a enfermedades y alto rendimiento son cruciales para la conservación del germoplasma.

V. CONCLUSIONES

Los 16 marcadores utilizados (nueve homólogos a marcadores históricamente utilizados y siete totalmente nuevos diseñados para este estudio), están distribuidos en el genoma del cacao y el cloroplasto y son suficientes para capturar la diversidad de las muestras analizadas, excepto el marcador INDES_mTC90 del cloroplasto por ser monomórfico.

El número promedio de alelos obtenidos con los marcadores nuevos diseñados en este estudio fue de 12.33 sin incluir el marcador del cloroplasto (INDES_mTC90), mientras que los marcadores históricamente utilizados mostraron un promedio de 6.00 alelos.

Cada uno de las 96 accesiones evaluadas presentaron un perfil genético único, asimismo las muestras estaban dispersas en los análisis de asociación, independientemente de la altitud de procedencia de las muestras. La altitud no es un factor que influye en la diversidad genética de cacao en Amazonas.

Las accesiones evaluadas de cacao derivaban de cuatro hipotéticas poblaciones ancestrales, presentes en Condorcanqui, a diferencia de las otras provincias donde se observó un menor número de poblaciones ancestrales de las que provenían sus muestras. En los índices de diversidad, Condorcanqui fue la provincia con mayor índice de heterocigosidad ($H_{exp} = 0.633$), lo que indica una estrecha relación al centro de origen del cacao.

Las accesiones de Luya, específicamente de los Bosques Tropicales Estacionalmente Secos del Marañón difieren a las de Condorcanqui, Bagua y Utcubamba, derivando de una población ancestral con presencia de reproducción sexual, lo que sugiere que estas muestras son nativas y únicas.

VI. RECOMENDACIONES

Realizar muestreos más homogéneos en las provincias cacaoteras de la región Amazonas, mediante la recolección sistemática de material genético por cada provincia evaluada, incluyendo accesiones no solo con características adecuadas recomendadas por los agricultores.

Realizar una evaluación completa a cada una de las accesiones de CNFA como: caracterización morfológica, productiva y organoléptica.

Para la caracterización molecular, corroborar los resultados con otros marcadores moleculares como los “*Single Nucleotide Polymorphism*” (SNPs).

Incluir muestras de los grupos genéticos de cacao conocidos para comparar los con las accesiones de este estudio.

Dado el estado único de las muestras en la provincia de Luya, diferentes a los que se han caracterizado, es necesario desarrollar estrategias de conservación del material genético silvestre en el ecosistema de los Bosques Tropicales Estacionalmente Secos del Marañón.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Agapow, P., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, 1, 101–102.
- Arevalo-Gardini, E., Meinhardt, L. W., Zuñiga, L. C., Arévalo-Gardni, J., Motilal, L., & Zhang, D. (2019). Genetic identity and origin of “Piura Porcelana”—a fine-flavored traditional variety of cacao (*Theobroma cacao*) from the Peruvian Amazon. *Tree Genetics and Genomes*, 15, 11. <https://doi.org/10.1007/s11295-019-1316-y>
- Argout, X., Martin, G., Droc, G., Fouet, O., Labadie, K., Rivals, E., Aury, J. M., & Lanaud, C. (2017). The cacao Criollo genome v2.0: An improved version of the genome for genetic and functional genomic studies. *BMC Genomics*, 18(1), 1–9. <https://doi.org/10.1186/s12864-017-4120-9>
- Azpeitia Morales, A., Gasca González, H. A., & Villordo Pineda, E. (2017). Relaciones de parentesco en híbridos F1 de cacao (*Theobroma cacao* L.) por marcadores moleculares. *Acta Universitaria*, 27(6), 66-77. doi: 10.15174/au.2017.1552
- Avendaño-Arrazate, C. H., López-Gómez, P., Iracheta-Donjuan, L., Vázquez-Ovando, A., Bouchan, R., Cortés-cruz, M., & Borrayo, E. (2018). DIVERSIDAD GENÉTICA Y SELECCIÓN DE UNA COLECCIÓN NÚCLEO PARA LA CONSERVACIÓN A LARGO PLAZO DE CACAO (*Theobroma cacao* L.). *Interciencia*, 43(November), 770–777.
- Barlaan, E. A., Sales, E. K., & Laurena, A. C. (2018). Functional genomics- and bioinformatics-assisted gene marker development for economically important traits in cacao. *Acta Horticulturae*, (1203), 85–92. doi:10.17660/actahortic.2018.1203.13
- Barrientos Felipa, P. (2015). El comercio del cacao y su influencia en la agricultura peruana. *Pensamiento Crítico*, 20(1), 007–023. <https://doi.org/10.15381/pc.v20i1.11481>
- Bartley, B. G. D. (2005). *The genetic diversity of cacao and its utilization*. CABI.
- Bekele FL, Bidaisee GG, Allegre M, Argout X, Fouet O, Boccara M, *et al.* (2022) Genome-wide association studies and genomic selection assays made in a large

- sample of cacao (*Theobroma cacao* L.) germplasm reveal significant marker-trait associations and good predictive value for improving yield potential. *PLoS ONE* 17(10): e0260907. <https://doi.org/10.1371/journal.pone.0260907>
- Bidot Martínez, I., Valdés de la Cruz, M., Riera Nelson, M., & Bertin, P. (2016). Establishment of a Core Collection of Traditional Cuban *Theobroma cacao* Plants for Conservation and Utilization Purposes. *Plant Molecular Biology Reporter*, 35(1), 47–60. doi:10.1007/s11105-016-0999-6
- Boza, E. J., Irish, B. M., Meerow, A. W., Tondo, C. L., Rodríguez, O. A., Ventura-López, M., ... Schnell, R. J. (2012). Genetic diversity, conservation, and utilization of *Theobroma cacao* L.: genetic resources in the Dominican Republic. *Genetic Resources and Crop Evolution*, 60(2), 605–619. doi:10.1007/s10722-012-9860-4
- Bustamante DE, Motilal LA, Calderon MS, Mahabir A and Oliva M (2022) Genetic diversity and population structure of fine aroma cacao (*Theobroma cacao* L.) from north Peru revealed by single nucleotide polymorphism (SNP) markers. *Front. Ecol. Evol.* 10:895056. doi: 10.3389/fevo.2022.895056
- Brown, A., Feldman, M., & Nevo, E. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, 96(10), 523–536. <https://doi.org/10.1007/s000360050102>
- Brown, J. S., Schnell, R.J., Motamayor, J.C., Lopes, U., Kuhn, D. N., & Borrone, J. W. (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 Horticultural Science* jashs, 130(3), 366-373. Retrieved May 22, 2023, from <https://doi.org/10.21273/JASHS.130.3.366>
- Bruvo, R., Michiels, N. K., D'Souza, T. G., & Schulenburg, H. (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, 13(7), 2101–2106. <https://doi.org/10.1111/j.1365-294X.2004.02209.x>
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*,

10, 421. <https://doi.org/10.1186/1471-2105-10-421>

Cattell, Raymond B. (1966). *The Scree Test For The Number Of Factors.* , 1(2), 245–276. doi:10.1207/s15327906mbr0102_10

Chandrakant, E. S., Tulshiram, W. S., Mathew, D., Minimol, J. S., Jose, S., Kurian, S. P., & Shylaja, M. R. (2019). Microsatellite and inter-microsatellite markers linked with resistance to vascular streak dieback in cocoa (*Theobroma cacao* L.). *The Journal of Horticultural Science and Biotechnology*, 1–9. doi:10.1080/14620316.2019.1677509

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 5, 1–9.

Ceccarelli, V., Lastra, S., Loo Solórzano, R. G., Chacón, W. W., Nolasco, M., Sotomayor Cantos, I. A., Plaza Avellán, L. F., López, D. A., Fernández Anchundia, F. M., Dessauw, D., Orozco-Aguilar, L. y Thomas, E. (2022). Conservación y uso de recursos genéticos de cacao (*Theobroma cacao* L.) por bancos de genes y viveros en seis países de América Latina. *Recursos genéticos y evolución de cultivos*, 69, 1283-1302. <https://doi.org/10.1007/s10722-021-01304-3>

Díaz-Valderrama, J. R., Leiva-Espinoza, S. T., & Aime, M. C. (2020). The History of Cacao and Its Diseases in the Americas. *Phytopathology®*, 110(10), 1604–1619. doi:10.1094/phyto-05-20-0178-rvw

Dray, S., & Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, 22(4), 1–20.

