

**GENOTIPIFICACION DE CACAO CRIOLLO COLOMBIANO USANDO  
MARCADORES DE ADN TIPO MICROSATELITES SSR**

**JESSIKA ANDREA HERNANDEZ GARCIA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE QUÍMICA  
MAESTRIA EN QUÍMICA  
BUCARAMANGA**

**2013**

**GENOTIPIFICACION DE CACAO CRIOLLO COLOMBIANO USANDO  
MARCADORES DE ADN TIPO MICROSATELITES SSR**

**JESSIKA ANDREA HERNANDEZ GARCIA**  
Trabajo de investigación para optar el título de:  
**MAGISTER EN QUÍMICA**

**DIRECTOR**  
**HERMINSUL CANO CALLE Ph.D**  
Escuela de Química

**CODIRECTOR**  
**ROBERTO CORONADO Ing. Agrónomo**  
Corporación Colombiana de Investigaciones Agropecuarias CORPOICA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER**  
**FACULTAD DE CIENCIAS**  
**ESCUELA DE QUÍMICA**  
**MAESTRIA EN QUIMICA**  
**BUCARAMANGA**

**2013**

## **DEDICATORIA**

**A Dios por darme la fortaleza para soportar y sobrellevar cada una de las adversidades y finalmente darme este triunfo.**

*A mis padres porque a pesar de la distancia siempre me dieron su apoyo incondicional, consejos y buena energía sin ellos no lo habría logrado.*

*A mis hermanas y sobrinos por ser el motor que mueve mi vida, la razón por la que cada día quiero ser mejor.*

*A la Familia Herrera por adoptarme y darme la oportunidad de tener unos hermanos maravillosos y una nueva mamá.*

***Jessika Andrea Hernández García***

## **AGRADECIMIENTOS**

A la Corporación Colombiana de Investigaciones Agropecuarias CORPOICA, por toda la colaboración prestada en la recolección de muestras y financiación de análisis, en especial al Dr. Roberto Coronado de la E.E la Suiza y al Dr. Víctor Núñez del laboratorio de Biotecnología en la sede Tibaitatá en Bogotá.

Al Dr. Herminsul cano por su dirección en esta investigación y al Grupo de Investigación GIBIM y GIFTEX.

A los calificadores por sus sabios consejos y la tomarse el tiempo de evaluar la investigación con el fin de aportarme conocimientos para mi formación.

A la Federación Nacional de Cacaoteros y en especial al Investigador Darwin Martínez por todo el apoyo recibido en el procesamiento de muestras y análisis.

Al laboratorio de Biotecnología de la Universidad Nacional en la escuela de agronomía y al Dr. Oscar Oliveros por darme la oportunidad de realizar los análisis moleculares en su grupo de investigación.

Al laboratorio de Genética de la Universidad Industrial de Santander por realizar el análisis de electroforesis capilar y en especial a las Dra. Adriana Gil y Adriana Castillo por su paciencia y consejos para el análisis.

A Viviana Chacón, primero por su colaboración en toda la parte de interpretación y análisis de los resultados y segundo por su valiosa amistad.

A Johanna Flórez, por cada una de las madrugadas y noches de traspasado ayudándome a extraer ADN, por su colaboración en la revisión del documento y por ser más que una compañera de trabajo una valiosa amiga.

A los amigos que hice gracias a la maestría, José Roper, Jenny Rodríguez, Ángela Vásquez, Andrés Escorcía, Christian castro, Luis Cortes. Por brindarme su amistad incondicional y apoyarme en esta etapa de mi vida.

A José Luis Pinto por todo el apoyo incondicional porque cada vez que quería desistir siempre tuvo una palabra de aliento para no dejarme caer.

A todas las personas que me apoyaron de una u otra manera a alcanzar este triunfo.

## TABLA DE CONTENIDO

	<b>Pág.</b>
INTRODUCCIÓN.....	14
OBJETIVOS .....	16
OBJETIVO GENERAL.....	16
OBJETIVO ESPECÍFICO .....	16
1. MARCO TEÓRICO Y ESTADO DEL ARTE .....	18
1.1 ORIGEN DEL CACAO.....	18
1.2 TIPOS DE CACAOS.....	20
1.2.1 Grupo de cacao criollo .....	20
1.2.2 Grupo forastero.....	21
1.2.3 Grupo de cacaos trinitarios .....	22
1.3 CARACTERÍSTICAS PRINCIPALES DEL CACAO CRIOLLO. ....	22
1.4 PROPAGACIÓN DEL CACAO. ....	26
1.5 COLECCIONES DE GERMOPLASMA.....	28
1.6 MARCADORES MOLECULARES.....	30
1.6.1 Microsatélites o SSR (Secuencias Simples Repetidas).....	32
1.7 ELECTROFORESIS CAPILAR .....	37
1.8 ANÁLISIS ESTADÍSTICO. ....	39
1.8.2 Número de alelos por locus .....	40
1.8.3 Heterocigocidad esperada (He) .....	40
1.8.4 Heterocigocidad observada (Ho).....	41
1.8.5 Índice de contenido polimórfico PIC .....	41
1.8.6 Test de Hardy Weinberg (HW).....	41
1.8.7 Estimación de las distancias genéticas .....	42
1.8.8 Estimación de endogamia.....	42
1.8.9 Probabilidad de exclusión (PE) .....	42
1.8.10 Análisis de varianza molecular (AMOVA).....	43

2. MATERIALES Y MÉTODOS.....	44
2.1 TOMA DE MUESTRAS .....	44
2.2 EXTRACCIÓN Y CUANTIFICACIÓN DE ADN.....	44
2.3 ANÁLISIS DE MICROSATÉLITES SSR. ....	45
2.4 ELECTROFORESIS CAPILAR .....	48
2.5 ANÁLISIS DE DATOS MOLECULARES.....	49
3. RESULTADOS Y ANÁLISIS .....	50
3.1 EXTRACCIÓN Y CUANTIFICACIÓN DE ADN.....	50
3.2 ANÁLISIS DE MICROSATÉLITES (SSR) .....	52
3.3 ANÁLISIS DE DATOS MOLECULARES.....	58
3.4 ANÁLISIS DE COMPONENTES PRINCIPALES .....	66
3.5 ANÁLISIS DE DISTANCIAS GENÉTICAS.....	67
3.5.1 DISTANCIAS ENTRE POBLACIONES. ....	67
3.5.2 Comparación con la base de datos ICGD .....	69
3.5.3 Distancias entre Individuos. ....	69
4. CONCLUSIONES .....	73
5. RECOMENDACIONES.....	75
BIBLIOGRAFICAS.....	82

## LISTA DE FIGURAS

	<b>Pág.</b>
Figura 1: Distribución mundial del cultivo de cacao	17
Figura 2: Cotiledones de cacao criollo.	23
Figura 3: Mazorca de cacao criollo	24
Figura 4: Producción nacional de cacao en toneladas por Departamentos en Colombia 2002-2010	26
Figura 5. Microsatélites, ejemplo de un di-nucleótido A-C(n).	31
Figura 6: Size standard LIZ 500	46
Figura 7: Electroforesis de extracción de ADN con el kit Quiagen.	47
Figura 8:electroforesis de cuantificación de ADN	48
Figura 9:Electroforesis vertical en poliacrilamida de productos de PCR.	50
Figura 10: Electroforesis en agarosa de diferentes productos de PCR usando condiciones en multiple.	51
Figura 11: Rango de alelos (bp) esperado de cada microsatélite con el color del fluorocromo con que se marcó.	52
Figura 12: Tamaño de alelos en pares de bases detectado por los 14 microsatélites evaluados para la accesión CR 46	53
Figura 13: Grafico de análisis de componentes principales	63
Figura14: Árbol de distancias de las 5 poblaciones DSA construido con el método UPGMA.	65
Figura15: Árbol de distancias individuales DSA construido con el método UPGMA	68

## LISTA DE TABLAS

	<b>Pag.</b>
Tabla 1: Diferencia entre el cacao criollo y forastero	19
Tabla 2: Países productores de cacao criollo.	21
Tabla 3: Valores de la relación teobromina/ cafeína en los diferentes grupos de cacao conocidos	22
Tabla 4: Producción de cacao en grano (miles de toneladas)	25
Tabla 5: Nombre, secuencias, temperaturas de anillamiento y marcaje con fluorocromos de cada uno de los microsatélites usados	45
Tabla 6: Resultados obtenidos de la cuantificación de ADN.	49
Tabla 7: Alelos identificados (bp) de los microsatélites amplificados en las diferentes accesiones de cacao criollo evaluadas.	55
Tabla 8: Parámetros genéticos para los marcadores SSRs.	58
Tabla 9: Probabilidad de Rst asumiendo el modelo de mutación stepwise	61
Tabla 10: Poder de exclusión combinado para las poblaciones de Guajira, Cesar y Magdalena	62
Tabla 11: Poder de exclusión combinado locus por locus para cada población	62
Tabla 12: Comparación entre los alelos de cacao criollo Colombiano y 2 originarios de Costa Rica.	69

## LISTA DE ANEXOS

	<b>Pág.</b>
Anexo A: Cultivares de cacao criollo recolectados en el Banco Nacional de Germoplasma	86
Anexo B: Protocolo de Extracción de ADN	89
Anexo C: protocolo de tinción con plata.	90
Anexo D: Alelos para cada locus (entre paréntesis frecuencias alelicas)	91
Anexo E: Matriz de distancias genéticas estándar de Nei entre las subpoblaciones (POB)	92

## RESUMEN

### TÍTULO: GENOTIPIFICACION DE CACAO CRIOLLO COLOMBIANO USANDO MARCADORES DE ADN TIPO MICROSATELITES (SSR)\*

AUTOR: HERNANDEZ GARCIA, Jessika Andrea\*\*

**PALABRAS CLAVES:** *Theobroma Cacao*, Marcadores Moleculares, Microsatélites, Análisis de poblaciones, Variabilidad genética.

#### DESCRIPCION

Los cacaos criollos son una de las variedades de *Theobroma cacao* que proporcionan chocolates finos altamente cotizados en la industria. En Colombia el cacao criollo está en proceso de extinción, debido principalmente a la introducción sucesiva de germoplasma cuyos descendientes híbridos presentan una mayor resistencia a plagas y enfermedades.

La caracterización actual de los cacaos criollos colombianos se hace utilizando metodologías tradicionales basadas en la combinación de aspectos morfoagronómicos que han sido cuestionados por basarse en características influenciadas por factores externos que pueden confundir la identificación de esta variedad. En el presente estudio se realizó la genotipificación de 51 accesiones de Cacao Criollo recolectadas de 3 regiones diferentes de Colombia (Guajira, Cesar, Magdalena) más una muestra de cacao forastero y otra de cacao trinitario conservadas en el banco nacional de germoplasma en la estación experimental la suiza de la Corporación Colombiana de Investigaciones Agropecuarias CORPOICA. Se usaron 14 microsatélites (SSR) empleando las técnicas de PCR y electroforesis capilar. Se evaluó la variabilidad genética interpoblacional mediante análisis estadísticos.

Los resultados obtenidos concluyen que las poblaciones de cacao criollo de los departamentos de Magdalena, Cesar y Guajira están estrechamente relacionados en su componente genético, mediante el cálculo de los índices de Wright se encontró que la población es altamente endogámica y presenta valores moderados de estructura genética. De los 14 sistemas SSR empleados, los microsatélites MlcCIR60, MlcCIR12, MlcCIR37, MlcCIR8 y MlcCIR33 confirmaron su informatividad y utilidad para la realización de estudios relacionados con cacao criollo colombiano, según los estándares internacionales, finalmente el patrón molecular de cada clon o individuo obtenido en este estudio sirve como base para establecer una metodología de certificación genotípica de árboles fuente de material de siembra como apoyo a la conformación de jardines clonales.

---

\*Trabajo de Investigación

\*\*Facultad de Ciencias. Escuela de Química. Programa de Química. Dir. Herminulde JesusCano. Calle PhD.Roberto Coronado, Ing.

## ABSTRACT

**TITLE: CRIOLLO COCOA GENOTYPING COLOMBIAN DNA MARKERS USING TYPE MICROSATELLITE (SSR)**

**AUTHOR: HERNANDEZ GARCIA, Jessika Andrea**

**KEYWORDS: *Theobroma Cacao*, molecular markers, microsatellites, population analysis, Genetic variability.**

### ABSTRACT

The criollo is one of the *Theobroma cacao* varieties of fine chocolates that highly valued in the industry. Criollo cacao in Colombia is in the process of extinction, mainly due to the successive introduction of germplasm whose hybrid offspring are more resistant to pests and diseases.

The actual characterization of Colombian criollo cacao is done by using traditional methods based on the combination of morpho-agronomic aspects that have been questioned based on characteristics influenced by external factors that may confound the identification of this variety. The present study was performed on 51 accessions of criollo cacao collected from 3 different regions of Colombia (Guajira, Cesar, Magdalena) plus a sample of trinitarian cocoa and another of forastero cocoa preserved in national germplasm bank at the experimental station "La suiza" affiliated to the Colombian Corporation of Agricultural Research CORPOICA. We used 14 microsatellite (SSR) using the PCR and capillary electrophoresis. Genetic variability was evaluated by statistical analysis interpopulational.

The results conclude that criollo cacao populations from the departments of Magdalena, Cesar and Guajira are closely related in their genetic composition, by calculating indices Wright it is found that the population is highly inbreeding and presents moderate values of genetic structure. Out of the 14 SSR microsatellites used, the MlcCIR60, MlcCIR12, MlcCIR37, MlcCIR33 MlcCIR8 confirmed their informativeness and usefulness in studies related to Colombian criollo cocoa, by international standards, finally the molecular pattern of each clone obtained in This study serves as a basis for establishing a certification methodology genotypic source tree planting material to support the formation of clonal gardens.

---

\*Investigación Work.

\*\*Faculty of Sciences, School of Chemistry. Program of Chemistry. Dir. Herminul de Jesus Cano. Calle PhD. Roberto Coronado, Ing.

## INTRODUCCIÓN

El género *Theobroma* comprende 22 especies de las cuales *Theobroma cacao* L. es el cultivo de mayor importancia comercial a nivel mundial [1]. El cacao, como producto de alto valor alimenticio y de otros usos para la humanidad, debe reunir algunos requisitos especiales para su consumo, es por esto que existen varios mercados para el cacao con precios diferentes, según su origen y la presentación del producto.[2] Clasificándolos en dos tipos de cacao como:

- I) Cacao corriente u ordinario que representa el 95 % de la producción anual, proviniendo principalmente de África y Brasil y se destina en gran parte para la producción de manteca
- II) Cacaos finos que tienen sabores y aromas distintivos, se utilizan en la elaboración de chocolates negros o chocolates tipo gourmet porque le confieren a los productos características de aroma y sabor especiales.[3]

Entre los cacao tipo II, se encuentran los criollos que son una de las variedades que proporcionan chocolates finos altamente cotizados en Estados Unidos y Europa donde se paga sobrepagos por almendras obtenidas del genotipo criollo.[3]. Esta situación, sumada a otras características de mercadeo, que premia los productos con denominación de origen y simplifica los requerimientos de comercio, estimula la conservación, cultivo y comercialización de genotipos de calidad superior. Sin embargo, debido a su bajo rendimiento agronómico y la susceptibilidad a enfermedades, la producción de cacao fino se estima en menos de 5% de la producción mundial de cacao.[4]

En Colombia el cacao criollo está en proceso de extinción, debido principalmente a la introducción sucesiva de germoplasma cuyos descendientes híbridos, con una mayor resistencia, han ido sustituyendo gradualmente las plantaciones nativas.

Como consecuencia se ha presentado una disminución en la calidad de este cacao.[5]

La caracterización actual de los cacaos criollos colombianos se hace utilizando metodologías tradicionales basadas en la combinación de aspectos morfoagronómicos y siguiendo 35 descriptores recomendados por la Unión Internacional de protección de nuevas variedades de plantas como el índice de mazorca, longitud de la hoja, número de flores, granos/mazorca etc.[6]

No obstante, esta caracterización ha sido cuestionada por basarse en caracteres fenotípicos directamente influenciados por la heredabilidad del rasgo, factores ambientales, herencia multigénica, herencia cuantitativa y por la dominancia parcial en algunos caracteres; lo cual puede enmascarar la expresión de un rasgo genético.[7]

La detección y análisis de la diversidad genética de los clones de cacao criollo colombianos se hace necesaria para generar información que puede ser utilizada en el mejoramiento del cultivo y en programas de selección y mejoramiento de productos, basados en su calidad y en el reconocimiento de su procedencia. Lo anterior, tiene implicaciones de propiedad industrial y comercial para aquellos cacaos nacionales con los que se podrían producir cacaos especiales; al igual que la selección de nuevas variedades de interés en cuanto a calidad y resistencia a patógenos. [8].

El estudio de la diversidad genética basada en marcadores moleculares se presenta como una alternativa más precisa para la identificación de los genomas. Estas técnicas permiten la discriminación de los individuos en función del genotipo, independientemente del efecto ambiental o de las interacciones entre genes[9]. Entre las metodologías desarrolladas, los microsátelites ofrecen la

ventaja adicional de ser una metodología rápida y sencilla, que no presenta el riesgo del uso de materiales radioactivos.[10]

Por lo antes expuesto, en el presente trabajo de investigación se realizó la genotipificación de 51 cultivares de cacao criollo colombiano usando 14 marcadores moleculares microsatélites (SSR). De esta manera se hizo posible caracterizar los individuos debido a que estos marcadores muestran un patrón de herencia codominante; es decir, siempre se puede identificar los individuos heterocigotos y homocigotos, así como evaluar relaciones filogenéticas entre los mismos de manera más eficiente.

Esta caracterización molecular es una herramienta valiosa para complementar la caracterización morfoagronómica y bioquímica de materiales nacionales, de esta manera se contribuye al conocimiento de la diversidad en recursos genéticos nativos proporcionando valor agregado, mayor capacidad de negociación y confiabilidad en cuanto a la procedencia del material.

La utilización de una técnica de certificación exclusiva de materiales colombianos, redundará en mejores ingresos para los productores y para la industria nacional, creando fuentes de trabajo y bienestar para los pequeños agricultores.

## **OBJETIVOS**

### **OBJETIVO GENERAL**

Genotipificar el cacao criollo presente en el Banco Nacional de Germoplasma ubicada en la estación experimental la suiza en Santander perteneciente a la Corporación Colombiana de Investigación Agropecuaria.

### **OBJETIVO ESPECÍFICO**

- ✓ Optimizar los protocolos de extracción de ADN para plantas de cacao (*Theobroma cacao*) tipo criollo.

- ✓ Determinar la eficiencia de los marcadores microsatélites moleculares para caracterizar la base genética del cacao tipo criollo del Banco Nacional de Germoplasma de la estación experimental de la Suiza.
- ✓ Determinar el grado de variabilidad genética presente en las accesiones de cacao tipo criollo de la estación experimental la Suiza.
- ✓ Agrupar a partir de las distancias genéticas los diferentes clones de cacao tipo criollo caracterizados en el presente estudio.
- ✓ Correlacionar los resultados con las bases de datos especializadas para identificar genéticamente los diferentes clones de cacao criollo Colombiano.

