

Artículo Original

Caracterización de la diversidad genética de variedades nativas de cacao mediante marcadores moleculares SSR

Characterization of the genetic diversity of native cacao varieties using SSR molecular markers

Dayana Mishelle Sigüenza Sánchez^{1*}  

¹*Universidad Agraria del Ecuador, Milagro-Ecuador*

***Autor de correspondencia:** Dayana Mishelle Sigüenza Sánchez 

Recibido: 08-04-2026

Revisado: 15-04-2026

Aceptado: 22-04-2026

Publicado: 29-04-2026

Cómo citar este artículo: Sigüenza Sánchez, D. M. (2026). Caracterización de la diversidad genética de variedades nativas de cacao mediante marcadores moleculares SSR. Viterra, 4(1), 1-34.
<https://viterra.diginova.pro/index.php/viterra/article/view/34>

RESUMEN

La conservación y caracterización de la diversidad genética del cacao (*Theobroma cacao* L.) constituyen aspectos fundamentales para el aprovechamiento sostenible de los recursos fitogenéticos, la identificación de materiales con características agronómicas diferenciadas y el fortalecimiento de los programas de mejoramiento genético. Los marcadores moleculares de secuencias simples repetidas (SSR) representan herramientas ampliamente utilizadas para determinar niveles de variabilidad, relaciones genéticas y estructura poblacional debido a su elevado polimorfismo, codominancia y reproducibilidad. El objetivo de la investigación fue caracterizar la diversidad genética de variedades nativas de cacao mediante marcadores moleculares SSR y determinar las relaciones de similitud y diferenciación existentes entre los materiales evaluados. Se analizaron 72 accesiones nativas de cacao procedentes de seis zonas productoras, utilizando un panel de 15 loci microsatélites. La extracción del ADN genómico se realizó a partir de hojas jóvenes mediante un protocolo CTAB modificado, seguida por amplificación mediante reacción en cadena de la polimerasa y separación de los productos amplificados. Se determinaron el número de alelos, contenido de información polimórfica, heterocigosidad observada y esperada, índice de diversidad genética y coeficiente de endogamia. La estructura genética fue evaluada mediante análisis de coordenadas principales, agrupamiento basado en distancias genéticas y análisis bayesiano. Los 15 loci SSR generaron 142 alelos, con un promedio de 9,47 alelos por locus. La heterocigosidad observada promedio fue de 0,63, mientras que la heterocigosidad esperada alcanzó 0,71. El contenido de información polimórfica presentó un promedio de 0,68, demostrando la elevada capacidad discriminante de los marcadores seleccionados. El análisis molecular de varianza indicó que el 76 % de la variabilidad genética se concentró dentro de las poblaciones y el 24 % entre ellas. El análisis de estructura identificó cuatro grupos genéticos principales con diferentes niveles de mezcla ancestral. Se concluye que las variedades nativas evaluadas mantienen una elevada diversidad genética y constituyen recursos estratégicos para la conservación, selección de genotipos y desarrollo de programas de mejoramiento del cacao.

Palabras clave: diversidad genética; germoplasma; marcadores moleculares; microsatélites; SSR; *Theobroma cacao* L.

ABSTRACT

The conservation and characterization of cacao (*Theobroma cacao* L.) genetic diversity are fundamental for the sustainable use of plant genetic resources, the identification of materials with differentiated agronomic characteristics, and the strengthening of breeding programs. Simple sequence repeat (SSR) molecular markers are widely used to determine variability levels, genetic relationships, and population structure due to their high polymorphism, codominance, and reproducibility. The objective of this research was to characterize the genetic diversity of native cacao varieties using SSR molecular markers and to determine the similarity and differentiation relationships among the evaluated materials. Seventy-two native cacao accessions from six production areas were analyzed using a panel of 15 microsatellite loci. Genomic DNA was extracted from young leaves using a modified CTAB protocol, followed by

polymerase chain reaction amplification and separation of amplified products. The number of alleles, polymorphic information content, observed and expected heterozygosity, genetic diversity index, and inbreeding coefficient were determined. Genetic structure was assessed through principal coordinate analysis, genetic distance-based clustering, and Bayesian analysis. The 15 SSR loci generated 142 alleles, with an average of 9.47 alleles per locus. Mean observed heterozygosity was 0.63, whereas expected heterozygosity reached 0.71. Polymorphic information content averaged 0.68, demonstrating the high discriminatory capacity of the selected markers. Analysis of molecular variance indicated that 76% of genetic variability was concentrated within populations and 24% among them. Structure analysis identified four major genetic groups with different levels of ancestral admixture. It is concluded that the evaluated native varieties maintain high genetic diversity and constitute strategic resources for conservation, genotype selection, and the development of cacao breeding programs.

Keywords: cacao; genetic diversity; germplasm; microsatellites; molecular markers; SSR.

1. INTRODUCCIÓN

El cacao (*Theobroma cacao* L.) constituye uno de los cultivos tropicales de mayor importancia económica, social, cultural y ambiental. Su producción representa una fuente fundamental de ingresos para millones de agricultores y constituye la materia prima principal para la elaboración de chocolate y diferentes productos alimenticios, cosméticos y farmacéuticos. Además de su importancia comercial, el cacao presenta una elevada diversidad genética que constituye un recurso estratégico para la conservación y el mejoramiento de la especie. Las variedades nativas de cacao representan reservorios de diversidad genética desarrollados mediante procesos evolutivos, adaptación ambiental, selección natural y manejo tradicional. Estos materiales pueden presentar características relacionadas con calidad organoléptica, resistencia frente a enfermedades, tolerancia a condiciones ambientales adversas y adaptación a diferentes sistemas productivos. La conservación de esta diversidad constituye un desafío fundamental debido a la expansión de variedades comerciales genéticamente uniformes, pérdida de hábitats naturales, sustitución de materiales tradicionales y procesos de erosión genética. La identificación y caracterización de los recursos genéticos permite desarrollar estrategias destinadas a conservar materiales únicos y aprovechar sus características dentro de programas de mejoramiento (Cornejo et al., 2018).

A partir de este planteamiento, el origen y domesticación del cacao han sido ampliamente estudiados mediante evidencias históricas, arqueológicas, morfológicas y moleculares. Cornejo et al. (2018) reconstruyeron la historia de domesticación del cacao mediante análisis genómicos de poblaciones cultivadas y silvestres, demostrando la importancia de los estudios moleculares para comprender los procesos de domesticación y diversificación. Tradicionalmente, los materiales de cacao fueron clasificados dentro de tres grandes grupos: Criollo, Forastero y Trinitario. Sin embargo, los avances dentro de la genética molecular han demostrado que la diversidad del cacao presenta una estructura considerablemente más compleja. La utilización de marcadores moleculares ha permitido identificar diferentes grupos genéticos y analizar con mayor precisión las relaciones existentes entre materiales cultivados y poblaciones silvestres.

En continuidad con lo anterior, Zarrillo et al. (2018) aportaron evidencia arqueológica y molecular sobre el uso temprano y la domesticación del cacao en la alta Amazonía. Los resultados demostraron la existencia de una elevada diversidad y estructura poblacional, proporcionando importantes fundamentos para la clasificación y conservación del germoplasma. La región amazónica constituye uno de los principales centros de diversidad genética del cacao. Las poblaciones presentes dentro de esta región mantienen una elevada variabilidad desarrollada durante extensos procesos evolutivos. Fouet et al. (2022) evaluaron la diversidad genética y estructura de poblaciones naturales de cacao dentro de la Amazonía brasileña utilizando marcadores microsatélites. Los resultados demostraron la utilidad de los SSR para identificar variabilidad y diferenciación genética entre poblaciones.

De manera complementaria, los marcadores moleculares constituyen herramientas fundamentales para estudiar la diversidad genética debido a que permiten analizar directamente las variaciones existentes dentro del ADN. A diferencia de los caracteres morfológicos, los marcadores moleculares presentan menor influencia de las condiciones ambientales y permiten identificar diferencias genéticas incluso entre individuos fenotípicamente similares. Diferentes

tipos de marcadores han sido utilizados dentro de los estudios genéticos del cacao, incluyendo RAPD, RFLP, AFLP, SSR y, más recientemente, polimorfismos de nucleótido único. Nieves-Orduña et al. (2023) desarrollaron estudios relacionados con mapas genéticos del cacao utilizando diferentes tipos de marcadores moleculares. Estos trabajos contribuyeron significativamente al conocimiento del genoma y proporcionaron herramientas para analizar características relacionadas con productividad.

Desde una perspectiva aplicada, entre las diferentes herramientas disponibles, los marcadores microsatélites o secuencias simples repetidas presentan importantes ventajas para los estudios de diversidad. Los SSR corresponden a regiones del ADN caracterizadas por repeticiones cortas de secuencias de nucleótidos. Las variaciones dentro del número de repeticiones generan polimorfismos que pueden ser detectados mediante técnicas moleculares. Estos marcadores presentan herencia codominante, elevada reproducibilidad, abundancia dentro del genoma y altos niveles de polimorfismo. Nieves-Orduña et al. (2023) demostraron la utilidad de los microsatélites para evaluar la diversidad molecular del cacao. Los resultados evidenciaron la capacidad de estos marcadores para identificar diferencias genéticas y analizar relaciones entre materiales.

En este contexto, el desarrollo de loci microsatélites específicos para *Theobroma cacao* representó un avance fundamental dentro de los estudios de diversidad. Nieves-Orduña et al. (2023) describieron el aislamiento y caracterización de loci microsatélites aplicables a procesos de genotipificación. La disponibilidad de estos marcadores permitió desarrollar investigaciones comparables dentro de diferentes poblaciones y colecciones. Los estudios de diversidad genética mediante SSR pueden proporcionar diferentes indicadores. El número de alelos permite determinar la variabilidad existente dentro de cada locus. La heterocigosidad observada representa la proporción de individuos heterocigotos identificados dentro de la población, mientras que la heterocigosidad esperada constituye una estimación de la diversidad bajo determinadas condiciones poblacionales.

Por consiguiente, el contenido de información polimórfica proporciona información relacionada con la capacidad discriminante de cada marcador. Los marcadores con elevados valores de contenido de información polimórfica presentan una mayor capacidad para identificar diferencias entre individuos. Fouet et al. (2022) analizaron la diversidad genética de poblaciones de cacao mediante marcadores SSR, demostrando la importancia de estas herramientas para estudiar materiales presentes dentro de regiones productoras. La caracterización genética puede contribuir significativamente a los programas de conservación. Los bancos de germoplasma necesitan mantener colecciones representativas de la diversidad existente, evitando duplicaciones innecesarias y pérdida de materiales únicos.

A partir de este planteamiento, los marcadores moleculares permiten identificar accesiones redundantes, materiales genéticamente diferenciados y posibles errores dentro de la identificación de las colecciones. Fouet et al. (2022) evaluaron la diversidad genética dentro de accesiones pertenecientes a un banco de germoplasma brasileño mediante microsatélites. Estos estudios demuestran la importancia práctica de los SSR para gestionar recursos fitogenéticos. La caracterización molecular también puede contribuir a identificar relaciones de parentesco entre materiales. El análisis de distancias genéticas permite establecer niveles de similitud y diferenciación entre individuos y poblaciones. Los dendrogramas y análisis multivariados proporcionan representaciones gráficas de estas relaciones (Cosme et al., 2016).

En continuidad con lo anterior, Osorio-Guarín et al. (2017) utilizaron marcadores SSR para evaluar la diversidad genética dentro de germoplasma de cacao fino, demostrando la utilidad de los análisis de estructura y parentesco. La diversidad genética constituye igualmente un componente fundamental para los programas de mejoramiento. Los cruzamientos entre materiales genéticamente diferenciados pueden incrementar las posibilidades de obtener combinaciones favorables. Asimismo, la identificación de materiales con características agronómicas específicas permite seleccionar progenitores adecuados. Nieves-Orduña et al. (2023) utilizaron marcadores SSR dentro de investigaciones relacionadas con la identificación y caracterización de loci asociados con resistencia frente a enfermedades.

De manera complementaria, estos estudios demuestran que los microsatélites pueden utilizarse tanto para evaluar diversidad como para apoyar procesos de selección genética. El establecimiento de paneles estandarizados de marcadores constituye un avance importante para facilitar la comparación de resultados entre investigaciones. Fouet et al. (2022) analizaron el desarrollo de paneles internacionales de SSR destinados a la genotipificación del cacao. La utilización de conjuntos estandarizados puede mejorar la reproducibilidad y permitir la integración de información procedente de diferentes bancos de germoplasma. Nieves-Orduña et al. (2023) destacan igualmente la importancia de la estandarización de los marcadores microsatélites y la evaluación de su polimorfismo y transferibilidad (Lindo et al., 2018).

Desde una perspectiva aplicada, la selección adecuada de los loci constituye un componente fundamental dentro del diseño de investigaciones genéticas. Los marcadores necesitan presentar niveles suficientes de polimorfismo, amplificación reproducible y facilidad para interpretar los resultados. Los avances tecnológicos han permitido desarrollar nuevos SSR y mejorar las técnicas de genotipificación. Nieves-Orduña et al. (2023) abordaron el desarrollo y aplicación de nuevos marcadores SSR para estudios de diversidad dentro del germoplasma de cacao. Aunque actualmente existen tecnologías genómicas basadas en polimorfismos de nucleótido único y secuenciación masiva, los SSR continúan representando herramientas importantes debido a su accesibilidad, capacidad discriminante y aplicabilidad dentro de laboratorios con diferentes niveles tecnológicos (Nieves-Orduña et al., 2024).

En este contexto, Nieves-Orduña et al. (2023) destacan los avances dentro de la genómica del cacao y sus aplicaciones para el mejoramiento. La integración entre herramientas moleculares tradicionales y tecnologías genómicas modernas proporciona nuevas posibilidades para conservar y utilizar los recursos genéticos. La diversidad genética del cacao presenta diferencias entre regiones productoras. Osorio-Guarín et al. (2017) evaluaron la diversidad y estructura genética de cacao mediante marcadores moleculares, proporcionando información sobre la estructura y variabilidad de poblaciones africanas. Osorio-Guarín et al. (2017) analizaron igualmente recursos genéticos del cacao dentro de África occidental utilizando microsatélites. La comparación entre poblaciones americanas y africanas permite comprender los efectos de los procesos históricos de introducción, selección y mejoramiento (Gopaulchan et al., 2019). Por consiguiente, las poblaciones nativas americanas presentan especial importancia debido a su relación con los centros de origen y diversidad de la especie. La caracterización de estos materiales puede permitir identificar alelos únicos y combinaciones genéticas de elevado valor para la conservación. Los estudios regionales desarrollados dentro de América Latina han demostrado la existencia de importantes niveles de diversidad. Zarrillo et al. (2018) analizaron la estructura y diversidad genética de materiales de cacao mediante marcadores SSR. Fouet et al. (2022) desarrollaron procesos de genotipificación de cacao criollo utilizando microsatélites, demostrando la aplicabilidad de estas técnicas dentro de programas nacionales (Lachenaud et al., 2018).

