



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**CARACTERIZACIÓN DEL PERFIL DE ANCESTRÍA EN
UNA MUESTRA DE LOS TRES PRINCIPALES GRUPOS
ÉTNICOS (AMERINDIOS, CAUCÁSICO-MESTIZOS Y
AFRODESCENDIENTES) DE LA POBLACIÓN
COLOMBIANA, UTILIZANDO MARCADORES
INFORMATIVOS DE ANCESTRÍA – INDEL**

Jair Arley Pérez Cárdenas

Universidad Nacional de Colombia
Facultad, Departamento de Morfología,
Maestría en Genética Humana
Bogotá, Colombia

2023

**CARACTERIZACIÓN DEL PERFIL DE ANCESTRÍA EN
UNA MUESTRA DE LOS TRES PRINCIPALES GRUPOS
ETNICOS (AMERINDIOS, CAUCASICO-MESTIZOS Y
AFRODESCENDIENTES) DE LA POBLACIÓN
COLOMBIANA, UTILIZANDO MARCADORES
INFORMATIVOS DE ANCESTRÍA – INDEL**

Jair Arley Pérez Cárdenas

Tesis presentada como requisito parcial para optar al título de:
Magister en Genética Humana

Director:

Juan José Yunis Londoño MD, MSc,
Profesor Titular

Departamento de Patología,
Facultad de Medicina e Instituto de Genética
Universidad Nacional de Colombia

Codirectora:

Luz Karime Yunis Hazbun MD, MSc, PhD (c).
Directora Médica área Citogenética y Microarreglos
Servicios Médicos Yunis Turbay y Cia SAS.

Universidad Nacional de Colombia
Facultad, Departamento de Morfología,
Maestría en Genética Humana
Bogotá, Colombia

2023

A mis padres, por su amor, apoyo incondicional y sacrificio incansable. A mi increíble hermana, por su constante aliento y compañerismo. A mis valiosos amigos y seres queridos, quienes han estado a mi lado en cada paso del camino. Al Dr. Juan, cuya paciencia, orientación y dedicación han sido fundamentales para el éxito de este trabajo. A todos ustedes, ¡Gracias! porque cada uno ha sido parte esencial en este camino hacia el conocimiento y el crecimiento personal.

"El mérito no está en lo que se logra, sino en la dificultad que se ha superado para lograrlo." - Marina Tsvietáieva

"El éxito en la vida no es solo acumular logros, sino también apreciar y reconocer a aquellos que nos han ayudado en el camino." - Byron Pulsifer

Resumen

CARACTERIZACIÓN DEL PERFIL DE ANCESTRÍA EN UNA MUESTRA DE LOS TRES PRINCIPALES GRUPOS ÉTNICOS (AMERINDIOS, CAUCASICO-MESTIZOS Y AFRODESCENDIENTES) DE LA POBLACIÓN COLOMBIANA, UTILIZANDO MARCADORES INFORMATIVOS DE ANCESTRÍA – INDEL

Colombia es uno de los países más diversos a nivel cultural y étnico, los diferentes procesos migratorios históricos y dinámicas sociopolíticas del país han permitido que en la actualidad exista una población con diferentes patrones de mezcla derivada de tres grupos étnicos principales: europea, nativo americana y africana. La mayoría de los estudios previos se han enfocado en caracterizar la población caucásicos mestiza colombiana utilizando diferentes tipos de marcadores informativos de ancestralidad (AIM), sin embargo, son pocos los estudios que han analizado las poblaciones amerindias y afrocolombianas. En el presente estudio, se analizaron un total de 153 individuos amerindios (AMR) de 28 comunidades indígenas de Colombia, 164 afrocolombianos (AFC) originarios de diferentes regiones del departamento del Chocó y 150 individuos de ancestralidad mixta (MIX) de la región andina central. Se utilizó un panel de 46 marcadores de tipo AIM-INDEL amplificados en una PCR multiplex única y se analizaron mediante electroforesis capilar con el software GeneMapper ID. Las frecuencias alélicas, el análisis del equilibrio de Hardy-Weinberg y el análisis de diversidad y distanciamiento genético (F_{st}) se calcularon utilizando el software Arlequin (v.3.5.2.2) y Genepop (v.4.6.) La proporción de ancestralidad se estimó a través del programa STRUCTURE (v.2.3.4.21.)