## 1. MARCO TEÓRICO Y ESTADO DEL ARTE

### 1.1 ORIGEN DEL CACAO

El cacao, desde hace más o menos 2000 años, es parte de las delicias del paladar y la economía de los pueblos del mundo; fue alimento y fuente de riqueza para los indígenas mayas y aztecas, radicados en Centroamérica. La especie vegetal cacao, ha sido denominada en términos científicos (*Theobromacacao L*) por el botánico **Lineo** quien la clasificó. [1]

Alrededor de 20 años después del descubrimiento de América, los conquistadores españoles de México se asombraron de las grandes cantidades de cacao que encontraron en los almacenes del emperador Moctezuma y de la gran popularidad que tenía en la corte la bebida que se hacía de él. Ésta es la razón por la que se estableció en Europa una firme creencia de relación entre los aztecas y el cacao.[11]

Según estudios de Pound y Cheesman (1934), el cacao es originario de América del Sur, del área del alto Amazonas, que comprende países como Colombia, Ecuador, Perú, y Brasil, donde se ha encontrado una alta variabilidad. Desde este lugar de origen, las especies se fueron difundiendo y evolucionando en dos grupos de cacao con características fenotípicas y genotípicas bien definidas, las cuales corresponden a los cacaos criollos y Forastero.[12]

Ya en el siglo XVI, en la era poscolombina, el cacao se dispersó a otros continentes, cuando Hernán Cortés reportó el hallazgo de una bebida amarga usada por los aztecas y envió las semillas y recetas a Europa[7]. Durante el siglo XIX, las recetas originales se refinaron, y se desarrollaron las tecnologías que facilitaron el tostado y molienda de los granos de cacao, con lo cual se originó el desarrollo de la industria del chocolate y se popularizó su consumo en el mundo. El cultivo de *T. cacao* en otros continentes se inició durante la era colonial entre

los siglos XVIII y XIX, y en 1900, el 80% de la producción se daba en el continente americano. Ya en el siglo XXI, América se convierte en el continente con la menor producción, contrastando con el continente africano, donde se encuentra 78% de la producción mundial [13] En la actualidad, el cacao es cultivado en la franja geográfica tropical húmeda (figura 1) ubicada desde los 18° norte hasta los 20° sur de la línea ecuatorial [14]

**Figura 1:**Distribución mundial del cultivo de cacao



Fuente[12]

Botánicamente al cacao se le ha asignado la siguiente clasificación:

División: Espermatofita  
Clase: Angiosperma  
Sub-clase: Dicotiledónea  
Orden: Malvales  
Sub-orden: Malvinae  
Familia: Esterculiáceas  
Tribu: Bitneria  
Género: *Theobroma*  
Especie: *cacao*

Todas las formas cultivadas están contenidas en la especie *cacao*, la cual ha sido dividida en 2 sub-especies: *cacao* y *sphaerocarpum*. El árbol de cacao es una planta diploide, con un genoma de 390 Mb ( $2n=20X=20$ ). [15]

## **1.2 TIPOS DE CACAOS**

La mayoría de los autores han dividido a los cacaos en tres grandes grupos sobre la base de dos elementos fundamentales, como son las características morfológicas del fruto y del grano, la cual se hizo por la necesidad de dividir el producto según la calidad para el mercado y dividir a las plantas de acuerdo a sus exigencias agrotécnicas. En la actualidad la clasificación genética se da en tres grandes grupos: criollos, Forasteros y una mezcla de los criollos y forasteros denominados Trinitarios.

### **1.2.1 Grupo de cacao criollo**

La variedad de cacao criollo, tiene un genotipo casi único y homocigoto, fue uno de los primeros en ser cultivados.

El término criollo (indígena) originalmente fue atribuido por los conquistadores españoles, al cacao que fue cultivado en esa época, en Venezuela. Actualmente, se ubica en este grupo todos los cacaos que muestran las características fenotípicas que son: mazorca cilíndrica, con 10 surcos profundos simples o en pares, cascara rugosa, delgada o gruesa, con una ligera capa lignificada en el centro del pericarpio, el color de la mazorca inmadura puede variar del verde al rojo. Los tipos con cotiledones color blanco; posee entre 20 y 30 semillas de color blanco o crema, alto contenido de grasa, sin astringencia y bastante aroma; son usados en la industria cosmética. Los principales tipos criollos incluyen cacao Pentágona, cacao Real y cacao Porcelana se cultivan principalmente en América Central, México, Colombia y parte de Venezuela.[16]

El hallazgo de descendientes aislados de cacaos criollos en la región poblada por los pueblos precolombinos, hasta nuestros días muestra que, pueden haber fungido como reservorio genético, y al recombinarse con las accesiones introducidas a mediados del siglo pasado, dieron origen a gran parte de la

variabilidad presente en los cultivos de Mesoamérica [13]. Lo anterior convierte a las plantaciones asentadas en esta parte del mundo, junto con las características edafoclimáticas, en candidatas a conservar y exhibir individuos con la calidad requerida por los mercados más exigentes.

### 1.2.2 Grupo forastero

Los cacao forasteros, conocidos también como cacao Amazónicos y/o amargos son originarios de América del Sur. Su centro de origen es la parte alta de la cuenca del Amazonas en el área comprendida entre los ríos Napo, Putumayo y Caquetá. Esta población es la más cultivada en las regiones cacaoteras de África y Brasil y proporcionan más del 80 % de la producción mundial. El cacao forastero es muy variable y se encuentra en forma silvestre en la alta (Perú, Ecuador y Colombia) y baja Amazonia (Brasil, Guyanas y a lo largo del río Orinoco en Venezuela), presenta estaminoides con pigmentación púrpura, mazorcas verdes con más de 30 semillas, de color púrpura, con alta astringencia y bajo contenido de grasa. En la tabla 1 se presentan las principales características morfológicas entre el cacao criollo y el forastero. Estos últimos son originarios del alto Amazonas y dispersados naturalmente, por dicha cuenca.[9]

**Tabla 1:** diferencia entre el cacao criollo y forastero

	<b>CRIOLLOS</b>	<b>FORASTEROS</b>
<b>Árbol</b>	Débil y pequeño	Robusto y grande
<b>Hojas</b>	Grandes, color verde oscuro	Pequeñas, Color verde claro
<b>Mazorcas</b>	Forma cundeamor y angoleta	Amelonado y calabacillo
<b>Cascaras</b>	Fina y Suave	Gruesa y dura
<b>Superficies</b>	Rugosa	Lisa
<b>Almendras</b>	Blancas, Violeta pálido (rosado) y de forma redondeada	Pigmentadas, violeta oscuro y de forma aplanada
<b>Plagas</b>	Susceptible	Tolerantes
<b>Sabor</b>	Fino	Ordinario
<b>Adaptación</b>	Pobre y limitada	Muy buena

### **1.2.3 Grupo de cacaos trinitarios**

Los cacaos Trinitarios son híbridos naturales (espontáneos) de las formas de criollos por los forasteros, como híbridos presentan las siguientes características: muestran gran variabilidad, heterosis o vigor híbrido, manifestado en el mayor tamaño de las mazorcas, crecimiento vigoroso del árbol, altas producciones, resistencia a los agentes adversos, la mazorca y los granos muestran características que los acercan por un lado a los criollos y por otro a los forasteros del tipo amelonado, con toda una gama de combinaciones superficiales

de colores. Debido a su heterogeneidad, los trinitarios constituyen un material excepcional al mejorador, quien podría seleccionar entre la multitud de combinaciones, aquellas que reúnan el mayor número de características interesantes y propagarlos vegetativamente.

El mercado mundial de cacao fino distingue dos amplias categorías de cacao engrano: el cacao en grano fino o de aroma y el cacao en grano corriente u ordinario.

Como una generalización hay que mencionar que el cacao fino o de aroma proviene de las variedades de tipo criollo y Trinitario, mientras que el cacao ordinario proviene de las variedades de tipo Forastero. [17]

### **1.3 CARACTERÍSTICAS PRINCIPALES DEL CACAO CRIOLLO.**

El cacao fino y de aroma representa el 5% de la producción mundial de cacao y según la organización Internacional del Cacao – ICCO, solo se produce en diecisiete países, ocho de los cuales tienen una producción exclusiva de este tipo de cacaos. Aunque su uso más conocido es como ingrediente principal del chocolate, existen en realidad cuatro productos intermedios que se derivan del

cacao en grano: pasta o licor de cacao, manteca de cacao, torta de cacao y cacao en polvo. Estos, además de ser materia prima para la industria alimenticia han incursionado en los mercados de productos cosméticos. [18]

**Tabla 2:** Países productores de cacao criollo.

<b>PAISES PRODUCTORES EXCLUSIVOS DE CACAO FINO Y DE AROMA</b>	<b>PAISES PRODUCTORES MIXTOS DE CACAO FINO DE AROMA</b>
República Dominicana	Ecuador
Jamaica	Venezuela
Santa Lucía	Costa Rica
Granadina	Colombia
Santa Lucía	Indonesia
San Vicente y Granadina	Papúa Nueva Guinea
Samoa	Panamá
Surinam	Sao Tomé y Príncipe
Trinidad y Tobago	Sri Lanka

Este tipo de cacao se encuentra principalmente entre Venezuela y Colombia, existiendo el mayor número de variedades Criollas en Venezuela como son los Porcelanas, Guasares, Pentágonas, Criollos Andino, Ocumares, Chuaos, entre otros.[19]

Los Cacaos criollos contienen más azúcar, su sabor amargo es usualmente bajo debido a la rápida fermentación, donde se elimina las purinas como son la cafeína y teobromina; los criollos poseen más cafeína que teobromina, a diferencia de los forasteros que sucede lo inverso. La teobromina y la cafeína constituyen más del 99% del contenido de los alcaloides en cacao, generalmente el contenido de alcaloides está determinado por el genotipo y la maduración de la semilla.[20] Braudeau (1970) señala que los cacaos criollos poseen menor valor de la relación teobromina/ cafeína en comparación con otros grupos de cacaos.[21]

Así mismo Cedeño (2008) referenciando a Hasing (2004), señala los valores de la relación Teobromina/Cafeína (T/C) para los tres grupos de cacaos. Obteniendo valores de 1.58 con muestras de cacao venezolano Guasaré, siendo ubicado este valor dentro del grupo de los criollos. Del mismo modo fue evaluada una muestra de cacao Forastero procedente de Ghana, la cual obtuvo un valor de 10 en T/C., de acuerdo a este mismo estudio se ubicó al clon CCN 51 dentro del grupo de los trinitarios, y a otros trinitarios como el ICS – 95 entre valores cercanos a 6 como se indica en la tabla 3.[20]

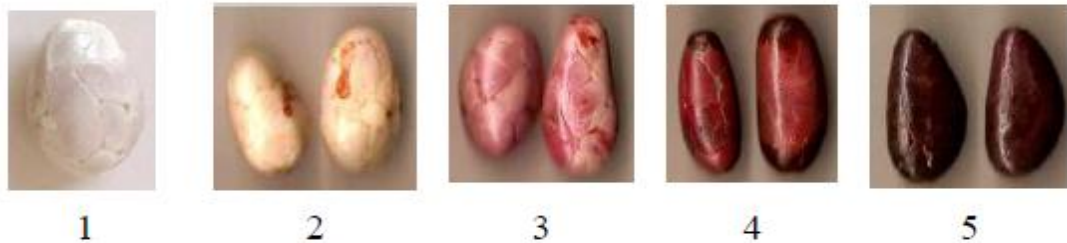
**Tabla 3:** Valores de la relación teobromina/ cafeína en los diferentes grupos de cacao conocidos.

GRUPOS	VALOR DE RELACION TEOBROMINA / CAFEINA
Cacao Forastero	15-10
Cacao Trinitario	10-5
Cacao Criollo	2-1

La relación Teobromina/cafeína es un parámetro de gran utilidad para diferenciar los tres grupos de cacaos más importantes.

Para la caracterización morfológica del cacao criollo se siguen 35 descriptores para diferenciar los distintos grupos de *Theobroma cacao* de acuerdo a la Unión Internacional de protección de nuevas variedades de plantas (UPOV).[6] Para el cacao criollo una de las características más relevantes son los cotiledones de color blanco, esto es debido a la ausencia de ciertos polifenoles como las Antocianinas, compuesto químico que produce el color violáceo o púrpura en los cacao Trinitarios y Forasteros.[20] ver Figura 2.

**Figura 2:** 1 y 2 Cotiledones de cacao criollo, 3 cotiledones de cacao trinitario, 4 y 5 cotiledones de cacao forastero



Fuente [6]

La fermentación del cacao criollo también es una característica relevante, ya que el proceso es más rápido que el de los otros grupos, esto se debe a que las células del cotiledón son más grandes y porosas, permitiendo la pérdida de compuestos químicos como Polifenoles y Purinas (compuestos químicos que producen la astringencia y amargura al cacao).[3]

Otra característica de identificación de los cacaos criollos es que las hojas nuevas de las plantas, son ligeramente velludas y las semillas son secas, éstas son usualmente grandes y abombadas. Así como la cáscara rugosa de la mazorca y su forma cilíndrica.(Figura3)

La principal limitante desde la perspectiva de los compradores de la industria del chocolate gourmet es garantizar la calidad cuando el grano proviene de empresas pequeñas que producen poca cantidad y en general, en condiciones de baja productividad, atribuida a las enfermedades y a inapropiadas técnicas de fermentación y almacenamiento de los granos de cacao. Los fabricantes y comercializadores que incorporan un cacao de origen a su gama de productos, corren el riesgo de incumplir con la entrega del producto, debido a la baja disponibilidad de grano de un solo origen. De otra parte, para los pequeños

agricultores es una carga demostrar su capacidad para entregar cacao de alta calidad y en cantidades consistentes durante varias temporadas.[22]

**Figura 3:** Mazorca de cacao criollo



Fuente UPOV.

#### **1.4 PROPAGACIÓN DEL CACAO.**

Tradicionalmente, el cacao se ha multiplicado a través de la semilla sexual, originada de la polinización de una flor ocurrida en la naturaleza por acción de un insecto o del ser humano, para formar lo que ya denominamos un *híbrido*. El productor suele tomar una semilla de las mejores plantas, sin conocer el origen del polen que le dio vida, ya que la mayoría de las plantas son *autoincompatibles*, esto es, necesitan del polen de otra planta diferente para quedar fecundadas. Este hecho hace que en un cultivo se den muchos tipos de plantas con características diferentes entre una y otra, incluso muchas veces muy diferentes a la planta madre que les dio origen. Esta variabilidad brinda a la plantas diferentes beneficios en cuanto al tamaño de los frutos, en la calidad de los granos y, muy seguramente, en su comportamiento frente a las plagas y enfermedades. [23, 24]Es por ello que en una plantación natural o de híbridos es tan difícil ver una misma producción en todos los árboles. Se dan plantas que pueden llegar a

formar 40 y50 frutos, pero hay otras que no producen nada o son contadas las mazorcas aprovechables en ellas.

En la tabla 4 se presenta la producción mundial de cacao en grano correspondiente al periodo 2002-2003 al 2006-2007. Se observa una producción con altas y bajas, con volúmenes que van desde 3.169.000 toneladas métricas en la cosecha 2002-2003 a 3.400.000 toneladas en el período 2006-2007. Sin embargo las enfermedades fúngicas, omicetos y virales, así como las plagas de insectos, son responsables de aproximadamente el 30% de las pérdidas de cosecha a nivel mundial [25]. Al igual que muchos otros cultivos tropicales, el conocimiento genético y del genoma del *T. cacao* es limitado.

**Tabla 4:** Producción de cacao en grano (miles de toneladas)

Continentes	2002/03	%	2003/04	%	2004/05	%	2005/06	%	2006/07	%
<b>África</b>	2231	70.4	2550	72.1	2,375	70.3	2642	71	2392	70.4
<b>América</b>	428	13.5	462	13.1	445	13.2	446	12	411	12.1
<b>Asia y Oceanía</b>	510	16.1	525	14.8	559	16.5	636	17.1	597	17.5
<b>Total</b>	3169		3537		3379		3724		3400	

FUENTE: ICCO, Boletín Cuatrimestral de Estadísticas de Cacao, Vol. XXXIII, año No.4 2006/07.

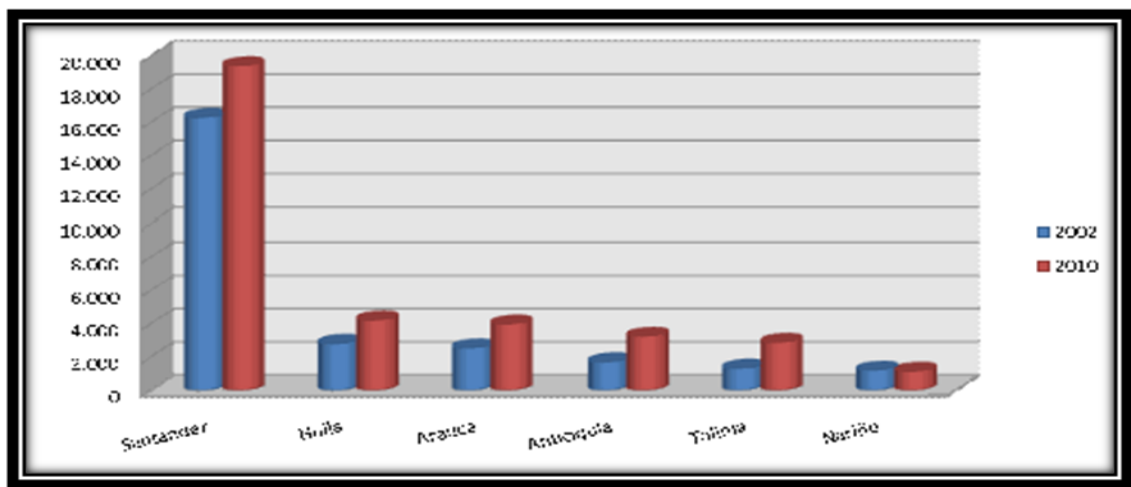
En Colombia la producción de Cacao es variable en los diferentes departamentos, siendo Santander el de mayor producción a nivel Nacional. (Figura 4)

Colombia es una de las cinco naciones mega diversas del mundo. Con una superficie continental de 114.17 millones de hectáreas, en 0,77% del área terrestre del mundo, alberga aproximadamente el 10% de las especies vegetales y animales conocidas. El país está catalogado entre las naciones con mayor riqueza por unidad de superficie de plantas vasculares y vertebrados.

Esta riqueza biológica es atribuida a la posición latitudinal y su localización en la franja tropical ecuatorial, lo que brinda características climáticas y geomorfológicas

que determinan su gran biodiversidad con respecto a otros países del mundo; posee más de 20 ecosistemas lo que brinda un valor económico y político de primera instancia al país, para su aprovechamiento social. Gracias a las características de sus suelos, Colombia presenta cultivos de cacao de los tres tipos conocidos.[26]

**Figura 4:** Producción Nacional de Cacao en toneladas por Departamentos en Colombia 2002-2010



fuentes[25]

## 1.5 COLECCIONES DE GERMOPLASMA

En Colombia los recursos genéticos agrícolas han sido caracterizados y evaluados de manera parcial, con mayor énfasis en la evaluación y la caracterización morfogenómicas. Al respecto se ha reportado, para un universo de 15840 accesiones del Sistema de Bancos de Germoplasma de la Nación colombiana, 10.181 de ellas con evaluación morfológica por atributos cuantitativos, es decir, 64,26% del total; 6057 están caracterizadas por caracteres cualitativos, un 38,23%

del conjunto; 355, por isoenzimas, 2,24% del conglomerado; 64 con información ecofisiológica, 0,40% de éstas; y 917 caracterizadas molecularmente, 5,80% de la Metapoblación. [27]

La utilización eficiente de bancos de germoplasma vegetales en gran medida dependiente de la disponibilidad y la precisión de los datos de pasaporte y otra información relacionada. La ausencia de conocimientos detallados sobre el origen, genética de identidad, las relaciones entre progenies individuales y estructura de la población ha dificultado la posible explotación de germoplasma en el mejoramiento de cultivos, y este ha sido el caso, con la mayoría de bancos de germoplasma de cacao.[26]

En la década de los 80, se realizaron colectas de cacao criollo en la finca Lolita Walt, ubicada en el área de Riofrío, municipio de Ciénaga, en donde se colectaron y seleccionaron 44 cultivares con características de criollos. La colección denominada “Lolita Walt”, la atiende CORPOICA y se encuentra ubicada en la Estación Experimental Caribia, en el municipio de Sevilla, departamento de Magdalena. Teniendo en cuenta estos estudios en el 2006 se propuso un proyecto para la Recolección, caracterización morfoagronómica y molecular de materiales criollos y de alto rendimiento en Colombia.

Se realizaron expediciones en la Sierra Nevada de Santa Marta y Serranía de Perijá de cacao criollo y se caracterizaron morfológicamente siguiendo descriptores como índice de mazorca, tamaño de grano, longitud de mazorca, resistencia a enfermedades entre otros.

Adicionalmente realizaron una caracterización molecular de los materiales nacionales incluyendo 2 cultivares criollos utilizando 23 microsatélites. Los resultados obtenidos generaron un dendrograma de UPGMA(método de pares agrupados usando promedio aritmético) con todas las accesiones, observaron que

los 2 cultivares de cacao criollos eran los genotipos que más se alejaban del resto, sin embargo entre ellos se determinó una similitud muy alta, cerca del 90%. [28]

La evaluación y caracterización de los cacaos tipo criollos recolectados en esas expediciones no han sido evaluados molecularmente y se han clasificado en este grupo según los datos de pasaporte siguiendo los descriptores morfológicos. En esta investigación se genotificarán la accesión de cacao criollo recolectada y almacenada por CORPOICA en el banco nacional de germoplasma.