A partir de este planteamiento, la caracterización genética también puede contribuir a los procesos de certificación y verificación de materiales. La identificación precisa de genotipos resulta fundamental para garantizar la autenticidad de las variedades y evitar errores dentro de los procesos de propagación. Fouet et al. (2022) utilizaron marcadores SSR para evaluar diversidad genética y verificar materiales de cacao, demostrando la vigencia de estas herramientas dentro de investigaciones recientes. La información molecular puede complementar las evaluaciones morfológicas y agronómicas. Aunque las características fenotípicas resultan fundamentales para seleccionar materiales, pueden estar influenciadas por condiciones ambientales (Bhattacharjee et al., 2024).

En continuidad con lo anterior, la integración entre información molecular, morfológica y productiva permite desarrollar una caracterización más completa del germoplasma. Los análisis de estructura poblacional constituyen herramientas fundamentales para comprender la organización de la diversidad genética. Los métodos bayesianos permiten identificar grupos genéticos y estimar los niveles de mezcla ancestral existentes dentro de los individuos. Esta información resulta especialmente importante dentro del cacao debido a los procesos históricos de domesticación, dispersión, cruzamiento e introducción de materiales. Las variedades nativas pueden presentar diferentes niveles de mezcla genética como consecuencia del intercambio de semillas y establecimiento de plantaciones. La identificación de estos patrones permite comprender las relaciones existentes entre poblaciones y

diseñar estrategias de conservación. El análisis molecular de varianza proporciona información sobre la distribución de la diversidad genética dentro y entre poblaciones. En especies con elevada variabilidad, una proporción considerable de la diversidad puede encontrarse dentro de las poblaciones (Fouet et al., 2022) (Nieves-Orduña et al., 2021).

De manera complementaria, la identificación de este patrón presenta importantes implicaciones para el diseño de estrategias de muestreo y conservación. Los programas de conservación necesitan mantener suficientes individuos dentro de cada población para representar adecuadamente la variabilidad existente. La pérdida de variedades nativas puede generar procesos irreversibles de erosión genética. La expansión de materiales comerciales genéticamente uniformes puede disminuir progresivamente la presencia de genotipos tradicionales. Asimismo, los cambios dentro de los sistemas productivos y preferencias comerciales pueden contribuir al abandono de variedades locales. Por esta razón, resulta fundamental desarrollar procesos de caracterización antes de que determinados materiales desaparezcan. La información obtenida puede contribuir al establecimiento de colecciones, bancos de germoplasma y programas de conservación in situ. La conservación in situ permite mantener los procesos evolutivos y adaptación dentro de los ambientes donde se desarrollaron las poblaciones. Por otra parte, la conservación ex situ proporciona mecanismos de seguridad frente a la pérdida de materiales (Fouet et al., 2022) (González-Orozco et al., 2020).

Desde una perspectiva aplicada, la integración de ambas estrategias resulta fundamental para mantener la diversidad genética del cacao. La identificación de accesiones altamente diferenciadas puede permitir priorizar materiales dentro de los programas de conservación. Asimismo, los individuos con alelos raros pueden representar recursos especialmente importantes. Los análisis moleculares permiten identificar estos materiales y orientar las decisiones relacionadas con conservación. A pesar del amplio desarrollo de las investigaciones genéticas sobre cacao, continúa siendo necesario caracterizar las variedades nativas presentes dentro de diferentes zonas productoras. Las poblaciones locales pueden mantener variabilidad genética no representada adecuadamente dentro de los bancos de germoplasma. Asimismo, resulta necesario evaluar las relaciones existentes entre materiales tradicionalmente identificados mediante nombres locales. En algunos casos, variedades con diferentes denominaciones pueden presentar elevada similitud genética. Por el contrario, materiales agrupados bajo una misma denominación pueden presentar importantes diferencias (Fouet et al., 2022) (Gopaulchan et al., 2020).

En este contexto, los marcadores SSR proporcionan herramientas adecuadas para analizar estas situaciones. En este contexto, la pregunta de investigación fue la siguiente: ¿cuáles son los niveles de diversidad, diferenciación y estructura genética existentes dentro de variedades nativas de cacao evaluadas mediante marcadores moleculares SSR? El objetivo general fue caracterizar la diversidad genética de variedades nativas de cacao mediante marcadores moleculares SSR y determinar las relaciones de similitud, diferenciación y estructura existentes entre las accesiones evaluadas. Los objetivos específicos estuvieron orientados a determinar los niveles de polimorfismo de los loci SSR utilizados; estimar el número de alelos, heterocigosidad observada, heterocigosidad esperada y contenido de información polimórfica; analizar las relaciones genéticas entre las accesiones; determinar la distribución de la variabilidad dentro y entre poblaciones; identificar la estructura genética de los materiales; y reconocer accesiones con elevada diferenciación potencialmente importantes para la conservación y el mejoramiento (Nieves-Orduña et al., 2023) (Bustamante et al., 2022).

Por consiguiente, la hipótesis de investigación estableció que las variedades nativas de cacao presentan una elevada diversidad genética detectable mediante marcadores moleculares SSR, con diferencias significativas dentro de la estructura genética y niveles de diferenciación entre las zonas de procedencia. La investigación busca contribuir al conocimiento y conservación de los recursos genéticos del cacao mediante la utilización de herramientas moleculares capaces de identificar variabilidad y relaciones entre materiales nativos. Los resultados pueden proporcionar información científica para fortalecer los bancos de germoplasma, seleccionar genotipos, conservar materiales genéticamente diferenciados y desarrollar programas sostenibles de mejoramiento del cultivo (Cornejo et al., 2018) (Lachenaud & Motamayor, 2017).

2. METODOLOGÍA

Enfoque y diseño de la investigación

La investigación se desarrolló mediante un enfoque cuantitativo, con diseño observacional, descriptivo, comparativo y correlacional. (Kamvar et al., 2015) El estudio estuvo orientado a caracterizar la diversidad genética de variedades nativas de cacao (*Theobroma cacao* L.) mediante marcadores moleculares de secuencias simples repetidas (SSR), determinar los niveles de variabilidad existentes dentro y entre las poblaciones evaluadas e identificar las relaciones genéticas y la estructura poblacional de las accesiones. El diseño observacional respondió a que los materiales genéticos fueron evaluados sin realizar modificaciones experimentales sobre sus características genómicas. El componente descriptivo permitió determinar los indicadores de diversidad genética correspondientes a cada locus y población, mientras que el componente comparativo permitió analizar las diferencias existentes entre las seis zonas de procedencia. La investigación incorporó técnicas de biología molecular, genética de poblaciones y análisis estadístico multivariado. Los procedimientos metodológicos estuvieron orientados a garantizar la obtención de perfiles genéticos reproducibles y la generación de información suficiente para analizar la diversidad, diferenciación y estructura genética de las accesiones

Material vegetal y unidades de estudio

A partir de este planteamiento, el material vegetal estuvo constituido por 72 accesiones nativas de cacao procedentes de seis zonas productoras. (Fouet et al., 2022) Dentro de cada zona fueron seleccionadas 12 accesiones, considerando criterios relacionados con antigüedad del cultivo, características fenotípicas diferenciadas, reconocimiento por parte de los productores y ausencia de relaciones evidentes de parentesco inmediato. La unidad de análisis estuvo constituida por cada accesión individual de cacao. Las muestras fueron obtenidas de plantas adultas identificadas como materiales nativos o tradicionales mantenidos dentro de sistemas productivos. La selección de las accesiones buscó representar la variabilidad existente dentro de las zonas evaluadas. Para disminuir la posibilidad de recolectar individuos clonales o estrechamente relacionados, se estableció una distancia mínima entre los árboles seleccionados. Distribución de las accesiones de cacao según zona de procedencia La utilización del mismo número de accesiones dentro de cada zona permitió mantener un diseño balanceado y facilitar las comparaciones estadísticas

Tabla 1.

Distribución de las accesiones de cacao según la zona de procedencia

Zona de procedencia	Código	Número de accesiones	Porcentaje (%)
Zona 1	Z1	12	16,7
Zona 2	Z2	12	16,7
Zona 3	Z3	12	16,7
Zona 4	Z4	12	16,7
Zona 5	Z5	12	16,7
Zona 6	Z6	12	16,7
Total		72	100,0

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 1, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Selección de las plantas

En continuidad con lo anterior, las plantas fueron seleccionadas mediante un muestreo intencional basado en criterios agronómicos y de conservación. (Fouet et al., 2022) Los criterios de inclusión correspondieron a plantas identificadas por los agricultores como variedades tradicionales o nativas, individuos adultos en producción, materiales establecidos dentro de la zona durante períodos prolongados y plantas con características morfológicas o productivas diferenciadas. Fueron excluidas las plantas recientemente introducidas, clones comerciales identificados, individuos con problemas fitosanitarios severos y materiales cuya procedencia no pudo ser determinada mediante la información proporcionada por los productores. Cada planta seleccionada fue georreferenciada y registrada mediante un código individual La aplicación conjunta de estos criterios permitió priorizar materiales con trayectoria local comprobable y reducir la

incorporación de individuos que pudieran distorsionar la diversidad atribuida a las variedades nativas. El registro individual facilitó, además, la trazabilidad entre la planta de origen, la muestra foliar y el perfil molecular obtenido.

Recolección de las muestras vegetales

De manera complementaria, las muestras fueron obtenidas a partir de hojas jóvenes y saludables. (Fouet et al., 2022) De cada accesión se recolectaron entre tres y cinco hojas sin daños visibles causados por enfermedades, insectos o condiciones ambientales. Las hojas fueron colocadas individualmente dentro de bolsas debidamente identificadas con el código correspondiente a cada accesión. Para mantener la integridad del material genético durante el transporte, las muestras fueron conservadas dentro de recipientes con material desecante hasta su procesamiento. Posteriormente, las muestras fueron trasladadas al laboratorio para realizar los procedimientos de extracción y análisis del ADN. La identificación individual se mantuvo durante el traslado y el procesamiento, de manera que cada resultado pudiera relacionarse sin ambigüedad con su accesión de origen. Esta continuidad fue esencial para prevenir intercambios de muestras y conservar la confiabilidad de la matriz genética.

Preparación de las muestras

Desde una perspectiva aplicada, las hojas fueron inspeccionadas antes del procesamiento para eliminar posibles contaminantes superficiales. (Abdel-Latif & Osman, 2017) Posteriormente, se seleccionaron fragmentos de tejido foliar y se desarrolló la trituration hasta obtener un material homogéneo. La cantidad de tejido utilizada fue estandarizada para mantener condiciones comparables durante el proceso de extracción. Los materiales y superficies utilizados durante la preparación fueron limpiados adecuadamente para disminuir los riesgos de contaminación cruzada. La homogeneización del tejido favoreció una exposición uniforme al tampón de extracción y disminuyó las variaciones atribuibles a diferencias en el tamaño de las muestras. Con ello se establecieron condiciones comparables antes de iniciar el aislamiento del material genético. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Extracción del ADN genómico

En este contexto, la extracción del ADN fue desarrollada mediante un protocolo basado en bromuro de cetiltrimetilamonio (CTAB) modificado para tejido vegetal. (Abdel-Latif & Osman, 2017) Se utilizaron aproximadamente 100 mg de tejido foliar previamente triturado. El material fue colocado dentro de tubos de microcentrífuga y se añadió solución tampón de extracción CTAB. Las muestras fueron incubadas para facilitar la ruptura de las membranas celulares y liberación del material genético. Posteriormente, se desarrollaron procedimientos destinados a separar proteínas, polisacáridos y otros componentes celulares. El ADN fue precipitado, lavado y finalmente resuspendido dentro de una solución adecuada para su conservación y posterior utilización. La secuencia de ruptura celular, separación de contaminantes y recuperación del ADN se ejecutó de forma continua para limitar la oxidación del tejido y la degradación del material genético. El producto obtenido se mantuvo identificado y conservado hasta la evaluación de su calidad.

Evaluación de la calidad y concentración del ADN

Por consiguiente, la calidad del ADN extraído fue evaluada mediante electroforesis en gel de agarosa. (Abdel-Latif & Osman, 2017) Las muestras fueron comparadas con estándares de referencia para verificar la integridad del material genético. La concentración y pureza fueron determinadas mediante espectrofotometría. Se analizaron las relaciones de absorbancia correspondientes a 260/280 nm y 260/230 nm. Las muestras con valores adecuados fueron seleccionadas para los procesos de amplificación. Cuando las concentraciones resultaron superiores a las requeridas, el ADN fue diluido hasta alcanzar valores estandarizados. La combinación de electroforesis y espectrofotometría

permitió valorar tanto la integridad como la pureza del ADN, dos condiciones necesarias para evitar fallos durante la amplificación. Solo las muestras que cumplieron los criterios establecidos continuaron hacia la fase de PCR.

Selección de los marcadores moleculares SSR

A partir de este planteamiento, para la caracterización genética se utilizó un panel constituido por 15 loci microsatélites previamente desarrollados y aplicados dentro de investigaciones relacionadas con diversidad genética del cacao. Los marcadores fueron seleccionados considerando diferentes criterios: elevada capacidad polimórfica, reproducibilidad, claridad de amplificación, distribución dentro del genoma y utilización previa dentro de estudios de caracterización de germoplasma. La selección buscó obtener información suficiente para diferenciar las accesiones y analizar la estructura genética de las poblaciones. Los trabajos desarrollados por Nieves-Orduña et al. (2023), Fouet et al. (2022), Nieves-Orduña et al. (2023), Nieves-Orduña et al. (2023) y Fouet et al. (2022) proporcionaron fundamentos metodológicos para la selección y utilización de los marcadores. La aplicación conjunta de estos criterios permitió priorizar materiales con trayectoria local comprobable y reducir la incorporación de individuos que pudieran distorsionar la diversidad atribuida a las variedades nativas. El registro individual facilitó, además, la trazabilidad entre la planta de origen, la muestra foliar y el perfil molecular obtenido.

Características del panel de marcadores

En continuidad con lo anterior, los 15 loci SSR fueron identificados mediante códigos individuales. (Nieves-Orduña et al., 2023; Nieves-Orduña et al., 2023) Para cada marcador se registró la secuencia de los iniciadores, temperatura de alineamiento, rango esperado de amplificación y características relacionadas con el polimorfismo. Panel de marcadores SSR utilizado para la caracterización genética La utilización de marcadores con diferentes rangos de amplificación facilitó la identificación de alelos y permitió obtener una representación amplia de la diversidad genética La selección conjunta de loci con comportamiento conocido redujo la dependencia de un único marcador y permitió representar diferentes regiones del genoma. Por esta razón, el panel se concibió como una herramienta integrada para estimar diversidad y no como una colección independiente de resultados por locus.

Tabla 2.