Everaert, H., Rottiers, H., Pham, P. H. D., Ha, L. T. V., Nguyen, T. P. D., Tran, P. D., De Wever, J., Maebe, K., Smagghe, G., Dewettinck, K., & Messens, K. (2017). Molecular characterization of Vietnamese cocoa genotypes (*Theobroma cacao* L.) using microsatellite markers. *Tree Genetics and Genomes*, 13(5). <https://doi.org/10.1007/s11295-017-1180-6>

Everaert, H., De Wever, J., Tang, T. K. H., Vu, T. L. A., Maebe, K., Rottiers, H., ... Messens, K. (2020). Genetic classification of Vietnamese cacao cultivars assessed

- by SNP and SSR markers. *Tree Genetics & Genomes*, 16(3). doi:10.1007/s11295-020-01439-x
- Ficiciyan, A., Loos, J., Sievers-Glotzbach, S. y Tschardtke, T. (2018). Más que rendimiento: Se revisan los servicios ecosistémicos de las variedades de cultivos tradicionales frente a las modernas. *Sostenibilidad*, 10, 2834. <https://doi.org/10.3390/su10082834>
- Fouet, O., Llor Solórzano, R. G., Rhoné, B., Subía, C., Calderón, D., Fernández, F., Sotomayor, I., Rivallan, R., Colonges, K., Vignes, H., Angamarca, F., Yaguana, B., Costet, P., Argout, X., & Lanaud, C. (2022). Collection of native *Theobroma cacao* L. accessions from the Ecuadorian Amazon highlights a hotspot of cocoa diversity. *Plants, People, Planet*, 1–13. <https://doi.org/10.1002/ppp3.10282>
- Frichot, E., & François, O. (2015). *LEA*: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925–929. <https://doi.org/10.1111/2041-210X.12382>
- García-Carrión, L. F. (2008). M & O CONSULTING S.A.C. Estudio de Caracterización del Potencial Genético del Cacao en el Perú. Obtenido de https://www.midagri.gob.pe/portal/download/pdf/direccionesyoficinas/dgca/estudio_caracterizacion.pdf
- Gerlach, G., Jueterbock, A., Kraemer, P., Deppermann, J. and Harmand, P. 2010 Calculations of population differentiation based on G(ST) and D: forget G(ST) but not all of statistics! *Molecular Ecology* 19, 3845-3852
- Goñas M, Rubio KB, Rojas Briceño NB, Pariente-Mondragón E, Oliva-Cruz M (2022) Tree diversity in agroforestry systems of native fine-aroma cacao, Amazonas, Peru. *PLoS ONE* 17(10): e0275994. <https://doi.org/10.1371/journal.pone.0275994>
- Gopaulchan, D., Motilal, L. A., Bekele, F. L., Clause, S., Ariko, J. O., Ejang, H. P., & Umaharan, P. (2019). Morphological and genetic diversity of cacao (*Theobroma cacao* L.) in Uganda. *Physiology and Molecular Biology of Plants*. doi:10.1007/s12298-018-0632-2

- Guiraud, B. S. H. B., Tahí, M. G., Fouet, O., Trebissou, C. I., Pokou, D., Rivallan, R., ... Lanaud, C. (2018). Assessment of genetic diversity and structure in cocoa trees (*Theobroma cacao* L.) in Côte d'Ivoire with reference to their susceptibility to Cocoa swollen shoot virus disease (CSSVD). *Tree Genetics & Genomes*, 14(4). doi:10.1007/s11295-018-1264-y
- Guzmán Bautista, J. H. (2020). Producción de cacao "*Theobroma cacao* L." fino o de aroma peruano: Agronegocio sostenible. *Alpha Centauri*, 1(1), 49–55. <https://doi.org/10.47422/ac.v1i1.6>
- Grünwald, N.J., Goodwin, S.B., Milgroom, M.G. & Fry, W.E. (2003) Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology*, 93, 738–746.
- Henao-Ramírez, AM, Salazar Duque, HJ, Calle Tobón, AF, & Urrea Trujillo, AI (2021). Determinación de la Estabilidad Genética en Plantas de Cacao (*Theobroma cacao* L.) Derivadas de Embriogénesis Somática Utilizando Marcadores Moleculares Microsatélites (SSR). *Revista Internacional de Ciencias de la Fruta*, 21(1), 284–298. doi:10.1080/15538362.2021.1873219
- Herrera-García, K., Aragón-Obando, E., & Aguilar-Bustamante, V. (2015). Diversidad genética en 105 accesiones de cacao (*Theobroma cacao* L.) colectadas en nicaragua, utilizando 10 marcadores moleculares tipo SSR. *La Calera*, 15(25), 54–62. Recuperado a partir de <https://lcalera.una.edu.ni/index.php/CALERA/article/view/260>
- INDECOPI (2016). *El Indecopi entrega décima Denominación de Origen Cacao Amazonas Perú que contribuirá al desarrollo económico de más de 1,200 productores*. <http://hdl.handle.net/11724/5443>
- Irish, B. M., Goenaga, R., Zhang, D., Schnell, R., Brown, J. S., & Motamayor, J. C. (2010). Microsatellite Fingerprinting of the USDA-ARS Tropical Agriculture Research Station Cacao (*Theobroma cacao* L.) Germplasm Collection. *Crop Science*, 50, 656–667. <https://doi.org/10.2135/cropsci2009.06.0299>
- Izzah, I., Sulistiyorini1, & Wicaksono, I.N.A. (2022). Identification of true male parents

- in F1 populations of cacao using SSR markers. IOP Conf. Ser.: Earth Environ. Sci. 974 012051. doi:10.1088/1755-1315/974/1/012051
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>
- Laliberté, B., End, M., Cryer, N., Daymond, A., Engels, J., Bernardus, . . . Weise, S. (2018). Conserving and exploiting cocoa genetic resources: the key challenges. doi:10.19103/as.2017.0021.02
- Lehmann, S. (2015). Estudio de mercado para el cacao fino de aroma proveniente de las regiones de Amazonas y San Martín en Perú. Lima: International Trade Center.
- Lins, A. B., Lawinsky, P. R., Barbosa, A. M. M., Gaiotto, F. A., & Corrêa, R. X. (2016). Molecular genetic diversity in a core of cocoa (*Theobroma cacao* L.) clones with potential for selection of disease resistance, plant height and fruit production. *African Journal of Biotechnology*, 15(44), 2517–2523. <https://doi.org/10.5897/AJB2015.15188>
- López, M.; Gori, M.; Bini, L.; Ordoñez, E.; Durán, E.; Gutierrez, O.; Masoni, A.; Giordani, E.; Biricolti, S.; Palchetti, E. Genetic Purity of Cacao Criollo from Honduras Is Revealed by SSR Molecular Markers. *Agronomy* 2021, 11, 225. <https://doi.org/10.3390/agronomy11020225>
- Megléc, E., Pech, N., Gilles, A., Dubut, V., Hingamp, P., Trilles, A., Grenier, R., & Martin, J. F. (2014). QDD version 3.1: A user-friendly computer program for microsatellite selection and primer design revisited: Experimental validation of variables determining genotyping success rate. *Molecular Ecology Resources*, 14(6), 1302–1313. <https://doi.org/10.1111/1755-0998.12271>
- Ministerio de Desarrollo Agrario y Riego. MIDAGRI. (2020). Producción nacional de cacao en grano creció en la última década a un promedio de 12.6% al año. Obtenido de <https://www.gob.pe/institucion/midagri/noticias/305143-produccion-nacional-de-cacao-en-grano-crecio-en-la-ultima-decada-a-un-promedio-de-12-6-al-ano>

- Ministerio de Desarrollo Agrario y Riego. MIDAGRI. (2022). Obtenido de Observatorio de Commodities. Cacao. Boletín Trimestral N°. 01-2022: <https://cdn.www.gob.pe/uploads/document/file/3561419/Commodities%20Cacao%3A%20ene-mar%202022.pdf>
- Motamayor, J. C., Lachenaud, P., da Silva e Mota, J. W., Loor, R., Kuhn, D. N., Brown, J. S., & Schnell, R. J. (2008). Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). *PLoS ONE*, 3(10), e3311. <https://doi.org/10.1371/journal.pone.0003311>
- Motamayor, Juan. C., Schnell, R., Kuhn, D., Phillips, W., Haiminen, N., Livingstone, D., Schmutz, J., Findley, S., Stack, J., Main, D., Scheffler, B., Zheng, P., Feltus, F., Parida, L., Gilbert, D., Cornejo, O., Amores, F., Zhao, M., Saski, C., ... Mustiga, G. (2013). The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology*, 14(6), r53.
- Nieves-Orduña, H.E.; Müller, M.; Krutovsky, K.V.; Gailing, O. Geographic Patterns of Genetic Variation among Cacao (*Theobroma cacao* L.) Populations Based on Chloroplast Markers. *Diversity* 2021, 13, 249. <https://doi.org/10.3390/d13060249>
- Oliva-Cruz, S.M. (2020). Caracterización socioeconómica de la diversidad biológica de cacao criollo fino de aroma en comunidades rurales de la región Amazonas. Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas. Recuperado de <https://hdl.handle.net/20.500.14077/2211>
- Oliva-Cruz M, Goñas M, Bobadilla LG, Rubio KB, Escobedo-Ocampo P, García Rosero LM, Rojas Briceño NB and Maicelo-Quintana JL (2022) Genetic Groups of Fine-Aroma Native Cacao Based on Morphological and Sensory Descriptors in Northeast Peru. *Front. Plant Sci.* 13:896332. doi: 10.3389/fpls.2022.896332
- Olivera-Núñez, Q. (2018). Jaén, Arqueología y Turismo (Yanapay Andina Consultores (ed.)). Municipalidad Provincial de Jaén.
- Peakall R, Smouse PE (2012) GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537-2539.