## **1.6 MARCADORES MOLECULARES**

Desde la prehistoria, el hombre ha seleccionado y mejorado especies vegetales, animales y microbianas basándose en el fenotipo. Las mejoras genéticas eran posible gracias a la variabilidad genética, a la heredabilidad del carácter que se quería aislar, a la eficacia e intensidad de la selección aplicada, y al tiempo necesario para realizar un ciclo de selección. Sin embargo, quedan muchos aspectos desconocidos, como son el número y efecto de los genes implicados en la expresión de un carácter, la localización de estos genes, y su función fisiológica. Por otra parte, la taxonomía siempre ha estudiado características morfológicas, lo cual requiere observaciones muy exhaustivas de los organismos en diferentes estados de desarrollo; además los criterios utilizados carecen muy a menudo de definición y objetividad, en cualquier caso son marcadores ambiguos debido a las influencias ambientales[29]

Afortunadamente la aparición de los marcadores moleculares está ayudando a eliminar tanto los inconvenientes de una selección basada en el análisis exclusivo del fenotipo, como la identificación de especies y variedades de una forma más rigurosa y repetitiva. [30]

Los primeros marcadores desarrollados a finales de los 70 se basaron en la identificación de proteínas e **isoenzimas** por electroforesis en geles de almidón o

poliacrilamida. Con ellos se abrió el conocimiento de la estructura y heterogeneidad genética entre diferentes especies, variedades, y poblaciones de distinto origen geográfico. Pero esta técnica tenía una limitación muy importante: no era capaz de detectar suficiente polimorfismo entre variedades o especies próximas debido a que las proteínas son el resultado de la expresión génica, que puede ser distinta de unos tejidos a otros, de una etapa de desarrollo a otra, de un medio ambiente a otro, y de una época del año a otra. Los avances de la tecnología del DNA recombinante han permitido el desarrollo de los **marcadores moleculares basados en el DNA**, consiguiendo estabilidad en la identificación de especies y variedades. Estos funcionan como señaladores de diferentes regiones del genoma y permiten evidenciar variaciones (**polimorfismos**) en la secuencia del ADN entre dos individuos, modifiquen éstas o no su fenotipo.[29]

La información genética de las plantas se encuentra almacenada en la secuencia del ácido desoxirribonucleico (ADN) de los cromosomas nucleares y en los genomas de las organelos. La definición de la estructura molecular del ADN por parte de Watson y Crick en los años 50, abrió todo un mundo de nuevas posibilidades científicas para el conocimiento y el mejor aprovechamiento de plantas, animales y microorganismos, contribuyendo en gran parte a lo que se ha dado en llamar la revolución de la biotecnología [9].

Las metodologías moleculares han revolucionado el análisis genético. Polimorfismos basados en ADN se han usado para construcción de mapas de ligamiento, estrategias de selección asistida por marcadores, pruebas de parentesco, identificación de especies y estudios de genética de poblaciones. Esta serie de técnicas moleculares de gran desarrollo en los últimos veinte años permiten conocer la información genética que los organismos portan. En relación a los vegetales son de utilidad en estudios evolutivos y de genética poblacional, manejo de bancos de germoplasma, identificación, protección legal de germoplasma, mapeo, selección asistida por marcadores y clonado de

genes[31].La caracterización molecular usando dichos marcadores tiene ventajas, tales como:

- ✓ No es influenciada por el ambiente
- ✓ El número de análisis es ilimitado
- ✓ Se requiere de pequeñas cantidades de material vegetal
- ✓ ADN es altamente ilimitado
- ✓ Alto polimorfismo (número de alelos/locus)
- ✓ Distribuido por todo el genoma.[16]

Azofeifa (2006) [10]en su revisión bibliográfica explica el uso de marcadores moleculares en plantas, con énfasis en el uso de los mismos en especies frutales del trópico. Comprende aspectos relativos al uso e importancia, ventajas, desventajas y características, evolución de las principales metodologías de marcaje a través del tiempo y una explicación de las técnicas de uso más frecuente. Concluye que Los microsatélites o (SSR) resultan ser altamente polimórficos entre variedades de la planta que se estudie, son útiles como herramientas de mapeo y para estudios de polimorfismos, representan un nivel adicional de información genómica en comparación con otros marcadores que se puede usar para los estudios de mapeo y para la determinación de polimorfismos.

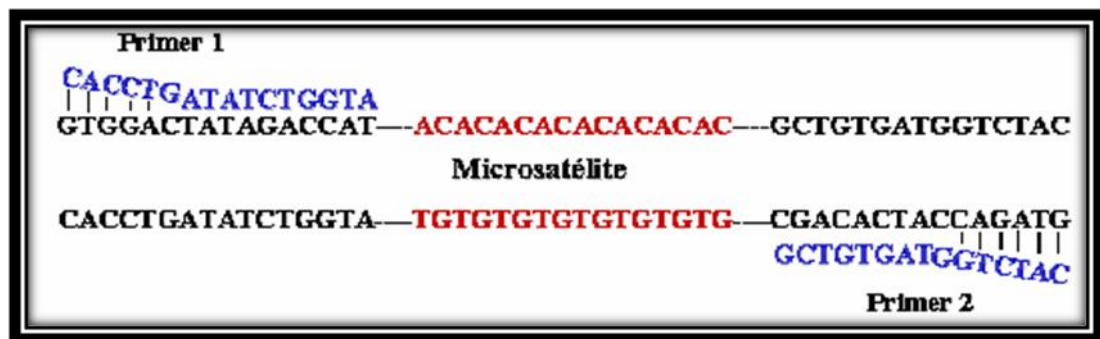
### **1.6.1 Microsatélites o SSR (Secuencias Simples Repetidas)**

Los marcadores de ADN, y en especial los microsatélites, están rápidamente reemplazando o complementando a otros marcadores o metodologías genéticas en las aplicaciones evolutivas y la conservación de las especies. Dado su elevado nivel de polimorfismo, resultan útiles para definir un único genotipo multilocus, de particular interés en estudios donde se requiera una escala muy fina de resolución y en los cuáles otros tipos de marcadores podrían presentar algunas limitaciones, siendo especialmente estos estudios los referidos a: análisis de paternidades, parentescos, etc.

Un buen marcador molecular debe reunir una serie de características para maximizar su utilidad, entre ellas, debe tener una buena distribución a lo largo del genoma y alto grado de polimorfismo, además la técnica para analizar el marcador debe ser rápida, práctica y debe repetirse con fiabilidad en otros laboratorios[30].

Los microsatélites son segmentos cortos de ADN de 1 a 6pares de bases (pb), que se repiten en tándem y de forma aleatoria en el genoma de los seres vivos (Figura 5). Algunas de las características de estos marcadores *versus* otros marcadores que son considerados por la mayoría de autores como una poderosa herramienta para estudios genéticos son: presentan un elevado grado de polimorfismo, su herencia es mendeliana simple, son codominates (pudiéndose diferenciar los individuos homocigotos de los heterocigotos), son fáciles de medir, la técnica para analizarlos es repetitiva y automatizable.[32].

**Figura 5.** Microsatélites, ejemplo de un di-nucleótido A-C(n).



Fuente [10].

El análisis de los microsatélites se hace usando el procedimiento PCR que consiste en la replicación del ADN en los organismos eucariotas realizada por la ADN polimerasa. Estas enzimas realizan la síntesis de una cadena complementaria de ADN en el sentido 5'→ 3' usando un molde de cadena

sencilla, pero a partir de una región de doble cadena. Para esto se utilizan iniciadores (llamados también cebadores o primers en inglés) específicos, que son fragmentos cortos de ADN de una sola cadena. [33]

El desarrollo de marcadores SSR es muy laborioso debido a que deben identificarse y secuenciarse regiones genómicas concretas (bordes del microsatélites), aunque una vez conseguido presenta un sistema muy informativo. Al respecto, Cervera y colaboradores [34]mencionan que aunque los microsatélites permiten analizar sólo un locus por experimento son bastante informativos ya que dejan diferenciar las variantes alélicas de los loci analizados y por lo tanto identificar grupos de ligamiento entre diferentes mapas genéticos. Sin embargo, para su desarrollo se precisa conocer la secuencia por lo que resultan ser menos numerosos que otros marcadores dominantes.

Una cantidad sustancial de trabajos han publicado sobre el uso demarcadores moleculares en la caracterización de bancos de germoplasma de cacao usando marcadores basados en ADN y en especial microsatélites (SSR). Sin embargo la mayoría de estos estudios se han realizado sobre la base genética de cacaos forasteros y trinitarios.

Saunders y colaboradores[35]evaluaron 25 microsatélites que había sido previamente descritos [36]y evaluados por su fiabilidad, amplia distribución en todo el genoma del cacao, numero de alelos producidos por los SSR y su capacidad para discriminar entre las accesiones de cacao. De estos 25 microsatélites 15 resultaron según criterios internacionales como los 15 loci que se pueden utilizar para evaluar la diversidad genética de las accesiones de *Theobroma cacao* en el mundo.

Siguiendo con el análisis genético, Santos [37]presentó un trabajo donde evaluaron la resistencia del árbol del cacao frente a una de las enfermedades que

tantos problemas ha causado a la cacaocultura como lo es la escoba de bruja, evaluaron las plantas de origen en el cruce entre la TSH 1188 y CCN 51 (segregación de la población), por medio de métodos métodos de inoculación en condiciones de campo, para identificar marcadores microsatélites específicos presente en las plantas que fueran resistentes y susceptibles a dicha enfermedad. Delos 18 microsatélites evaluados, 15 fueron polimórficos entre los progenitores y seis fueron polimórficos entre los dos grupos de plantas evaluadas para la resistencia a escoba de bruja.

Por otra parte Zhang [38]evaluó el banco de germoplasma internacional que se encuentra en el Centro Agronómico Tropical de Investigación y Enseñanza CATIE en Costa Rica que contiene uno de las colecciones de Cacao más completas que existe, los resultados de ese estudio indicaron que la colección del CATIE contiene un alto nivel de redundancia genética. Se demostraron que los SSR marcadores, junto con las herramientas estadísticas para la identificación y evaluación de la redundancia, son técnicas prácticas y suficientemente informativas para ayudar a la gestión de colecciones en un banco de germoplasma. Con este estudio se permitió identificar, clasificar y evaluar las diferentes accesiones internacionales existentes, los resultados obtenidos mostraron un dendrograma de las distancias genéticas de 12 de las accesiones internacionales presentes en el CATIE donde se observan la variabilidad genética de las muestras de diferentes países.

Uno de los problemas que frecuentemente se encuentra es la falta de información en la estructura de la población, las relaciones de variedades y el etiquetado incorrecto de las accesiones. Estas son las principales preocupaciones para la conservación y el uso eficiente del germoplasma de cacao. En otro estudio Zhang [39] utilizó los microsatélites para realizar análisis de paternidad, evaluando la identidad individual, la hermandad, y la estructura de la población de cacao recogido en la Región Loreto, Perú en el 1930 -1940.

Por otro lado Ruiz *et al.* (2011), [40] realizó un estudio en Mesoamérica, usando diez microsatélites analizando 70 accesiones de cacao presentes en Nicaragua, encontraron deficiencia de heterocigosis. Pese a esto se encontró un nivel moderado de diversidad genética teniendo en cuenta el corto historial de cultivo extensivo en ese país reportado por Trognitz *et al.* (2011)[41]. Esto fue atribuido a un posible muestreo entre consanguíneos o derivados mayoritariamente de autofecundaciones o, como resultado de diferente frecuencia alélica entre poblaciones.

Del mismo modo, Johnson *et al.* (2009)[42], referenciado por Ovando 2012 encontró una estrecha relación genética entre genotipos “Trinitarios selectos” y los posibles parentales, colectados 70 años antes. Además obtuvieron, usando 35 microsatélites una diferenciación espacial de los genotipos Trinitarios y criollos (agrupados en un *clúster*) al compararlos con poblaciones del Alto y Bajo Amazonas y de Ecuador obtenidos de banco de germoplasma.

Sánchez 2007[8] estudió la diversidad genética presente en el banco de germoplasma de *Theobroma cacao* L. que se conserva en la Estación Experimental ‘La Suiza’ de la Corporación Colombiana de Investigación Agropecuaria –Corpoica– en el departamento de Santander. A tal fin se caracterizaron 100 genotipos de cacao, ninguno de estos criollos, mediante isoenzimas, RFLPs, RAPDs y SSR, utilizando 25 microsatélites publicados previamente en el GenBank. El porcentaje de amplificación obtenido fue del 100% lo que permitió identificar 168 alelos. Los niveles de polimorfismo variaron entre 2 y 14 alelos por locus con un promedio de 6,72. Mediante la construcción de un dendrograma obtuvieron valores de similaridad superiores a 0,45 y en el dendrograma se identificaron dos grupos genéticos principales y varios subgrupos internos. Los resultados obtenidos se consideraron un avance importante en el conocimiento de la diversidad genética de accesiones de *Theobroma cacao* L. conservadas en bancos de germoplasma en Colombia.

En el trabajo de Motamayoret *al.* (2002)[43] se evaluó la composición alélica en las variedades de América Central (genotipos de Guayana, la Amazonia y el Orinoco). Los resultados apoyan la hipótesis de que el cacao se originó en el Alto Amazonas y sugieren la probable ruta de dispersión mediada por humanos desde el Amazonas hasta América Central y México para establecer el cacao “Criollo” que los Olmecas domesticaron hace casi cuatro mil años.

Así mismo Motamayor 2008 evaluó la clasificación y diferenciación de la población, en 1241 accesiones de cacao cubriendo un muestreo geográfico grande, el estudio lo realizó usando 106 marcadores de microsatélites y concluyo que aparte de los 10 grupos genéticos tradicionalmente conocidos existen dos grupos genéticos más de *T. cacao*, la evaluación la realizo usando estadística bayesiana. Esto conlleva a proponer una nueva clasificación del germoplasma de cacao. Los resultados también proporcionan nuevos conocimientos sobre la diversificación de las especies amazónicas en general. [19]

Basándose en las anteriores investigaciones, se ratifica que los marcadores microsatélites, se han convertido en los marcadores de elección para estudios tanto de genética básica como de mejora vegetal. La información generada al caracterizar mediante marcadores moleculares el cacao es de gran utilidad para valorizar los recursos genéticos locales del mismo, precisar mejor lo que se conoce como cacao criollo colombiano en Santander y orientar la selección de nuevos materiales de siembra para ser usados en las nuevas plantaciones colombianas.

## **1.7 ELECTROFORESIS CAPILAR**

El empleo de capilares para la separación de sustancias neutras o iones cargados eléctricamente, apareció en 1967 en una experiencia desarrollada por Hjerten empleando capilares milimétricos, los que eran rotados a través de su sección longitudinal para evitarlos efectos de la convección. Virtanen y Mikkers en 1979

desarrollaron la separación empleando electroforesis en capilares de 200  $\mu\text{m}$  de diámetro interno en vidrio y teflón respectivamente. Más adelante, en 1980, Jorgenson y Lukacs empleando técnicas avanzadas en la obtención de capilares de sílica fundida emplean diámetros de 7  $\mu\text{m}$  y Jorgenson clarifica teóricamente las relaciones entre los parámetros operacionales y las cualidades de la separación revelando el elevado potencial analítico de esta técnica.[44]

El principio fisicoquímico de esta técnica se basa en la separación de moléculas por las fuerzas del campo eléctrico aplicado. La electroforesis ha sido definida como el movimiento diferencial de especies cargadas (iones) o no, por atracción o repulsión en un campo eléctrico. Las muestras separadas en el capilar son monitoreadas por el detector.[44] Los capilares de sílica dejan pasar la luz ultravioleta visible a distintas longitudes de onda, generando por arreglos de diodos, los espectros para la identificación de los compuestos separados por electroforesis capilar.[45]

Las separaciones se realizan empleando mecanismos tradicionales, en un ámbito capilar, que además ofrece más facilidad y velocidad que la cromatografía líquida de alta performance (HPLC). Mientras elimina el problema de los solventes de la HPLC, la toxicidad de los mismos y su costo, pues emplea soluciones acuosas en su gran mayoría con muy baja concentración iónica, incorpora los principios de la automatización a través de un hardware creado especialmente con un software altamente optimizado. Las separaciones se obtienen en pocos minutos, obteniéndose simultáneamente resultados cuantitativos, en oposición a los procedimientos tradicionales que utilizan horas o días.[46]

El explosivo avance de la electroforesis capilar se ha extendido al área **biomédica**, en el campo de las proteínas, péptidos, ADN, análisis de líquidos de perfusión, monitoreo de drogas, marcadores genéticos tumorales y neurobioquímicos, drogas xenobióticas, de abuso, y pericias forenses.

El uso de marcadores moleculares para la identificación de genotipos de interés mediante microsatélites es susceptible de ser analizadas a través de esta instrumentación. La Electroforesis capilar se encarga de la etapa final de esta técnica la cual consiste en la cuantificación del tamaño de los fragmentos de DNA generados en el desarrollo de las PCR. Dicha cuantificación se realiza en relación a estándares internos que se añaden a las muestras. Se pueden utilizar hasta 5 marcadores fluorescentes simultáneamente (uno de ellos reservado para el marcador de tamaños moleculares) y combinar diferentes tamaños alélicos en la misma muestra por lo que por cada capilar en una misma electroforesis se pueden analizar multitud de fragmentos.[47]

### **1.8 ANÁLISIS ESTADÍSTICO.**

Los métodos y procedimientos estadísticos de utilidad en ensayos biotecnológicos son agrupados en aquellos que tienen como propósito evaluar la diversidad, clasificar la estructura y composición genética de las poblaciones, y los desarrollados para la construcción de mapas cromosómicos. Durante los últimos años se han realizado muchos estudios que han utilizados los microsatelites para análisis filogenéticos, concluyendo que con un buen número de loci analizados y con apropiadas tasas de mutación, los microsatelites pueden dar una buena aproximación de la filogenia.[48]

En los estudios de genética de poblaciones, los marcadores permiten la identificación de cada alelo por locus, la obtención de datos poblacionales, y el cálculo de frecuencias alélicas. De esta manera se pueden estimar las distancias genéticas entre poblaciones o entre individuos; así como también realizar análisis filogenéticos y de estructura de poblaciones.[32]

La diversidad o variabilidad genética se puede definir como «la capacidad genética para variar», y por ende, la capacidad a responder tanto a variaciones de índole

ambiental como a cambios en los objetivos de selección. Es así, como la variabilidad genética constituye la base del progreso genético.[49]

Existe una gran diversidad de estadísticos para cuantificar la variabilidad genética y resumir la información a términos manejables. Los estadísticos más empleados son: número de alelos por locus, la heterocigocidad esperada ( $H_e$ ) y observada ( $H_o$ ), el índice de contenido polimórfico (PIC), test de equilibrio de Hardy Weinberg (HW), la estimación de las distancias genéticas poblacionales, análisis de componentes principales y coeficiente de endogamia. Estos estadísticos son los evaluados en el presente estudio y se explican a continuación:

### **1.8.2 Número de alelos por locus**

Esta estadística indica el número medio de alelos que presenta un *locus* en una población. Sin embargo, dicha medida depende mucho del número de individuos analizados, ya que cuando el número es grande, mayor es la probabilidad de detectar alelos. No obstante, este estadístico es útil para estudiar la existencia de variabilidad críptica en los *loci*. [32]

### **1.8.3 Heterocigocidad esperada ( $H_e$ )**

La  $H_e$ , desde el punto de vista matemático, es la probabilidad de que dos alelos tomados al azar de la población sean diferentes [50]. Según Zapata (1987), referenciado por Araguren la  $H_e$  es un buen estimador de la variabilidad, dado que se aplica a cualquier especie, independientemente de su estructura reproductiva o genética, pudiéndose por tanto realizar comparaciones entre ellas. [51]

#### **1.8.4 Heterocigocidad observada (Ho)**

La  $H_o$  se define como la frecuencia relativa de individuos heterocigotos observados en la muestra para cualquiera de los *loci* y se calcula por cómputo directo.

#### **1.8.5 Índice de contenido polimórfico PIC**

El PIC es similar al valor de heterocigosidad y varía entre 0 y 1. Este índice evalúa la informatividad de un marcador en la población de acuerdo a las frecuencias de los alelos. Para su cálculo se multiplica la probabilidad de cada posible cruzamiento (estimado a partir de las frecuencias alélicas) por la probabilidad que sean informativos, es decir, que se pueda identificar al progenitor del que procede el alelo[52].