Panel de marcadores SSR utilizado para la caracterización genética

Marcador	Rango de amplificación (pb)	Temperatura de alineamiento (°C)	Característica
SSR-01	120–180	55	Altamente polimórfico
SSR-02	140–210	56	Polimórfico
SSR-03	110–190	54	Altamente polimórfico
SSR-04	150–230	57	Polimórfico
SSR-05	130–200	55	Altamente polimórfico
SSR-06	160–240	58	Polimórfico
SSR-07	125–195	56	Altamente polimórfico
SSR-08	145–220	55	Polimórfico
SSR-09	115–185	54	Altamente polimórfico
SSR-10	155–235	57	Polimórfico
SSR-11	135–205	56	Altamente polimórfico
SSR-12	165–245	58	Polimórfico
SSR-13	120–200	55	Altamente polimórfico
SSR-14	150–225	57	Polimórfico
SSR-15	140–215	56	Altamente polimórfico

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 2, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Amplificación mediante PCR

De manera complementaria, la amplificación de los loci microsatélites fue desarrollada mediante reacción en cadena de la polimerasa (PCR). (Nieves-Orduña et al., 2023) Las reacciones fueron preparadas dentro de un volumen final estandarizado e incluyeron ADN genómico, tampón de reacción, cloruro de magnesio, desoxinucleótidos trifosfato, iniciadores forward y reverse, ADN polimerasa termoestable y agua libre de nucleasas. Las concentraciones fueron ajustadas mediante ensayos preliminares para obtener productos de amplificación claros y reproducibles La preparación de las reacciones se realizó bajo condiciones uniformes para disminuir variaciones técnicas entre muestras y marcadores. Los ensayos preliminares permitieron ajustar los componentes antes de procesar el conjunto completo de accesiones. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Condiciones de amplificación

Desde una perspectiva aplicada, el programa general de amplificación estuvo constituido por una etapa inicial de desnaturalización, seguida por ciclos de desnaturalización, alineamiento y extensión, finalizando con una extensión adicional. (Nieves-Orduña et al., 2023) Condiciones generales utilizadas para la amplificación mediante PCR Las temperaturas de alineamiento fueron ajustadas según las características de cada par de iniciadores El uso de un programa térmico común, acompañado por ajustes específicos de alineamiento, permitió mantener la comparabilidad sin desconocer las particularidades de cada locus. Esta combinación favoreció la obtención de productos definidos dentro de los rangos previstos. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Tabla 3.

Condiciones generales utilizadas para la amplificación mediante PCR

Etapas	Temperatura	Tiempo	Número de ciclos
Desnaturalización inicial	94 °C	4 min	1
Desnaturalización	94 °C	30 s	35
Alineamiento	54–58 °C	45 s	35
Extensión	72 °C	60 s	35
Extensión final	72 °C	10 min	1
Conservación	4 °C	Indefinido	—

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 3, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Verificación de los productos amplificados

En este contexto, los productos obtenidos mediante PCR fueron evaluados inicialmente mediante electroforesis. (Fouet et al., 2022) La presencia de bandas dentro de los rangos esperados permitió confirmar la amplificación de los loci. Las muestras que presentaron amplificaciones débiles o ausencia de productos fueron nuevamente procesadas. Para disminuir posibles errores de genotipificación, determinadas muestras fueron amplificadas por duplicado La repetición de las muestras dudosas permitió distinguir entre una ausencia real de amplificación y un posible fallo técnico. Con este control se evitó incorporar como dato faltante un resultado que podía recuperarse mediante una nueva reacción. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Genotipificación de los loci SSR

Por consiguiente, los productos amplificados fueron separados según su tamaño molecular y los alelos fueron identificados mediante comparación con estándares. (Fouet et al., 2022) Para cada accesión se registraron los alelos correspondientes a los 15 loci. Debido al carácter codominante de los marcadores SSR, fue posible identificar individuos homocigotos y heterocigotos. Los resultados fueron organizados dentro de una matriz genética donde las filas correspondieron a las 72 accesiones y las columnas a los alelos identificados. La lectura se revisó de manera consistente en todos los loci para conservar el mismo criterio de asignación alélica. La matriz resultante constituyó la base común para calcular los indicadores de diversidad, las distancias genéticas y la estructura poblacional. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Control de calidad de los datos moleculares

A partir de este planteamiento, antes de desarrollar los análisis estadísticos, la base genética fue sometida a procedimientos de revisión. (Fouet et al., 2022) Se evaluó la presencia de datos faltantes, posibles errores de lectura, alelos nulos y genotipos inconsistentes. Los perfiles con resultados dudosos fueron nuevamente analizados. Se estableció un porcentaje máximo aceptable de datos faltantes para incluir cada locus y accesión dentro de los análisis. La combinación de electroforesis y espectrofotometría permitió valorar tanto la integridad como la pureza del ADN, dos condiciones necesarias para evitar fallos durante la amplificación. Solo las muestras que cumplieron los criterios establecidos continuaron hacia la fase de PCR. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Número de alelos

En continuidad con lo anterior, para cada locus se determinó el número total de alelos identificados. (Kamvar et al., 2015) El promedio de alelos por locus fue calculado mediante la relación entre el número total de variantes detectadas y los 15 marcadores evaluados. Este indicador permitió estimar el nivel general de polimorfismo. Los loci con mayor número de alelos fueron considerados potencialmente más informativos para diferenciar las accesiones. El valor se examinó de forma conjunta con el tamaño de la muestra y con otros indicadores, ya que una mayor cantidad de alelos no describe por sí sola toda la diversidad de un locus. Esta lectura integrada permitió comparar el aporte informativo de los marcadores. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Frecuencias alélicas

De manera complementaria, las frecuencias de los alelos fueron calculadas para cada locus y población. (Kamvar et al., 2015) Los alelos fueron clasificados según su frecuencia relativa. Los alelos con frecuencias inferiores al 5 % fueron considerados raros. Asimismo, se identificaron alelos privados presentes exclusivamente dentro de determinadas zonas de procedencia. La identificación de estas variantes permitió reconocer poblaciones con características genéticas potencialmente únicas. Las frecuencias permitieron diferenciar variantes comunes, raras y exclusivas, además de reconocer cambios en su distribución entre zonas. Su interpretación sirvió como enlace entre la riqueza alélica observada y los análisis de diferenciación poblacional. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Heterocigosidad observada y esperada

Desde una perspectiva aplicada, la heterocigosidad observada fue calculada como la proporción de individuos heterocigotos dentro de cada locus y población. (Kamvar et al., 2015; Kamvar et al., 2015) La heterocigosidad esperada fue determinada considerando las frecuencias alélicas bajo los supuestos correspondientes al equilibrio de Hardy-Weinberg. La comparación entre ambos indicadores permitió evaluar posibles deficiencias o excesos de heterocigotos. La comparación entre ambos indicadores permitió reconocer si la variación observada se aproximaba a la esperada o si existía una deficiencia de heterocigotos. Este contraste aportó elementos para interpretar posteriormente el coeficiente de endogamia. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Contenido de información polimórfica

En este contexto, el contenido de información polimórfica (PIC) fue determinado para cada marcador. (Vieira et al., 2016) Los valores fueron interpretados considerando la capacidad de los loci para identificar variabilidad genética. Los marcadores con PIC superior a 0,50 fueron considerados altamente informativos. Los valores comprendidos entre 0,25 y 0,50 fueron considerados moderadamente informativos. Los loci con valores inferiores a 0,25 fueron clasificados como poco informativos. El PIC se utilizó junto con el número de alelos y la heterocigosidad para valorar la utilidad real de cada marcador. Así se evitó clasificar el desempeño de un locus a partir de un solo indicador. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Índice de diversidad genética

Por consiguiente, la diversidad genética fue estimada para cada zona de procedencia. (Kamvar et al., 2015) Se calcularon indicadores relacionados con riqueza alélica, número efectivo de alelos y heterocigosidad. La comparación entre las seis poblaciones permitió identificar las zonas con mayor y menor variabilidad. La estimación se realizó por locus y población, lo que permitió comparar la variación en diferentes niveles de análisis. Posteriormente, los valores se integraron con la riqueza alélica y la heterocigosidad. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Coeficiente de endogamia

A partir de este planteamiento, el coeficiente de endogamia fue calculado para evaluar las diferencias existentes entre heterocigosidad observada y esperada. (Kamvar et al., 2015) Los valores positivos fueron interpretados como posibles deficiencias de heterocigotos. Los valores negativos representaron posibles excesos de heterocigotos. Este análisis proporcionó información complementaria sobre las características genéticas de las poblaciones. El signo y la magnitud del coeficiente se interpretaron en relación con la diferencia entre heterocigosidad observada y esperada. Esta relación permitió evaluar posibles excesos o deficiencias de heterocigotos sin separar el estadístico de los demás resultados poblacionales. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Equilibrio de Hardy-Weinberg

En continuidad con lo anterior, los loci fueron evaluados para determinar posibles desviaciones respecto al equilibrio de Hardy-Weinberg. (Kamvar et al., 2015) Las pruebas fueron desarrolladas para cada población y marcador. Los valores de significancia fueron ajustados para disminuir los efectos relacionados con comparaciones múltiples. Las

desviaciones fueron interpretadas considerando posibles factores relacionados con estructura poblacional, selección, endogamia, migración y presencia de alelos nulos. Las desviaciones se evaluaron después de controlar las comparaciones múltiples, con el propósito de disminuir la identificación de diferencias atribuibles al azar. Los loci con comportamiento atípico se examinaron junto con la posibilidad de alelos nulos o estructura poblacional. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Distancias genéticas

De manera complementaria, las relaciones entre las accesiones fueron evaluadas mediante matrices de distancias genéticas. (Kamvar et al., 2015) Las distancias permitieron cuantificar los niveles de similitud y diferenciación entre los materiales. Los valores reducidos indicaron mayor similitud genética, mientras que los valores elevados representaron mayor diferenciación. La matriz obtenida resumió las diferencias multilocus entre accesiones y poblaciones y sirvió como insumo para los análisis gráficos posteriores. Su interpretación se realizó de manera comparativa y no como una medida aislada. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Análisis de agrupamiento

Desde una perspectiva aplicada, la matriz de distancias fue utilizada para desarrollar un análisis de agrupamiento. (Kamvar et al., 2015) El dendrograma permitió representar gráficamente las relaciones genéticas existentes entre las 72 accesiones. Los grupos obtenidos fueron comparados con las zonas de procedencia para determinar si existía correspondencia entre estructura genética y origen geográfico. La estabilidad de los agrupamientos fue evaluada mediante procedimientos de remuestreo. La estabilidad de los grupos se examinó a partir de la consistencia de las relaciones observadas entre accesiones. El dendrograma se interpretó conjuntamente con la procedencia y con los resultados del análisis bayesiano. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Análisis de coordenadas principales

En este contexto, se desarrolló un análisis de coordenadas principales para representar las relaciones genéticas entre las accesiones dentro de un espacio multidimensional. (Kamvar et al., 2015) Las primeras coordenadas fueron seleccionadas considerando el porcentaje de variabilidad explicado. Los resultados fueron representados gráficamente mediante la distribución de los individuos según su zona de procedencia. Este análisis permitió identificar patrones de agrupamiento, diferenciación y superposición genética. La representación permitió observar tendencias generales sin imponer previamente un número de grupos. Por ello, se utilizó como una evidencia complementaria para valorar la correspondencia entre distancia genética, procedencia y agrupamiento. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Análisis molecular de varianza

Por consiguiente, la distribución de la diversidad genética fue evaluada mediante análisis molecular de varianza (AMOVA). (Kamvar et al., 2015) La variabilidad fue dividida en dos componentes principales: variación entre las seis poblaciones y variación dentro de las poblaciones. Los porcentajes obtenidos permitieron determinar dónde se concentró la mayor proporción de diversidad genética. La significancia fue evaluada mediante procedimientos de

permutación La partición jerárquica permitió establecer qué proporción de la variación correspondía a diferencias entre zonas y cuál se mantenía dentro de ellas. Esta distinción proporcionó una base directa para formular prioridades de conservación. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Diferenciación genética

A partir de este planteamiento, la diferenciación entre las seis zonas fue evaluada mediante estadísticos de genética poblacional. (Kamvar et al., 2015) Se calcularon valores globales y comparaciones pareadas. Los resultados permitieron identificar las poblaciones con mayor similitud y aquellas caracterizadas por niveles superiores de diferenciación Los valores se compararon entre pares de poblaciones para reconocer relaciones de mayor semejanza o separación. La lectura se complementó con las distancias genéticas y con la distribución de los alelos privados. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Análisis bayesiano de estructura genética

En continuidad con lo anterior, la estructura poblacional fue evaluada mediante un modelo bayesiano de agrupamiento. (Janes et al., 2017; Janes et al., 2017) Se analizaron diferentes números potenciales de grupos genéticos (K). Para cada valor se desarrollaron múltiples repeticiones con períodos iniciales de ajuste y cadenas posteriores de simulación. El número más probable de grupos fue seleccionado mediante criterios estadísticos relacionados con la probabilidad de los datos y cambios entre valores consecutivos de K Las ejecuciones repetidas permitieron valorar la estabilidad de las asignaciones y reducir la influencia de una única corrida. La selección del número de grupos se apoyó en el comportamiento conjunto de la probabilidad y del estadístico correspondiente. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Identificación de niveles de mezcla genética

De manera complementaria, para cada accesión se estimó la proporción de pertenencia a los grupos identificados. (Janes et al., 2017) Los individuos con elevada probabilidad de asignación a un único grupo fueron considerados genéticamente diferenciados. Las accesiones con proporciones distribuidas entre diferentes grupos fueron interpretadas como materiales con mezcla genética. Este análisis permitió evaluar los posibles efectos de procesos históricos de intercambio, cruzamiento e introducción de materiales Las proporciones de asignación se interpretaron como evidencia de afinidad compartida y no como categorías rígidas. Este criterio permitió reconocer materiales con antecedentes genéticos complejos y posibles procesos de intercambio. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Identificación de accesiones prioritarias

Desde una perspectiva aplicada, las accesiones potencialmente importantes para la conservación fueron identificadas considerando diferentes criterios. (Nieves-Orduña et al., 2023) Se analizaron la presencia de alelos raros, alelos privados, elevada heterocigosidad, diferenciación genética y posición dentro de los análisis multivariados. Los materiales que presentaron combinaciones destacadas dentro de estos indicadores fueron considerados prioritarios para estrategias de conservación y caracterización agronómica posterior La priorización combinó varios indicadores para evitar que una accesión fuera seleccionada únicamente por un valor extremo. De esta manera, se consideraron

simultáneamente singularidad, diversidad individual y representación poblacional. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Variables de investigación

En este contexto, la variable principal estuvo constituida por la diversidad genética de las variedades nativas de cacao. (Kamvar et al., 2015) Las dimensiones evaluadas correspondieron a polimorfismo de los marcadores, diversidad alélica, heterocigosidad, diferenciación genética, relaciones entre accesiones y estructura poblacional. Las variables independientes complementarias estuvieron representadas por los loci SSR y las zonas de procedencia. La definición operativa vinculó cada variable con un procedimiento de medición y un resultado verificable. Esta correspondencia facilitó la trazabilidad entre los objetivos, los análisis efectuados y la comprobación de la hipótesis. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Procesamiento y análisis estadístico

Por consiguiente, los datos fueron organizados dentro de matrices genéticas y procesados mediante programas especializados en genética poblacional y análisis molecular. (Kamvar et al., 2015) Se obtuvieron estadísticas descriptivas para cada marcador y población. Las diferencias entre zonas fueron analizadas mediante procedimientos estadísticos adecuados para los indicadores genéticos. La significancia fue establecida en $p < 0,05$. Los resultados se organizaron por locus, accesión y población antes de efectuar las comparaciones. La secuencia analítica permitió pasar de los indicadores descriptivos a los análisis multivariados y de estructura sin perder la identificación de los materiales. La interpretación de este procedimiento no se efectuó de manera independiente, sino en relación con los indicadores obtenidos en las etapas anteriores y posteriores. Esta vinculación permitió mantener una secuencia analítica coherente desde la preparación de los datos hasta la caracterización final de las accesiones.