- Ramasamy, R.K., Ramasamy, S., Bindroo, B.B. & Naik, V.G. (2014) STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. Springerplus, 3, 1–3.
- Ricaño-Rodríguez, Jorge, Hipólito-Romero, Enrique, Ramos-Prado, José M., & Cocolletzi-Vásquez, Eliezer. (2019). Genotipado por secuenciación de variedades nativas de *Theobroma cacao* (Malvaceae) de los Estados de Tabasco y Chiapas, México. *Botanical Sciences*, 97(3), 381-397. <https://doi.org/10.17129/botsci.2258>
- Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, 18(2), 233–234. <https://doi.org/10.1038/72708>
- Sistema Integrado de Estadísticas Agrarias* . (2022). Obtenido de (SIEA): <https://siea.midagri.gob.pe/portal/publicaciones/informacion-estadistica>
- Slatkin, M. Desequilibrio de vinculación: comprensión del pasado evolutivo y mapeo del futuro médico. *Nat Rev Genet* 9, 477–485 (2008). <https://doi.org/10.1038/nrg2361>
- Suazo-Ubieta, T., Miranda-Calero, S., Rivers-Carcache, E., Lacayo-Romero, M., & Tenorio-Juárez, D. (2020). Evaluación de metodologías de extracción de ADN de plantas recalcitrantes. *Revista Torreón Universitario*, 9(24), 45-57. <https://doi.org/10.5377/torreon.v9i24.9723>
- Solórzano, R. G. L., Fouet, O., Lemainque, A., Pavék, S., Boccara, M., Argout, X., Amores, F., Courtois, B., Risterucci, A. M. y Lanaud, C. (2012). Información sobre el origen silvestre, la migración y la historia de domesticación de la variedad de sabor fino Nacional *Theobroma cacao* L. de Ecuador. *PLoS One*, 7, e48438. <https://doi.org/10.1371/journal.pone.0048438>
- Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G., & van Etten, J. (2012). Present Spatial Diversity Patterns of *Theobroma cacao* L. in the Neotropics Reflect Genetic Differentiation in Pleistocene Refugia Followed by Human-Influenced Dispersal. *PLOS One*, 7(10), e47676. <https://doi.org/10.1371/journal.pone.0047676>
- Tscharntke, T., Ocampo-Ariza, C., Vansynghel, J., Ballesteros, B. I., Aycart, P., Rodríguez, L., Ramírez, M., Steffan-Dewenter, I., Maas, B., & Thomas, E. (2023). Socio-ecological benefits of fine-flavor cacao in its center of origin. *Conservation*

Letters,16,e12936.<https://doi.org/10.1111/conl.12936>

Turnbull CJ, Hadley P (2016) International Cocoa Germplasm Database (ICGD). [Online Database]. UK: CRA Ltd./ICE Futures Europe/ University of Reading, Available at <http://www.icgd.reading.ac.uk>.

Vázquez-Ovando, A., Escalona-Buendía, H., Ovando-Medina, I., & Salvador-Figueroa, M. (2018). *Association Between SSR Molecular Markers and Chemical and Sensory Traits of Cacao Samples Using Multiple Regression Analysis. Biochemical Genetics.* doi:10.1007/s10528-018-9891-4

Van Inghelandt, D., Melchinger, A. E., Lebreton, C., & Stich, B. (2010). Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. *Theoretical and Applied Genetics*, 120(7), 1289–1299. doi:10.1007/s00122-009-1256-2

Viena-Vásquez, L. A. (2020). Cacao nativo, una oportunidad de biocomercio para los cacaoteros de la provincia de Satipo. Obtenido de <http://hdl.handle.net/20.500.12404/17129>

Wang, B., Motilal, L. A., Meinhardt, L. W., Yin, J., & Zhang, D. (2020). *Molecular Characterization of a Cacao Germplasm Collection Maintained in Yunnan , China Using Single Nucleotide Polymorphism (SNP) Markers.*

Winter, D. J. (2012). MMOD: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, 12(6), 1158–1160. <https://doi.org/10.1111/j.1755-0998.2012.03174.x>

Yang, J. Y., Motilal, L. A., Dempewolf, H., Maharaj, K., & Cronk, Q. C. B. (2011). Chloroplast microsatellite primers for cacao (*Theobroma cacao*) and other Malvaceae. *American Journal of Botany*, 98(12), e372–e374. doi:10.3732/ajb.1100306

Zarrillo, S., Gaikwad, N., Lanaud, C., Powis, T., Viot, C., Lesur, I., Fouet, O., Argout, X., Guichoux, E., Salin, F., Solorzano, R. L., Bouchez, O., Vignes, H., Severt, P., Hurtado, J., Yopez, A., Grivetti, L., Blake, M., & Valdez, F. (2018). The use and domestication of *Theobroma cacao* during the mid-Holocene in the upper Amazon.

- Nature Ecology and Evolution*, 2(12), 1879–1888. <https://doi.org/10.1038/s41559-018-0697-x>
- Zhang, D., Arevalo-Gardini, E., Mischke, S., Zúñiga-Cernades, L., Barreto-Chávez, A., & Adriaola del Aguila, 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. <https://doi.org/10.1093/aob/mcl146>
- Zhang, D., Boccara, M., Motilal, L., Mischke, S., Johnson, E. S., Butler, D. R., 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. <https://doi.org/10.1007/s11295-009-0212-2>
- Zhang, D., Martínez, W. J., Johnson, E. S., Somarriba, E., Phillips-Mora, W., Astorga, C., Mischke, S., & Meinhardt, L. W. (2012). Genetic diversity and spatial structure in a new distinct *Theobroma cacao* L. population in Bolivia. *Genetic Resources and Crop Evolution*, 59(2), 239–252. <https://doi.org/10.1007/s10722-011-9680-y>

ANEXOS

Anexo 1. Muestras representativas de frutos de cacao a nivel de las provincias evaluadas.

Provincia de Chachapoyas

TCCH-14



Provincia de Luya

TCL-12

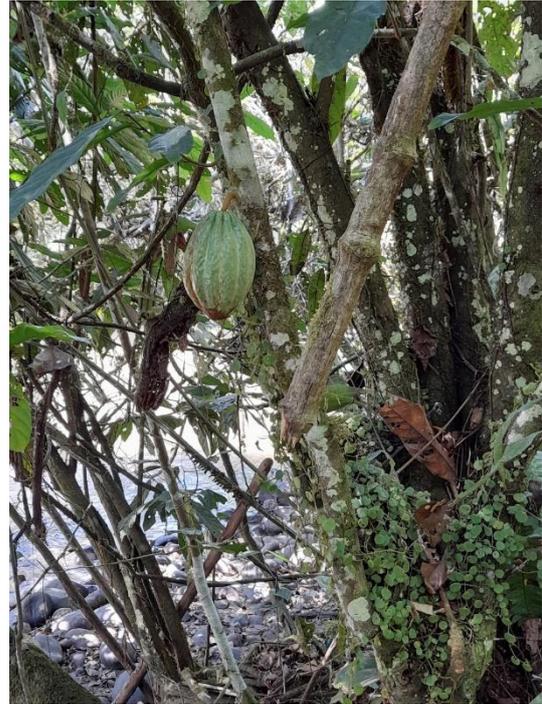


Provincia de Rodríguez de Mendoza

TCRM-11

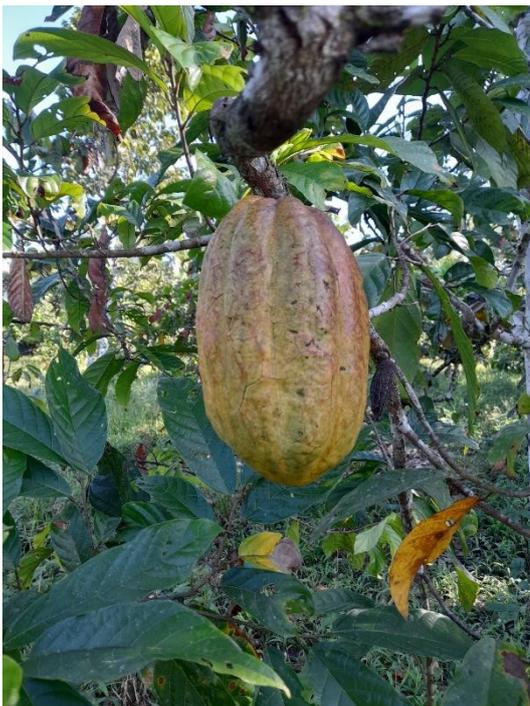


TCRM-15



Provincia de Condorcanqui

TCC-11



TCC-12



Provincia de Bagua y Utcubamba

INDES-64



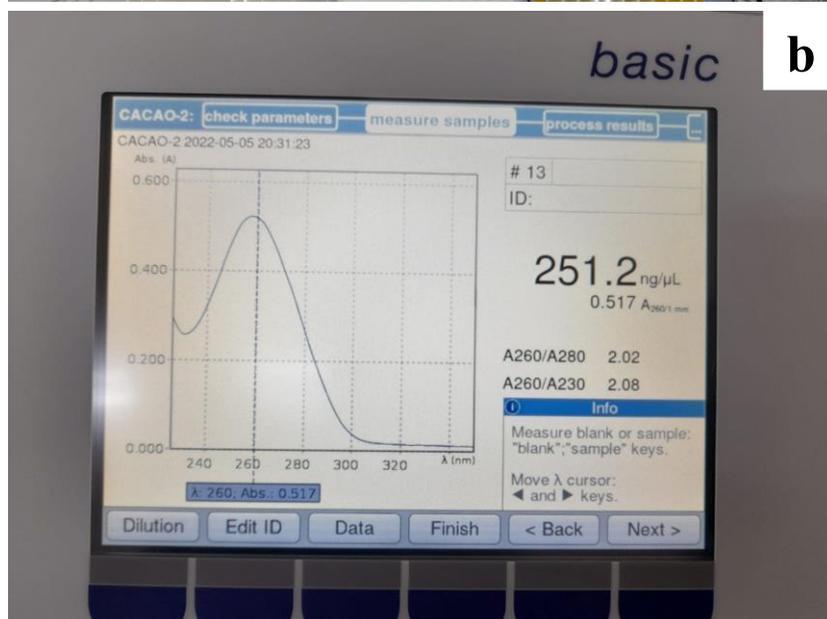
Anexo 2. Actividades de campo. **a)** Identificación y selección de accesiones de CNFA. **b)** Georreferenciación y etiquetado de la accesión de cacao. **c)** Colecta de muestras de tejido foliar.