#### **1.8.6 Test de Hardy Weinberg (HW)**

Para describir la constitución genética de una población es necesario tener en cuenta la constitución genética de los individuos y la transmisión de alelos de una generación a la siguiente; se considera que una población se encuentra en equilibrio de Hardy- Weinberg, si los alelos del locus están distribuidos al azar en la población y no existe asociación entre el par de alelos que un hijo recibe de sus padres.[53]

La ley de Hardy-Weinberg representa a una población grande de individuos diploides, con reproducción sexual aleatoria, sin selección, mutación y migración, las frecuencias génicas y genotípicas permanecen constantes de generación en generación y, además, existe una relación simple entre ambas.[54] Así, una población con frecuencias génicas y genotípicas constantes, se dice que está en equilibrio H-W. Cuando la población se desvía de manera significativa de estas proporciones se habla de desequilibrio H-W, y se puede medir mediante el índice de fijación  $F$ , el cual se expresa para un *locus* cualquiera como:

$F = (H_e - H_o) / H_e$ ; Siendo  $H_e$  y  $H_o$  la heterocigosidad esperada y observada para ese *locus*, respectivamente. Cuando el índice de fijación  $F$  es igual a cero se indica que la población está en equilibrio; mientras que si  $F$  es diferente de cero, ya sea en forma positiva o negativa, indicaría que existe un déficit o exceso de heterocigotos, respectivamente.[55]

### **1.8.7 Estimación de las distancias genéticas**

Para estudiar las diferencias entre poblaciones se han de comparar las frecuencias alélicas observadas de cada *locus* mediante una prueba adecuada y calcularlas distancias genéticas entre ellas. En este sentido, la distancia genética mide el grado de diferenciación existente entre poblaciones de una misma especie, o entre especies, pudiéndose cuantificar esa diferencia mediante diferentes índices de distancias. Las distancias genéticas ayudan a entender las relaciones evolutivas entre poblaciones y nos permite obtener información para caracterización de razas. De las distancias genéticas es importante estudiar el tipo de distancia, el método de construcción del dendrograma o algoritmos.[56]

### **1.8.8 Estimación de endogamia**

El coeficiente de endogamia mide la proporción en que disminuye la frecuencia de heterocigotos de una población en equilibrio HW al someterla a endogamia. El coeficiente de endogamia ( $F_{is}$ ) es la proporción en que aumenta la frecuencia de homocigotos por cada ciclo de apareamiento endogámico. [55]

### **1.8.9 Probabilidad de exclusión (PE)**

Probabilidad de Exclusión (PE), definida como la fracción de individuos que tienen un perfil de ADN que es diferente del de un individuo seleccionado aleatoriamente en una muestra, representadas por,  $h = nh/n$  y  $PE = h^2(1 - 2hH^2)$ , donde  $nh$  es el número de individuos observados con dos alelos,  $n$  es el número total de individuos,  $H$  y  $h$  es el número de homocigotos y heterocigotos, respectivamente. [57, 58]

### 1.8.10 Análisis de varianza molecular (AMOVA)

Permite calcular, además de los componentes de la varianza, el estadístico  $F_{ST}$  (grado de diferenciación genética), considerando la identidad de alelos bajo el modelo de alelos infinitos (IAM), y el estadístico  $R_{ST}$ , a partir de los tamaños alélicos bajo un modelo gradual de mutación (SMM). Esta medida de distancia adaptada para microsatélites  $R_{ST}$ , análoga al  $F_{ST}$  de Wright (1965), fue propuesta por Slatkin (1995):

$$R_{ST} = \frac{S_{bar} - S_w}{S_{bar}}$$

Donde  $S_w$  es la suma sobre todos los loci de dos veces el peso medio de la varianza, dentro de poblaciones  $V(A)$  y  $V(B)$ , y  $S_{bar}$  es la suma sobre todos los loci de dos veces la varianza de las poblaciones combinadas  $V(A+B)$ . [59-61]

Los R-estadísticos son equivalentes a los F-estadísticos, pero sobre la base de tamaños de los alelos en lugar de la identidad del alelo (Slatkin, 1995; Rousset, 1997) [66,67]. Se pueden definir como los coeficientes de correlación intra clase de los tamaños de los alelos de los genes dentro de los individuos en relación con todas las poblaciones ( $R_{IT}$ ), los genes dentro de los individuos en relación con una población ( $R_{IS}$ ), y los genes en las poblaciones con respecto a todas las poblaciones ( $R_{ST}$ ). Fueron desarrollados para los loci en un proceso gradual de mutación. En virtud de un proceso de mutación al azar (IAM, KAM), las expectativas para los R-estadísticos son equivalentes a los correspondientes F-estadísticos, pero sufren mayores variaciones de muestreo [61].

## **2. MATERIALES Y MÉTODOS**

### **2.1 TOMA DE MUESTRAS**

Para realizar el presente estudio fue empleada una muestra de 51 cultivares de cacao criollo (*Theobroma cacao* L.) provenientes de Cesar, Magdalena y la Guajira (Anexo 1) disponibles en las colecciones del banco Nacional de germoplasma de la Estación Experimental “La Suiza” de CORPOICA ubicada a 12 Km del municipio de Río Negro- Santander; a 500msnm y con temperaturas de aproximadamente 26°C. Adicionalmente se tomaron 2 clones diferentes, el IMC-67 tipo forastero y el ICS-60 trinitario.

Como muestra para procesar se tomaron foliolos de cacao criollo de los más jóvenes encontrados en el momento de recolección y se procedió a guardar en bolsas de papel manila marcadas con el número de accesión y árbol respectivo. Seguidamente se almacenaron en frío con el fin de mantenerlas en un ambiente fresco hasta el traslado al laboratorio de Bioquímica y microbiología de la Universidad Industrial de Santander donde fueron almacenadas a -20°C para su procesamiento. Posteriormente las hojas se maceraron en Nitrógeno líquido hasta obtener un polvo fino que se transfirió a tubos ependorff de 2 ml previamente fríos y rotulados para evitar la hidratación del macerado. Finalmente fueron nuevamente almacenadas a -70°C hasta el momento de su empleo en la extracción de ADN.

### **2.2 EXTRACCIÓN Y CUANTIFICACIÓN DE ADN**

Se evaluaron diferentes protocolos para la extracción de ADN de plantas,[36, 62, 63] así como también el Kit de extracción marca MoBiol.

El Kit DNeasy Plant System (Quiagen Inc., Valencia, CA, USA) para la obtención de ADN de plantas, fue el que presento los resultados más consistentes en cuanto

a estabilidad y eficiencia, en el anexo 2 se muestra el protocolo de extracción descrito por el fabricante.

Para la cuantificación de ADN se realizó un método comparativo, teniendo como patrón ADN de fago  $\lambda$  de invitrogen. El método se desarrolló a través de electroforesis en geles de agarosa al 0.8 % en una cámara de electroforesis horizontal marca BioRad. Seguidamente se sembraron en los pozos 12 $\mu$ L de preparación de ADN.(2  $\mu$ L de ADN, 2  $\mu$ L de agua destilada y 8  $\mu$ L del búfer 1X).

Las muestras se colocaron en el gel junto al marcador de peso testigo (ADN fago  $\lambda$ ) de diferentes concentraciones. Se dejaron migrar por un tiempo de 40 minutos a 100 watts. Finalmente se visualizó usando tinción con bromuro de etidio en un equipo de foto documentación marca BioRad con luz UV.

Adicionalmente se cuantificó el ADN mediante espectrofotometría usando un equipo NanoDrop marca Shimatzu, con 1 micro litro de cada muestra y midiendo la relación de absorbancia a 260y 280 nm.

### **2.3 ANÁLISIS DE MICROSATÉLITES SSR.**

La amplificación de ADN se realizó utilizando la técnica de reacción en cadena de la polimerasa (PCR), para ello se usaron diferentes iniciadores con el fin de amplificar 14microsatélites reportados por Saunders et al 2004[35] y descritos en la tabla 5. Los iniciadores fueron sintetizados por IDT (Integrated DNA Technologies, Estados Unidos).

Las condiciones estandarizadas para la amplificación de estos fragmentos sedescriben a continuación: 10ng de ADN, 0,2  $\mu$ M de cada cebador, 2 mM de MgCl<sub>2</sub>, 0,2mM de dNTPs mix, 10mM de Buffer de PCR y 1U de Taq-polimerasa,

los ciclos de amplificación se llevaron a cabo en un termociclador PTC 100 (MJ Research) con un paso de denaturalización a 94°C durante 4 minutos, seguidamente 35 ciclos de 94°C por 30 segundos, 1 minuto a la temperatura de anillamiento de cada par de cebadores, 1 minuto a 72°C y finalmente un paso de extensión de 5 minutos a 72°C.

Los productos de PCR en monoplex(un par de primers por reacción) fueron visualizados en electroforesis vertical de geles de poliacrilamida al 8%. Se sembraron en los pozos 5µL del preparado (3µL de producto de PCR y 2µL de buffer de carga).

Se realizó la corrida usando como buffer TBE 1X con una pre-corrida de 15min a 60V, seguido de 30min a 90V y finalmente 30min a 120V. La tinción del gel se realizó usando plata como colorante, en el anexo 3 se especifica el protocolo de tinción.

Adicionalmente, se amplificaron fragmentos usando reacciones de PCR en multiplex(varios pares de primers por reacción) para correr en electroforesis capilar, teniendo en cuenta las temperaturas de anillamiento, diferencia en tamaños y marcaje de los primers. Se usaron las condiciones reportadas por Motamayor 2008[19] que se describen a continuación: una reacción de multiplex con un volumen final de 25µL con 6.25ng/µL de ADN, 0.05U/µL de Taq polimerasa, 0.2mM de dNTPs, 0.4µM de cada primer forward y reverse, 2mg/ml de BSA, 1.5mM de MgCl<sub>2</sub> y 1x de buffer de PCR. El programa utilizado para las reacciones en multiplex es el mismo descrito anteriormente, solo que la temperatura de acción de la polimerasa fue cambiada a 65°C y el tiempo de extensión se amplió a 7min.

**TABLA 5:** Nombre, secuencias, temperaturas de anillamiento y marcaje con fluorocromos de cada uno de los microsatélites usados.

Nombre del marcador	Cromosoma	Secuencia Primer (5'-3')	Secuencia reversa	Ta (°C)	Secuencia repetida	Marcaje
<b>MTcCIR7</b>	7	ATGCGAATGACAACCTGGT	GCTTTCAGTCCTTTGCTT	51	(GA)11	VERDE (VIC)
<b>MTcCIR18</b>	4	GATAGCTAAGGGGATTGAGGA	GGTAATTCATCATTGAGGATA	51	(GA)12	AZUL (6-FAM)
<b>MTcCIR40</b>	3	AATCCGACAGTCTTTAATC	CCTAGGCCAGAGAATTGA	51	(AC)15	VERDE (VIC)
<b>MTcCIR22</b>	1	ATTCTCGCAAAAACCTTAG	GATGGAAGGAGTGAAATAG	46	(TC)12N146(CT)10	AMAR (NED)
<b>MTcCIR24</b>	9	TTTGGGGTGATTTCTTCTGA	TCTGTCTCGTCTTTTGGTGA	46	(AG)13	AZUL (6-FAM)
<b>MTcCIR15</b>	1	CAGCCGCCTCTTGTTAG	TATTTGGGATTCTTGATG	46	(TC)19	AZUL (6-FAM)
<b>MTcCIR33</b>	4	TGGGTTGAAGATTTGGT	CAACAATGAAAATAGGCA	51	(TG)11	ROJO (PET)
<b>MTcCIR1</b>	8	GCAGGGCAGGCTCAGTGAAGCA	TGGGCAACCAGAAAACGAT	51	(CT)14	AZUL (6-FAM)
<b>MTcCIR60</b>	2	CGCTACTAACAACATCAAAA	AGAGCAACCATCACTAATCA	51	(CT)7(CA)20	VERDE (VIC)
<b>MTcCIR11</b>	2	TTTGGTGATTATTAGCAG	GATTTCGATTTGATGTGAG	46	(TC)13	VERDE (VIC)
<b>MTcCIR12</b>	4	TCTGACCCCAAACCTGTA	ATTCCAGTTAAAGCACAT	46	(CATA)4N18(TG)6	ROJO (PET)
<b>MTcCIR26</b>	8	GCATTCATCAATACATTC	GCACTCAAAGTTCATACTAC	46	(TC)9C(CT)4TT(CT)11	AZUL (6-FAM)
<b>MTcCIR37</b>	10	CTGGGTGCTGATAGATAA	AATACCCTCCACACAAAT	46	(GT)15	AMAR (NED)
<b>MTcCIR6</b>	6	TTCCCTCTAAACTACCCTAAAT	TAAAGCAAAGCAATCTAACATA	46	(TG)7(GA)13	AMAR (NED)
<b>MTcCIR8</b>	9	CTAGTTTCCATTTACCA	TCCTCAGCATTTTCTTTC	46	(TC)5TT(TC)17TTT(CT)4	ROJO (PET)

Ta= Temperatura de anillamiento

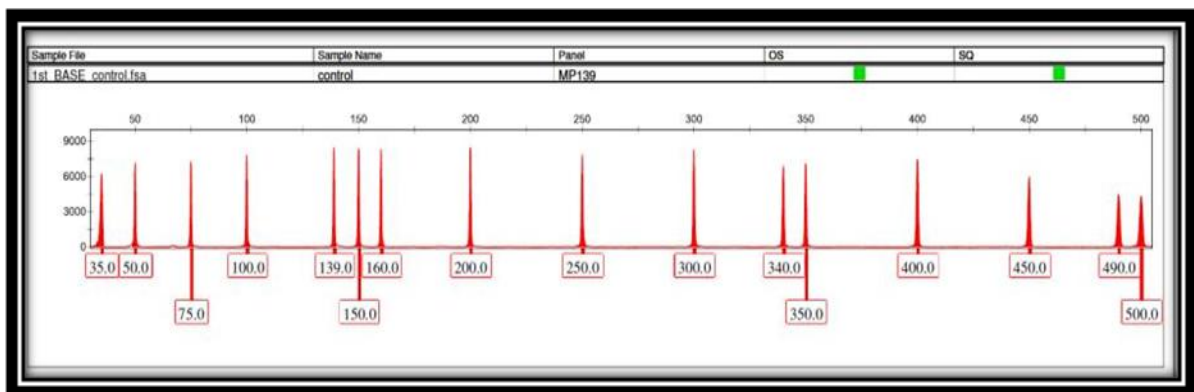
## 2.4 ELECTROFORESIS CAPILAR

Una vez estandarizadas las condiciones de PCR en multiplex se cambiaron los iniciadores por oligonucleótidos forward marcados en el extremo 5' con fluorocromos de 4 colores diferentes (NED, 6-FAM, PET, VIC) ver tabla 5 sintetizados por Applied Bio systems. y se corrieron nuevamente las PCR.

Finalmente se procedió a separar los fragmentos amplificados de los microsatélites por electroforesis capilar. Las muestras se prepararon usando 0,5  $\mu$ L del amplificado de PCR, 12  $\mu$ L de formamida como agente denaturante y, 0,5  $\mu$ L del marcador de peso LIZ 500 de color naranja de 50 a 500pb. En la figura 6 se observa el electroferograma marcador de peso utilizado

Se empleó para la separación de los fragmentos por electroforesis capilar el equipo ABI PRISM® 310 GeneticAnalyzer, secuenciador de ADN de Applied Bio systems con filtro 65 (1mL) 65 mds, soporte restrictivo, polímero POP 4 con un capilar de 47cmx 50 $\mu$ m pln: 402839. Los electroferogramas se visualizaron utilizando el software GENEMAPPER ID versión 3.2.

**Figura 6:** Size standard LIZ 500



## 2.5 ANÁLISIS DE DATOS MOLECULARES

Para los estudios poblacionales tales como: el cálculo de la prueba de equilibrio de Hardy-Weinberg para los 14 marcadores moleculares siguiendo el procedimiento Guo y Thompson (1992) con el algoritmo en Cadena de Monte Carlo Markov, se realizó con el programa ARLEQUIN versión 3.1[64]. Para el cálculo de parámetros poblacionales tales como frecuencias alélicas, parámetros de polimorfismo, análisis de componentes principales, índices de fijación  $F$  de Wright o estadísticos  $F$  denominados  $F_{is}$ ,  $F_{it}$  y  $F_{st}$  [53] son calculados por el programa GENETIX versión 4.05, mediante la metodología de Weir y Cockerham (1984) [65, 66].

Se calculó el contenido de información polimórfica (PIC) a partir de la fórmula propuesta por Botstein y col. (1980) con el complemento The Excel Microsatellite Toolkit.[67] Utilizando el programa MS Excel 2010. El índice de paternidad de poder de exclusión (PE) fue calculado con el programa GenAlex versión 6. [68]

### 3. RESULTADOS Y ANÁLISIS

#### 3.1 EXTRACCIÓN Y CUANTIFICACIÓN DE ADN.

El trabajo con marcadores moleculares involucra el análisis de gran cantidad de individuos y por lo tanto requiere de la capacidad de extraer ADN de buena calidad y en cantidades adecuadas, de forma rápida y eficiente[69]. En la figura 7 se muestra mediante la electroforesis de geles de agarosa, que con el procedimiento de extracción, se obtuvo ADN de buena calidad ya que la banda se presenta una única banda.

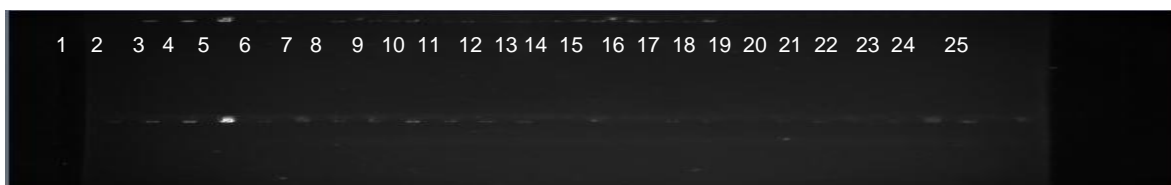
**Figura 7:** Electroforesis de verificación de extracción de ADN con el kit Quiagen en geles de agarosa al 0.8% teñida con bromuro de etidio. Carril 1 al 10 son muestras de ADN de cacao criollo. El carril 11 contiene un marcador de peso hyperladder IV de 50 a 2000bp.



La cuantificación se realizó comparando ADN de las muestras con diferentes diluciones y concentraciones conocidas de ADN del bacteriófago Lambda; las muestras fueron observadas en geles de agarosa al 0,8% teñidos con bromuro de etidio. Se obtuvieron aproximadamente 10 ng/ $\mu$ L de ADN. Esta técnica se basa en la cuantificación de fluorescencia producida por el bromuro de etidio en el ADN a través de un transiluminador de luz ultravioleta.

**Figura 8:** electroforesis de cuantificación de ADN en gel de agarosa. carril1: dilución de 10ng del marcador. Carril 2: dilución de 25ng. Carril 3: dilución de

50ng. Carril 4: dilución de 100ng. Carriles del 5 al 25 diferentes muestras de ADN de cacao criollo extraídas con el kit marca Quiagen.



Por otra parte se realizó la cuantificación utilizando espectrofotometría en un equipo Nanodrop marca Shimatzu. Las bases nitrogenadas en oligonucleótidos tienen su máxima absorbancia a 260nm y la absorbancia máxima de las proteínas es aproximadamente de 280nm; esto se debe principalmente a residuos aromáticos como el triptófano, fenilalanina y tirosina.

La relación de 260/280, por lo tanto es una relación de pureza y debe permanecer aproximadamente entre 1.5 y 1.9, valores inferiores a esto sugieren una contaminación de proteínas.

Los resultados obtenidos visualizados en la tabla 6 sugieren que la concentración obtenida de ADN fue entre 4 y 10 ng y de buena pureza para todas las muestras.

**Tabla 6:** Resultados obtenidos de la cuantificación midiendo la absorbancia en un equipo Nanodrop de los cacaos criollos 53, 57 y 46, 4, 24 y 26. También se incluyó el ICS-60 que es un cacao tipo trinitario.

#	Muestra	Concentración en ng/ $\mu$ L	Relación 260/280
1	CR 53	8	1.5
2	CR 57	4.2	1.5
3	ICS 60	5.6	1.5
4	CR 46	10	1.5
5	CR 4	4.2	1.5
6	CR 24	9.2	1.7
7	CR 26	10	1.7

La baja concentración de ADN obtenida en la extracción comparada con las obtenidas por otros autores [8],[36]quienes reportan concentraciones alrededor de

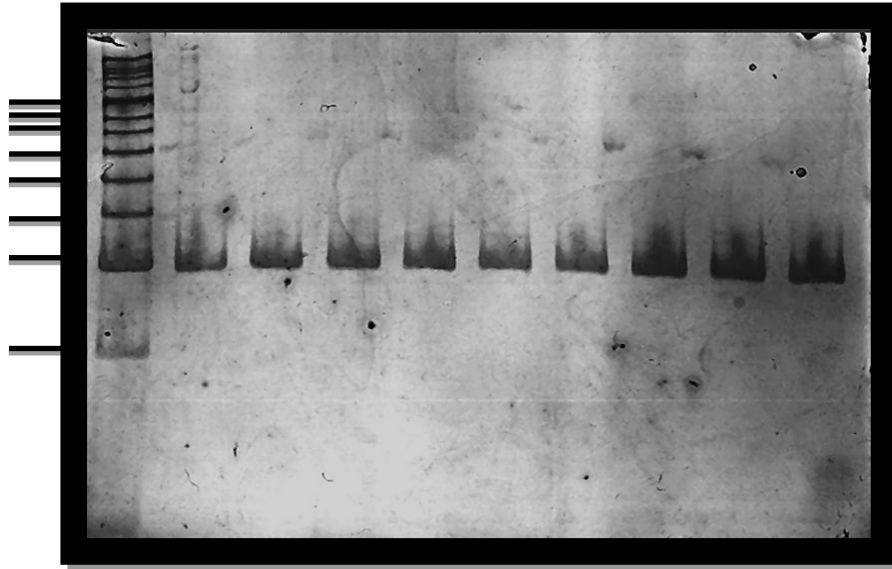
los 100ng se debe principalmente a que en la colecta de las muestras fue difícil conseguir foliolos en un estado verdaderamente joven, a pesar de que se hicieron muestreos en diferentes épocas del año con la finalidad de conseguir las hojas más jóvenes posibles. Es importante resaltar que esto también pudo ser una de las dificultades a la hora de optimizar el protocolo de extracción de ADN, ya que se evaluaron alrededor de 4 protocolos previamente estandarizados[8, 62, 63, 70] para otros grupos de cacao y 1 kit de extracción marca MoBiol sin tener resultados eficientes. Finalmente el Kit de marca Quiagen fue el único que arrojó resultados positivos de extracción de ADN, con concentraciones bajas pero suficientes para el análisis molecular.