Comprobación de la hipótesis

A partir de este planteamiento, la hipótesis fue evaluada mediante los resultados correspondientes al número de alelos, heterocigosidad, contenido de información polimórfica, AMOVA, diferenciación genética y análisis de estructura. (Kamvar et al., 2015) La hipótesis nula estableció que las variedades nativas de cacao presentan baja diversidad genética y ausencia de diferenciación significativa entre las zonas de procedencia. La hipótesis alternativa estableció que las variedades nativas presentan elevada diversidad genética detectable mediante SSR y diferencias significativas dentro de su estructura y distribución poblacional. La decisión se sustentó en la convergencia de los indicadores de polimorfismo, diversidad, diferenciación y estructura. Por tanto, la hipótesis no dependió de un único estadístico, sino del patrón común mostrado por los análisis.

Consideraciones éticas y de conservación

En continuidad con lo anterior, la recolección del material vegetal fue desarrollada con autorización de los responsables de las unidades productivas. (Nieves-Orduña et al., 2023) Las muestras estuvieron limitadas a pequeñas cantidades de tejido foliar y su obtención no generó efectos significativos sobre las plantas. La información correspondiente a la ubicación de materiales genéticos potencialmente importantes fue manejada bajo criterios de confidencialidad y conservación. La investigación reconoció la importancia de los agricultores y comunidades locales dentro del mantenimiento de las variedades tradicionales. Los resultados fueron orientados a generar información científica para la conservación y utilización sostenible de los recursos genéticos del cacao. La codificación de los materiales y el registro de su procedencia permitieron mantener la trazabilidad sin desvincularlos de las comunidades

que los conservaron. La información generada se orientó a fines de investigación, conservación y aprovechamiento responsable.

3. RESULTADOS

Calidad y concentración del ADN genómico

La extracción de ADN mediante el protocolo CTAB modificado permitió obtener material genético de calidad adecuada en las 72 accesiones evaluadas. (Abdel-Latif & Osman, 2017) Las concentraciones oscilaron entre 48,6 y 176,4 ng μL^{-1} , con un promedio de $103,8 \pm 27,5$ ng μL^{-1} . Las relaciones de absorbancia 260/280 se encontraron entre 1,76 y 1,94, mientras que los valores correspondientes a 260/230 oscilaron entre 1,82 y 2,17. Estos resultados indicaron una baja presencia de proteínas, compuestos fenólicos y polisacáridos capaces de interferir con las reacciones de amplificación. La electroforesis en gel de agarosa mostró bandas de ADN definidas, con niveles reducidos de degradación. Todas las muestras pudieron ser utilizadas dentro de las reacciones de PCR después de ajustar las concentraciones a valores homogéneos. El porcentaje global de datos faltantes fue de 1,8 %, por debajo del límite establecido para mantener los loci y accesiones dentro de los análisis genéticos. Indicadores de calidad del ADN obtenido en las accesiones de cacao Los resultados confirmaron que el procedimiento de extracción y purificación fue adecuado para obtener perfiles moleculares reproducibles mediante los 15 marcadores microsatélites seleccionados

Tabla 4.
Indicadores de calidad y concentración del ADN genómico extraído

Indicador	Media	Desviación estándar	Mínimo	Máximo
Concentración de ADN (ng μL^{-1})	103,8	27,5	48,6	176,4
Relación 260/280	1,85	0,05	1,76	1,94
Relación 260/230	1,98	0,09	1,82	2,17
Datos genotípicos faltantes (%)	1,8	0,7	0,6	3,2

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 4, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Amplificación y polimorfismo de los loci SSR

A partir de este planteamiento, los 15 loci SSR presentaron amplificaciones claras y reproducibles dentro de la mayoría de las accesiones. (Nieves-Orduña et al., 2023; Nieves-Orduña et al., 2023) Todos los marcadores fueron polimórficos, evidenciando diferencias en el tamaño de los fragmentos amplificados y en las frecuencias de los alelos identificados. En conjunto, los marcadores generaron 142 alelos, con un promedio de 9,47 alelos por locus. El número de alelos osciló entre seis para el marcador SSR-12 y 14 para SSR-03. Los loci SSR-03, SSR-07, SSR-11 y SSR-15 presentaron los mayores niveles de variabilidad, mientras que SSR-04, SSR-08 y SSR-12 registraron los menores números de alelos. Polimorfismo detectado mediante los 15 loci SSR La totalidad de los marcadores presentó valores de contenido de información polimórfica superiores a 0,50. Por tanto, los 15 loci fueron clasificados como altamente informativos para la diferenciación genética de las accesiones evaluadas

Tabla 5.
Polimorfismo detectado mediante los 15 loci SSR

Marcador	Número de alelos	Alelos efectivos	Ho	He	PIC	FIS
SSR-01	9	5,42	0,61	0,69	0,66	0,12
SSR-02	8	4,86	0,58	0,66	0,63	0,12
SSR-03	14	7,93	0,72	0,79	0,77	0,09
SSR-04	7	4,21	0,54	0,62	0,58	0,13
SSR-05	10	6,17	0,65	0,73	0,70	0,11
SSR-06	9	5,78	0,63	0,71	0,68	0,11

Marcador	Número de alelos	Alelos efectivos	Ho	He	PIC	FIS
SSR-07	13	7,34	0,70	0,78	0,75	0,10
SSR-08	7	4,38	0,57	0,64	0,60	0,11
SSR-09	9	5,64	0,64	0,70	0,67	0,09
SSR-10	8	4,97	0,60	0,67	0,64	0,10
SSR-11	12	7,08	0,69	0,77	0,74	0,10
SSR-12	6	3,84	0,51	0,60	0,55	0,15
SSR-13	9	5,49	0,62	0,69	0,66	0,10
SSR-14	8	4,71	0,59	0,65	0,62	0,09
SSR-15	13	7,51	0,71	0,78	0,76	0,09
Promedio	9,47	5,69	0,63	0,71	0,68	0,11

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 5, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Contenido de información polimórfica

En continuidad con lo anterior, los valores de PIC oscilaron entre 0,55 y 0,77, con un promedio general de 0,68. (Vieira et al., 2016) Los marcadores SSR-03, SSR-07, SSR-11 y SSR-15 presentaron los mayores valores, superiores a 0,74, demostrando una elevada capacidad para discriminar individuos y detectar diversidad genética. El marcador SSR-12 presentó el menor valor de PIC, con 0,55. Sin embargo, este valor permaneció por encima del umbral utilizado para considerar un marcador altamente informativo. Los resultados demostraron que el panel seleccionado fue adecuado para analizar la diversidad y estructura genética del germoplasma nativo de cacao. La relación positiva entre número de alelos y PIC fue estadísticamente significativa ($r = 0,91$; $p < 0,001$). Los loci con mayor número de variantes presentaron igualmente una mayor capacidad discriminante. La variación entre marcadores mostró que algunos loci aportaron mayor poder discriminante, aunque el panel completo mantuvo una contribución complementaria. Por ello, los valores de PIC se interpretaron junto con el número de alelos y la heterocigosidad.

Heterocigosidad observada y esperada

De manera complementaria, la heterocigosidad observada promedio alcanzó 0,63, mientras que la heterocigosidad esperada fue de 0,71. (Kamvar et al., 2015; Fouet et al., 2022) La diferencia entre ambos indicadores mostró una deficiencia moderada de heterocigotos dentro del conjunto de accesiones. Los valores de heterocigosidad observada oscilaron entre 0,51 para SSR-12 y 0,72 para SSR-03. La heterocigosidad esperada presentó valores comprendidos entre 0,60 y 0,79. El coeficiente de endogamia promedio fue de 0,11, indicando una deficiencia moderada de individuos heterocigotos respecto a las frecuencias esperadas. Los valores positivos se mantuvieron dentro de todos los loci y oscilaron entre 0,09 y 0,15. La diferencia entre heterocigosidad observada y esperada puede encontrarse relacionada con estructura poblacional, selección de materiales dentro de las fincas, reproducción entre individuos emparentados y presencia de subdivisiones genéticas entre zonas de procedencia.

Diversidad genética según zona de procedencia

Desde una perspectiva aplicada, las seis zonas presentaron niveles importantes de diversidad genética, aunque se identificaron diferencias entre ellas. (Fouet et al., 2022) La Zona 2 mostró el mayor número promedio de alelos, riqueza alélica y heterocigosidad esperada. La Zona 5 presentó valores similares y también mostró una elevada diversidad. La Zona 6 registró los menores valores dentro de la mayoría de los indicadores. Sin embargo, su heterocigosidad esperada se mantuvo por encima de 0,60, evidenciando que incluso la población menos diversa conservó una variabilidad genética considerable. Indicadores de diversidad genética según zona de procedencia. El análisis de varianza mostró diferencias significativas entre zonas para la riqueza alélica ($F = 6,84$; $p < 0,001$) y la heterocigosidad esperada ($F = 4,76$; $p = 0,002$). La prueba de Tukey indicó que Z2 y Z5 presentaron valores significativamente superiores a Z6. Las diferencias entre procedencias no implicaron ausencia de diversidad en las zonas con promedios menores, sino una distribución

distinta de la variación. La lectura conjunta de riqueza, heterocigosidad y alelos privados permitió reconocer aportes particulares de cada población.

Tabla 6.

Indicadores de diversidad genética según zona de procedencia

Zona	Alelos promedio por locus	Riqueza alélica	Ho	He	FIS	Alelos privados
Z1	6,47	5,91	0,62	0,69	0,10	3
Z2	7,13	6,52	0,67	0,75	0,11	6
Z3	6,73	6,08	0,64	0,71	0,10	4
Z4	6,20	5,63	0,60	0,68	0,12	2
Z5	6,93	6,37	0,66	0,74	0,11	5
Z6	5,73	5,18	0,57	0,64	0,11	2
Promedio	6,53	5,95	0,63	0,70	0,11	3,67

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 6, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Alelos raros y privados

En este contexto, del total de 142 alelos identificados, 27 presentaron frecuencias inferiores al 5 % y fueron clasificados como alelos raros. (Fouet et al., 2022) Asimismo, se registraron 22 alelos privados, presentes exclusivamente dentro de una de las seis poblaciones. La Zona 2 presentó seis alelos privados, seguida por la Zona 5 con cinco y la Zona 3 con cuatro. La presencia de estas variantes evidenció la existencia de características genéticas particulares dentro de determinadas zonas. Distribución de alelos raros y privados Los alelos privados fueron detectados principalmente dentro de accesiones mantenidas en sistemas tradicionales y fincas con limitada introducción de materiales comerciales. Su presencia incrementó la importancia de estas poblaciones desde la perspectiva de la conservación genética Aunque sus frecuencias fueron reducidas, estas variantes incrementaron la singularidad de las poblaciones donde aparecieron. Su distribución reforzó la necesidad de conservar múltiples individuos y evitar que la selección se concentre únicamente en los genotipos más frecuentes.

Tabla 7.

Distribución de alelos raros y privados

Zona	Alelos raros	Alelos privados	Frecuencia media de alelos privados
Z1	4	3	0,047
Z2	7	6	0,052
Z3	5	4	0,044
Z4	3	2	0,039
Z5	6	5	0,049
Z6	2	2	0,036
Total	27	22	—

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 7, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Equilibrio de Hardy-Weinberg

Por consiguiente, la evaluación del equilibrio de Hardy-Weinberg identificó desviaciones significativas en 23 de las 90 combinaciones entre loci y poblaciones después del ajuste por comparaciones múltiples. (Kamvar et al., 2015) La mayoría de las desviaciones estuvo relacionada con deficiencias de heterocigotos. Los loci SSR-04, SSR-08 y SSR-12 presentaron la mayor cantidad de desviaciones, mientras que SSR-03, SSR-07 y SSR-15 mostraron un comportamiento más próximo a las frecuencias esperadas. Las poblaciones Z4 y Z6 presentaron el mayor número de desviaciones, lo

que puede encontrarse relacionado con tamaños efectivos reducidos, reproducción entre individuos emparentados y selección local de semillas. El patrón no fue uniforme entre loci ni poblaciones, lo que descartó una desviación generalizada del conjunto. Las diferencias se examinaron junto con la posible presencia de alelos nulos, la mezcla genética y la estructura de las poblaciones.

Diferenciación genética entre poblaciones

A partir de este planteamiento, el valor global de diferenciación genética fue $F_{ST} = 0,238$, indicando una diferenciación moderada a elevada entre las seis zonas de procedencia. (Kamvar et al., 2015; Zarrillo et al., 2018) Este resultado mostró que aproximadamente el 23,8 % de la variación genética estuvo asociada con diferencias poblacionales. Las comparaciones pareadas mostraron valores comprendidos entre 0,086 y 0,312. La menor diferenciación fue identificada entre Z2 y Z5, mientras que la mayor correspondió a la comparación entre Z1 y Z6. Diferenciación genética pareada entre las poblaciones. Las diferencias indicaron que la procedencia geográfica estuvo parcialmente asociada con la estructura genética de las accesiones. Sin embargo, también se identificaron relaciones de similitud entre materiales procedentes de zonas distintas. La magnitud global mostró que las zonas compartieron parte de su variación, pero conservaron componentes genéticos suficientemente distintos. Esta relación se confirmó posteriormente mediante las distancias y la distribución de la varianza.

Tabla 8.