Anexo 3. Actividades en laboratorio. a) Extracción de ADN de muestras de tejido foliar.
b) Cuantificación de ADN genómico de accesiones de cacao.



a



b

Anexo 4. Ejemplos de resultados de la búsqueda del BLASTN 2.8.1+ (grupo genético Criollo)

```
File Edit Selection Find View Goto Tools Project Preferences Help
blast_primer_genoma_cacao.out x SSR_primers_blastn.out
499 # BLASTN 2.8.1+
500 # Query: INDES_mTC12_F
501 # Database: criollo_genome_all_new_ids.fasta
502 # Fields: query id, subject id, evalue, % identity, s. start, s. end
503 # 10 hits found
504 INDES_mTC12_F chromosome_IV 0.012 100.000 24533112 24533129
505 INDES_mTC12_F chromosome_IV 3.0 100.000 945932 945919
506 INDES_mTC12_F chromosome_I 0.19 100.000 29873451 29873436
507 INDES_mTC12_F chromosome_I 3.0 100.000 35484774 35484787
508 INDES_mTC12_F chromosome_V 0.75 100.000 13417168 13417154
509 INDES_mTC12_F chromosome_II 0.75 100.000 12626319 12626305
510 INDES_mTC12_F chromosome_II 0.75 100.000 13571324 13571310
511 INDES_mTC12_F chromosome_III 3.0 100.000 7879296 7879283
512 INDES_mTC12_F chromosome_III 3.0 100.000 35275275 35275262
513 INDES_mTC12_F chromosome_III 3.0 100.000 35510671 35510658
514 # BLASTN 2.8.1+
515 # Query: INDES_mTC12_R
516 # Database: criollo_genome_all_new_ids.fasta
517 # Fields: query id, subject id, evalue, % identity, s. start, s. end
518 # 10 hits found
519 INDES_mTC12_R chromosome_IV 0.012 100.000 24533299 24533282
520 INDES_mTC12_R chromosome_IV 0.75 100.000 16601842 16601828
521 INDES_mTC12_R chromosome_IV 3.0 100.000 7991532 7991519
522 INDES_mTC12_R chromosome_VII 0.75 100.000 19550624 19550638
523 INDES_mTC12_R chromosome_V 0.75 100.000 21443228 21443214
524 INDES_mTC12_R chromosome_IX 3.0 100.000 567804 567791
525 INDES_mTC12_R chromosome_III 3.0 100.000 2797686 2797699
526 INDES_mTC12_R chromosome_II 3.0 100.000 2996224 2996237
527 INDES_mTC12_R chromosome_II 3.0 100.000 25424756 25424769
528 INDES_mTC12_R chromosome_I 3.0 100.000 34734274 34734287
529
530
Line 499, Column 16
```

Anexo 5. Ejemplos de resultados de la búsqueda del BLASTN 2.11.0+ (grupo genético Amelonado)

```
File Edit Selection Find View Goto Tools Project Preferences Help
blast_primer_genoma_cacao.out  SSR_primers_blastn.out
168 # Query: mTcINDES6_F
169 # Database: cacao11allasm_nomito.fa
170 # Fields: query id, subject id, evalue, % identity, s. start, s. end
171 # 9 hits found
172 mTcINDES6_F scaffold_6 2.09e-07 100.000 194868 194894
173 mTcINDES6_F scaffold_6 3.0 94.737 7695556 7695538
174 mTcINDES6_F scaffold_6 3.0 100.000 18448251 18448237
175 mTcINDES6_F scaffold_5 0.77 100.000 27249901 27249886
176 mTcINDES6_F scaffold_5 0.77 95.000 32028571 32028590
177 mTcINDES6_F scaffold_2 0.77 100.000 12862721 12862736
178 mTcINDES6_F scaffold_9 3.0 100.000 11999969 11999955
179 mTcINDES6_F scaffold_8 3.0 94.737 17164493 17164475
180 mTcINDES6_F scaffold_4 3.0 100.000 2322856 2322842
181 # BLASTN 2.11.0+
182 # Query: mTcINDES6_R
183 # Database: cacao11allasm_nomito.fa
184 # Fields: query id, subject id, evalue, % identity, s. start, s. end
185 # 16 hits found
186 mTcINDES6_R scaffold_6 3.24e-05 100.000 195060 195038
187 mTcINDES6_R scaffold_6 7.6 100.000 23970681 23970668
188 mTcINDES6_R scaffold_9 0.031 100.000 22766789 22766772
189 mTcINDES6_R scaffold_9 7.6 100.000 37428358 37428371
190 mTcINDES6_R scaffold_5 1.9 94.737 1706871 1706889
191 mTcINDES6_R scaffold_5 7.6 100.000 3543370 3543383
192 mTcINDES6_R scaffold_8 7.6 100.000 20145291 20145304
193 mTcINDES6_R scaffold_2 7.6 100.000 10536194 10536207
194 mTcINDES6_R scaffold_2 7.6 100.000 19621356 19621369
195 mTcINDES6_R scaffold_2 7.6 100.000 30876140 30876153
196 mTcINDES6_R scaffold_2 7.6 100.000 28854559 28854546
197 mTcINDES6_R scaffold_1 7.6 94.444 5096481 5096498
198 mTcINDES6_R scaffold_1 7.6 100.000 18217294 18217307
199 mTcINDES6_R scaffold_1 7.6 100.000 8026463 8026450
200 mTcINDES6_R scaffold_1 7.6 100.000 33006369 33006356
201 mTcINDES6_R scaffold_4 7.6 100.000 23106123 23106110
202
```

Line 168, Column 21

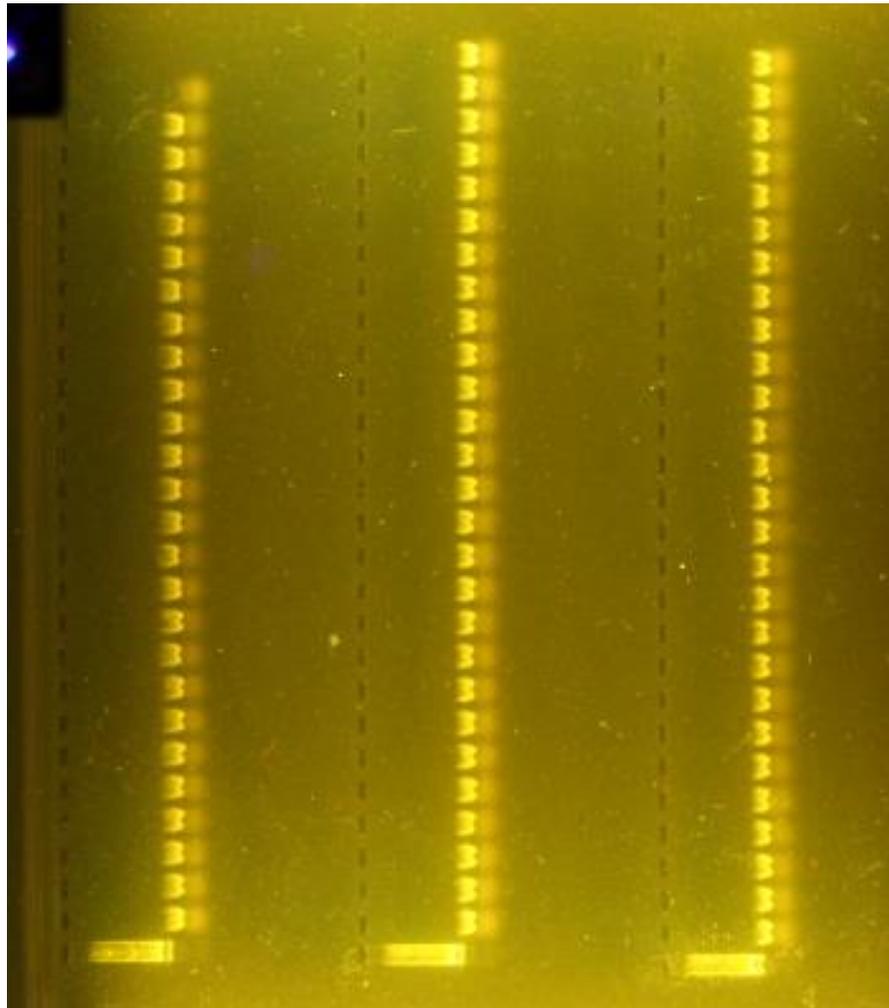
Anexo 6. *Primers* utilizados para la amplificación de regiones SSR de *Theobroma cacao*.