### **3.2 ANÁLISIS DE MICROSATÉLITES (SSR)**

Para visualizar los productos de PCR se eligió la electroforesis vertical en poliacrilamida, ya que gracias a su poder de resolución con respecto a los geles de agarosa, permitió una mejor separación de las bandas y una buena distinción entre homocigotos y heterocigotos.

En la figura 9, se observa un ejemplo de un gel de electroforesis en poliacrilamida donde se visualiza la amplificación del marcador mTcCIR22 con 9 genotipos diferentes, utilizando como marcador de peso hyperladder II de 50 a 2000 bp. Se observa que para todas las muestras se amplifica una banda a 300 bp aproximadamente, este valor coincide cercanamente con el rango de alelo reportado por Motamayor 2008, para este SSR el cual va desde 289-300 bp.

**Figura 9:** Electroforesis vertical en poliacrilamida de productos de PCR usando el microsatélite 22 con 9 muestras y tinción con plata. El carril 1 contiene el marcador de peso de 200 a 100pb.

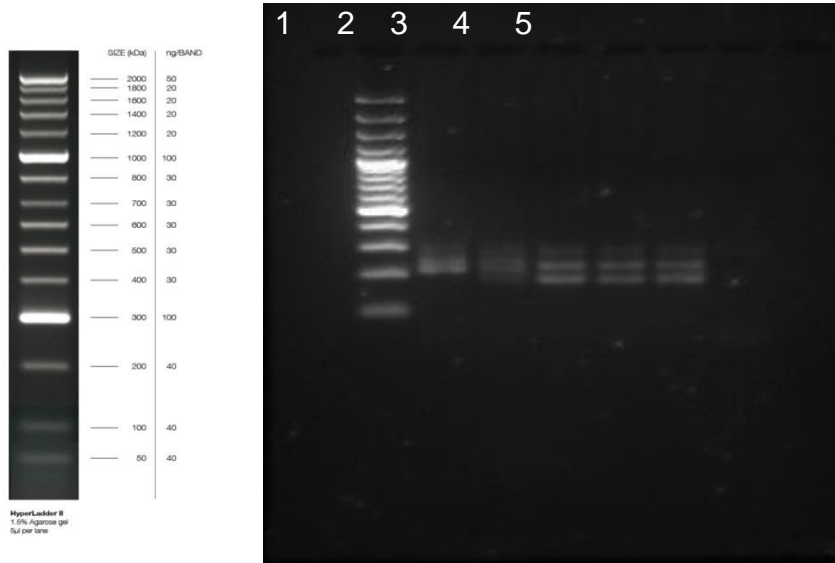


Los diferentes alelos detectados por locus de los microsatélites evaluados fueron organizados según el respectivo valor en pares de bases estimado a partir de los geles de poliacrilamida utilizando el programa QuantityOne de Biorad. Obteniendo una matriz de genotipos individuales para cada clon de cacao criollo que se utilizara para estimar los diferentes estadísticos descritos más adelante.

Como se observa en la tabla 5, la mayoría de los cebadores utilizados en el presente estudio amplifican microsatélites con repeticiones de dinucleótidos, los cuales al ser detectados en geles de poliacrilamida pueden presentar bandas dobles, si se presenta separación de un par de dinucleótidos se amplifica como una sola banda en el gel, dificultando la genotipificación sobretodo en individuos heterocigotos. De este modo buscando minimizar esta ambigüedad se decide usar electroforesis capilar.

Seguidamente se procedió a estandarizar las condiciones de PCR en multiplex usando oligonucleótidos sin marcar y corriendo los productos de PCR en electroforesis horizontal de agarosa los resultados se observan en la figura 10.

**Figura 10:** Electroforesis en agarosa de diferentes productos de PCR usando condiciones en multiplex. M: marcador de peso hyperladder II de 50 a 2000bp. Carril 1 al 5 producto de PCR con reacciones en multiplex.

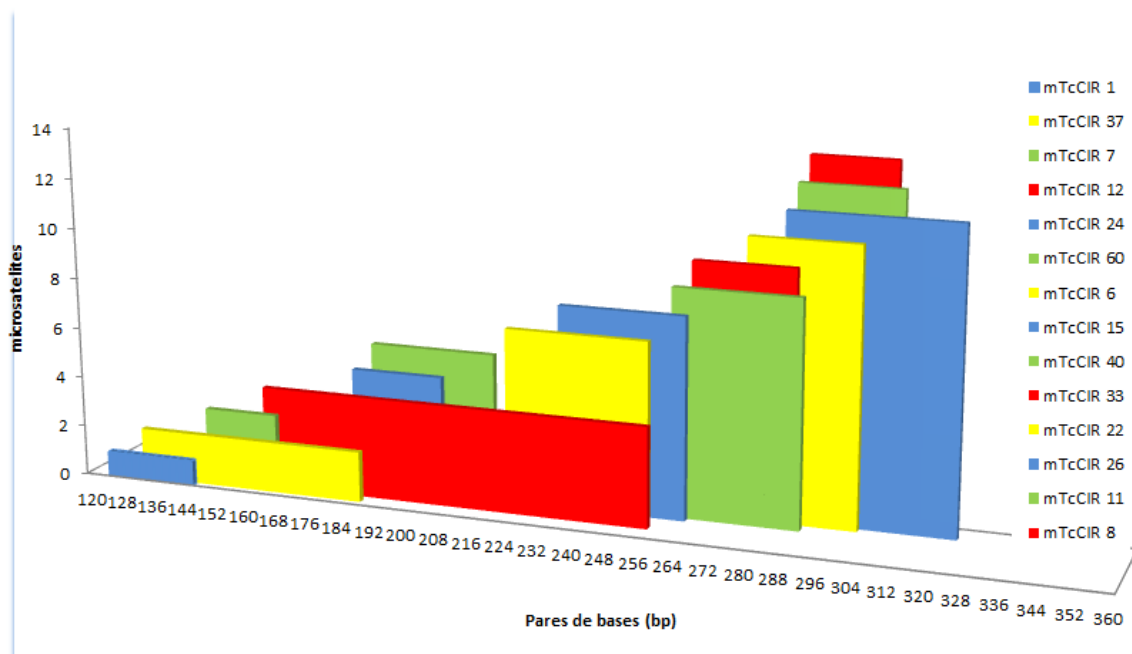


Finalmente al contar con las condiciones estandarizadas se procedió a realizar las amplificaciones de todas las muestras en multiplex cambiando los cebadores por los marcados con fluorocromos, teniendo en cuenta que en la preparación de reacciones debían mezclarse cebadores con las mismas temperaturas de anillamiento.

En la figura 11, se evalúa el rango de pares de bases según Saunders *et al*, 2004 de cada microsatélite y el color del fluorocromo con que se marcó el extremo 5' del primer forward. Los microsatélites que amplifican en el mismo rango de pares de bases se marcaron con diferentes colores para mezclarlos en la reacción de multiplex y así poderlos diferenciar en la electroforesis capilar.

**Figura 11:** Rango de alelos (bp) esperado de cada microsatélite con el color del fluorocromo con que se marcó.

## RANGO DE ALELOS MARCADORES



En la figura 12, se muestran los electroferogramas obtenidos de la separación por electroforesis capilar del genotipo CRFC-46, donde de evidencia la amplificación de 10 microsatélites de los 14 esperados. Cada pico representa un alelo amplificado en pares de bases. Los marcadores MTCcIR1 y MTCcIR60 fueron amplificados en todas las muestras, mientras que MTCcIR 22 y MTCcIR 24 no mostraron una amplificación completa, lo cual pudo deberse a que en la reacción de multiplex se debió ajustar la concentración de los primers en los que se incluían estos dos marcadores moleculares o alterar los valores de  $MgCl_2$ .

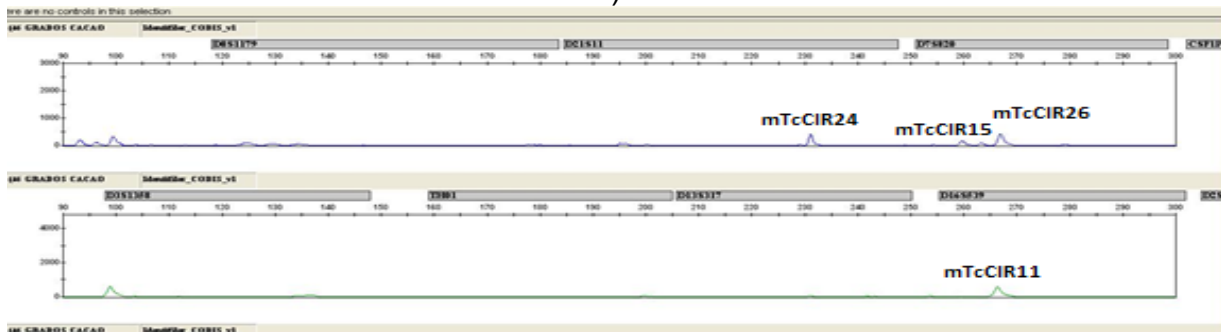
Comparando los resultados anteriores con los obtenidos por Martínez 2007, se muestra que en dicho estudio el marcador mTcCIR40 fue excluido por presentar poca información, a diferencia del presente trabajo en el cual fue de buen aporte gracias a su amplificación en todas las muestras. No solo este marcador sino los 14 microsatélites analizados en este estudio han presentado en estudios anteriores el mayor polimorfismo en cacao.[9] Sin embargo no han sido evaluados sobre bases genéticas de cacao tipo criollo.

Los resultados de la electroforesis capilar se compararon con los obtenidos en electroforesis en poliacrilamida donde se obtuvieron mayor número de datos amplificados en la forma monoplex de las PCR. Al evaluar la matriz de datos obtenida por pares de bases de las dos metodologías se evidencia que concuerdan perfectamente el rango esperado de pares de bases para cada microsatélite.

Para los análisis de datos moleculares se utilizó la matriz de datos obtenidos a partir de la electroforesis en poliacrilamida debido a que presentaron menos del 5% de datos perdidos

**Figura 12:** Tamaño de alelos en pares de bases detectado por los 14 microsatélites evaluados para la accesión CR 46 ilustrando 10 alelos a) electroferograma del genotipo CR- 46 perfiles de una reacción de multiplex con los microsatélites con temperaturas de anillamiento de 46°C. b) electroferograma del genotipo CR-46 perfiles de una reacción de multiplex con los microsatélites con temperaturas de anillamiento de 51°C.

a)





b)

Por otra parte, Santos 2007 evaluó 18 marcadores microsatélites relacionados con resistencia a una de las enfermedades que causan problemas en la caca o cultura como lo es la escoba de bruja, ocasionada por un hongo denominado *Moniliophthora perniciosa*. Los resultados indicaron que mTcCIR 24 y el mTcCIR33 fueron potencialmente ligados a genes de resistencia de esta enfermedad.

Como se ha descrito anteriormente el cacao tipo criollo presenta baja resistencia a plagas y enfermedades, sin embargo las plantaciones de este grupo de cacao caracterizadas morfológicamente según los datos de pasaporte y la evaluación hecha por el Banco Nacional de Germoplasma no presentan indicios de esta enfermedad.

El microsatélites mTcCIR33 evaluados en el presente estudio amplificó en todas las muestras evaluadas de cacao criollo. Teniendo en cuenta lo anterior, la razón por la cual estas accesiones de cacao criollo no exhiben síntomas asociados a la escoba de bruja podría deberse a la presencia de dichos microsatélites.

### 3.3 ANÁLISIS DE DATOS MOLECULARES

Se determinaron los alelos para cada uno de los 14 marcadores SSR en las 51 accesiones. El número promedio de alelos por locus para todos los marcadores analizados fue 4.57. El mayor número de alelos en los loci analizados fue 9, presentándolos mTcCIR60 y mTcCIR12; mientras que el más bajo fue de dos alelos en los marcadores mTcCIR18, mTcCIR24 y mTcCIR26.

Para algunos microsatélites, el número de pares de bases (bp) encontrados son iguales a los reportados en los estudios de Motamayor2008, Saunder 2004, Lanaud 1999 en otros casos no presentaron el mismo número de bp (Tabla 7), posiblemente debido a las diferencias genéticas y orígenes de los materiales evaluados.

**Tabla 7:** Alelos identificados (bp) de los microsatélites amplificados en las diferentes accesiones de cacao criollo evaluadas.

No Locus	cromosoma	Alelo reportado	ALELOS IDENTIFICADOS									No. de Alelos			
MTcCIR18	4	345	300	310											2
MTcCIR40	3	287	266	270	277	279	266	300							6
MTcCIR22	1	289	300	320											2
MTcCIR24	9	198	186	200											2
MTcCIR15	1	254	231	235	240	250									4
MTcCIR33	4	285	273	274	275	267	280								5
MTcCIR1	8	143	125	126	138	139									4
MTcCIR60	2	207	187	188	189	191	192	193	210	211	212				9
MTcCIR11	2	308	280	300	346										3
MTcCIR12	4	188	187	188	189	190	192	193	215	231	258				9
MTcCIR26	8	298	300	320											2
MTcCIR37	10	150	169	170	180	189	190	191	192	250					8
MTcCIR6	6	231	250	260	290										3
MTcCIR8	9	301	289	290	294	298	300	301							6

En el anexo 4, se encuentran los valores de las frecuencias alélicas para cada locus evaluado. La heterocigosidad esperada ( $H_e$ ) promedio para todos los loci evaluados fue 0.3326 y la heterocigosidad observada ( $H_o$ ) promedio fue 0.1481; mostrando que en los marcadores estudiados hay una baja diversidad genética observada con respecto a la esperada, evidenciada por un déficit de heterocigotos en la mayoría de los loci (tabla 8).

En cuanto a la  $H_o$  comparado con  $H_e$  locus por locus, se muestra que los loci mcTCIR33 (0.001), mcTCIR22 (0.023), mcTCIR8 (0.024) y mcTCIR6 (0.047) presentan la menor diversidad, mientras que los loci mcTCIR60 (0.45) y mcTCIR37 (0.44) presentaron el mayor valor, es decir, son los sistemas más diversos.

**Tabla 8:** Parámetros genéticos para los marcadores SSRs: Número de alelos, Heterocigocidad observada (Ho), Heterocigocidad esperada (He), (Gst) índice de diferenciación genética, índice de diversidad, estadísticos-F o de Wright (F<sub>is</sub>, F<sub>st</sub>) y PIC (Contenido de información polimorfica)

Microsatélites	No. Alelos	Ho	He	H-W	p-valor (H-W)	F <sub>is</sub>	F <sub>st</sub>	Gst	PIC
<b>MTcCIR18</b>	2	0.11908	0.11331	0,99	0,01	-0.053	0.15423	0.1266	0.1079
<b>MTcCIR40</b>	5	0.05128	0.48085	0,02	0,05	0.899	0.14151	0.1985	0.4125
<b>MTcCIR22</b>	2	0.02381	0.02381	0,99	0,05	0			0
<b>MTcCIR24</b>	2	0.14286	0.13425	0,99	0,05	-0.067	0.19306	0.1538	0.1264
<b>MTcCIR15</b>	4	0.04762	0.13741	0,009	0,01	0.656	0.00508	0.0428	0.1353
<b>MTcCIR33</b>	5	0.0000	0.73111	0,001	0,01	1	0.06606	0.076	0.6774
<b>MTcCIR1</b>	4	0.2619	0.38009	0,002	0,05	0.248	0.12418	0.0781	0.3255
<b>MTcCIR60</b>	9	0.45238	0.64085	0,005	0,01	0.32	0.081	0.1087	0.5777
<b>MTcCIR11</b>	3	0.2619	0.251	0,11	0,05	-0.024	-	0.00494	0.0269
<b>MTcCIR12</b>	9	0.19048	0.66839	0,001	0,01	0.711	0.04201	0.0723	0.6056
<b>MTcCIR26</b>	2	0.11905	0.11331	0,79	0,05	-0.053	-	0.0014	0.1079
<b>MTcCIR37</b>	8	0.3659	0.74992	0,003	0,01	0.509	0.00707	0.0374	0.7093
<b>MTcCIR6</b>	3	0.04762	0.04733	0,99	0,05	-0.006	-	0.00298	0.0473
<b>MTcCIR8</b>	6	0.02381	0.73867	0,001	0,01	0.967	0.11828	0.2091	0.6722

La ley de equilibrio de Hardy Weinberg permite observar si existe variación genética en una población, a partir de las heterocigosidades observadas y esperadas en cada locus; por lo tanto, permite ver cómo está actuando la población en términos evolutivos. En el análisis de cada sistema se encontraron varios de estos con diferencias significativas (mTcCIR 40, mTcCIR15, mTcCIR 33, mTcCIR1, mTcCIR60, mTcCIR12, mTcCIR37 y mTcCIR8) lo cual indica que están en desequilibrio (Tabla 8), algunos marcadores mostraron un déficit de heterocigotos lo que puede deberse principalmente a 3 factores: que el locus esté bajo selección o a la presencia de alelos nulos en esa población; dando una falsa lectura de un exceso de homocigotos y finalmente a que en la población se están presentando altos niveles de endogamia, producto de cruces entre individuos emparentados.

La endogamia se puede estimar calculando el descenso relativo en los niveles de variación genética que no deberían ocurrir en una población en donde los apareamientos fueran al azar. Para medir esta desviación en la frecuencia de heterocigos en una población, Wright (1965) desarrolló lo que se conoce como el coeficiente de endogamia ( $F_{is}$ ), definido como:[55]

$$F_{is} = 1 - (H_o/H_e)$$

En donde  $H_o$  es la heterocigosis observada y  $H_e$  es la heterocigosis esperada cuando los apareamientos se dan al azar. Este coeficiente puede tomar un valor de  $-1$  a  $1$ . Si es de  $0$  no hay endogamia y la población está en equilibrio de Hardy Weinberg. Cuando el valor es negativo, significa que hay un exceso de heterocigosis y si es positivo que hay un exceso de homocigosis, ya que cuando  $H_o < H_e$  se asume que en esa población se produjeron más apareamientos endogámicos de lo que se espera por azar.

Los resultados obtenidos del efecto de endogamia en la población que fueron estimados por los valores  $F_{is}$ (coeficiente de endogamia) para cada locus coinciden con el tercer factor mencionado anteriormente.

El valor de  $F_{is}$  para la población en estudio fue de 0.5700 lo que indica que la población presenta déficit de heterocigocidad asumiendo que los individuos comparten alelos idénticos por descendencia. En la tabla 8, se observan los diferentes valores de  $F_{is}$  para cada sistema estudiado.

Cabe resaltar que el cacao criollo muestra signos de depresión endogámica [71] que es la pérdida de adaptación (vigor, viabilidad, fecundidad) producida por la pérdida de variación genética debido a la homocigosidad, que impide la supervivencia de la especie. Según Barret y Kohn (1991)[72], la endogamia permite la expresión de alelos recesivos perjudiciales procedentes de ambos progenitores, ya que suele producirse por el cruzamiento genético entre parientes próximos. Precisamente esta expresión de alelos recesivos podría ser una de las razones por las cuales el cacao criollo presenta tan baja resistencia a plagas y enfermedades.

El valor de  $F_{it}$  indica el grado de desequilibrio que detecta cada marcador derivado de la diferenciación genética que hay entre las muestras. Posiblemente esté relacionado con la localización del mismo en zonas del genoma más o menos conservadas. Depende esta capacidad del  $F_{is}$  que se haya manifestado[73].