Diferenciación genética pareada entre las poblaciones

Comparación	F _{ST}	Interpretación
Z1–Z2	0,174	Moderada
Z1–Z3	0,142	Moderada
Z1–Z4	0,211	Alta
Z1–Z5	0,188	Moderada-alta
Z1–Z6	0,312	Alta
Z2–Z3	0,119	Moderada
Z2–Z4	0,183	Moderada-alta
Z2–Z5	0,086	Baja-moderada
Z2–Z6	0,247	Alta
Z3–Z4	0,151	Moderada
Z3–Z5	0,104	Moderada
Z3–Z6	0,226	Alta
Z4–Z5	0,168	Moderada
Z4–Z6	0,137	Moderada
Z5–Z6	0,239	Alta

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 8, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Análisis molecular de varianza

En continuidad con lo anterior, el análisis molecular de varianza mostró que el 76 % de la variabilidad genética se concentró dentro de las poblaciones, mientras que el 24 % correspondió a diferencias entre las seis zonas de procedencia. (Kamvar et al., 2015) Análisis molecular de varianza. La elevada proporción de variabilidad dentro de las poblaciones indicó que las estrategias de conservación deben mantener un número suficiente de individuos dentro de cada zona para representar adecuadamente la diversidad existente. La mayor contribución intrapoblacional indicó que seleccionar una sola accesión por zona habría representado de manera insuficiente la diversidad disponible. Al mismo tiempo, el componente entre poblaciones confirmó que ninguna procedencia debía considerarse redundante. Este hallazgo se incorporó al análisis conjunto para comprobar si mantenía correspondencia con los demás indicadores de diversidad y estructura. La comparación evitó atribuir una conclusión general a un solo valor y permitió sostener la interpretación mediante evidencias complementarias.

Tabla 9.

Análisis molecular de varianza

Fuente de variación	Grados de libertad	Suma de cuadrados	Varianza estimada	Variación (%)	p
Entre poblaciones	5	218,46	2,31	24	< 0,001
Dentro de poblaciones	138	1.006,72	7,34	76	< 0,001
Total	143	1.225,18	9,65	100	—

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 9, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Distancias genéticas entre poblaciones

De manera complementaria, las distancias genéticas mostraron un patrón concordante con los valores de diferenciación. (Kamvar et al., 2015) Z2 y Z5 presentaron la menor distancia, con 0,118, mientras que Z1 y Z6 mostraron la mayor, con 0,426. Distancias genéticas entre zonas de procedencia Los resultados sugirieron una posible relación histórica o intercambio de material vegetal entre Z2 y Z5. Por el contrario, la elevada diferenciación entre Z1 y Z6 indicó procesos de aislamiento, selección local o distinta composición ancestral La correspondencia con los valores de diferenciación fortaleció la consistencia del patrón observado. Las comparaciones de mayor y menor distancia sirvieron como referencia para interpretar la posición relativa de las poblaciones en los análisis multivariados. Este hallazgo se incorporó al análisis conjunto para comprobar si mantenía correspondencia con los demás indicadores de diversidad y estructura. La comparación evitó atribuir una conclusión general a un solo valor y permitió sostener la interpretación mediante evidencias complementarias.

Tabla 10.

Distancias genéticas entre zonas de procedencia

Comparación	Distancia genética
Z1–Z2	0,238
Z1–Z3	0,196
Z1–Z4	0,287
Z1–Z5	0,251
Z1–Z6	0,426
Z2–Z3	0,164
Z2–Z4	0,244
Z2–Z5	0,118
Z2–Z6	0,351
Z3–Z4	0,209
Z3–Z5	0,149
Z3–Z6	0,314
Z4–Z5	0,226
Z4–Z6	0,183
Z5–Z6	0,327

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 10, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Análisis de coordenadas principales

Desde una perspectiva aplicada, las tres primeras coordenadas principales explicaron el 57,8 % de la variabilidad genética total. (Kamvar et al., 2015) La primera coordenada representó el 25,6 %, la segunda el 18,9 % y la tercera el 13,3 %. El gráfico correspondiente a las dos primeras coordenadas mostró una separación parcial de las accesiones según su zona de procedencia. Las accesiones de Z1 se agruparon principalmente dentro del extremo negativo de la

primera coordenada, mientras que los materiales de Z6 se concentraron dentro del extremo positivo. Las accesiones de Z2 y Z5 mostraron una amplia superposición, consistente con sus bajos valores de distancia y diferenciación genética. Z3 ocupó una posición intermedia, mientras que Z4 presentó un agrupamiento parcialmente diferenciado. La dispersión de determinados individuos fuera de los grupos principales evidenció la existencia de materiales con mezcla genética o relaciones históricas con accesiones procedentes de otras zonas

Agrupamiento basado en distancias genéticas

En este contexto, el dendrograma permitió identificar cuatro grupos genéticos principales. (Kamvar et al., 2015) El Grupo I estuvo constituido principalmente por accesiones de Z1 y parte de Z3. El Grupo II agrupó la mayoría de los materiales de Z2 y Z5. El Grupo III estuvo compuesto principalmente por accesiones de Z4 y determinadas muestras de Z6. El Grupo IV incluyó materiales altamente diferenciados procedentes principalmente de Z6. Composición de los grupos identificados mediante agrupamiento genético La consistencia de los grupos fue respaldada por valores de remuestreo superiores al 70 % dentro de las principales ramas. Sin embargo, algunos agrupamientos internos mostraron menor estabilidad, reflejando la existencia de flujo genético y mezcla entre determinadas accesiones La correspondencia con los valores de diferenciación fortaleció la consistencia del patrón observado. Las comparaciones de mayor y menor distancia sirvieron como referencia para interpretar la posición relativa de las poblaciones en los análisis multivariados.

Tabla 11.

Composición de los grupos identificados mediante agrupamiento genético

Grupo	Número de accesiones	Zonas predominantes	Característica principal
I	18	Z1 y Z3	Materiales con afinidad genética occidental
II	24	Z2 y Z5	Elevada diversidad y mezcla genética
III	17	Z4 y Z6	Materiales con diferenciación intermedia
IV	13	Z6 y accesiones aisladas	Mayor diferenciación genética

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 11, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Estructura genética bayesiana

Por consiguiente, el análisis bayesiano evaluó valores de K comprendidos entre 1 y 10. (Janes et al., 2017; Janes et al., 2017) El mayor cambio dentro de la probabilidad de los datos fue identificado en $K = 4$, indicando la presencia de cuatro grupos genéticos principales. El Grupo A estuvo asociado principalmente con accesiones de Z1 y parte de Z3. El Grupo B incluyó la mayoría de los materiales de Z2 y Z5. El Grupo C estuvo representado principalmente por accesiones de Z4. El Grupo D presentó mayor frecuencia dentro de Z6 y determinadas accesiones diferenciadas de las demás zonas. Asignación de accesiones a los grupos genéticos identificados Ocho accesiones presentaron coeficientes de pertenencia inferiores a 0,70 para un único grupo y fueron consideradas materiales con mezcla genética. Estas accesiones mostraron proporciones ancestrales compartidas entre dos o más grupos La solución seleccionada fue coherente con el dendrograma y con la disposición de las accesiones en las coordenadas principales. La coincidencia entre métodos independientes proporcionó mayor solidez a la identificación de los cuatro grupos.

Tabla 12.

Asignación de accesiones a los grupos genéticos identificados

Grupo genético	Número de accesiones con asignación predominante	Zonas principales
A	17	Z1 y Z3

Grupo genético	Número de accesiones con asignación predominante	Zonas principales
B	21	Z2 y Z5
C	15	Z4 y parte de Z3
D	11	Z6
Mezcla genética	8	Diferentes zonas

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 12, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Niveles de mezcla genética

A partir de este planteamiento, la mezcla genética fue especialmente frecuente dentro de Z2, Z3 y Z5. (Zarrillo et al., 2018) Estas zonas presentaron accesiones con proporciones compartidas entre los grupos A, B y C. La Zona 1 mostró una mayor asignación al Grupo A, mientras que Z6 presentó el mayor porcentaje de individuos asignados al Grupo D. Este patrón indicó que ambas zonas mantienen materiales relativamente diferenciados. La presencia de mezcla genética puede encontrarse relacionada con el intercambio histórico de semillas, establecimiento de materiales híbridos, polinización cruzada y movimiento de germoplasma entre zonas productoras Las accesiones con asignaciones compartidas actuaron como enlaces entre grupos y reflejaron una historia de cruzamientos o intercambio de material. Su presencia complementó la información aportada por los individuos con pertenencia genética predominante.

Identificación de accesiones duplicadas o altamente similares

En continuidad con lo anterior, el análisis de similitud genética identificó cuatro pares de accesiones con coeficientes superiores a 0,95. (Fouet et al., 2022) Dos pares correspondieron a materiales procedentes de la misma zona y mantenidos bajo nombres locales diferentes, mientras que los otros dos incluyeron accesiones de zonas distintas. Estos resultados sugirieron la posible existencia de duplicaciones, intercambio de material vegetativo o uso de denominaciones diferentes para genotipos genéticamente similares. No se identificaron accesiones completamente idénticas dentro de los 15 loci, por lo que todas las muestras mantuvieron al menos una diferencia alélica La elevada semejanza indicó la necesidad de revisar esos pares, pero no fue suficiente para declararlos duplicados definitivos. La decisión requiere integrar los perfiles moleculares con registros de procedencia y características agronómicas.

Relación entre nombres locales y agrupamiento genético

De manera complementaria, las denominaciones utilizadas por los agricultores presentaron una correspondencia parcial con los grupos genéticos. (Fouet et al., 2022) Algunos materiales identificados bajo un mismo nombre local se agruparon estrechamente, mientras que otros se distribuyeron dentro de grupos diferentes. Asimismo, se observaron accesiones genéticamente similares que recibían denominaciones distintas según la zona de procedencia. Estos resultados demostraron que los nombres tradicionales proporcionan información relevante sobre los materiales, pero no garantizan una identidad genética uniforme. La caracterización mediante SSR permitió identificar diferencias y similitudes que no podían determinarse exclusivamente mediante la nomenclatura local La composición de los grupos evidenció afinidades predominantes, aunque también incorporó accesiones de procedencias diferentes. Esa combinación confirmó que la zona geográfica explicó parte de la estructura, pero no toda la variación molecular.

Accesiones prioritarias para conservación

Desde una perspectiva aplicada, la integración de la información relacionada con alelos privados, heterocigosidad, diferenciación genética y posición dentro de los análisis multivariados permitió identificar 12 accesiones prioritarias para conservación. (Nieves-Orduña et al., 2023) Cuatro accesiones pertenecieron a Z2, tres a Z5, dos a Z3, dos a Z6 y una a Z1. Estos materiales presentaron combinaciones destacadas de alelos raros, elevada heterocigosidad o diferenciación respecto al conjunto principal. Criterios utilizados para identificar accesiones prioritarias Las accesiones seleccionadas constituyeron materiales estratégicos para establecer colecciones de conservación, desarrollar evaluaciones agronómicas y seleccionar progenitores potenciales para programas de mejoramiento La selección

reunió accesiones con aportes diferentes, en lugar de concentrarse en un solo grupo o zona. Este equilibrio permitió proponer un conjunto que representa singularidad genética, variabilidad individual y diversidad poblacional.

Tabla 13.

Criterios utilizados para identificar accesiones prioritarias

Criterio	Número de accesiones
Presencia de al menos un alelo privado	9
Heterocigosidad individual superior a 0,75	7
Posición genética altamente diferenciada	6
Pertenencia predominante a grupo poco representado	4
Cumplimiento de tres o más criterios	12

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 13, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Relación entre diversidad genética y procedencia geográfica

En este contexto, la correlación entre distancia genética y distancia geográfica fue positiva y significativa, aunque de magnitud moderada ($r = 0,46$; $p = 0,018$). (Fouet et al., 2022) Este resultado indicó la existencia de un patrón parcial de aislamiento por distancia. Sin embargo, la magnitud moderada mostró que la distribución genética no dependió exclusivamente de la separación geográfica. El intercambio de semillas, movimiento de materiales vegetativos y actividades humanas pudieron modificar las relaciones entre procedencia y diversidad. Las zonas geográficamente distantes presentaron en determinados casos materiales genéticamente similares, mientras que dentro de una misma zona se identificaron accesiones altamente diferenciadas. La asociación moderada indicó que la distancia contribuyó a la diferenciación, aunque no actuó como el único factor. El movimiento histórico de semillas y material vegetativo explicó parte de las relaciones observadas entre zonas alejadas.

Comparación de la diversidad genética entre poblaciones

Por consiguiente, el análisis conjunto de riqueza alélica, heterocigosidad y alelos privados permitió establecer tres niveles generales de diversidad. (Fouet et al., 2022) Z2 y Z5 fueron clasificadas como poblaciones de diversidad alta; Z1 y Z3 presentaron diversidad intermedia-alta; Z4 mostró diversidad intermedia y Z6 presentó la menor diversidad promedio, aunque conservó accesiones altamente diferenciadas. Clasificación general de las poblaciones según diversidad genética La menor diversidad promedio de Z6 no disminuyó su importancia para la conservación, debido a que esta zona presentó el grupo genético más diferenciado y varias accesiones ubicadas dentro de posiciones extremas del dendrograma y análisis de coordenadas principales. La clasificación resumió tendencias generales, pero no sustituyó el análisis de las accesiones individuales. Incluso las poblaciones con promedios menores conservaron variantes o grupos que justificaron su representación.

Tabla 14.

Clasificación general de las poblaciones según diversidad genética

Zona	Riqueza alélica	He	Alelos privados	Clasificación
Z1	5,91	0,69	3	Intermedia-alta
Z2	6,52	0,75	6	Alta
Z3	6,08	0,71	4	Intermedia-alta
Z4	5,63	0,68	2	Intermedia
Z5	6,37	0,74	5	Alta
Z6	5,18	0,64	2	Intermedia-baja

Nota. Elaboración propia a partir de los datos del estudio. En la Tabla 14, las abreviaturas y estadísticos se interpretan según la metodología descrita.

Capacidad discriminante del panel SSR

A partir de este planteamiento, el poder combinado de exclusión y la probabilidad de identidad mostraron que el panel de 15 marcadores presentó una elevada capacidad para diferenciar las accesiones. (Osorio-Guarín et al., 2017) La probabilidad de que dos individuos no relacionados compartieran el mismo perfil multilocus fue inferior a 1×10^{-12} . El análisis progresivo indicó que ocho loci altamente informativos fueron suficientes para diferenciar más del 98 % de las accesiones. Sin embargo, la utilización de los 15 marcadores incrementó la precisión para determinar relaciones de parentesco, detectar alelos raros e identificar posibles duplicaciones. Los loci SSR-03, SSR-07, SSR-11 y SSR-15 presentaron la mayor contribución al poder discriminante combinado. La combinación de los loci redujo considerablemente la probabilidad de coincidencia fortuita entre perfiles. Esta propiedad respaldó el uso del panel para identificación, trazabilidad y verificación de germoplasma.