N°	Primer name	Forward and reverse sequences (5'–3')		Primer name Motomayor <i>et al.</i> (2008)	Chromosome		Approx. Allele size	Repeat motif	Dye
					Number	Position			
1	INDES_mTC4_F INDES_mTC4_R	TGTA AACGACGGCCAGT	GTTGATTTGCCAGCACAGA TTAATCCATGCCTGGACCTT	mTcCIR8	9	22510859	217	(AG)8	VIC
2	INDES_mTC6_F INDES_mTC6_R	TGTA AACGACGGCCAGT	TTGCCTCATATTTCTGACCC TCTGCCATCCTTCTTGATCTC	mTcCIR12	4	24533034	284	(AC)6	PET
3	INDES_mTC12_F INDES_mTC12_R	TGTA AACGACGGCCAGT	CCCTTATGGTAAGATTCTCGCA TGCATGTGAGGCATTTCTGT	mTcCIR24	9	6355004	191	(AG)14	6-FAM
4	INDES_mTC22_F INDES_mTC22_R	TGTA AACGACGGCCAGT	TCAGAAGATGCTCAGCCTTG CGGCGTGTGTCTATAACCATT	mTcCIR60	2	19251737	215	(AC)12	VIC
5	INDES_mTC23_F INDES_mTC23_R	TGTA AACGACGGCCAGT	AAGGCGAAGTGCCACATATT TTTGCTCATTGCCACTCAA	mTcCIR61	10	14256196	223	(AC)12	PET
6	INDES_mTC49_F INDES_mTC49_R	TGTA AACGACGGCCAGT	TTTGCAATTCTTCTCTCTCTTT TATGGAGCTCACATGGACCC	mTcCIR147	7	9135080	204	(AG)6	VIC
7	INDES_mTC59_F INDES_mTC59_R	TGTA AACGACGGCCAGT	TGACTACTACATCATGAGATTTCC AGAAACAACACACCAACAACCC	mTcCIR184	1	871383	163	(AG)8	NED
8	INDES_mTC77_F INDES_mTC77_R	TGTA AACGACGGCCAGT	ATCTGCAGGACTTGGACCAC TCAAGCCTTACAGCATTCCC	mTcCIR255	6	23520517	204	(AC)11	VIC
9	INDES_mTC84_F INDES_mTC84_R	TGTA AACGACGGCCAGT	TCACTATACATAACAGCAGCCATT GTTACGTTTCTTCCGCCACT	mTcCIR289	3	29547629	134	(AG)8	6-FAM
10	INDES_mTC90*_F INDES_mTC90*_R	TGTA AACGACGGCCAGT	TCGATATTTAGTCATTGGAGGG CGTTCTATTTTCATCTGATATTTCCG	ND	Cloroplasto	44393	128	(AT)8	6-FAM
11	INDES_mTC93*_F INDES_mTC93*_R	TGTA AACGACGGCCAGT	AGAGGGTTCAGGTTTCTCCC TTTGGTAACGCTTTGGCTCT	ND	1	35571464	283	(AG)27	6-FAM
12	INDES_mTC98*_F INDES_mTC98*_R	TGTA AACGACGGCCAGT	TCATCCAAACTATGACCAAGTGA TTGGGTGCCAGCTTACTCTT	ND	3	789419	298	(AG)104	PET

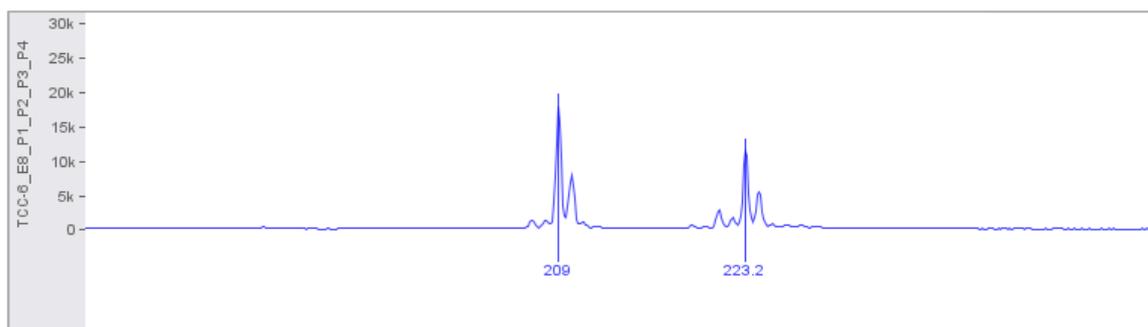
13	INDES_mTC104*_F	TGTA AACGACGGCCAGT TGTTGGTTGGTCTTGCTGAG	ND	5	33872115	263	(AG)30	PET
	INDES_mTC104*_R	AAAGGACGGCACATAGATCG						
14	INDES_mTC107*_F	TGTA AACGACGGCCAGT TGTGTCTATTTACCTCGCAG	ND	6	6686311	250	(AT)24	NED
	INDES_mTC107*_R	TGGGAATTGCATACCCAATAA						
15	INDES_mTC113*_F	TGTA AACGACGGCCAGT TTGATCTTGTGGAATTCCTG	ND	8	411586	219	(AT)33	NED
	INDES_mTC113*_R	TGTACATACCTTGCTCCGAAT						
16	INDES_mTC125*_F	TGTA AACGACGGCCAGT TATCATGACAAGCAGCGGTC	ND	8	9830566	248	(AT)24	NED
	INDES_mTC125*_R	CCCTCAACATTTGAAAGTTCC						

Nota: ND = No definido

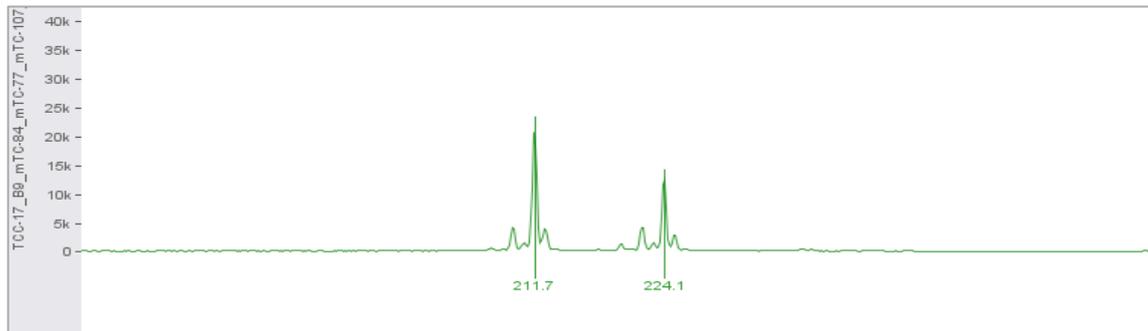
Anexo 7. Amplificación de muestras. Visualización de productos PCR en gel de agarosa al 1%.



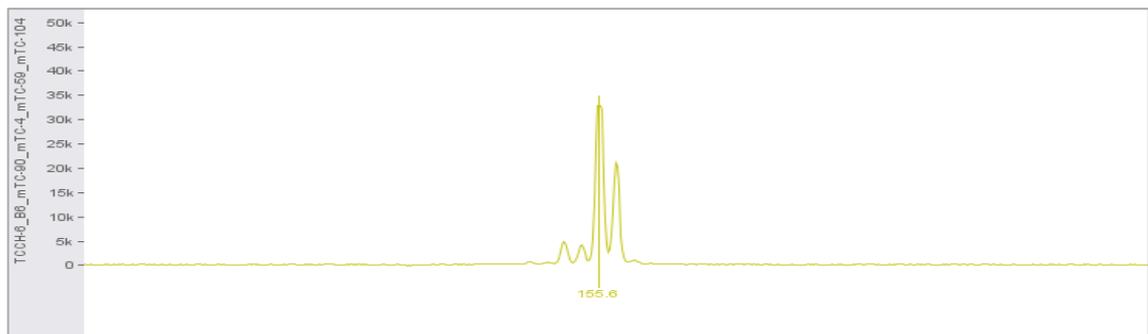
Anexo 8. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra alelos heterocigotos en la accesión de cacao TCC-6 (colorante 6-FAMTM, azul), analizados con el marcador INDES-mTC12.



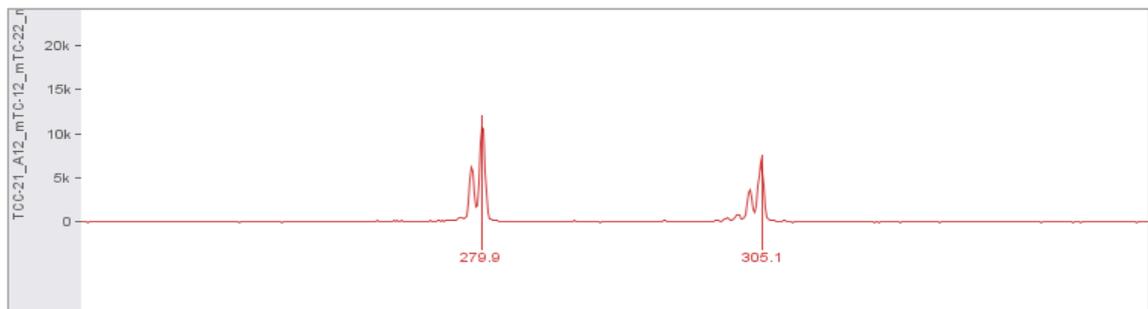
Anexo 9. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra alelos heterocigotos en la accesión de cacao TCC-17 (colorante VIC[®], verde), analizados con el marcador INDES-mTC77.