El  $F_{st}$  contempla el grado de diferenciación total  $F_{it}$  y el  $F_{is}$ , es en definitiva, el que indica el grado de diferenciación genética que se ha obtenido con cada marcador. En este caso el mTcCIR24 es el más destacado ya que presenta valores de  $F_{st}$  que oscilan entre 0.15 – 0.25 lo que indica que este marcador permite una diferenciación genética grande.[74] Sin embargo el mTcCIR22, mTcCIR15, mTcCIR33, mTcCIR1, mTcCIR12 son los que menos variabilidad detectan. Este tipo de análisis tiene interés para diseñar paneles más reducidos demarcadores

que proporcionen la máxima información cuando se abordan trabajos masivos de tipificación de cacaos criollos.

La diferenciación genética entre las subpoblaciones de cacao criollo (Cesar, Magdalena y Guajira) evaluadas es relativamente moderada ya que el  $F_{ST}$  es de 0.06338, este valor indica que existe una baja diversidad por locus, aunque la población es heterocigota, la variabilidad es muy estrecha entre los loci.

Nei (1977), reformuló los índices de  $F$ , basándose en los valores de  $H_o$  y  $H_e$ . En este modelo, el cual es independiente del número de alelos presentes en cada locus, la diferenciación o estructura de la población se mide por un parámetro denominado  $G_{st}$  análogo al  $F_{st}$ .

El índice de diversidad para la población ( $G_{st}$ ) es de 0.067, prácticamente igual al de  $F_{st}$  aunque se utilice otra estrategia para su cálculo. Sugiriendo que las poblaciones son homogéneas, presentando una estructura moderadamente baja heterocigota, que posiblemente comparte alelos entre las subpoblaciones (tabla 8).

Así mismo en la tabla 8 se representan los valores del índice de contenido polimórfico (PIC) que es un parámetro introducido por Botstein *et al.* en 1980, como un indicador de la calidad de un marcador en estudios de cartografía génica.[52] Los marcadores con valores de PIC superiores a 0.5 se consideran muy informativos, los que tienen valores entre 0.25 y 0.5 medianamente informativos y los que muestran valores inferiores a 0.25 poco informativos[52].

Los valores de PIC más altos obtenidos en los resultados evidencian que los microsatélites mTcCIR8, mTcCIR37, mTcCIR12, mTcCIR60, mTcCIR33 se encuentran dentro del rango teórico del contenido de información polimórfica mayor a 0.5 el cual indica que dichos microsatélites son muy informativos y por lo tanto útiles para detectar variabilidad genética en esta especie.

En los últimos años se ha popularizado el cálculo del PIC a fin de obtener una valoración de la calidad de un marcador para estudios genéticos (de segregación, de identificación y control de paternidad, de población) pues refleja el polimorfismo detectado. No obstante, dada su dependencia del número de alelos y de sus frecuencias, la información que aporta no es suficiente para basar en ella la elección de un marcador u otro.[75]

Se calculó el parámetro poblacional poder de exclusión combinado (PE) de importancia en la identificación individual. El cual para las poblaciones mencionadas (Magdalena, Cesar y Guajira) se encontró valores altos en el poder de discriminación combinado 0,9771 (tabla 9) para los 14 sistemas de STR's caracterizados. Los marcadores STR's que dan un mayor aporte para análisis de identificación, por su poder de exclusión combinado son MTcCIR 60, MTcCIR12, MTcCIR37 y MTcCIR8 (tabla 10)

**Tabla 9:** Probabilidad de Rst asumiendo el modelo de mutación stepwise

<b>Estadístico</b>	<b>Valor</b>	<b>P</b>
<b>R</b>		
<b>Rst</b>	0,034	0,026
<b>Ris</b>	0,173	0,005
<b>Rit</b>	0,201	0,001

Diversos estudios empíricos han mostrado, por ejemplo, que las plantas asexuales o que se autofecundan (autocompatibles) como en el caso del cacao criollo presentan menos alelos por locus, tienen menores niveles de diversidad genética dentro de sus poblaciones y están más diferenciadas entre sí que especies con fecundación cruzada [76, 77]Es por ello que las consecuencias evolutivas de promover la fecundación cruzada y el mantenimiento de altos niveles de variación genética, ha sido un tema ampliamente debatido en la literatura.[78-80]

**Tabla10:** poder de exclusión combinado para las poblaciones de Guajira, Cesar y Magdalena

Poder de exclusión combinado															
Pop	N	1	1+2	1+2+3	1+2+3+4	1+2+..5	1+2+3+..+6	1+2+3+4+..+7	1+2+3+4...+8	1+2+3+4..+9	1+2+3+4+10	1+2+3..+11	1+2+3+4..+11	1+2+3+4+....	1+2+3+4.....+
Guajira	10	0,00E+00	2,20E-01	2,20E-01	2,20E-01	2,20E-01	5,50E-01	6,20E-01	8,00E-01	8,20E-01	8,90E-01	9,00E-01	9,60E-01	9,60E-01	9,80E-01
Cesar	17	0,00E+00	1,60E-01	1,60E-01	1,60E-01	2,70E-01	6,90E-01	7,70E-01	8,70E-01	9,00E-01	9,40E-01	9,50E-01	9,80E-01	9,80E-01	9,90E-01
Magdalena	14	2,00E-01	5,10E-01	5,10E-01	6,20E-01	6,90E-01	9,00E-01	9,30E-01	9,70E-01	9,80E-01	9,90E-01	1,00E+00	1,00E+00	1,00E+00	1,00E+00

**Tabla 11:**poder de exclusión combinado locus por locus para cada población.

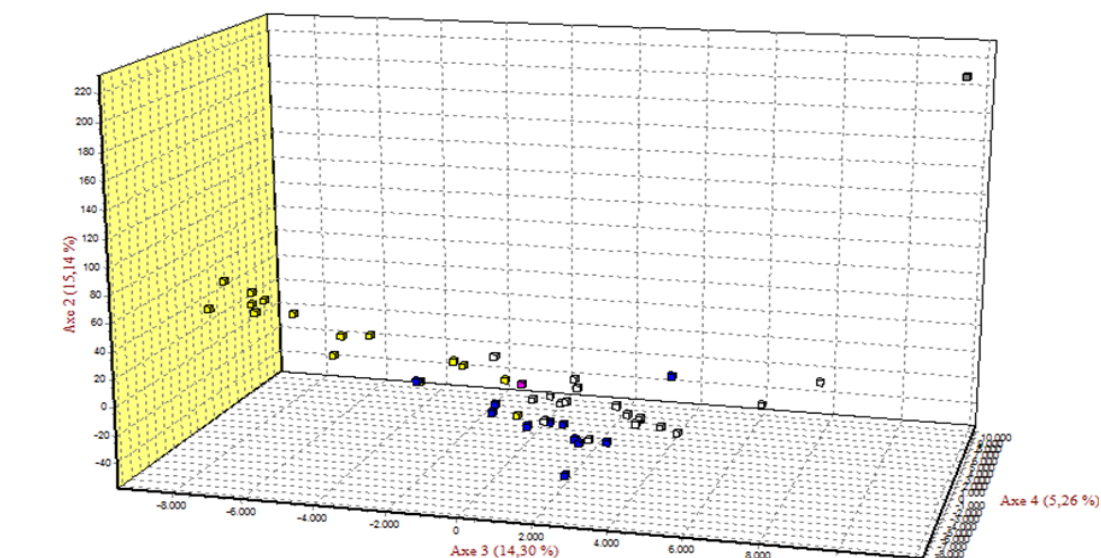
Poder de exclusión combinado															
Pop	N	MTcCIR18	MTcCIR40	MTcCIR22	MTcCIR24	MTcCIR15	MTcCIR33	MTcCIR1	MTcCIR60	MTcCIR11	MTcCIR12	MTcCIR26	MTcCIR37	MTcCIR6	MTcCIR8
Guajira	10	0,00E+00	2,80E-01	0,00E+00	0,00E+00	0,00E+00	5,20E-01	2,80E-01	7,70E-01	2,80E-01	5,20E-01	2,80E-01	7,70E-01	0,00E+00	5,20E-01
Cesar	17	0,00E+00	2,80E-01	0,00E+00	0,00E+00	2,80E-01	6,80E-01	5,20E-01	6,80E-01	5,20E-01	7,70E-01	2,80E-01	7,70E-01	5,20E-01	7,70E-01
Magdalena	14	2,80E-01	5,20E-01	0,00E+00	2,80E-01	5,20E-01	6,80E-01	5,20E-01	8,30E-01	2,80E-01	8,30E-01	2,80E-01	8,30E-01	0,00E+00	5,20E-01

### 3.4 ANÁLISIS DE COMPONENTES PRINCIPALES

En el análisis de correspondencia con la subpoblaciones como fuente de variación (componente principales por subpoblaciones), cada uno de los 3 ejes (dimensiones 1, 2 y 3) contribuyen con 14.30, 15.14 y 5.26% de la inercia total (mide la dispersión de la información) (Figura 13).

En el gráfico de la figura 13 se observa como la subpoblación 2 que se trata del clon IMC-67 originario de Perú y que pertenece al grupo de cacaos forasteros, forma un grupo separado bastante alejado de los demás, . La subpoblación 1 que está representada por el clon ICS-60 originario de Trinidad y que está clasificado dentro del grupo de los cacaos Trinitarios se ubicó espacialmente cerca del centroide donde se encuentran las subpoblaciones 3 y 5 que corresponden a los clones evaluados de cacao criollo de dos zonas del país (Guájira y Cesar). Lo cual se podría explicar porque comparten una porción importante de sus alelos, en frecuencias similares.

**Figura 13:** Grafico de Análisis de Componentes principales



■ Subpoblación 1 ■ Subpoblación 2 ■ Subpoblación 3 ■ Subpoblación 4 ■ Subpoblación 5

Cabe destacar que los cacaos trinitarios son híbridos provenientes de los grupos criollos ancestrales. Johnson et al. (2009), sugiere mediante un análisis bayesiano que el cacao criollo y el Trinitario ICS pertenecen al mismo grupo. Del mismo modo Ciferri y Ciferri (1957) menciona que la distinción entre lo que fue llamado cacao de Trinidad y cacao criollo fue realmente basado en el sistema de clasificación de grano comercial en Venezuela y que las dos poblaciones en realidad comparten una genealogía común.[42, 81]

Bartley, 2005 referenciado por Lambert 2009, señala que es altamente probable que el cacao cultivado en Trinidad antes de 1727 era un material criollo y por introsgresión es ahora conocido como criollo puro.

Por otro lado, la subpoblación 4 se ubicó medianamente alejada del centroide, a pesar de tratarse también de cacaos criollos pero estos originarios de Magdalena lo cual indica posiblemente que estas accesiones son las que aportan más a la diversidad total de la población.

### **3.5 ANÁLISIS DE DISTANCIAS GENÉTICAS**

#### **3.5.1 Distancias entre poblaciones.**

A partir del análisis del cálculo de las frecuencias alélicas se estimaron las distancias genéticas estándar (Ds) (Anexo. 5) utilizando la medida propuesta por Nei (1972).

En la figura 14 se observa la representación gráfica de las distancias genéticas mediante la técnica de agrupamiento generando un árbol de distancias considerado en algunos casos como una representación de la filogenia.

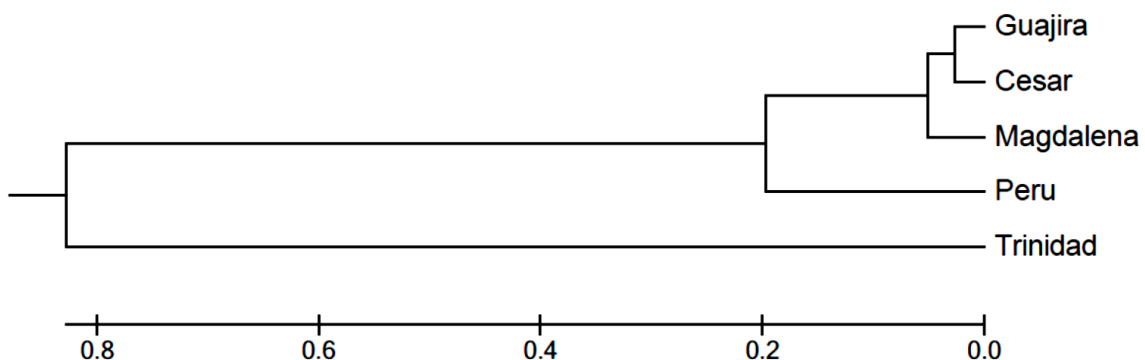
Los programas informáticos basados en este modelo utilizan para construir los árboles filogenéticos una matriz de distancias y algoritmos de agrupamiento simples como el UPGMA (unweighted pair group method using arithmetic averages) (Sneath y Sokal 1973) y el NJ (neighbor-joining) (Saitou y Nei 1987).

Estos árboles se construyen agrupando primero parejas de poblaciones con las mínimas diferencias alélicas, sumando las siguientes más distantes y repitiendo este proceso hasta que todas las poblaciones están incluidas.

En este estudio se muestra que las poblaciones de Guajira y Cesar presenta la menor media de distancia genética calculada entre ellas (0.05) y se encuentran agrupadas a la población Magdalena por deberse precisamente a poblaciones de cacao criollo, todos pertenecientes a Colombia. Se observa que la población de Perú le sigue con una media de aproximadamente (0.37) población de cacao forastero. Además, el bajo valor de distancia genética y el alto valor estimado de migración interpoblacional ( $Nm=0,719$ ), indican que existe un constante flujo génico entre dichas poblaciones, corroborado con los valores moderados de subestructura en la población total.

Cuando dos poblaciones presentan aislamiento genético, los procesos de mutación y deriva conllevan al cambio en las frecuencias alélicas diferenciándose genéticamente cada población; al incrementarse el aislamiento, las frecuencias alélicas también varían.

**Figura14:** Árbol de distancias de las 5 poblaciones DSA construido con el método UPGMA



### **3.5.2 Comparación con la base de datos ICGD**

La Base de Datos Internacional de Germoplasma de Cacao (ICGD) es un servicio de información para la comunidad de investigación del cacao. Con base en la Universidad de Reading, el proyecto tiene como objetivo recopilar información sobre el origen, las características y la disponibilidad de germoplasma de cacao y ponerla a disposición de los investigadores y cultivadores. Para 1990, los objetivos del proyecto se extendieron para hacer más accesible la información publicada de los clones de cacao, sus orígenes, sus características y donde actualmente se cultivan.

La información contenida en la base de datos se ha obtenido de publicaciones, actas de reuniones e informes de investigación enviados directamente al ICGD de investigadores individuales o de instituciones de investigación y bancos de germoplasma.

La información acerca de la caracterización molecular en el ICGD de los clones de cacao criollo es relativamente escasa en comparación con los otros grupos de cacao. Sin embargo dentro de la base de datos se encuentra disponible la huella dactilar (fingerprint) realizada en el estudio de Zhang (2010) para los “CRIOLLO 3” y “CRIOLLO4” utilizando los mismos 14 microsatélites evaluados en el presente estudio. Como se mencionó anteriormente dichos marcadores han sido elegidos por presentar mayor polimorfismo en la base Genética de cacao. En la tabla 12 se muestran los datos extraídos de la ICGD para los 2 clones criollos junto con los alelos obtenidos en la genotipificación del cacao criollo colombiano.

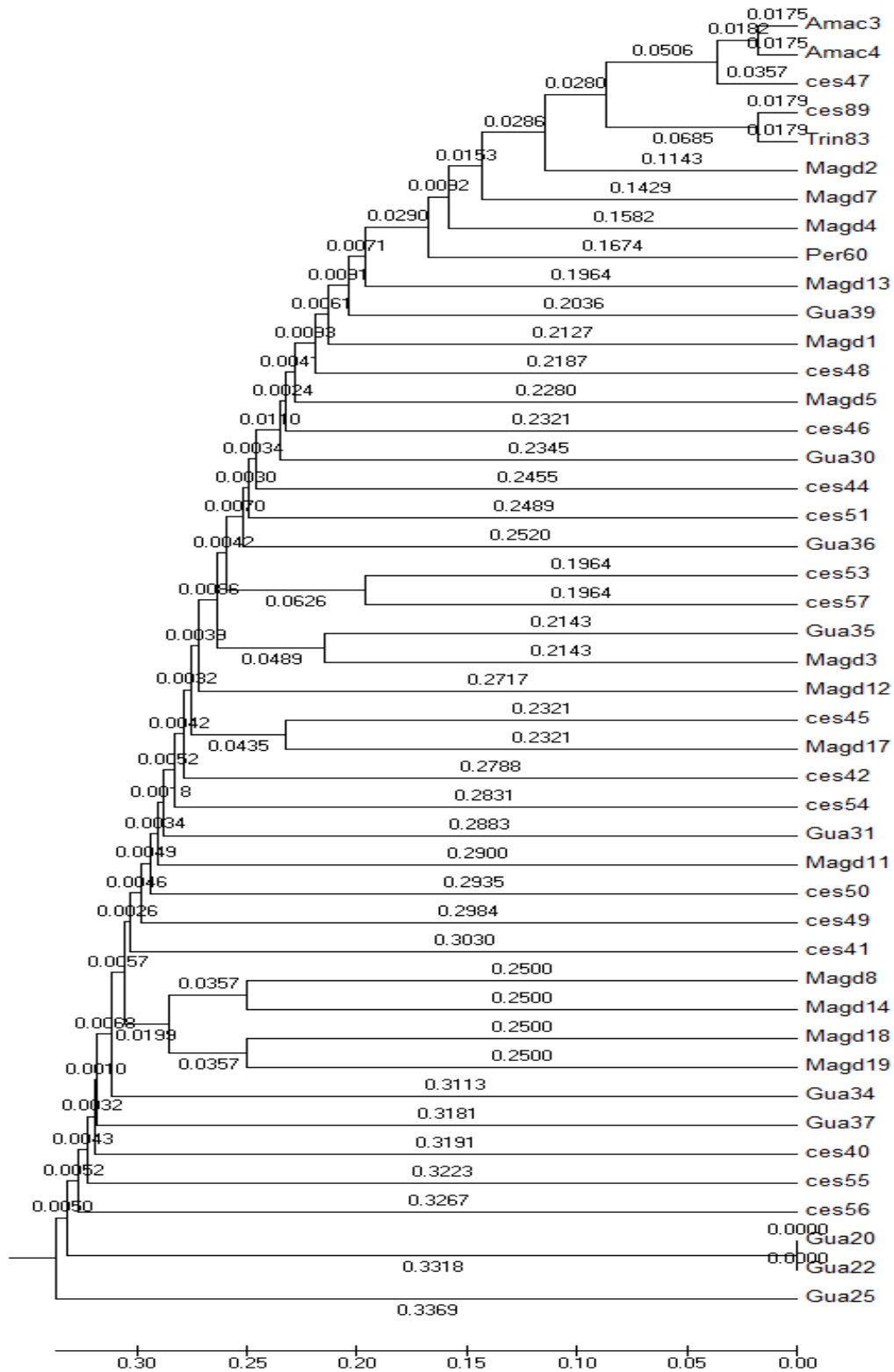
### **3.5.3 Distancias entre Individuos.**

La estructura de diversidad obtenida con el análisis de componentes principales es confirmada por el análisis de distancias genéticas entre individuos (figura 15). Básicamente, se muestra una tendencia a la formación de varios subgrupos donde las accesiones más similares tienen una relación de origen.

En el dendrograma se puede observar que valores bajos de distancias genéticas entre los individuos de la población estudiada (criollos, forasteros y trinitarios), presentan cercanía desde una perspectiva genética. Se observa que el cacao criollo 3 y 4 (Amec 3 y Amec 4) de Costa Rica evaluados por Zhang y registrado en la base de datos ICGD son cercanos genéticamente con la accesión del criollo 47 ( $Ces47= 0,05$ ). Por otro lado el criollo 89 fue el que presentó valores de cercanía más altos con respecto al cacao trinitario. Lo que sería de esperarse debido a que el cacao trinitario resulta ser un híbrido entre forastero y criollo.

Se observa claramente ciertas diferencias entre el cacao criollo colombiano separadas en su gran mayoría por el lugar de la recolección. En general, la población muestra una formación de grupos, posiblemente debido al origen y a las diferencias genéticas entre clones, sin embargo en comparación con el estudio de Sánchez 2006 el cual realizó una caracterización de todo el banco de germoplasma, los valores de similaridad son menores, este resultado era de esperarse ya que se evaluó el jardín clonal del banco de germoplasma solo de cacaos criollos.

**Figura15:** Árbol de distancias Individuales DSA construido con el método UPGMA



**Tabla12:** Comparación entre los alelos obtenidos en el presente estudio (3 y 4) y los obtenidos por Zhang 2010 para los cacaos criollo 3 y 4.