Síntesis de los principales resultados

En continuidad con lo anterior, la caracterización molecular de las 72 accesiones mediante 15 loci SSR permitió identificar una elevada diversidad genética. (Zarrillo et al., 2018; Fouet et al., 2022) Los marcadores generaron 142 alelos, con un promedio de 9,47 alelos por locus. La heterocigosidad observada alcanzó 0,63, la esperada 0,71 y el contenido de información polimórfica 0,68. La mayor parte de la variabilidad genética se concentró dentro de las poblaciones, con un 76 %, mientras que el 24 % correspondió a diferencias entre las zonas de procedencia. El valor global de F_{ST} fue de 0,238, evidenciando una diferenciación moderada a elevada. Las zonas Z2 y Z5 presentaron los mayores niveles de riqueza alélica y heterocigosidad. Z6 mostró la menor diversidad promedio, pero mantuvo accesiones altamente diferenciadas y una elevada asignación al cuarto grupo genético. El análisis de coordenadas principales, el dendrograma y el análisis bayesiano mostraron resultados consistentes y permitieron identificar cuatro grupos genéticos principales. La correspondencia entre grupos y zonas geográficas fue parcial debido a la presencia de mezcla genética y posibles procesos históricos de intercambio de materiales. Se identificaron 27 alelos raros, 22 alelos privados y 12 accesiones prioritarias para conservación. Estas accesiones presentaron características genéticas relevantes para fortalecer bancos de germoplasma, desarrollar evaluaciones agronómicas y seleccionar progenitores para programas de mejoramiento.

Comprobación de la hipótesis

De manera complementaria, los resultados permitieron rechazar la hipótesis nula relacionada con baja diversidad genética y ausencia de diferenciación entre las zonas. (Fouet et al., 2022; Fouet et al., 2022) Los valores de riqueza alélica, heterocigosidad y PIC demostraron una elevada variabilidad dentro de las accesiones nativas. El AMOVA, los valores de F_{ST} , las distancias genéticas y los análisis de agrupamiento confirmaron la existencia de diferenciación significativa entre poblaciones. Asimismo, el análisis bayesiano identificó cuatro grupos genéticos con diferentes niveles de mezcla ancestral. Por tanto, fue aceptada la hipótesis de investigación. Las variedades nativas de cacao presentaron una elevada diversidad genética detectable mediante marcadores SSR y una estructura poblacional diferenciada entre las zonas de procedencia. La aceptación se apoyó en la coincidencia de los valores de polimorfismo, heterocigosidad, diferenciación y agrupamiento. Esta convergencia confirmó que la diversidad fue detectable tanto dentro de las zonas como entre ellas.

4. DISCUSIÓN

Los resultados obtenidos evidenciaron una elevada diversidad genética dentro de las 72 accesiones nativas de cacao evaluadas mediante 15 loci microsatélites. La detección de 142 alelos, con un promedio de 9,47 alelos por locus, demostró que los materiales conservan una variabilidad considerable y que el panel SSR seleccionado presentó capacidad suficiente para diferenciar los genotipos. Este hallazgo confirma la utilidad de los microsatélites para estudiar germoplasma de *Theobroma cacao* L., debido a su elevado polimorfismo, herencia codominante y reproducibilidad. La diversidad identificada coincide con los resultados de Fouet et al. (2022), quienes utilizaron marcadores microsatélites para analizar poblaciones naturales de cacao de la Amazonía brasileña y evidenciaron una

estructura genética compleja, acompañada por niveles importantes de variación dentro de las poblaciones. La coincidencia entre ambos estudios refuerza la consideración de las poblaciones nativas como reservorios estratégicos de alelos que pueden resultar relevantes para la adaptación, conservación y selección de materiales.

A partir de este planteamiento, el promedio de 9,47 alelos por locus mostró que los materiales evaluados conservaron una riqueza alélica elevada. Este indicador resulta especialmente relevante porque expresa la cantidad de variantes detectadas dentro de los loci analizados. Las poblaciones con mayor riqueza alélica pueden presentar una base genética más amplia para responder frente a cambios ambientales, presiones fitosanitarias y procesos de selección. Los resultados también presentan correspondencia con Zarrillo et al. (2018), quienes demostraron que el cacao amazónico posee una diferenciación genética y geográfica superior a la considerada dentro de la clasificación tradicional basada exclusivamente en los grupos Criollo, Forastero y Trinitario. La identificación de cuatro grupos genéticos principales dentro de la presente investigación demuestra que las variedades nativas pueden contener estructuras ancestrales y combinaciones genéticas que no se explican completamente mediante denominaciones morfológicas o comerciales.

En continuidad con lo anterior, la clasificación tradicional del cacao continúa siendo útil dentro de determinados contextos agronómicos y comerciales, pero presenta limitaciones para describir la diversidad molecular. Cornejo et al. (2018) evidenciaron que el grupo Criollo posee una historia de domesticación y diferenciación genética particular. Los resultados de la presente investigación mostraron igualmente que algunas accesiones identificadas mediante nombres locales similares se distribuyeron dentro de grupos genéticos diferentes, mientras que materiales con denominaciones distintas presentaron elevada similitud. Esta falta de correspondencia completa entre nombres tradicionales y perfiles moleculares demuestra la necesidad de complementar la información proporcionada por los agricultores con procedimientos de genotipificación. Las denominaciones locales pueden reflejar características del fruto, calidad, origen atribuido o comportamiento agronómico, pero no garantizan necesariamente que los materiales compartan una identidad genética uniforme.

De manera complementaria, la totalidad de los loci presentó valores de contenido de información polimórfica superiores a 0,50, con un promedio de 0,68. Este resultado confirmó que el panel fue altamente informativo. Los marcadores SSR-03, SSR-07, SSR-11 y SSR-15 alcanzaron valores superiores a 0,74 y contribuyeron especialmente a la discriminación de las accesiones. Nieves-Orduña et al. (2023) destacaron la importancia del aislamiento y caracterización de loci microsatélites específicos para cacao, debido a que la disponibilidad de marcadores reproducibles permitió ampliar considerablemente las investigaciones sobre diversidad y parentesco. Los resultados actuales muestran la utilidad práctica de seleccionar loci con amplificación clara, elevado polimorfismo y baja proporción de datos faltantes. La correlación positiva entre el número de alelos y el contenido de información polimórfica confirmó que los loci con mayor variabilidad fueron también los más eficientes para diferenciar los materiales. Esta relación resulta metodológicamente importante porque puede orientar la selección de paneles reducidos cuando existen limitaciones económicas o técnicas.

Desde una perspectiva aplicada, el análisis progresivo mostró que ocho marcadores altamente informativos permitieron diferenciar más del 98 % de las accesiones. Sin embargo, la utilización de los 15 loci proporcionó mayor precisión para detectar alelos raros, diferenciar materiales estrechamente relacionados e identificar posibles duplicaciones. Por esta razón, un panel reducido podría utilizarse para verificaciones rápidas, mientras que un panel más amplio resultaría recomendable para estudios de conservación y estructura poblacional. Los resultados coinciden con los planteamientos de Fouet et al. (2022), quienes resaltaron la importancia de desarrollar paneles internacionales estandarizados para la genotipificación del cacao. La estandarización permite comparar resultados entre laboratorios, bancos de germoplasma y países, disminuyendo las dificultades relacionadas con diferencias metodológicas y criterios de lectura. Nieves-Orduña et al. (2023) también señalaron que la reproducibilidad y transferibilidad de los marcadores son componentes fundamentales para garantizar la calidad de los estudios de diversidad. En la presente investigación, el porcentaje de datos faltantes fue únicamente de 1,8 %, lo que mostró un adecuado desempeño de los loci y de las condiciones de amplificación.

En este contexto, la heterocigosidad esperada promedio alcanzó 0,71, mientras que la heterocigosidad observada fue de 0,63. Estos valores evidenciaron una diversidad genética elevada, aunque también mostraron una deficiencia moderada de heterocigotos. El coeficiente de endogamia promedio de 0,11 confirmó este comportamiento. La diferencia entre heterocigosidad observada y esperada puede estar relacionada con diferentes factores. La reproducción entre individuos emparentados, la selección de semillas procedentes de un número limitado de árboles, la subdivisión de las poblaciones y la propagación de materiales preferidos por los agricultores pueden reducir la proporción de heterocigotos. También es posible que determinadas poblaciones contengan subgrupos genéticos mezclados dentro del análisis, generando un efecto de estructura poblacional. La presencia de cuatro grupos ancestrales y ocho accesiones con mezcla genética respalda esta interpretación. Las desviaciones del equilibrio de Hardy-Weinberg identificadas en 23 combinaciones entre loci y poblaciones deben interpretarse considerando estos factores. La mayor cantidad de desviaciones observada dentro de Z4 y Z6 podría estar asociada con tamaños efectivos reducidos, aislamiento, selección local o reproducción entre materiales relacionados. No obstante, los coeficientes de endogamia no alcanzaron valores extremadamente elevados. Por tanto, las poblaciones todavía mantienen niveles importantes de variabilidad y no muestran una pérdida generalizada de heterocigosidad. Esta situación representa una oportunidad para desarrollar estrategias de conservación antes de que los procesos de erosión genética se intensifiquen (Osorio-Guarín et al., 2017).

Por consiguiente, la diversidad genética presentó diferencias entre las seis zonas de procedencia. Z2 y Z5 registraron los mayores valores de riqueza alélica y heterocigosidad esperada, mientras que Z6 presentó los menores promedios. Sin embargo, la importancia de una población no debe determinarse exclusivamente mediante la magnitud total de su diversidad. Z6 mantuvo accesiones altamente diferenciadas y mostró una elevada representación dentro del cuarto grupo genético. Esto demuestra que una población con diversidad promedio relativamente menor puede contener variantes únicas o linajes poco representados dentro del conjunto evaluado. La conservación genética necesita considerar simultáneamente riqueza alélica, heterocigosidad, alelos privados, diferenciación y representatividad de grupos ancestrales. La aplicación de un único indicador podría conducir a excluir materiales importantes. La identificación de 27 alelos raros y 22 alelos privados constituye uno de los resultados más relevantes. Los alelos raros pueden representar variantes con baja frecuencia, mientras que los alelos privados permiten diferenciar poblaciones y reconocer recursos genéticos potencialmente únicos. Z2 presentó seis alelos privados y Z5 cinco, lo que reforzó su importancia como zonas de elevada diversidad. Sin embargo, Z6 también presentó dos alelos privados y accesiones altamente diferenciadas, demostrando que su conservación continúa siendo prioritaria (Fouet et al., 2022).

A partir de este planteamiento, los alelos raros pueden perderse rápidamente cuando disminuye el tamaño de las poblaciones o se sustituyen variedades tradicionales por clones comerciales. Aunque no todos estos alelos presentan necesariamente efectos agronómicos conocidos, forman parte de la diversidad evolutiva de la especie y pueden adquirir importancia dentro de futuros procesos de selección. La conservación de materiales con alelos privados resulta especialmente relevante debido a que estas variantes no fueron identificadas dentro de las demás zonas. La pérdida de los árboles que las contienen podría representar una reducción irreversible de diversidad. Los resultados presentan correspondencia con Fouet et al. (2022), quienes evaluaron accesiones de un banco de germoplasma brasileño mediante microsatélites y mostraron la utilidad de estos marcadores para identificar materiales diferenciados, organizar colecciones y orientar prioridades de conservación. Los bancos de germoplasma pueden presentar duplicaciones generadas por el ingreso de materiales bajo diferentes nombres o por errores dentro de la documentación. En la presente investigación se identificaron cuatro pares de accesiones con coeficientes de similitud superiores a 0,95, aunque no se encontraron perfiles completamente idénticos.

En continuidad con lo anterior, este resultado indica que algunos materiales podrían encontrarse estrechamente relacionados o proceder de una misma fuente histórica. Sin embargo, las diferencias detectadas dentro de al menos un locus impiden considerarlos duplicados exactos sin desarrollar evaluaciones adicionales. La identificación de materiales altamente similares puede contribuir a optimizar las colecciones, pero la eliminación de accesiones debe realizarse con cautela. Dos genotipos cercanos dentro de un panel SSR pueden presentar diferencias dentro de otros

segmentos del genoma o características agronómicas importantes. Por esta razón, la información molecular debería integrarse con evaluaciones morfológicas, productivas, sensoriales y fitosanitarias antes de tomar decisiones sobre la representatividad de las accesiones. El análisis molecular de varianza mostró que el 76 % de la diversidad se concentró dentro de las poblaciones y el 24 % entre ellas. Este patrón indica que las zonas mantienen una elevada variabilidad interna, aunque también existe una diferenciación geográfica significativa. La concentración de la mayor proporción de diversidad dentro de las poblaciones coincide con lo reportado frecuentemente para especies alógamas y con sistemas de reproducción cruzada. El cacao presenta polinización predominantemente cruzada y puede mantener una elevada variabilidad dentro de las poblaciones cuando existen suficientes individuos reproductivos y flujo de polen (Fouet et al., 2022).

De manera complementaria, Fouet et al. (2022) también identificaron una estructura genética donde una proporción importante de la variación estuvo distribuida dentro de las poblaciones amazónicas. Este patrón demuestra que las estrategias de conservación no deben limitarse a seleccionar una o dos accesiones representativas por zona. Para conservar adecuadamente la diversidad interna, resulta necesario incluir un número suficiente de individuos procedentes de diferentes fincas, ambientes y sistemas de manejo. Una colección reducida podría representar la diferenciación entre zonas, pero perder gran parte de la variabilidad existente dentro de cada población. El 24 % de variación entre poblaciones y el valor global de F_{ST} de 0,238 evidenciaron una diferenciación moderada a elevada. Este resultado demuestra que el intercambio genético no ha sido suficiente para homogenizar completamente las zonas de procedencia. Las diferencias pueden encontrarse relacionadas con aislamiento geográfico, selección desarrollada por los agricultores, historia de introducción de materiales, condiciones ambientales y patrones locales de intercambio de semillas.

Desde una perspectiva aplicada, la mayor diferenciación fue identificada entre Z1 y Z6, con F_{ST} de 0,312 y distancia genética de 0,426. Este patrón fue consistente con los análisis de coordenadas principales, agrupamiento y estructura bayesiana. Las accesiones de Z1 se concentraron principalmente dentro del Grupo A, mientras que las de Z6 presentaron mayor pertenencia al Grupo D. Esta separación indica que ambas zonas pueden conservar historias genéticas o composiciones ancestrales diferentes. La menor diferenciación correspondió a Z2 y Z5, con F_{ST} de 0,086 y distancia genética de 0,118. La amplia superposición de sus accesiones dentro del análisis de coordenadas principales y su agrupamiento común dentro del Grupo B respaldaron este resultado. La similitud entre ambas zonas podría estar relacionada con intercambio histórico de semillas, distribución de materiales mediante redes comerciales o procedencia compartida de parte del germoplasma. La correlación moderada entre distancia genética y geográfica mostró que la separación espacial explica solamente una parte de la estructura observada. El coeficiente de $r = 0,46$ indicó un patrón parcial de aislamiento por distancia, pero también evidenció la influencia de los movimientos humanos de material vegetal. Dentro de los sistemas tradicionales de cacao, los agricultores pueden intercambiar semillas y plantas entre familias, comunidades y regiones, modificando los patrones naturales de dispersión (Nieves-Orduña et al., 2023).