Anexo 10. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra alelos homocigotos en la accesión de cacao TCCH-6 (colorante NED[™], amarillo), analizados con el marcador INDES-mTC59.



Anexo 11. Análisis de fragmentos de la amplificación de marcadores SSR, evaluados en el Software Geneious Prime Versión 2023.0.4. Se muestra alelos heterocigotos en la accesión de cacao TCC-21 (colorante PET[®], rojo), analizados con el marcador INDES-mTC6.



Anexo 12. Tabla de tamaño de alelos de muestras de cacao, evaluado por cada marcador microsatélite SSR (SSR homólogos a los utilizados por Motomayor *et al.* (2008)).

Sample	Province	INDES_mTC12		INDES_mTC22		INDES_mTC6		INDES_mTC84		INDES_mTC77		INDES_mTC4		INDES_mTC59		INDES_mTC49		INDES_mTC23	
TCCH-1	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	224	224	222	240
TCCH-2	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	224	224	240	240
TCCH-3	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	240	240
TCCH-4	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	240
TCCH-5	Chachapoyas	225	225	212	212	297	297	149	149	212	212	236	236	156	156	222	222	240	240
TCCH-6	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	240	240
TCCH-7	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	240
TCCH-8	Chachapoyas	225	225	204	210	295	303	149	149	212	212	236	236	156	182	222	222	240	242
TCCH-9	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	222
TCCH-10	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	222
TCCH-11	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	241
TCCH-12	Chachapoyas	209	209	210	230	291	295	149	161	212	224	234	236	156	182	222	222	222	240
TCCH-13	Chachapoyas	209	225	212	212	295	303	149	161	220	224	236	236	158	182	222	222	222	240
TCCH-14	Chachapoyas	209	221	206	212	295	303	151	161	214	220	236	236	158	182	222	222	234	241
TCCH-15	Chachapoyas	209	221	206	230	291	303	151	161	220	224	234	236	156	182	222	222	224	230
TCCH-16	Chachapoyas	209	221	212	230	295	303	149	161	214	224	234	236	156	182	222	222	240	240
TCCH-17	Chachapoyas	209	209	212	212	279	303	147	151	214	214	234	236	172	182	222	222	230	230
TCCH-18	Chachapoyas	209	225	212	212	279	295	149	149	212	212	234	236	156	156	222	222	224	230
TCCH-19	Chachapoyas	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	241
TCL-1	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	222
TCL-2	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	224	241
TCL-3	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	224	224
TCL-4	Luya	225	225	212	212	295	303	149	149	212	212	236	236	156	156	222	222	222	240

TCL-5	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	241
TCL-6	Luya	225	225	210	210	295	295	149	149	212	212	236	236	156	156	222	222	240	240
TCL-7	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	240
TCL-8	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	240	240
TCL-9	Luya	223	223	212	212	295	295	149	149	212	212	236	236	156	156	222	222	222	241
TCL-10	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	244	246
TCL-11	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	224	241
TCL-12	Luya	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-1	Rodríguez de Mendoza	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-2	Rodríguez de Mendoza	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-3	Rodríguez de Mendoza	225	225	210	210	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-4	Rodríguez de Mendoza	209	209	206	230	295	295	149	153	214	220	236	236	182	182	222	222	234	241
TCRM-5	Rodríguez de Mendoza	209	225	212	230	295	295	149	153	212	214	236	236	156	182	222	222	241	241
TCRM-6	Rodríguez de Mendoza	209	225	212	230	295	295	149	149	212	214	236	236	156	182	222	222	234	241
TCRM-7	Rodríguez de Mendoza	209	225	212	230	295	295	149	149	212	214	236	236	156	182	222	222	234	241
TCRM-8	Rodríguez de Mendoza	209	225	212	230	295	295	149	149	212	220	236	236	156	182	222	222	234	241
TCRM-9	Rodríguez de Mendoza	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-10	Rodríguez de Mendoza	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-11	Rodríguez de Mendoza	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
TCRM-12	Rodríguez de Mendoza	223	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241
TCRM-13	Rodríguez de Mendoza	225	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241

TCRM-14	Rodríguez de Mendoza	225	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241
TCRM-15	Rodríguez de Mendoza	225	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241
TCRM-16	Rodríguez de Mendoza	225	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241
TCC-1	Condorcanqui	209	221	212	228	293	297	161	161	214	220	234	236	158	182	224	224	242	252
TCC-2	Condorcanqui	209	209	212	212	297	297	149	151	212	224	236	236	158	158	224	224	242	242
TCC-3	Condorcanqui	225	227	206	206	299	305	147	149	212	224	234	236	156	182	224	224	236	241
TCC-4	Condorcanqui	209	209	212	230	279	279	151	161	212	214	236	236	158	182	224	224	244	244
TCC-5	Condorcanqui	209	209	212	228	293	297	159	161	224	224	236	236	156	182	224	224	242	242
TCC-6	Condorcanqui	209	223	206	212	293	305	149	161	212	212	236	236	182	182	224	224	234	241
TCC-7	Condorcanqui	209	223	212	212	297	297	149	149	212	214	236	236	156	158	224	224	234	252
TCC-8	Condorcanqui	209	223	212	226	279	299	149	149	212	224	234	236	156	156	224	224	241	252
TCC-9	Condorcanqui	209	225	206	212	297	297	149	161	224	224	236	236	156	182	224	224	241	246
TCC-10	Condorcanqui	209	217	212	212	297	297	149	161	212	214	236	236	156	156	222	222	240	240
TCC-11	Condorcanqui	209	217	212	230	297	305	149	149	212	214	236	236	156	158	222	222	240	252
TCC-12	Condorcanqui	225	225	212	230	305	305	149	161	212	224	234	234	156	182	222	222	230	240
TCC-13	Condorcanqui	209	225	212	212	293	297	149	161	214	224	234	234	156	182	222	222	230	234
TCC-14	Condorcanqui	221	225	212	228	293	297	149	161	212	224	234	236	156	182	222	222	240	240
TCC-15	Condorcanqui	209	209	206	212	293	297	149	161	212	214	236	236	182	182	222	222	241	242
TCC-16	Condorcanqui	217	225	212	230	297	297	149	149	214	224	234	234	156	156	222	222	241	252
TCC-17	Condorcanqui	225	225	212	230	297	305	149	149	212	224	236	236	156	182	222	222	234	241
TCC-18	Condorcanqui	225	225	212	230	297	299	149	149	212	212	236	236	156	182	222	222	240	240
TCC-19	Condorcanqui	209	221	212	230	279	297	149	151	212	214	236	236	156	172	222	222	224	230
TCC-20	Condorcanqui	209	209	206	212	279	305	149	161	212	214	236	236	156	158	222	222	240	246
TCC-21	Condorcanqui	209	209	212	212	279	305	147	151	214	214	234	236	182	182	222	222	230	230
INDES-4	Bagua y Utcubamba	225	225	212	212	297	297	149	149	212	224	236	236	156	156	222	222	230	230

INDES-32	Bagua y Utcubamba	225	225	212	212	297	297	149	151	212	212	236	236	156	156	224	224	241	241
INDES-35	Bagua y Utcubamba	225	225	212	212	297	305	149	149	212	212	236	236	156	156	224	224	241	241
INDES-80	Bagua y Utcubamba	209	225	212	212	297	297	149	149	212	224	234	236	156	156	224	224	224	241
INDES-14	Bagua y Utcubamba	221	225	212	212	297	297	149	149	212	212	236	236	156	156	222	222	241	241
INDES-27	Bagua y Utcubamba	209	225	212	230	293	297	149	151	212	214	234	236	156	182	222	222	241	246
INDES-50	Bagua y Utcubamba	209	225	212	230	293	305	147	151	212	212	234	236	156	182	224	224	241	246
INDES-63	Bagua y Utcubamba	209	209	212	212	279	279	149	149	214	224	236	236	156	182	224	224	224	241
INDES-6	Bagua y Utcubamba	209	225	210	210	279	295	149	161	212	214	236	236	156	158	222	222	241	246
INDES-31	Bagua y Utcubamba	209	225	212	230	297	305	149	151	212	224	236	236	156	156	222	222	241	246
INDES-53	Bagua y Utcubamba	209	223	206	212	293	305	159	161	212	214	234	236	158	166	224	224	246	252
INDES-70	Bagua y Utcubamba	209	221	210	212	279	297	159	161	214	214	236	236	158	182	224	224	246	246
INDES-64	Bagua y Utcubamba	209	225	210	230	297	303	149	149	212	214	236	236	156	156	224	224	228	241
INDES-98	Bagua y Utcubamba	225	225	212	212	297	305	149	151	212	214	236	236	156	156	224	224	241	241
INDES-99	Bagua y Utcubamba	209	225	210	230	279	297	147	161	212	214	234	236	156	182	224	224	241	246
INDES-100	Bagua y Utcubamba	209	225	212	212	297	305	149	151	212	214	234	236	156	182	224	224	224	230
INDES-12	Bagua y Utcubamba	223	223	212	212	297	313	149	149	212	224	236	236	156	156	222	222	241	241
INDES-38	Bagua y Utcubamba	225	225	212	212	297	297	149	149	212	212	236	236	156	156	224	224	241	241
INDES-78	Bagua y Utcubamba	209	225	210	212	297	297	149	161	212	214	236	236	158	158	224	224	241	246

INDES-81	Bagua y Utcubamba	225	225	212	230	297	305	149	151	220	220	236	236	156	156	224	224	224	224
TCU-1	Bagua y Utcubamba	225	225	212	212	297	297	149	149	212	212	236	236	156	156	222	222	241	241
TCU-2	Bagua y Utcubamba	225	225	212	212	297	297	149	149	212	212	236	236	156	156	222	222	241	241
TCU-3	Bagua y Utcubamba	225	225	212	212	295	295	149	149	212	212	236	236	156	156	222	222	241	241
CAP-6	Bagua y Utcubamba	225	225	212	230	297	297	147	149	212	224	236	236	156	156	222	222	224	238
CAP-10	Bagua y Utcubamba	225	225	212	230	297	297	149	149	212	212	236	236	156	156	222	222	241	241
CAP-126	Bagua y Utcubamba	209	225	212	230	297	297	149	149	212	212	236	236	156	156	222	222	241	241
CCN-51	Bagua y Utcubamba	209	209	212	230	279	293	151	161	212	214	234	236	158	182	222	222	244	246
CCN-51-C	Condorcanqui	209	209	212	230	279	293	151	161	212	214	234	236	158	182	222	222	244	246

Anexo 13. Tabla de tamaño de alelos de muestras de cacao, evaluado por cada marcador microsatélite SSR (SSR diseñados en este estudio).