MUESTRAS	MTcCIR18		MTcCIR40		MTcCIR22		MTcCIR24		MTcCIR15		MTcCIR33		MTcCIR1		MTcCIR60		MTcCIR11		MTcCIR12		MTcCIR26		MTcCIR37		MTcCIR6		MTcCIR8	
CRIOLLO 3	333	347	274	288	292	294	187	187	239	257	274	274	128	128	196	210	317	317	190	190	300	300	165	179	250	250	290	306
CRICF3	300																											
	310	270	277	300	300	186	200	250	250	275	275	125	138	188	188	280	300	187	187	300	300	300	300	300	250	250	290	290
CRIOLLO 4	303	355	282	285	290	294	187	187	253	255	275	275	128	128	196	214	303	317	190	203	298	306	142	165	250	250	290	292
CRICF4	300																											
	310	270	277	300	300	186	200	250	250	280	280	125	138	188	188	280	300	190	190	300	300	300	170	170	250	250	290	290

#### 4. CONCLUSIONES

Los resultados obtenidos concluyen que las poblaciones de cacao criollo de los departamentos de Magdalena, Cesar y Guajira están estrechamente relacionados en su componente genético, hecho que se ve reflejado en los análisis poblacionales planteados en este trabajo.

Los valores obtenidos en los parámetros poblacionales de probabilidad de exclusión combinados y Contenido información polimórfico ( $>0,50$ ) calculados a partir de los 14 sistemas SSR empleados, y en especial MlcCIR60, MlcCIR12, MlcCIR37, MlcCIR8 y MlcCIR33 confirman su informatividad y utilidad para la realización de estudios relacionados con cacao criollo colombiano, según los estándares internacionales que se han planteado.

Los sistemas SSR analizados permitieron verificar el estado de la diversidad genética conservada de los clones de cacao criollo del Banco Nacional de germoplasma, y obtener información pertinente para toma de decisiones adecuadas en cuanto al tamaño y la diversidad de la colección.

El análisis de estructura genética arrojó valores moderadamente bajos tanto en la población total como para todos los marcadores utilizados en el estudio, indicando que las poblaciones de Cesar y Guajira se podrían estar comportando como una sola población, concordante con el flujo génico presente en ellas, pues en la entrada o salida de los alelos del reservorio génico se pueden introducir nuevos alelos o alterar las proporciones de los alelos ya presentes. Normalmente se da una pérdida de los alelos menos frecuentes y una fijación (frecuencia próxima al 100%) de los más frecuentes, lo que podría mostrar una disminución en la diversidad genética de la población.

La población de cacao criollo está moderadamente estructurada según los valores encontrados de los índices  $F$  de Wright, lo cual se ratificó con la generación del dendrograma de distancias por individuos; en donde no se observaron grupos definidos.

El patrón molecular de cada clon o individuo obtenido en el estudio sirve como base para establecer una metodología de certificación genotípica de árboles fuente de material de siembra como apoyo a la conformación de jardines clonales.

## 5. RECOMENDACIONES

Se recomienda estandarizar las condiciones de las reacciones en multiplex con el fin de ahorrar considerablemente tiempo y recursos en la caracterización de bancos de germoplasma, ya que se amplificarían varias combinaciones de cebadores en una sola reacción de PCR.

Para la extracción de ADN de plantas es necesario que se utilicen hojas frescas y jóvenes para obtener mayor cantidad de ADN.

Los alelos 191, 192, y 193 amplificados por el microsatélite mTCIR 60 no han sido reportado en trabajos anteriores por lo que se recomienda secuenciarlos de manera de que se corrobore si se trata de alelos nuevos encontrados.

Sería importante evaluar los 14 Microsatélites presentes en el cacao criollo en cuanto a la resistencia a diferentes enfermedades como escoba de bruja y monilia ya que estos resultados podrían contribuir al mejoramiento genético del cultivo.

En el caso de utilizar genotipos de esta colección como fuente de material de propagación o de siembra se debe tener en cuenta la identidad genética a nivel molecular para hacerle seguimiento al material clonado que se derive de él.

## REFERENCIAS BIBLIOGRAFICAS

1. SCHNELL, R; et al. Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. J. Am. Soc. Hortic. Sci. , (2005). 130(2): p. 181-190.
2. MORENO, L; Et al. Beneficio del Cacao. Fundación Hondureña de Investigaciones Agrícolas., 1989. Fascículo N° p. 6. 26 p.
3. SÁNCHEZ, C. Caracterización organoléptica del cacao (*Theobroma cacao* L.), para la selección de árboles con perfiles desabor de interés comercial. facultad de ciencias agrarias escuela de ingeniería agronómica. Universidad técnica estatal de Quevedo Quito 2007. p. 13-24.
4. ARGOUT, X; et al. The genome of *Theobroma cacao*. Nat Genet. 2011 43(2):101-8.
5. MOTILAL, L. et al. The relic Criollo cacao in Belize genetic diversity and relationship with Trinitario and other cacao clones held in the International Cocoa Genebank, Trinidad. Plant Genetic Resources. 2009(02): p. 106-115.
6. UPOV, *Theobroma cacao* L. guidelines for the conduct of test for distinctness, uniformity and stability. International union for the protection of new varieties of plants, 2010.
7. BHATTACHARJEE R. An Improved Semiautomated Rapid Method of Extracting Genomic DNA for Molecular Marker Analysis in Cocoa, *Theobroma cacao* L. Plant Molecular Biology Reporter International Society for Plant Molecular Biology. 2004. 22: p. 435a–435h.
8. SÁNCHEZ, I. et al. Análisis de la diversidad genética de accesiones de *Theobroma cacao* L del banco de conservación a cargo de Corpoica. Revista Corpoica – Ciencia y Tecnología Agropecuaria 2007. 8(2): p. 26-31.
9. MARTINEZ, W.J. Caracterización morfológica y molecular del Cacao Nacional Boliviano y de selecciones élites del Alto Beni, Bolivia. Programa de educación para el desarrollo y la conservación Escuela de Posgrado. Centro Agronómico Tropical de Investigación y Enseñanza.: Costa Rica. (2007)
10. AZOFEIFA, D. Uso de marcadores moleculares en plantas; aplicaciones en frutales del trópico. Agronomía mesoamericana, 2006. 17(2): p. 221-242.
11. BATISTA, L. Serie Cultivos Guía Técnica El Cultivo de Cacao. Primera ed. Centro para el Desarrollo Agropecuario y Forestal. Santo Domingo- República Dominicana. 2009
12. CHEESMAN, E. et al. Further Notes on Criterion of Selection in Cacao. I.C.T.A. . Trinidad Annual Report on Cocoa Research. 1934.
13. PLOETZ, R.C. Cacao Diseases: Important Threats to Chocolate Production Worldwide. Phytopathology. 2007. 97(12): p. 1634-1639.
14. SCARPARI, L.M. et al. Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by

- Crinipellis pernicioso. Journal of Experimental Botany, 2005. 56(413): p. 865-877.
15. AIKKAL RIJU, M. Mining of expressed sequence tag libraries of cacao for microsatellite markers using five computational tools. Journal of Genetics, 2009. 88(2): p. 217-225.
  16. QUIROZ, J. Caracterización Molecular y Morfológica de genotipos superiores con características de cacao nacional (*Theobroma cacao* L.) de Ecuador. Programa de educación para el desarrollo y la conservación escuela de Posgrado. Centro Agronómico Tropical de Investigación y Enseñanza.: Costa Rica.(2002).
  17. MENDA, Z. Programa de desarrollo del cacao en el estado Lara (Proyecto). Consultado 15 feb. 2010]; <http://www.laraenred.com/proyectolaracamb/PROYECTO%20DEL%20CACAO.doc>].
  18. BOLETIN DE ANALISIS POR PRODUCTOS CACAO, B. 4, Editor. 2010, Ministerio de Agricultura y Desarrollo Rural
  19. MOTAMAYOR J; Et al. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). PLoS ONE (2008). 3:e: p. 3311.
  20. ANGELA., C. Establecimientos de parámetros (físicos, químicos y Organolépticos) para diferenciar y valorizar el cacao (*Theobroma Cacao* L.) Producido en dos zonas identificadas al Norte y sur del litoral ecuatoriano, in facultad de ingeniería agronomica. universidad técnica de manabi: santa Ana - Ecuador.2008
  21. BRAUDEAU, J. El cacao. Técnicas agrícolas y producciones tropicales, ed. E. Blumé. 1970 Barcelona -España.
  22. JAIMES, Y; ARANZAZU, F. Manejo de las enfermedades Del cacao (*Theobroma cacao* L) en Colombia, con énfasis en monilia (*Moniliophthora roreri*). Corporación Colombiana de Investigación Agropecuaria, Corpoica E.E. La Suiza, 2010.
  23. ECHEVERRI, R.J. El injerto en la producción de cacao orgánico. Manejo Integrado de Plagas y Agroecología, 2006. 78: p. p. 101-105.
  24. MONGUE-NAJERA, J; et al. Biología General. EUNED ed. 1990.
  25. PRESENTACION CNC CACAO 2011 Consultado febrero 2013 [https://docs.google.com/file/d/1X-la-iwttcoRXhMdeWu3LVmdbContext-dn\\_XhK4GVRPCRpmUnSnZ2kFy5gyxFkx/edit?hl=es&pli=1](https://docs.google.com/file/d/1X-la-iwttcoRXhMdeWu3LVmdbContext-dn_XhK4GVRPCRpmUnSnZ2kFy5gyxFkx/edit?hl=es&pli=1)].
  26. VALENCIA A; Et al. Estado del arte de los recursos genéticos vegetales en Colombia: Sistema de Bancos de Germoplasma. Corpoica Cienc. Tecnol. Agropecu, 2010. 11(1): p. 85-94.
  27. LOBO, M. Conservación, conocimiento y utilización de los recursos fitogenéticos en Colombia, con énfasis en alimentación y agricultura. trabajo presentado al Instituto Alexander von Humboldt. bogotá. s.p, (2003).
  28. ARANZAZU, F; VALENCIA, G; CORONADO, R; RINCON, D. Mejoramiento Genético para incrementar la producción y productividad del sistema de cacao en Colombia. Union Temporal de Colombia Cacao Uno, (2009): p. Pag 36-51 .

29. CLAROS, M. Marcadores moleculares: Qué son, cómo se obtienen y para qué valen. [cited Octubre 2012].
30. CHENG, H; et al. Markers for genetic mapping in the chicken. Poultry Science. 1994. 73: p. 539-546.
31. PICCA, A; Et al. Marcadores Moleculares.II.-Capitulo 4 Biotecnología y Mejoramiento Vegetal. 2007: p. Pág. 61.
32. ARANGUREN-MÉNDEZ, J.Los microsatélites (STRs), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. Arch Latinoam Prod Anim, 2005. 13: p. 30-42.
33. CORTAZAR, M; Rincón, E. Instituto de biotecnología métodos físico-químicos en biotecnología PCR. UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. 2004.
34. CERVERA, M; MARTÍNEZ, J. Análisis genético de la vida (en línea). Consultado 1 feb. 2010 [http://www.rubes.es/ace/ciencia56\\_1.htm](http://www.rubes.es/ace/ciencia56_1.htm)].
35. SAUNDERS, J; LEAMY, S; HEMEIDA, A. Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. Theor Appl Genet, 2004. 110: p. 41-47.
36. LANAUD, C; et al. Isolation and characterization of microsatellites in *Theobroma cacao* L. P J L Mol Ecol, (1999). 8: p. 2141–2143.
37. SANTOS, G; et al. Marcadores microsatélites relacionados com a resistência à vassoura-de-bruxa do cacau. Pesq. agropec. bras., Brasília, 2007. 42 n.8: p. 1137-1142.
38. ZHANG, D; et al. Molecular characterization of an international cacao collection using microsatellite markers. Tree Genetics & Genomes, (2009). 5: p. 1–10.
39. ZHANG, D; et al. Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers Tree Genetics & Genomes, (2009). 5: p. 595–607.
40. RUIZ, J, et al. Molecular ecology of genetic diversity of cacao cultivated in the south-east region of Nicaragua. International Research Journal Agricultural Science., 2011. 1(1): p. 6-13.
41. TROGNITZ, B; et al. Genetic Population Structure of Cacao Plantings within a Young Production Area in Nicaragua. PLoS ONE. 6(1): p. e16056.
42. JOHNSON, E; et al. Population structure and genetic diversity of the Trinitario cacao (*Theobroma cacao* L.) from Trinidad and Tobago. Crop Sci., (2009). 49: p. 564-572.
43. VAZQUEZ A; et al. Potencial de los Marcadores Moleculares para el Rescate de Individuos de *Theobroma cacao* L. de Alta Calidad. Revista de la Sociedad Mexicana de Biotecnología y Bioingeniería A.C., 2012. Vol. 16 No 1.
44. CASTAGNINO, J. Electroforesis capilar. Reimpresión del artículo publicado en Acta Bioquímica Clínica Latinoamericana, 1999. Vol. XXXIII, No. 3: p. 297-329.
45. HEIGER, D. High Performance Capillary Electrophoresis- An Introduction., ed. Hewlett-Packard. 1992: 12-5091-6199E.

46. ALBILLOS, S. Aplicación de la electroforesis capilar para el análisis y seguimiento del grado de maduración de quesos de oveja y mezcla de la provincia de Burgos., in Departamento de Biotecnología y Ciencia de los Alimentos. 2003, Universidad de Burgos.: Burgos. p. 52-71.
47. MANUAL DEL GENOTIPADO. [CONSULTADO 20 de Noviembre 2012. <http://www.uco.es/servicios/scai/impresos/GEN/Manual%20de%20Genotipado.pdf>
48. MARTÍNEZ, M; et al. and Identificación Molecular de Patrones Genéticos en Distintas Muestras de Arándanos (*Vaccinium* sp.). En: RIA., . (Agosto de 2007). Vol. 36. Nº 2: p. p. 3-15.
49. ROCHAMBEAU, H; et al. Measuring and managing genetic variability in small populations. *Annales de Zootechnie.*, 2001. 49: p. 77-93.
50. CROW, J; et al. An introduction to population genetics theory. Harper and Row publisher, 1970: p. 83-98.
51. ZAPATA, C. La variabilidad Genetica de las poblaciones. Espinosa de los monteros, J. *genetica de la acuicultura.* 1987: p. 33-57.
52. BOSTSEIN, D; et al. Construction of a genetic linkage map in man using restriction fragments length polymorphisms. *American Journal of Human Genetic*, 1980. 32: p. 314-331.
53. FALCONER, D; ;ACKAY, T. Introduction to Quantitative Genetics. 4 ed. E. 1996, Harlow, Essex, UK: Longmans Green.
54. Hardy, G.H., Mendelian proportions in a mixed population (letter). *Science*, (1908). 28: p. 49–50.
55. WRIGTH, S. Interpretation of population structure by F-statistics with special regard to system of mating. *Evolution.*, 1965. 19: p. 395-420.
56. NAGAMINE, Y; et al. Genetic distance and classification of domestic animals using genetic markers. *journal of Animal Breeding and Genetics.*, 2001. 118: p. 101-109.
57. HUSTON, K; et al. Statistical Analysis of STR Data. Profiles in DNA. Defining microsatellite alleles by genotyping global indigenous human populations and nonhuman primates. *Forensic Sci*, 1998. 1(3): p. 14-15.
58. MORONI, R; et al. Effects of reference population and number of STR markers on paternity testing. . *Forensic Science International-Genetics. Supplement Series*, 2008. 10: p. 102 p.
59. SLATKIN, M; et al. measure of population subdivision based on microsatellite allele frequencies. (1995). 139: p. 1463.
60. ROUSSET, F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, (1997). 145: p. 1219-1228.
61. BALLOUX, F; et al. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology.* (2002). 11: p. 771-783.
62. GIRALDO, C; et al. Determination of borojo sex (*Borojoa patinoi*, Cuatrecasas) through molecular markers;. *Revista Colombiana de Biotecnología* (2004). 4 No. 2p. p. 9-14.
63. MOTAMAYOR , J; et al. Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. *Heredity* (2003). 91: p. 322-330.

64. EXCOFFIER, L; SLATKIN M. Incorporating genotypes of relatives into a test of linkage disequilibrium. *Am. J. Hum. Genet*, 1998: p. 171-180.
65. BELKHIR, K; BORSA, P ; et al. GENETIX, logiciel sous Windows TM pour la génétique des populations., L.d.G. et and C.U. Populations, Editors. 1996-1998, Université de Montpellier II, Montpellier (France).
66. WEIR, B; Et al. Estimating F-statistics for the analysis of population structure. *Evolution.*, 1984.38.: p. 1358-1370.
67. PARK, S. Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection (2001), University of Dublin.
68. PEAKALL, R. GENALEX 6 genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, (2006) 6: p. 288-295.
69. ROCHA; P. Marcadores moleculares, una herramienta útil para la selección genética de palma de aceite. *Palmas.*, 2003. 24 No 2: p. 11-15.
70. MOTAMAYOR, J; et al. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity*, 2001. 89: p. 386.
71. RUFINO, L. Fortalecimiento de la cadena productiva de cacao, en las provincias de San Ignacio, Jaén y Celendín. Expediente técnico. Cajamarca. 2008.
72. BARRETT, S. The genetic and evolutionary consequences of small population size in plant: implications for conservation. *Genetics and Conservation of Rare Plants* (Eds. D. Falk & K.E. Holsinger), 1991. Oxford University Press: p. pp. 3–30.
73. AVILES, D. Estudio de la influencia de los bovinos andaluces en la formación de las razas bovinas criollas de Latinoamérica, in *Zootecnia y Gestion sostenible*. 2012, Universidad de Cordoba.: Cordoba. p. 19.
74. HARTL, D. A primer of population genetics,. Sinauer associates ed. Vol. 3rd edition. 2000., Sunderland, Massachusetts, U.S.A.
75. MOAZAMI, K. et al. Analysis of genetic diversity in French cattle breeds by the use of microsatellites - preliminary-results. *Genetics Selection Evolution*.1994. 26: p. S155-S165.
76. LIU, F; et al. The effect of mating system differences on the nucleotide diversity at the phosphoglucose isomerase in the plant genus *Leavenworthia*. *Genetics and Conservation of Rare Plants* (Eds. D. Falk & K.E. Holsinger), 1999. 151: p. 343-357.
77. SAVOLAINEN, et al. Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. *Molecular Biology and Evolution* 2000. 17(4): p. 645-655.
78. CHARLESWORTH, D; YANG, Z. Allozyme diversity in *Lavenworthia* populations with different inbreeding levels. *Genetics.*, 1998. . 81: p. 453-461.
79. FISHER, R. The genetical theory of natural selection. Oxford: Clarendon, 1930.
80. MULLER, H. Some genetic aspects of sex. *American Naturalist*, 1932. 66: p. 118-138.

81. CIFERRI, R ; CIFERRI, F. The evolution of cultivated cacao. *Evolution International. Journal of Organic Evolution* (1957). 9 p. 381–397.

## BIBLIOGRAFICAS

AIKKAL RIJU, M. Mining of expressed sequence tag libraries of cacao for microsatellite markers using five computational tools. *Journal of Genetics*, 2009. 88(2): p. 217-225.

ALBILLOS, S. Aplicación de la electroforesis capilar para el análisis y seguimiento del grado de maduración de quesos de oveja y mezcla de la provincia de Burgos., in Departamento de Biotecnología y Ciencia de los Alimentos. 2003, Universidad de Burgos.: Burgos. p. 52-71.

ANGELO., C. Establecimientos de parámetros (físicos, químicos y Organolépticos) para diferenciar y valorizar el cacao (*Theobroma Cacao L.*) Producido en dos zonas identificadas al Norte y sur del litoral ecuatoriano, in facultad de ingeniería agronomica. universidad técnica de manabi: santa Ana - Ecuador.2008

ARANGUREN-MÉNDEZ, J.Los microsatélites (STRs), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. *Arch Latinoam Prod Anim*, 2005. 13: p. 30-42.

ARANZAZU, F; PALENCIA, G; CORONADO, R; RINCON, D. Mejoramiento Genético para incrementar la producción y productividad del sistema de cacao en Colombia. *Union Temporal de Colombia Cacao Uno*, (2009): p. Pag 36-51

ARGOUT, X; et al. The genome of *Theobroma cacao*. *Nat Genet*. 2011 43(2):101-8.

AVILES, D. Estudio de la influencia de los bovinos andaluces en la formación de las razas bovinas criollas de Latinoamérica, in *Zootecnia y Gestion sostenible*. 2012, Universidad de Cordoba.: Cordoba. p. 19.

AZOFEIFA, D. Uso de marcadores moleculares en plantas; aplicaciones en frutales del trópico. *Agronomía mesoamericana*, 2006. 17(2): p. 221-242.

BALLOUX, F; et al. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology*. (2002). 11: p. 771-783.

BARRETT, S. The genetic and evolutionary consequences of small population size in plant: implications for conservation. *Genetics and Conservation of Rare*.