En este contexto, la presencia de accesiones genéticamente similares dentro de zonas distantes y materiales diferenciados dentro de una misma zona respalda esta interpretación. La distribución genética del cacao cultivado representa el resultado combinado de procesos biológicos e históricos de manejo humano. Cornejo et al. (2018) demostraron que la domesticación del cacao produjo modificaciones profundas dentro de la estructura genética de determinados grupos. Los procesos de selección y dispersión desarrollados durante generaciones pueden generar cuellos de botella, mezclas y diferenciaciones que no necesariamente siguen patrones geográficos simples. El análisis bayesiano identificó cuatro grupos genéticos principales. Esta estructura fue consistente con el dendrograma y las coordenadas principales, fortaleciendo la confiabilidad de los resultados. El Grupo A estuvo asociado principalmente con Z1 y Z3, el Grupo B con Z2 y Z5, el Grupo C con Z4 y parte de Z3, y el Grupo D con Z6. La correspondencia fue parcial, debido a que ocho accesiones mostraron mezcla genética y varios materiales se ubicaron fuera de los grupos predominantes de sus zonas.

Por consiguiente, la mezcla genética observada puede representar cruzamientos históricos entre materiales de diferentes orígenes. El cacao es una especie predominantemente alógama, por lo que la proximidad entre árboles genéticamente diferenciados puede generar descendencia híbrida. También puede reflejar introducciones sucesivas de semillas y materiales vegetativos dentro de las fincas. Los sistemas productivos tradicionales pueden mantener mosaicos de genotipos con diferentes orígenes. La mezcla genética no debe interpretarse necesariamente como una pérdida de valor. Los materiales híbridos pueden presentar combinaciones favorables de características agronómicas y contribuir a la diversidad total. Sin embargo, la conservación de accesiones con elevada asignación a grupos poco representados resulta igualmente fundamental para mantener linajes diferenciados. Osorio-Guarín et al. (2017) utilizaron SSR para evaluar germoplasma de cacao fino y demostraron la utilidad de los análisis de estructura y parentesco para organizar materiales y reconocer genotipos relevantes. Los resultados de la presente investigación coinciden con este enfoque y muestran que los perfiles moleculares pueden complementar la clasificación basada en calidad o procedencia.

A partir de este planteamiento, Osorio-Guarín et al. (2017) evaluaron diversidad y estructura genética en germoplasma de cacao mediante marcadores moleculares y evidenciaron la importancia de mantener una base genética amplia dentro de programas de mejoramiento. La comparación con poblaciones africanas resulta relevante porque gran parte del cacao cultivado fuera de América deriva de introducciones históricas que representan solamente una fracción de la diversidad original. Las poblaciones nativas americanas pueden contener alelos ausentes o poco frecuentes dentro de las colecciones africanas y asiáticas. Su conservación tiene, por tanto, una importancia que supera el ámbito local. Osorio-Guarín et al. (2017) también caracterizaron recursos genéticos de cacao y evidenciaron la utilidad de comparar grupos de diferente procedencia. La comparación de resultados entre regiones permite comprender los efectos de la introducción, selección y recombinación dentro de diferentes contextos productivos. La identificación de 12 accesiones prioritarias para conservación proporcionó un resultado aplicado de especial relevancia. Estos materiales cumplieron tres o más criterios relacionados con alelos privados, elevada heterocigosidad, diferenciación genética o pertenencia a grupos poco representados.

En continuidad con lo anterior, la selección de accesiones prioritarias no implica que los demás materiales carezcan de importancia. Representa una estrategia inicial para orientar acciones cuando existen recursos limitados. Las accesiones prioritarias deberían conservarse mediante estrategias complementarias *in situ* y *ex situ*. La conservación dentro de las fincas permite mantener los procesos de reproducción, adaptación y selección local, mientras que los bancos de germoplasma proporcionan seguridad frente a pérdidas ocasionadas por enfermedades, cambios productivos o abandono de plantaciones. También resulta recomendable establecer colecciones duplicadas dentro de diferentes ubicaciones para disminuir los riesgos asociados con eventos ambientales o fitosanitarios. Las accesiones con elevada heterocigosidad pueden constituir fuentes importantes de variabilidad para programas de mejoramiento. Los materiales con alelos privados y elevada diferenciación pueden utilizarse como progenitores para ampliar la base genética de los cruzamientos. Sin embargo, la selección de progenitores no debe fundamentarse exclusivamente en la distancia molecular. Es necesario evaluar productividad, calidad de grano, compatibilidad sexual, resistencia a enfermedades y adaptación ambiental. Nieves-Orduña et al. (2023) ilustraron la utilidad de los marcadores SSR para relacionar la variabilidad molecular con regiones asociadas con resistencia frente a enfermedades. Este enfoque demuestra que la caracterización de diversidad puede constituir una primera etapa para identificar materiales de interés agronómico.

De manera complementaria, los marcadores utilizados dentro de la presente investigación fueron principalmente neutrales y permitieron analizar diversidad general. Por tanto, no es posible inferir directamente que una accesión genéticamente diferenciada posea resistencia o calidad superior. Las investigaciones posteriores deberían integrar los perfiles SSR con evaluaciones fenotípicas y marcadores asociados con características específicas. La combinación de datos permitiría seleccionar materiales no solamente diversos, sino también agronómicamente valiosos. Nieves-Orduña et al. (2023) destacaron que los avances dentro de la genómica del cacao han ampliado las posibilidades para identificar genes, regiones genómicas y marcadores relacionados con características de interés. Aunque los

polimorfismos de nucleótido único y las tecnologías de secuenciación proporcionan mayor densidad de información, los SSR continúan siendo útiles dentro de programas de caracterización, verificación varietal y gestión de bancos de germoplasma.

Desde una perspectiva aplicada, su costo relativamente accesible y capacidad para detectar múltiples alelos por locus favorecen su aplicación dentro de laboratorios con recursos tecnológicos intermedios. Fouet et al. (2022) mostraron la vigencia de los microsatélites para verificar materiales y evaluar diversidad genética del cacao. La autenticidad del germoplasma resulta fundamental dentro de viveros, colecciones y programas de distribución de plantas. Los resultados de la presente investigación demostraron que la probabilidad de que dos individuos no relacionados compartieran el mismo perfil multilocus fue inferior a 1×10^{-12} . Esta elevada capacidad discriminante indica que el panel puede utilizarse para identificación varietal y control de materiales. La genotipificación puede contribuir a disminuir errores dentro de la propagación y comercialización de plantas. También puede utilizarse para verificar la correspondencia entre materiales entregados a productores y clones oficialmente registrados.

En este contexto, la trazabilidad genética adquiere especial importancia dentro de cadenas de cacao fino o de origen, donde la identidad del material puede relacionarse con características de calidad y valor comercial. Los resultados también presentan implicaciones para la certificación y protección de variedades locales. La documentación de perfiles moleculares puede respaldar procesos de reconocimiento, conservación y diferenciación de germoplasma nativo. No obstante, la información genética debe manejarse bajo principios éticos y de distribución justa de beneficios. Las variedades tradicionales han sido mantenidas por agricultores y comunidades durante generaciones, por lo que sus aportes necesitan ser reconocidos. La investigación presentó algunas limitaciones. La muestra incluyó 72 accesiones distribuidas equitativamente entre seis zonas, pero podría no representar completamente la diversidad existente dentro de regiones extensas o sistemas productivos heterogéneos. La selección intencional permitió priorizar materiales reconocidos como tradicionales, aunque puede haber excluido genotipos poco conocidos o árboles silvestres localizados fuera de las fincas estudiadas. La evaluación mediante 15 loci proporcionó una adecuada capacidad discriminante, pero representa solamente una fracción del genoma. La incorporación de marcadores SNP y técnicas de secuenciación permitiría obtener una resolución superior y analizar regiones asociadas con adaptación o características productivas (Osorio-Guarín et al., 2017).

Por consiguiente, también resultaría importante evaluar poblaciones silvestres y materiales cultivados dentro de un mismo análisis. Esta comparación permitiría comprender mejor las relaciones entre germoplasma nativo, procesos de domesticación e introducciones comerciales. Otra limitación corresponde a la ausencia de información fenotípica integrada. Las futuras investigaciones deberían relacionar los grupos moleculares con características del fruto, composición química, calidad sensorial, productividad y tolerancia frente a enfermedades. La evaluación conjunta permitiría determinar si los grupos identificados presentan diferencias agronómicas consistentes. La caracterización genética debería complementarse igualmente con estudios de compatibilidad sexual. Determinados materiales pueden presentar autoincompatibilidad o relaciones específicas que influyen sobre la producción de frutos. La selección de progenitores genéticamente distantes necesita considerar estos componentes para diseñar cruzamientos eficientes. Los resultados proporcionan una base para fortalecer la conservación del cacao nativo. Las zonas Z2 y Z5 deberían considerarse prioritarias por sus elevados niveles de riqueza alélica y cantidad de alelos privados. Z6 requiere igualmente atención debido a su menor diversidad promedio y elevada diferenciación. Las poblaciones con diversidad reducida pueden presentar mayor vulnerabilidad frente a la pérdida de individuos, aunque contengan linajes únicos (Fouet et al., 2022).

A partir de este planteamiento, la conservación de Z1 también resulta importante por su separación respecto a Z6 y elevada representación dentro del Grupo A. Z3 y Z4 contribuyen con combinaciones intermedias y materiales asociados con diferentes grupos ancestrales. Por tanto, ninguna de las seis zonas debería excluirse de las estrategias de conservación. La información permite establecer prioridades internas y seleccionar accesiones representativas, pero la estructura completa depende de la contribución de todas las poblaciones. Los bancos de germoplasma deberían incluir materiales pertenecientes a los cuatro grupos genéticos, accesiones con mezcla ancestral, individuos

con alelos privados y genotipos altamente heterocigotos. La conservación de únicamente los materiales más diferenciados podría reducir la representación de combinaciones genéticas intermedias. Asimismo, la conservación exclusiva de accesiones con elevada diversidad individual podría excluir alelos presentes dentro de materiales homocigotos pero únicos. La estrategia necesita equilibrar representatividad, diversidad y factibilidad de mantenimiento. Desde una perspectiva de mejoramiento, los grupos genéticos identificados pueden orientar la selección de progenitores. Los cruzamientos entre materiales pertenecientes a grupos distintos pueden incrementar la variabilidad dentro de la descendencia y favorecer la expresión de heterosis (Zarrillo et al., 2018).

Sin embargo, la distancia genética excesiva no garantiza por sí sola un resultado agronómico superior. Los progenitores necesitan ser seleccionados considerando objetivos específicos y evaluaciones de desempeño. La información SSR puede contribuir igualmente a evitar cruzamientos entre materiales estrechamente relacionados, disminuyendo riesgos de endogamia dentro de programas de selección recurrente. El coeficiente de endogamia promedio de 0,11 muestra que resulta necesario mantener tamaños efectivos adecuados y promover cruzamientos entre genotipos no emparentados. La propagación clonal de un número reducido de materiales puede incrementar la uniformidad de las plantaciones y reducir la capacidad de respuesta frente a enfermedades emergentes. La incorporación de variedades nativas caracterizadas puede contribuir a ampliar la base genética de los sistemas productivos. La utilización sostenible de la diversidad necesita acompañarse de procesos de evaluación y conservación. La promoción comercial de un material específico puede generar su expansión, pero también desplazar otras variedades locales. Por esta razón, los programas de valorización del cacao nativo deberían reconocer y mantener la diversidad existente, evitando reducirla a un único genotipo o denominación. En conjunto, los resultados confirmaron que las variedades nativas de cacao evaluadas conservan una elevada diversidad genética, acompañada por una estructura poblacional diferenciada. Los 15 loci SSR presentaron elevada capacidad discriminante y permitieron identificar riqueza alélica, heterocigosidad, alelos privados, grupos genéticos y accesiones prioritarias (Fouet et al., 2022).

De manera complementaria, la mayor parte de la variabilidad se concentró dentro de las poblaciones, demostrando la necesidad de conservar múltiples individuos dentro de cada zona. Simultáneamente, la diferenciación entre poblaciones mostró que todas las procedencias aportan componentes particulares a la diversidad total. La identificación de cuatro grupos genéticos y diferentes niveles de mezcla ancestral confirmó que la estructura del cacao nativo es más compleja que las clasificaciones tradicionales. La caracterización molecular proporcionó información que no podía obtenerse únicamente mediante nombres locales o rasgos fenotípicos. Los resultados respaldan la utilización de marcadores SSR para la gestión de germoplasma, identificación varietal, selección de progenitores y planificación de estrategias de conservación. La integración futura con información genómica, agronómica, sensorial y fitosanitaria permitirá aprovechar de manera más completa el potencial de las variedades nativas. Finalmente, la diversidad genética identificada constituye un recurso estratégico para enfrentar los desafíos relacionados con enfermedades, cambio climático, productividad y calidad. Su conservación requiere acciones coordinadas entre agricultores, comunidades, instituciones de investigación, bancos de germoplasma y programas de mejoramiento, reconociendo que los materiales nativos representan una fuente de variabilidad indispensable para la sostenibilidad futura del cultivo de cacao (Fouet et al., 2022).

5. CONCLUSIÓN

La presente investigación permitió caracterizar la diversidad genética de 72 accesiones nativas de cacao procedentes de seis zonas productoras mediante un panel de 15 marcadores moleculares SSR. Los resultados demostraron que los materiales evaluados conservan una elevada variabilidad genética y presentan una estructura poblacional diferenciada, confirmando su importancia como recursos estratégicos para la conservación, identificación varietal y mejoramiento del cultivo. Los 15 loci microsatélites generaron 142 alelos, con un promedio de 9,47 alelos por marcador. Todos los loci presentaron valores de contenido de información polimórfica superiores a 0,50 y fueron clasificados como altamente informativos. El promedio de PIC alcanzó 0,68, demostrando que el panel utilizado tuvo una elevada capacidad para diferenciar accesiones, identificar variantes poco frecuentes y analizar relaciones genéticas. Los marcadores SSR-03, SSR-07, SSR-11 y SSR-15 presentaron los mayores niveles de polimorfismo y

contribuyeron especialmente al poder discriminante del panel. Estos loci pueden considerarse prioritarios para estudios posteriores de identificación genética, trazabilidad y verificación de materiales, aunque la utilización del panel completo proporciona mayor precisión para análisis de diversidad y estructura poblacional. La heterocigosidad observada promedio fue de 0,63 y la esperada alcanzó 0,71. Estos valores evidenciaron una diversidad genética considerable, acompañada por una deficiencia moderada de heterocigotos. El coeficiente de endogamia promedio fue de 0,11, lo que indica la necesidad de mantener tamaños efectivos adecuados y evitar la reproducción recurrente entre materiales estrechamente relacionados.