Para la población con mezcla, la proporción del componente europeo fue del 57,8%, un 36,3% para el nativo americano y 5,9% africano. En la población AMR, la proporción promedio nativo americana fue estimada en un 95%; el componente europeo fue de 3,1% y africano de 1,5%, observándose algunas diferencias entre las comunidades según localización geográfica. En población AFC, la proporción africana fue la principal con 81,1%, seguida de la proporción europea 10,5% y un 8,4% para la nativo americana, sin embargo, la estimación europea para algunas poblaciones afrodescendientes fue mayor que el promedio, entre un 25-30%.

Estos resultados amplían y destacan la diversidad genética que las poblaciones amerindias y afrocolombianas pueden tener entre sí, reflejando sus distintos procesos históricos y culturales. Nuestros resultados indican que ocurrieron diferentes dinámicas de mezcla dentro del departamento del Chocó, lo que sugiere, que no se debería de utilizar una población única afrocolombiana como referencia, ya que grandes diferencias en la proporción de ancestralidad pueden encontrarse en áreas geográficamente cercanas y culturalmente relacionadas. El trabajo presentado aquí contribuye al esfuerzo continuo por documentar la variabilidad genética humana en Colombia, al enfocarse en explorar poblaciones poco estudiadas en análisis previos de estructura genética colombiana.

Palabras Clave: Ancestría, Colombia, genética de poblaciones, estructura genética

Abstract

Characterization of Ancestry Profile in a Sample Representing the Three Main Population Groups in Colombia (Amerindians, Caucasian-Mestizos, and Afrodescendants), employing Ancestry Informative Markers - INDEL.

Colombia is one of the most culturally and ethnically diverse countries, with various historical migratory processes and socio-political dynamics have led to a population exhibiting diverse patterns of genetic admixture from three main ethnic groups: European, native American and African. While previous studies have predominantly focused on characterizing the Colombian European admix population using various Ancestry Informative Markers (AIM), there have been limited investigations into the Amerindian and Afro-Colombian populations. In this study, a total of 153 Amerindian individuals (AMR) from 28 indigenous communities in Colombia, 164 Afro-Colombians (AFC) originating from different regions of the Chocó department, and 150 Admixture ancestry individuals (MIX) from the central Andean region were analyzed. A panel of 46 AIM-INDEL markers was amplified in a single multiplex PCR and analyzed on an ABI3500 using GeneMapper ID. Allele frequencies, Hardy-Weinberg equilibrium analysis, and genetic diversity and distance analysis (F_{st}) were calculated using Arlequin (v.3.5.2.2) and Genepop (v.4.6.). Ancestry proportions were estimated using STRUCTURE (v.2.3.4.21).

In the MIX population, the European proportion was 57.8%, Native American 36.3%, and African 5.9%. Among AMR individuals, the average Native American proportion was estimated at 95%, the European and African proportions were 3.1% and 1.5%, respectively. Nevertheless, some variations were observed among communities based on geographical location. Respecting AFC population, the African proportion was predominant at 81.1%, followed by European at 10.5%, and Native American at 8.4%. However, some Afro-descendant populations exhibited significant proportions from European ancestry ranging to 25-30%.

These results underscore the genetic diversity present among Amerindian and Afro-Colombian populations, reflecting their distinct historical and cultural processes. Our findings indicate diverse mixing dynamics within the Chocó department, suggesting that a single reference population of Afrocolombian should not be used due to significant differences in ancestry proportions in geographically proximate areas. This study contributes to ongoing efforts to document human genetic variability in Colombia by focusing on exploring populations understudied in previous structure genetic analyses.

Keywords: Ancestry, Colombia, Population genetics, genetic structure

Tabla de contenido

Pág.