BATISTA, L. Serie Cultivos Guía Técnica El Cultivo de Cacao. Primera ed. Centro para el Desarrollo Agropecuario y Forestal. Santo Domingo- República Dominicana. 2009

BELKHIR, K; BORSA, P ; et al. GENETIX, logiciels sous Windows TM pour la génétique des populations., L.d.G. et al. C.U. Populations, Editors. 1996-1998, Université de Montpellier II, Montpellier (France).

Plants (Eds. D. Falk & K.E. Holsinger), 1991. Oxford University Press: p. pp. 3–30.

BHATTACHARJEE R. An Improved Semiautomated Rapid Method of Extracting Genomic DNA for Molecular Marker Analysis in Cocoa, *Theobroma cacao* L. *Plant Molecular Biology Reporter International Society for Plant Molecular Biology*. 2004. 22: p. 435a–435h.

BOLETIN DE ANALISIS POR PRODUCTOS CACAO, B. 4, Editor. 2010, Ministerio de Agricultura y Desarrollo Rural.

BOSTEIN, D; et al. Construction of a genetic linkage map in man using restriction fragments length polymorphisms. *American Journal of Human Genetic*, 1980. 32: p. 314-331.

BRAUDEAU, J. El cacao. Técnicas agrícolas y producciones tropicales, ed. E. Blumé. 1970 Barcelona -España.

CASTAGNINO, J. Electroforesis capilar. Reimpresión del artículo publicado en *Acta Bioquímica Clínica Latinoamericana*, 1999. Vol. XXXIII, No. 3: p. 297-329

CERVERA, M; MARTÍNEZ, J. Análisis genético de la vida (en línea). Consultado 1 feb. 2010 [http://www.rubes.es/ace/ciencia56\\_1.htm](http://www.rubes.es/ace/ciencia56_1.htm)].

CHARLESWOTH, D; YANG, Z. Allozyme diversity in *Lavenworthia* populations with different inbreeding levels. *Genetics*., 1998. . 81: p. 453-461

CHEESMAN, E. et al. Further Notes on Criterion of Selection in Cacao. I.C.T.A. . Trinidad Annual Report on Cocoa Research. 1934.

CHENG, H; et al. Markers for genetic mapping in the chicken. *Poultry Science*. 1994. 73: p. 539-546.

CIFERRI, R ; CIFERRI, F. The evolution of cultivated cacao. *Evolution International. Journal of Organic Evolution* (1957). 9 p. 381–397

CLAROS, M. Marcadores moleculares: Qué son, cómo se obtienen y para qué valen. [cited Octubre 2012].

CORTAZAR, M; Rincón, E. Instituto de biotecnología métodos físico-químicos en biotecnología PCR. UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. 2004.

CROW, J; et al. An introduction to population genetics theory. Harper and Row publisher, 1970: p. 83-98.

ECHEVERRI, R.J. El injerto en la producción de cacao orgánico. *Manejo Integrado de Plagas y Agroecología*, 2006. 78: p. p. 101-105.

EXCOFFIER, L; SLATKIN M. Incorporating genotypes of relatives into a test of linkage disequilibrium. *Am. J. Hum. Genet*, 1998: p. 171-180.

FALCONER, D ;ACKAY, T. *Introduction to Quantitative Genetics*. 4 ed. E. 1996, Harlow, Essex, UK: Longmans Green.

FISHER, R. *The genetical theory of natural selection*. Oxford: Clarendon, 1930.

GIRALDO, C; et al. Determination of borojo sex (*Borojoa patinoi*, Cuatrecasas) through molecular markers;. *Revista Colombiana de Biotecnología* (2004). 4 No. 2p. p. 9-14.

HARTL, D. *A primer of population genetics*., Sinauer associates ed. Vol. 3rd edition. 2000., Sunderland, Massachusetts, U.S.A.

HARDY, G.H., Mendelian proportions in a mixed population (letter). *Science*, (1908). 28: p. 49–50.

HEIGER, D. *High Performance Capillary Electrophoresis- An Introduction*., ed. Hewlett-Packard. 1992: 12-5091-6199E.

HUSTON, K; et al. Statistical Analysis of STR Data. *Profiles in DNA. Defining microsatellite alleles by genotyping global indigenous human populations and nonhuman primates*. *Forensic Sci*, 1998. 1(3): p. 14-15.

JAIMES, Y; ARANZAZU, F. Manejo de las enfermedades Del cacao (*Theobroma cacao L*) en Colombia, con énfasis en monilia (*Moniliophthora roreri*). Corporación Colombiana de Investigación Agropecuaria, Corpoica E.E. La Suiza, 2010.

JOHNSON, E; et al. Population structure and genetic diversity of the Trinitario cacao (*Theobroma cacao* L.) from Trinidad and Tobago. *Crop Sci.*, (2009). 49: p. 564-572.

LANAUD, C; et al. Isolation and characterization of microsatellites in *Theobroma cacao* L. *PJL Mol Ecol*, (1999). 8: p. 2141–2143.

LOBO, M. Conservación, conocimiento y utilización de los recursos fitogenéticos en Colombia, con énfasis en alimentación y agricultura. trabajo presentado al Instituto Alexander von Humboldt. bogotá. s.p, (2003).

LIU, F; et al. The effect of mating system differences on the nucleotide diversity at the phosphoglucose isomerase in the plant genus *Leavenworthia*. *Genetics and Conservation of Rare Plants* (Eds. D. Falk & K.E. Holsinger), 1999. 151: p. 343-357.

MANUAL DEL GENOTIPADO. [CONSULTADO 20 de Noviembre 2012. <http://www.uco.es/servicios/scai/impresos/GEN/Manual%20de%20Genotipado.pdf>

MARTINEZ, W.J. Caracterización morfológica y molecular del Cacao Nacional Boliviano y de selecciones élites del Alto Beni, Bolivia. Programa de educación para el desarrollo y la conservación Escuela de Posgrado. Centro Agronómico Tropical de Investigación y Enseñanza.:Costa Rica. (2007)

MARTÍNEZ, M; et al. and Identificación Molecular de Patrones Genéticos en Distintas Muestras de Arándanos (*Vaccinium* sp.). En: *RIA.*, . (Agosto de 2007). Vol. 36. N° 2: p. p. 3-15.

MENDA, Z. Programa de desarrollo del cacao en el estado Lara (Proyecto). Consultado 15 feb. 2010]; <http://www.laraenred.com/proyectolaracamb/PROYECTO%20DEL%20CACAO.doc>].

MOAZAMI, K. et al. Analysis of genetic diversity in French cattle breeds by the use of microsatellites - preliminary-results. *Genetics Selection Evolution*.1994. 26: p. S155-S165.

MONGUE-NAJERA, J; et al. *Biología General*. EUNED ed. 1990.

MORENO, L; Et al. *Beneficio del Cacao*. Fundación Hondureña de Investigaciones Agrícolas., 1989. Fascículo N° p. 6. 26 p.

MORONI, R; et al. Effects of reference population and number of STR markers on paternity testing. . Forensic Science International-Genetics. Supplement Series, 2008. 10: p. 102 p.

MOTAMAYOR J; Et al. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). PLoS ONE (2008). 3:e: p. 3311.

MOTAMAYOR, J; et al. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity, 2001. 89: p. 386.

MOTAMAYOR , J; et al. Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. Heredity (2003). 91: p. 322-330.

MOTILAL, L. et al. The relic Criollo cacao in Belize genetic diversity and relationship with Trinitario and other cacao clones held in the International Cocoa Genebank, Trinidad. Plant Genetic Resources. 20098(02): p. 106-115.

MULLER, H. Some genetic aspects of sex. American Naturalist, 1932. 66: p. 118-138.

NAGAMINE, Y; et al. Genetic distance and classification of domestic animals using genetic markers. journal of Animal Breeding and Genetics., 2001. 118: p. 101-109.

PARK, S. Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection (2001), University of Dublin.

PEAKALL, R. GENALEX 6 genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, (2006) 6: p. 288-295.

PICCA, A; Et al. Marcadores Moleculares. II.-Capítulo 4 Biotecnología y Mejoramiento Vegetal. 2007: p. Pág. 61.

PRESENTACION CNC CACAO 2011 Consultado febrero 2013  
[https://docs.google.com/file/d/1X-la-iwttcoRXhMdeWu3LVmdbXcC-dn\\_XhK4GVRPCRpmUnSnZ2kFy5gyxFkx/edit?hl=es&pli=1](https://docs.google.com/file/d/1X-la-iwttcoRXhMdeWu3LVmdbXcC-dn_XhK4GVRPCRpmUnSnZ2kFy5gyxFkx/edit?hl=es&pli=1)].

PLOETZ, R.C. Cacao Diseases: Important Threats to Chocolate Production Worldwide. Phytopathology. 2007. 97(12): p. 1634-1639.

QUIROZ, J. Caracterización Molecular y Morfológica de genotipos superiores con características de cacao nacional (*Theobroma cacao* L.) de Ecuador. Programa de educación para el desarrollo y la conservación escuela de Posgrado. Centro Agronómico Tropical de Investigación y Enseñanza.: Costa Rica. (2002).

- ROCHA; P. Marcadores moleculares, una herramienta útil para la selección genética de palma de aceite. Palmas., 2003. 24 No 2: p. 11-15.
- ROCHAMBEAU, H; et al. Measuring and managing genetic variability in small populations. Annales de Zootechnie., 2001. 49: p. 77-93.
- ROUSSET, F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics, (1997). 145: p. 1219-1228.
- RUIZ , J,et. al. Molecular ecology of genetic diversity of cacao cultivated in the south-east region of Nicaragua. International Research Journal Agricultural Science., 2011. 1(1): p. 6-13.
- RUFINO, L. Fortalecimiento de la cadena productiva de cacao, en las provincias de San Ignacio, Jaén y Celendín. Expediente técnico. Cajamarca. 2008.
- SÁNCHEZ, I.et al. Análisis de la diversidad genética de accesiones de *Theobroma cacao* L del banco de conservación a cargo de Corpoica. Revista Corpoica – Ciencia y Tecnología Agropecuaria 2007. 8(2): p. 26-31.
- SCARPARI, L.M. et al. Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis pernicios*a. Journal of Experimental Botany, 2005. 56(413): p. 865-877.
- SLATKIN, M; et al. measure of population subdivision based on microsatellite allele frequencies. (1995). 139: p. 1463.
- SANTOS, G; et al. Marcadores microsatélites relacionados com a resistência à vassoura-de-bruxa do cacau. Pesq. agropec. bras., Brasília, 2007. 42 n.8: p. 1137-1142.
- SAUNDERS, J; LEAMY, S; HEMEIDA, A. Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. Theor Appl Genet, 2004. 110: p. 41-47.
- SAVOLAINEN, et al. Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. Molecular Biology and Evolution 2000. 17(4): p. 645-655.
- TROGNITZ, B; et al. Genetic Population Structure of Cacao Plantings within a Young Production Area in Nicaragua. PLoS ONE. 6(1): p. e16056.

UPOV, *Theobroma cacao* L. guidelines for the conduct of test for distinctness, uniformity and stability. International union for the protection of new varieties of plants, 2010.

VALENCIA A; Et al. Estado del arte de los recursos genéticos vegetales en Colombia: Sistema de Bancos de Germoplasma. *Corpoica Cienc. Tecnol. Agropecu*, 2010. 11(1): p. 85-94.

VAZQUEZ A; et al. Potencial de los Marcadores Moleculares para el Rescate de Individuos de *Theobroma cacao* L. de Alta Calidad. *Revista de la Sociedad Mexicana de Biotecnología y Bioingeniería A.C.*, 2012. Vol. 16 No 1.

WEIR, B; Et al. Estimating F-statistics for the analysis of population structure. *Evolution.*, 1984.38:: p. 1358-1370.

WRIGHT, S. Interpretation of population structure by F-statistics with special regard to system of mating. *Evolution.*, 1965. 19: p. 395-420.

ZAPATA, C. La variabilidad Genética de las poblaciones. Espinosa de los monteros, J. *genética de la acuicultura*. 1987: p. 33-57.

ZHANG, D; et al. Molecular characterization of an international cacao collection using microsatellites markers. *Tree Genetics & Genomes*, (2009). 5: p. 1–10.

ZHANG, D; et al. Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellites DNA markers *Tree Genetics & Genomes*, (2009). 5: p. 595–607.

**Anexo A:** Cultivares de cacao criollo recolectados en el Banco Nacional de Germoplasma de la estación experimental la Suiza. (CRICF) CRIOLLO CORPOICA FEDECACAO Colectores (Coronado, Chacón y Baéz)

<b>Cultivar</b>	<b>Origen</b>	<b>Tipo</b>
CRICF 1	Magdalena (Colombia)	Criollo
CRICF 2	Magdalena (Colombia)	Criollo
CRICF 3	Magdalena (Colombia)	Criollo
CRICF 4	Magdalena (Colombia)	Criollo
CRICF 5	Magdalena (Colombia)	Criollo
CRICF 6	Magdalena (Colombia)	Criollo
CRICF 7	Magdalena (Colombia)	Criollo
CRICF 8	Magdalena (Colombia)	Criollo
CRICF 9	Magdalena (Colombia)	Criollo
CRICF 10	Magdalena (Colombia)	Criollo
CRICF 11	Magdalena (Colombia)	Criollo
CRICF 12	Magdalena (Colombia)	Criollo
CRICF 13	Magdalena (Colombia)	Criollo
CRICF 14	Magdalena (Colombia)	Criollo
CRICF 15	Magdalena (Colombia)	Criollo
CRICF 16	Magdalena (Colombia)	Criollo
CRICF 18	Magdalena (Colombia)	Criollo
CRICF 19	Magdalena (Colombia)	Criollo
CRICF 20	Guajira (Colombia)	Criollo
CRICF 24	Guajira (Colombia)	Criollo
CRICF 25	Guajira (Colombia)	Criollo

<b>Cultivar</b>	<b>Origen</b>	<b>Tipo</b>
CRICF 26	Guajira (Colombia)	Criollo
CRICF 27	Guajira (Colombia)	Criollo
CRICF 28	Guajira (Colombia)	Criollo
CRICF 30	Guajira (Colombia)	Criollo
CRICF 31	Guajira (Colombia)	Criollo
CRICF 32	Guajira (Colombia)	Criollo
CRICF 33	Guajira (Colombia)	Criollo
CRICF 35	Guajira (Colombia)	Criollo
CRICF 37	Guajira (Colombia)	Criollo
CRICF 38	Guajira (Colombia)	Criollo
CRICF 39	Guajira (Colombia)	Criollo
CRICF 40	Cesar (Colombia)	Criollo
CRICF 41	Cesar (Colombia)	Criollo
CRICF 42	Cesar (Colombia)	Criollo
CRICF 43	Cesar (Colombia)	Criollo
CRICF 44	Cesar (Colombia)	Criollo
CRICF 45	Cesar (Colombia)	Criollo
CRICF 46	Cesar (Colombia)	Criollo
CRICF 47	Cesar (Colombia)	Criollo
CRICF 48	Cesar (Colombia)	Criollo
CRICF 49	Cesar (Colombia)	Criollo
CRICF 50	Cesar (Colombia)	Criollo
CRICF 51	Cesar (Colombia)	Criollo
CRICF 52	Cesar (Colombia)	Criollo

<b>Cultivar</b>	<b>Origen</b>	<b>Tipo</b>
CRICF 53	Cesar (Colombia)	Criollo
CRICF 54	Cesar (Colombia)	Criollo
CRICF 55	Cesar (Colombia)	Criollo
CRICF 56	Cesar (Colombia)	Criollo
CRICF 57	Cesar (Colombia)	Criollo
CRICF 58	Cesar (Colombia)	Criollo
ICS 60	<b>TRINIDAD</b>	Trinitario
IMC 67	<b>PERÚ</b>	Amazónico

## **Anexo B: Protocolo de Extracción de ADN**

- 1.)** Adicionar 400µL de buffer AP1 y 4 µL de Rnasa de (100mg/ml) a 20mg de tejido pulverizado de la planta y dar vortex vigorosamente.
- 2.)** Incubar la mezcla por 10 min a 65°C Mezclar 2 o 3 veces durante la incubación por inversión. Este paso es para la lisis de la célula.
- 3.)** Adicionar 130 µL de buffer AP2 al lisado anterior, mezclar e incubar en hielo por 5min.
- 4.)** Centrifugar por 5 min a 20000g o 14000rpm.
- 5.)** Pasar el sobrenadante a la columna lila QIA shredde mini spin y centrigugar 2 min a 20000 g o 14000 rpm.
- 6.)** Transferir el flujo que pasa a través de la columna en el paso anterior a un nuevo tubo evitando el pellet, se recolectan aproximadamente 450 µL de sobrenadante.
- 7.)** Adicionar 1.5 volúmenes de buffer AP3/E del sobrenadante.
- 8.)** Adicione 650 µL de la mezcla del paso anterior, incluyendo el precipitado que se pudo haber formado en la columna Dneasy mini spin y centrifugar por 1 min a 6000 g o 8000 rpm y descartar el flujo que pase a través.
- 9.)** Repetir el paso anterior en el mismo tubo.
- 10.)** Colocar la mini spin columna en un nuevo tubo y adicionar 500 µL de Buffer AW y centrifugar por 1 min a 6000g o 8000rpm descartar el flujo y reusar el tubo en el paso 11.
- 11.)** Adiciones 500µL de Buffer AW y centrifugue por 2 min a 20000g o 14000rpm para secar la membrana.
- 12.)** Transferir la mini spin columna a un tubo de 2 ml y pipetear 50 µL de Buffer AE directamente sobre la mambrana de la columna. Incubar por 5 min a temperatura ambiente y centrifugar por 1 min a 6000g o 8000rpm.
- 13.)** Repetir el paso 12.

**Anexo C:** protocolo de tinción con plata.

Colocar el gel de poliacrilamida en una bandeja conteniendo solución de fijación (Solución al 10% de ácido acético glacial) e incubar en agitación constante durante 30 min alternativamente, el gel puede dejarse en solución de fijación toda la noche.

Retirar el gel de la solución de fijación y colocar en un recipiente con agua destilada.

Lavar el gel con agua desionizada agitando constantemente durante 2min, repetir esta operación 2 veces.

Retirar el gel del recipiente y dejar secar por 10 o 20 seg.

Colocar el gel en la solución de tinción (0.1g de  $\text{AgNO}_3$  en 100mL de  $\text{H}_2\text{O}$ ) y dejar en agitación constante durante 30 minutos.

Retirar el gel de la solución de tinción y lavar con agua desionizada.

Añadir en el momento 150  $\mu\text{L}$  de Formaldehído al 37% a la solución de revelado (solución de NaOH al 10%).

Inmediatamente colocar el gel en la solución de revelado y dejar en agitación constante hasta que empiece a aparecer las bandas.

Cuando las bandas hayan alcanzado la intensidad deseada agregar una solución fría 4°C de reactivo de parada (el mismo de fijación) y dejar en agitación hasta que no aparezcan burbujas.

**Anexo D:** Alelos para cada locus (entre paréntesis frecuencias alelicas)

	VARIANZA DE LA FRECUENCIAS ALELICAS (%)								
MTcCIR18	300(93.90)	310(6.10)							
MTcCIR40	266(4.88)	270 ( 68.29)	277 (24.39)	279 (2.44)					
MTcCIR22	300 (100)								
MTcCIR24	186 (92.68)	200 (7.32)							
MTcCIR15	231(3.66)	235 (2.44)	240 (1.22)	250 (92.68)					
MTcCIR33	267(2.44)	273 (29.27)	274 (26.83)	275 (34.15)	280 (7.32)				
MTcCIR1	125 (79.27)	126(7.32)	138 (12.20)	139 (1.22)					
MTcCIR60	187(1.22)	188(32.93)	189(6.10)	191 (1.22)	192 (2.44)	193(50.0)	210(1.22)	211(2.44)	212(2.44)
MTcCIR11	280(86.59)	300(12.20)	346(1.22)						
MTcCIR12	187(1.88)	188(34.15)	189(6.10)	190(2.44)	192(1.22)	193(46.34)	215(2.44)	231(1.22)	258(1.22)
MTcCIR26	300(93.90)	320 (6.10)							
MTcCIR37	169(13.41)	170(42.68)	180(4.88)	189(2.44)	19(18.29)	191(14.63)	192(2.44)	250(1.22)	
MTcCIR6	250(97.56)	260(1.22)	290(1.22)						
MTcCIR8	289(37.80)	290(31.71)	294(1.22)	298(17.07)	301(4.88)				

**Anexo E:** Matriz de distancias genéticas estándar de Nei entre las subpoblaciones (POB)

	Magdalena	Guajira	Cesar	Peru	Trinidad
Magdalena	0,00				
Guajira	0,09	0,00			
Cesar	0,11	0,05	0,00		
Peru	0,37	0,38	0,44	0,00	
Trinidad	1.669	1.610	1.888	1.452	0,00