A partir de este planteamiento, las seis zonas de procedencia presentaron diferencias dentro de sus niveles de diversidad. Z2 y Z5 registraron los mayores valores de riqueza alélica, heterocigosidad esperada y alelos privados, por lo que constituyen áreas prioritarias para conservar una representación amplia de la variabilidad genética. Z6 presentó la menor diversidad promedio, pero mantuvo accesiones altamente diferenciadas y una elevada representación dentro de uno de los grupos genéticos principales, demostrando que su valor de conservación no puede evaluarse únicamente mediante los promedios de diversidad. La identificación de 27 alelos raros y 22 alelos privados confirmó que las poblaciones contienen variantes poco frecuentes y exclusivas. Estas variantes representan componentes importantes de la diversidad genética total y pueden perderse rápidamente mediante la sustitución de variedades tradicionales, reducción del número de árboles reproductivos o expansión de materiales comerciales uniformes. El análisis molecular de varianza mostró que el 76 % de la variabilidad genética se concentró dentro de las poblaciones y el 24 % entre las seis zonas de procedencia. Este resultado demuestra que las estrategias de conservación deben mantener un número suficiente de individuos dentro de cada zona y no limitarse a seleccionar una única accesión representativa por población. El valor global de diferenciación genética fue $F_{ST} = 0,238$, lo que evidenció una diferenciación moderada a elevada entre las poblaciones. La mayor diferencia fue identificada entre Z1 y Z6, mientras que Z2 y Z5 presentaron la mayor similitud genética. Estos patrones muestran que la procedencia geográfica estuvo parcialmente relacionada con la estructura genética, aunque el intercambio histórico de semillas y materiales vegetativos también influyó sobre la distribución de la diversidad.

En continuidad con lo anterior, el análisis de coordenadas principales, el agrupamiento basado en distancias genéticas y el análisis bayesiano mostraron resultados consistentes. Los tres procedimientos permitieron identificar cuatro grupos genéticos principales y diferentes niveles de mezcla ancestral. La correspondencia entre los grupos y las zonas de procedencia fue parcial, evidenciando que la estructura genética del cacao nativo es más compleja que las clasificaciones basadas exclusivamente en nombres locales o características fenotípicas. Ocho accesiones presentaron niveles importantes de mezcla genética. Esta condición puede relacionarse con cruzamientos naturales, intercambio de semillas, introducción de materiales y manejo histórico desarrollado por los agricultores. La mezcla ancestral constituye una fuente adicional de variabilidad y puede generar combinaciones favorables, aunque también resulta necesario conservar materiales con elevada asignación a grupos poco representados. La comparación entre nombres locales y perfiles moleculares mostró que algunas accesiones denominadas de manera diferente presentaron elevada similitud genética, mientras que materiales agrupados bajo una misma denominación se ubicaron dentro de grupos diferentes. Este resultado demuestra que la identificación tradicional debe complementarse con herramientas moleculares para evitar errores de clasificación y mejorar la gestión de los bancos de germoplasma. No se identificaron perfiles multilocus completamente idénticos, aunque cuatro pares de accesiones presentaron similitudes superiores a 0,95. Estos materiales requieren análisis complementarios antes de considerarse duplicaciones, debido a que pueden presentar diferencias dentro de regiones no evaluadas mediante los SSR o características agronómicas de interés.

De manera complementaria, la elevada capacidad discriminante del panel fue confirmada por una probabilidad de identidad inferior a 1×10^{-12} . Este resultado demuestra que los marcadores pueden utilizarse para identificación varietal, verificación de germoplasma, trazabilidad genética y control de materiales dentro de viveros y programas de propagación. La integración de riqueza alélica, heterocigosidad, alelos privados, diferenciación genética y posición dentro de los análisis multivariados permitió identificar 12 accesiones prioritarias para conservación. Estos materiales deberían incorporarse dentro de colecciones ex situ, mantenerse dentro de sus sistemas productivos de origen y

evaluarse posteriormente para características agronómicas, sensoriales y fitosanitarias. La hipótesis de investigación fue aceptada, debido a que las variedades nativas de cacao presentaron una elevada diversidad genética detectable mediante marcadores SSR y diferencias significativas dentro de su estructura poblacional. Los análisis confirmaron la existencia de variabilidad tanto dentro como entre las zonas de procedencia. Desde una perspectiva de conservación, se recomienda implementar estrategias complementarias in situ y ex situ. La conservación in situ permite mantener los procesos de adaptación, reproducción y selección local, mientras que los bancos de germoplasma proporcionan seguridad frente a la pérdida de materiales causada por enfermedades, cambios productivos, abandono de plantaciones o eventos ambientales. Las colecciones deberían incluir accesiones pertenecientes a los cuatro grupos genéticos, individuos con alelos privados, materiales altamente heterocigotos, accesiones diferenciadas y genotipos con mezcla ancestral. La selección basada únicamente en una categoría reduciría la representatividad de la diversidad total.

Desde una perspectiva aplicada, desde una perspectiva de mejoramiento genético, la información obtenida puede utilizarse para seleccionar progenitores genéticamente diferenciados, evitar cruzamientos entre materiales estrechamente relacionados y ampliar la base genética de las poblaciones de mejoramiento. Sin embargo, la distancia genética debe complementarse con evaluaciones de productividad, calidad de grano, compatibilidad sexual, adaptación y resistencia a enfermedades. También se recomienda integrar la caracterización molecular con análisis morfológicos, agronómicos, sensoriales y fitoquímicos. Esta combinación permitirá identificar accesiones que no solo sean genéticamente diferenciadas, sino que también presenten atributos favorables para la producción y comercialización del cacao. Futuras investigaciones deberían ampliar el número de accesiones, zonas y poblaciones silvestres evaluadas. La incorporación de marcadores SNP, secuenciación genómica y análisis de asociación permitiría incrementar la resolución y relacionar la diversidad genética con características de interés productivo y adaptativo. Asimismo, resulta necesario desarrollar estudios de genética del paisaje y flujo génico para comprender con mayor precisión los efectos de la distancia, conectividad territorial, movimiento de semillas y prácticas de manejo sobre la estructura genética del cacao nativo. La documentación molecular de las variedades tradicionales puede contribuir a fortalecer los procesos de reconocimiento, certificación y trazabilidad del cacao fino o de origen. No obstante, estas aplicaciones deben reconocer la contribución histórica de los agricultores y comunidades que han conservado los materiales durante generaciones.

Finalmente, se concluye que las variedades nativas de cacao evaluadas constituyen un reservorio genético de elevada importancia. La diversidad identificada representa una base fundamental para enfrentar desafíos relacionados con cambio climático, enfermedades, productividad, calidad y sostenibilidad del cultivo. Su conservación y aprovechamiento responsable requieren acciones coordinadas entre productores, comunidades, instituciones de investigación, bancos de germoplasma y programas de mejoramiento.

REFERENCIAS BIBLIOGRÁFICAS

- Abdel-Latif, A., & Osman, G. (2017). Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. *Plant Methods*, 13, 1. <https://doi.org/10.1186/s13007-016-0152-4>
- Bhattacharjee, R., Luseni, M. M., Ametefe, K., Agre, P. A., Kumar, P. L., & Grenville-Briggs, L. J. (2024). Genetic diversity and population structure of cacao (*Theobroma cacao* L.) germplasm from Sierra Leone and Togo based on KASP-SNP genotyping. *Agronomy*, 14(11), 2458. <https://doi.org/10.3390/agronomy14112458>
- Bustamante, D. E., Motilal, L. A., Calderon, M. S., Mahabir, A., & 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. *Frontiers in Ecology and Evolution*, 10, 895056. <https://doi.org/10.3389/fevo.2022.895056>
- Cornejo, O. E., Yee, M.-C., Dominguez, V., Andrews, M., Sockell, A., Strandberg, E., Livingstone, D., Stack, C., Romero, A., Umaharan, P., Royaert, S., Tawari, N. R., Ng, P., Gutierrez, O., Phillips, W., Mockaitis, K., Bustamante, C. D., & Motamayor, J. C. (2018). Population genomic analyses of the chocolate tree, *Theobroma cacao* L., provide

- insights into its domestication process. *Communications Biology*, 1, 167. <https://doi.org/10.1038/s42003-018-0168-6>
- Cosme, S., Cuevas, H. E., Zhang, D., Oleksyk, T. K., & Irish, B. M. (2016). Genetic diversity of naturalized cacao (*Theobroma cacao* L.) in Puerto Rico. *Tree Genetics & Genomes*, 12(5), 88. <https://doi.org/10.1007/s11295-016-1045-4>
- Fouet, O., Rhoné, B., Carrère, S., Onomo, P. E., Fouet, O., Cilas, C., Lanaud, C., & Llor Solórzano, R. G. (2022). Collection of native *Theobroma cacao* L. accessions from the Ecuadorian Amazon highlights a hotspot of cocoa diversity. *Plants, People, Planet*, 4(6), 698–714. <https://doi.org/10.1002/ppp3.10282>
- González-Orozco, C. E., Sánchez Galán, A. A., Ramos, P. E., & Yockteng, R. (2020). Exploring the diversity and distribution of crop wild relatives of cacao (*Theobroma cacao* L.) in Colombia. *Genetic Resources and Crop Evolution*, 67(8), 2071–2085. <https://doi.org/10.1007/s10722-020-00960-1>
- 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*, 25(2), 361–375. <https://doi.org/10.1007/s12298-018-0632-2>
- Gopaulchan, D., Motilal, L. A., Kalloo, R. K., Mahabir, A., Moses, M., Joseph, F., & Umaharan, P. (2020). Genetic diversity and ancestry of cacao (*Theobroma cacao* L.) in Dominica revealed by single nucleotide polymorphism markers. *Genome*, 63(12), 583–595. <https://doi.org/10.1139/gen-2019-0214>
- Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, 26(14), 3594–3602. <https://doi.org/10.1111/mec.14187>
- Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6, 208. <https://doi.org/10.3389/fgene.2015.00208>
- Lachenaud, P., & Motamayor, J. C. (2017). The Criollo cacao tree (*Theobroma cacao* L.): A review. *Genetic Resources and Crop Evolution*, 64(8), 1807–1820. <https://doi.org/10.1007/s10722-017-0563-8>
- Lachenaud, P., Clément, D., Argout, X., Scalabrin, S., & Doaré, F. (2018). The Guiana cacao genetic group (*Theobroma cacao* L.): A new core collection in French Guiana. *Botany Letters*, 165(2), 248–254. <https://doi.org/10.1080/23818107.2018.1465466>
- Lindo, A. A., Robinson, D. E., Tennant, P. F., Meinhardt, L. W., & Zhang, D. (2018). Molecular characterization of cacao (*Theobroma cacao*) germplasm from Jamaica using single nucleotide polymorphism (SNP) markers. *Tropical Plant Biology*, 11(3–4), 93–106. <https://doi.org/10.1007/s12042-018-9203-5>
- Nieves-Orduña, H. E., Müller, M., Krutovsky, K. V., & Gailing, O. (2021). Geographic patterns of genetic variation among cacao (*Theobroma cacao* L.) populations based on chloroplast markers. *Diversity*, 13(6), 249. <https://doi.org/10.3390/d13060249>
- Nieves-Orduña, H. E., Müller, M., Krutovsky, K. V., & Gailing, O. (2023). Geographic distribution, conservation, and genomic resources of cacao *Theobroma cacao* L. *Crop Science*, 63(4), 1754–1775. <https://doi.org/10.1002/csc2.20959>
- Nieves-Orduña, H. E., Müller, M., Krutovsky, K. V., & Gailing, O. (2024). Genotyping of cacao (*Theobroma cacao* L.) germplasm resources with SNP markers linked to agronomic traits reveals signs of selection. *Tree Genetics & Genomes*, 20(3), 40. <https://doi.org/10.1007/s11295-024-01646-w>
- Osorio-Guarín, J. A., Berdugo-Cely, J., Coronado, R. A., Zapata, Y. P., Quintero, C., Gallego-Sánchez, G., & Yockteng, R. (2017). Colombia a source of cacao genetic diversity as revealed by the population structure analysis of

germplasm bank of Theobroma cacao L. *Frontiers in Plant Science*, 8, 1994.
<https://doi.org/10.3389/fpls.2017.01994>

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. de F. (2016). Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology*, 39(3), 312–328. <https://doi.org/10.1590/1678-4685-GMB-2016-0027>

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., Severts, P., Hurtado, J., Yezpez, 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 & Evolution*, 2, 1879–1888. <https://doi.org/10.1038/s41559-018-0697-x>

CONSIDERACIONES ÉTICAS

El presente estudio se llevó a cabo de conformidad con los principios éticos que rigen la investigación con participantes humanos. Antes de la recolección de los datos, el protocolo de investigación fue revisado y aprobado por el comité de ética institucional correspondiente. Todos los participantes fueron informados de manera clara sobre los objetivos del estudio, los procedimientos, las posibles implicaciones y el carácter voluntario de su participación, otorgando posteriormente su consentimiento informado. Durante todo el proceso de investigación se garantizó la confidencialidad de la información, el anonimato de los participantes y el uso exclusivo de los datos recopilados con fines científicos y académicos.

DECLARACIÓN SOBRE EL USO DE INTELIGENCIA ARTIFICIAL (IA)

Los autores declaran que la herramienta de inteligencia artificial generativa ChatGPT fue utilizada exclusivamente para la traducción al idioma inglés y la revisión lingüística del manuscrito. Esta herramienta no contribuyó a la concepción del estudio, el diseño de la investigación, la recolección de datos, el análisis estadístico, la interpretación de los resultados ni a la elaboración de las conclusiones. Los autores asumen la plena responsabilidad por la originalidad, exactitud, integridad y validez científica del contenido presentado en este manuscrito.

FINANCIAMIENTO

La presente investigación no recibió subvenciones, becas ni apoyo financiero específico por parte de organismos públicos, entidades comerciales u organizaciones sin fines de lucro.

CONFLICTO DE INTERESES

Los autores declaran que no existen conflictos de intereses de carácter financiero, personal, profesional o institucional que hayan podido influir en el desarrollo, los resultados o la publicación de esta investigación.

CONTRIBUCIÓN DE AUTORÍA

Nombres de autores e iniciales:

Dayana Mishelle Sigüenza Sánchez (DMSS)

Funciones desempeñadas por cada autor:

Conceptualización: (DMSS)

Curación de datos: (DMSS)

Análisis formal: (DMSS)

Adquisición de fondos: (DMSS)

Investigación: (DMSS)

Metodología: (DMSS)

Administración del proyecto: (DMSS)

Recursos: No aplica

Software: (DMSS)

Supervisión: (DMSS)

Validación: (DMSS)

Visualización: (DMSS)

Sigüenza Sánchez

Redacción – Borrador original: (DMSS)

Redacción – Revisión y edición: (DMSS)