ii. Introducción	7
1. Marco Teórico	10
1.1 Análisis de la Ancestría Genética en el estudio de la Sub-Estructura poblacional. .	10
1.2 Estudio del poblamiento americano y Colombia	15
1.3 Estudios previos de análisis de ancestralidad en Colombia.....	17
1.4 Importancia del estudio de la subestructura poblacional a través de la ancestralidad genética en la población colombiana.....	24
2. Objetivos	29
2.1 Objetivo general:	29
2.2 Objetivos Específicos:	29
3. Metodología.....	30
3.1 Población de Estudio.....	30
3.2 Extracción de ADN genómico.....	30
3.3 Amplificación PCR Multiplex.....	31
3.3.1 Primers	31
3.3.2 Condiciones de Reacción para la PCR	33
3.3.3 Condiciones de Termociclado.....	34
3.4 Electroforesis Capilar	34
3.5 Análisis Estadístico	34
4. RESULTADOS	37
4.1 Genotipificación y Descripción de la Población.....	37
4.2 Análisis Genético Descriptivo y Diversidad Poblacional:.....	39
4.3 Análisis de la Subestructura y Distanciamiento Genético:	44
4.4 Estimación de Componentes Ancestrales:	50
5. DISCUSIÓN:.....	62
6. CONCLUSIONES:.....	70
7. LISTADO DE TABLAS Y FIGURAS.....	72
7.1 TABLAS	72
7.2 FIGURAS	73
8. Bibliografía	74

Introducción

Colombia fue reconocido como país independiente desde 1819, cuando inicialmente recibió el nombre de La Gran Colombia, posteriormente, debido a diversos cambios Político-administrativo que llevaron a separaciones topográficas, quedo constituida la nación colombiana actual. Por tanto, se hace imperativo describir y entender los procesos históricos, antropológicos y biológicos que han constituido lo que conocemos hoy como La República de Colombia. Las dinámicas políticas y etnográficas asociadas a los diferentes procesos de conquista, colonia, y procesos migratorios ancestrales y recientes (socio-políticos y laborales) han generado que Colombia presente una de las poblaciones más diversas, lo cual es evidente entre sus distintas regiones biogeográficas, dado la presencia de diferentes grupos étnicos y una notable multiculturalidad interpoblacional. (Rodríguez et al., 2007).

En el ámbito genético y cultural Colombia recibe el aporte de tres principales grupos étnicos ancestrales: nativos americanos, europeos y africanos. (Mesa et al., 2000). Diversos estudios han estado enfocados en analizar la diversidad genética y estructura de algunas poblaciones colombianas utilizando diferentes tipos de marcadores informativos de ancestralidad.

Para evaluar la estructura genética y el nivel de mezcla étnica de las poblaciones, inicialmente se usaron marcadores como el sistema HLA y los grupos sanguíneos. (Sandoval, De la Hoz, & Yunis, 1993; Yunis et al., 1994, 2001). Posteriormente, se incorporaron otros marcadores como el ADN mitocondrial, los Haplotipos STR de Cromosoma Y, el uso de marcadores STR autosómicos entre otros. (Mesa et al., 2000; Rojas et al., 2010; Yunis, 2013; Yunis et al., 2005, 2013; Alonso et al., 2012; Lamprea, N., 2009; Rojas et al., 2013; Alonso et al., 2018; Moncada J, 2018). Con los avances de la tecnología y la identificación de nuevos marcadores, entre ellos, los marcadores InDel, se demostró que estos marcadores eran muy útiles para el análisis de ancestralidad con lo cual se convirtieron en los marcadores predilectos para caracterizar el grado de mezcla de las poblaciones. (Pereira et al., 2009; N. P. Santos et al., 2010)

Los marcadores de tipo InDel se utilizan actualmente para los estudios de mezcla debido a que son muy frecuentes en el genoma humano y presentan diversas ventajas que los hace útiles a la hora de estudiar la variación y mezcla de las poblaciones humanas. Dentro de sus ventajas, permiten mejorar aquellas características no deseadas de los marcadores de tipo STR y SNV, tales como, bajas tasas de mutación y amplicones pequeños respecto a los STRs, y una fácil interpretación y requerimiento de menor número de marcadores comparados con los SNV. Adicionalmente, la posibilidad de analizar múltiples marcadores simultáneamente al utilizar PCR multiplex. (Da Costa Francez, Rodrigues, de Velasco, & dos Santos, 2012; Manta et al., 2013; Pereira et al., 2012; C. Santos et al., 2015)