Sample	Province	INDES_mTC125		INDES_mTC107		INDES_mTC98		INDES_mTC90		INDES_mTC104		INDES_mTC93		INDES_mTC113	
TCCH-1	Chachapoyas	191	191	254	254	149	163	145	145	156	182	161	161	166	168
TCCH-2	Chachapoyas	191	191	238	254	149	167	145	145	182	196	161	161	166	166
TCCH-3	Chachapoyas	191	191	254	254	149	167	145	145	182	182	161	161	166	168
TCCH-4	Chachapoyas	191	191	238	254	149	163	145	145	196	196	161	161	168	170
TCCH-5	Chachapoyas	191	191	238	254	149	167	145	145	196	196	161	161	166	168
TCCH-6	Chachapoyas	191	191	254	254	149	163	145	145	196	196	161	161	156	166
TCCH-7	Chachapoyas	191	191	238	254	149	163	145	145	196	196	161	161	164	164
TCCH-8	Chachapoyas	191	191	236	236	123	149	145	145	180	180	161	161	158	164
TCCH-9	Chachapoyas	191	191	238	254	167	167	145	145	196	196	161	161	166	168
TCCH-10	Chachapoyas	191	191	238	254	149	167	145	145	196	196	161	161	166	166
TCCH-11	Chachapoyas	191	191	238	254	149	167	145	145	182	196	161	161	166	168
TCCH-12	Chachapoyas	191	191	238	258	145	157	145	145	180	180	169	169	164	166
TCCH-13	Chachapoyas	191	191	238	272	157	163	145	145	182	192	169	169	152	164
TCCH-14	Chachapoyas	193	193	272	272	121	157	145	145	156	156	169	183	164	164
TCCH-15	Chachapoyas	191	191	238	272	121	151	145	145	182	182	169	169	164	172
TCCH-16	Chachapoyas	191	191	238	272	143	157	145	145	180	180	169	183	164	164
TCCH-17	Chachapoyas	205	205	234	272	123	173	145	145	160	192	187	187	164	182
TCCH-18	Chachapoyas	191	207	238	238	149	163	145	145	178	178	161	161	166	166
TCCH-19	Chachapoyas	191	191	238	254	149	167	145	145	196	196	161	161	166	166
TCL-1	Luya	191	191	238	238	163	163	145	145	182	196	161	161	166	170
TCL-2	Luya	191	191	238	254	163	163	145	145	182	196	161	161	166	166
TCL-3	Luya	191	191	238	238	167	167	145	145	196	196	161	161	168	170
TCL-4	Luya	191	191	238	254	151	173	145	145	196	196	161	161	164	170
TCL-5	Luya	191	191	238	238	167	167	145	145	182	196	161	161	166	166
TCL-6	Luya	191	191	238	238	149	167	145	145	156	156	161	161	166	166

TCL-7	Luya	191	191	238	238	149	167	145	145	182	196	161	161	166	166
TCL-8	Luya	191	191	238	238	149	167	145	145	196	196	161	161	166	168
TCL-9	Luya	191	191	238	238	149	163	145	145	196	196	161	161	168	168
TCL-10	Luya	191	191	234	238	149	163	145	145	182	182	161	161	166	166
TCL-11	Luya	191	191	238	238	149	167	145	145	196	196	161	161	166	170
TCL-12	Luya	191	191	238	254	163	163	145	145	182	182	161	161	168	170
TCRM-1	Rodríguez de Mendoza	191	191	238	238	151	167	145	145	182	182	161	161	164	170
TCRM-2	Rodríguez de Mendoza	191	191	238	238	167	167	145	145	182	182	161	161	164	170
TCRM-3	Rodríguez de Mendoza	191	191	238	254	151	167	145	145	182	182	161	161	164	170
TCRM-4	Rodríguez de Mendoza	187	187	258	272	133	145	145	145	180	200	169	169	184	192
TCRM-5	Rodríguez de Mendoza	187	191	258	258	133	167	145	145	196	196	161	169	164	192
TCRM-6	Rodríguez de Mendoza	187	191	254	272	145	167	145	145	200	200	161	169	166	192
TCRM-7	Rodríguez de Mendoza	187	191	258	258	133	167	145	145	180	180	161	169	166	192
TCRM-8	Rodríguez de Mendoza	187	191	258	258	143	167	145	145	500	500	161	169	166	184
TCRM-9	Rodríguez de Mendoza	191	191	238	238	149	149	145	145	182	196	161	161	170	170
TCRM-10	Rodríguez de Mendoza	191	191	238	254	149	149	145	145	182	196	161	161	170	170
TCRM-11	Rodríguez de Mendoza	191	191	238	238	149	149	145	145	182	196	161	161	166	170
TCRM-12	Rodríguez de Mendoza	191	191	256	268	149	167	145	145	196	196	161	161	170	170
TCRM-13	Rodríguez de Mendoza	191	191	254	254	149	167	145	145	182	196	161	161	170	170
TCRM-14	Rodríguez de Mendoza	191	191	238	238	149	167	145	145	196	196	161	161	170	170
TCRM-15	Rodríguez de Mendoza	191	191	238	238	149	167	145	145	182	196	161	161	166	170
TCRM-16	Rodríguez de Mendoza	191	191	238	254	149	149	145	145	196	196	161	161	170	170
TCC-1	Condorcanqui	187	187	234	278	123	123	145	145	178	178	169	169	152	186
TCC-2	Condorcanqui	193	193	262	262	123	151	145	145	174	174	161	161	168	168
TCC-3	Condorcanqui	191	191	258	258	145	173	145	145	500	500	169	185	164	166
TCC-4	Condorcanqui	193	207	234	262	173	173	145	145	192	192	189	189	164	186
TCC-5	Condorcanqui	191	191	232	232	167	167	145	145	180	192	169	169	166	166
TCC-6	Condorcanqui	191	191	258	258	123	167	145	145	156	156	169	199	166	182

TCC-7	Condorcanqui	191	191	258	272	123	167	145	145	158	174	169	189	164	166
TCC-8	Condorcanqui	191	191	258	268	167	167	145	145	500	500	161	161	170	170
TCC-9	Condorcanqui	191	193	230	238	123	167	145	145	178	178	161	169	164	176
TCC-10	Condorcanqui	215	227	258	258	121	151	145	145	158	180	161	169	151	151
TCC-11	Condorcanqui	207	207	258	258	123	133	145	145	196	196	185	189	161	161
TCC-12	Condorcanqui	191	191	238	258	149	167	145	145	158	196	161	161	166	166
TCC-13	Condorcanqui	211	215	230	258	149	167	145	145	156	168	169	169	164	164
TCC-14	Condorcanqui	191	191	258	258	149	167	145	145	178	178	169	183	166	166
TCC-15	Condorcanqui	207	215	240	258	133	161	145	145	156	180	161	169	164	164
TCC-16	Condorcanqui	191	191	238	254	143	161	145	145	180	196	169	189	164	164
TCC-17	Condorcanqui	191	193	238	272	161	161	145	145	178	178	161	189	166	166
TCC-18	Condorcanqui	218	218	230	238	149	167	145	145	156	156	161	161	166	166
TCC-19	Condorcanqui	191	191	234	238	143	149	145	145	156	156	161	195	170	170
TCC-20	Condorcanqui	193	193	234	262	123	149	145	145	178	178	189	189	164	170
TCC-21	Condorcanqui	207	207	234	272	123	123	145	145	160	192	189	189	166	166
INDES-4	Bagua y Utcubamba	191	191	234	238	149	149	145	145	182	182	161	161	166	170
INDES-32	Bagua y Utcubamba	191	191	238	254	123	151	145	145	160	182	161	161	166	166
INDES-35	Bagua y Utcubamba	191	207	238	256	167	167	145	145	500	500	161	161	166	166
INDES-80	Bagua y Utcubamba	191	191	238	238	123	167	145	145	196	196	161	161	166	170
INDES-14	Bagua y Utcubamba	191	191	238	238	149	149	145	145	156	156	161	183	166	168
INDES-27	Bagua y Utcubamba	191	191	234	254	167	173	145	145	192	192	161	189	164	170
INDES-50	Bagua y Utcubamba	191	191	238	262	123	149	145	145	174	180	161	189	164	164
INDES-63	Bagua y Utcubamba	191	191	234	238	123	167	145	145	160	160	161	189	170	170
INDES-6	Bagua y Utcubamba	191	191	234	238	163	173	145	145	182	182	161	189	170	186
INDES-31	Bagua y Utcubamba	191	191	234	238	163	173	145	145	178	178	161	189	166	166
INDES-53	Bagua y Utcubamba	193	193	234	258	123	123	145	145	168	168	169	203	146	186
INDES-70	Bagua y Utcubamba	191	191	234	238	123	123	145	145	180	192	169	189	164	164
INDES-64	Bagua y Utcubamba	191	229	234	234	143	151	145	145	160	196	161	189	166	182

INDES-98	Bagua y Utcubamba	191	191	254	268	123	149	145	145	500	500	161	161	166	166
INDES-99	Bagua y Utcubamba	191	191	254	262	123	167	145	145	500	500	195	195	164	164
INDES-100	Bagua y Utcubamba	191	191	238	272	151	173	145	145	160	182	161	189	166	170
INDES-12	Bagua y Utcubamba	191	191	238	238	121	167	145	145	182	182	161	161	166	166
INDES-38	Bagua y Utcubamba	191	191	238	254	149	167	145	145	182	182	161	161	170	170
INDES-78	Bagua y Utcubamba	191	191	238	238	123	149	145	145	196	196	161	169	164	168
INDES-81	Bagua y Utcubamba	191	191	238	238	123	167	145	145	178	196	161	189	166	166
TCU-1	Bagua y Utcubamba	191	191	238	254	163	163	145	145	156	196	161	183	166	170
TCU-2	Bagua y Utcubamba	191	191	238	254	149	167	145	145	182	196	161	161	168	170
TCU-3	Bagua y Utcubamba	191	191	238	238	167	167	145	145	500	500	161	161	166	170
CAP-6	Bagua y Utcubamba	191	191	238	272	149	167	145	145	178	196	161	161	166	170
CAP-10	Bagua y Utcubamba	191	191	238	238	121	163	145	145	154	154	161	161	168	170
CAP-126	Bagua y Utcubamba	191	191	256	256	149	167	145	145	182	196	161	161	166	166
CCN-51	Bagua y Utcubamba	193	205	234	262	123	173	145	145	192	192	169	189	164	186
CCN-51-C	Condorcanqui	193	207	234	262	123	173	145	145	192	192	169	189	163	163

Nota: El número 500 en el marcador SSR INDES-mTC104 representa un alelo nulo

Anexo 14. Tabla de datos de coeficiente de mezcla genética (*admixture*) de las accesiones de cacao basado en cuatro hipotéticas poblaciones ($K = 4$).

Sample	Province	V _{K1}	V _{K2}	V _{K3}	V _{K4}
TCCH-1	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-2	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-3	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-4	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-5	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-6	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-7	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-8	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-9	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-10	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-11	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-12	Chachapoyas	0.00010	0.00010	0.00010	0.99970
TCCH-13	Chachapoyas	0.00010	0.18509	0.00010	0.81471
TCCH-14	Chachapoyas	0.00010	0.00010	0.00010	0.99970
TCCH-15	Chachapoyas	0.00010	0.00010	0.00010	0.99970
TCCH-16	Chachapoyas	0.00010	0.00010	0.00010	0.99970
TCCH-17	Chachapoyas	0.99970	0.00010	0.00010	0.00010
TCCH-18	Chachapoyas	0.00010	0.99970	0.00010	0.00010
TCCH-19	Chachapoyas	0.00010	0.99970	0.00010	0.00010
	Promedio	0.05271	0.69377	0.00010	0.25342
TCL-1	Luya	0.00010	0.99970	0.00010	0.00010
TCL-2	Luya	0.00010	0.99970	0.00010	0.00010
TCL-3	Luya	0.00010	0.99970	0.00010	0.00010
TCL-4	Luya	0.00010	0.99970	0.00010	0.00010
TCL-5	Luya	0.00010	0.99970	0.00010	0.00010
TCL-6	Luya	0.00010	0.99970	0.00010	0.00010
TCL-7	Luya	0.00010	0.99970	0.00010	0.00010
TCL-8	Luya	0.00010	0.99970	0.00010	0.00010
TCL-9	Luya	0.00010	0.99970	0.00010	0.00010
TCL-10	Luya	0.00010	0.99970	0.00010	0.00010
TCL-11	Luya	0.00010	0.99970	0.00010	0.00010
TCL-12	Luya	0.00010	0.99970	0.00010	0.00010
	Promedio	0.00010	0.99970	0.00010	0.00010
TCRM-1	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-2	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-3	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-4	Rodríguez de Mendoza	0.00010	0.00010	0.00010	0.99970
TCRM-5	Rodríguez de Mendoza	0.00010	0.11845	0.00010	0.88136
TCRM-6	Rodríguez de Mendoza	0.00010	0.01157	0.00010	0.98823
TCRM-7	Rodríguez de Mendoza	0.00010	0.02214	0.00010	0.97766
TCRM-8	Rodríguez de Mendoza	0.00010	0.04564	0.00010	0.95416
TCRM-9	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010

TCRM-10	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-11	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-12	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-13	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-14	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-15	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
TCRM-16	Rodríguez de Mendoza	0.00010	0.99970	0.00010	0.00010
	Promedio	0.00010	0.69966	0.00010	0.30014
TCC-1	Condorcanqui	0.99970	0.00010	0.00010	0.00010
TCC-2	Condorcanqui	0.28108	0.71872	0.00010	0.00010
TCC-3	Condorcanqui	0.00010	0.17676	0.40880	0.41434
TCC-4	Condorcanqui	0.99970	0.00010	0.00010	0.00010
TCC-5	Condorcanqui	0.13972	0.03094	0.00010	0.82924
TCC-6	Condorcanqui	0.00010	0.00010	0.00010	0.99970
TCC-7	Condorcanqui	0.00010	0.43459	0.51662	0.04869
TCC-8	Condorcanqui	0.00010	0.58539	0.41441	0.00010
TCC-9	Condorcanqui	0.00010	0.00010	0.62806	0.37174
TCC-10	Condorcanqui	0.00010	0.22576	0.00010	0.77404
TCC-11	Condorcanqui	0.58108	0.41872	0.00010	0.00010
TCC-12	Condorcanqui	0.00010	0.93968	0.00010	0.06012
TCC-13	Condorcanqui	0.00010	0.03091	0.00010	0.96889
TCC-14	Condorcanqui	0.00010	0.23340	0.00010	0.76640
TCC-15	Condorcanqui	0.00010	0.00010	0.00010	0.99970
TCC-16	Condorcanqui	0.00010	0.63801	0.36179	0.00010
TCC-17	Condorcanqui	0.00010	0.41115	0.58865	0.00010
TCC-18	Condorcanqui	0.00010	0.99970	0.00010	0.00010
TCC-19	Condorcanqui	0.00010	0.20170	0.79811	0.00010
TCC-20	Condorcanqui	0.99970	0.00010	0.00010	0.00010
TCC-21	Condorcanqui	0.99970	0.00010	0.00010	0.00010
CCN-51-C	Condorcanqui	0.00010	0.99970	0.00010	0.00010
	Promedio	0.22737	0.32026	0.16900	0.28337
INDES-4	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
INDES-32	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
INDES-35	Bagua y Utcubamba	0.00010	0.87534	0.12447	0.00010
INDES-80	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
INDES-14	Bagua y Utcubamba	0.00010	0.00010	0.99970	0.00010
INDES-27	Bagua y Utcubamba	0.00010	0.00010	0.99970	0.00010
INDES-50	Bagua y Utcubamba	0.00010	0.28393	0.71587	0.00010
INDES-63	Bagua y Utcubamba	0.00010	0.12126	0.87854	0.00010
INDES-6	Bagua y Utcubamba	0.00010	0.15562	0.84419	0.00010
INDES-31	Bagua y Utcubamba	0.99970	0.00010	0.00010	0.00010
INDES-53	Bagua y Utcubamba	0.99970	0.00010	0.00010	0.00010
INDES-70	Bagua y Utcubamba	0.00010	0.00010	0.99970	0.00010
INDES-64	Bagua y Utcubamba	0.00010	0.94161	0.05819	0.00010
INDES-98	Bagua y Utcubamba	0.00010	0.00010	0.99970	0.00010
INDES-99	Bagua y Utcubamba	0.00010	0.00010	0.99970	0.00010

INDES-100	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
INDES-12	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
INDES-38	Bagua y Utcubamba	0.00010	0.34352	0.65628	0.00010
INDES-78	Bagua y Utcubamba	0.00010	0.61091	0.38889	0.00010
INDES-81	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
TCU-1	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
TCU-2	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
TCU-3	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
CAP-6	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
CAP-10	Bagua y Utcubamba	0.00010	0.99970	0.00010	0.00010
CAP-126	Bagua y Utcubamba	0.99970	0.00010	0.00010	0.00010
CCN-51	Bagua y Utcubamba	0.99970	0.00010	0.00010	0.00010
Promedio		0.14819	0.53073	0.32098	0.00010

Nota: K_1 , K_2 , K_3 y K_4 representan a las poblaciones ancestrales