En estudios poblacionales previos utilizando marcadores de ancestría en poblaciones colombianas, no se han utilizado muestras seleccionadas con base en criterios que permitan descartar un componente de mezcla previo entre los 3 grupos étnicos ancestrales, en especial de la población Amerindia y Afrodescendiente. En este contexto, varios estudios han caracterizado poblaciones caucásico-mestizas en el país. (Ibarra, Freire-Aradas, et al., 2014; Noguera et al., 2014; Ossa et al., 2016; W. Rojas et al., 2010; Xavier et al., 2015) Sin embargo, no se cuenta con estudios de muestras Amerindias y Afrodescendientes con bajo grado de mezcla que permitan tener de base, poblaciones de referencia autóctonas que sean representativas de Colombia para estudios de mezcla posteriores. La mayoría de dichos estudios, han utilizado poblaciones mezcladas para caracterizar el acervo amerindio, como es el caso de la utilización de poblaciones Wayuu, Zenú, poblaciones del Tolima y otras, incluidas las muestras de amerindios utilizadas en el proyecto de 1000 genomas. (Alonso Morales et al., 2018; Ibarra, Freire-Aradas, et al., 2014; Mesa et al., 2000; Moncada J, 2018; Ossa et al., 2016; Rondon et al., 2008; Salzano et al., 2014; Xavier et al., 2015). Muy pocas como en el caso de Ossa et al., 2015 han analizado una población pequeña de etnias conservadas, entre las cuales están Desano, Puinave, Cubeo y Tucano.

Por consiguiente, se busca a través de diferentes metodologías (árbol genealógico, ubicación geográfica y análisis previos de marcadores uniparentales y autosómicos) determinar individuos con menor grado de mezcla que puedan ser utilizados como representación de la población amerindia y afrodescendiente Colombia para servir como futuras poblaciones de referencia.

Este estudio, hace parte de un proyecto Macro titulado “*Caracterización molecular y farmacogenética en una muestra de pacientes con Leucemia Mieloide Aguda y su correlación con la estratificación del riesgo y respuesta al tratamiento. Una aproximación hacia los patrones moleculares de la Leucemia Mieloide Aguda Pediátrica en población colombiana*” que tiene como uno de sus objetivos analizar si el componente de ancestría de los pacientes con leucemia mieloide aguda (LMA) pediátrica se relaciona con variantes farmacogenéticas que ayuden a determinar el riesgo y la respuesta al tratamiento de estos pacientes, para lo cual es necesario tener estudios de poblaciones propias con bajo grado de mezcla que puedan posteriormente utilizarse como referencias adecuadas.

Por lo tanto, en el presente estudio se tiene como objetivo caracterizar el perfil de ancestría y analizar el nivel de subestructura genética de los 3 principales grupos étnicos Amerindios, Afrodescendientes y caucásicos mezclados en una muestra de población colombiana, en donde a través de una metodología de muestreo, encuestas demográficas y socio-culturales e información de ancestralidad previa (análisis previos de marcadores uniparentales y autosómicos), se pueda realizar un selección de individuos con bajo grado de mezcla previa en los cuales se pueda evaluar si se presentan diferencias en la estimación de la proporción de ancestría, comparándolos con las muestras de individuos utilizados como de referencia para poblaciones ancestrales en las bases de datos internacionales.

1. Marco Teórico

1.1 Análisis de la Ancestría Genética en el estudio de la Sub-Estructura poblacional.

La extensa variabilidad genética que existe entre los individuos es una de las principales características de las poblaciones humanas. La gran variabilidad es favorecida por las diversas mutaciones que se generan en el genoma humano y es forjada a través de la dinámica de diferentes fuerzas de cambio evolutivo como la selección, el flujo génico (migración) y la deriva genética. Por tanto, cada individuo presentará un genoma específico con diversas características y variantes. (Hartl & Grant, 2007)

El estudio de la variabilidad genética dentro de las poblaciones humanas, no solo nos permite conocer la evolución de la especie, sino también nos ayuda a dilucidar y correlacionar los aspectos históricos de las poblaciones. (Wallace D.C., 1995) Adicionalmente, el conocimiento de la estructura poblacional es relevante para estudios de asociaciones genómicas que nos permitan a futuro detectar variantes asociadas con enfermedades de alto impacto en la salud pública, como el cáncer, entre otras. (Winkler, Nelson, & Smith, 2010)

Para el estudio de la subestructura genética, uno de los métodos más comúnmente utilizados es el análisis del estadístico F , planteado por Sewall Wright, quien mostro un interés principal en el estudio de la estructura demográfica-poblacional y sus efectos en la evolución. Wright se propuso estudiar la relación entre la demografía y la estructura genética, donde infirió el método del índice de fijación F . (Slatkin, 1994)

Los índices de fijación F , se pueden emplear para determinar grados de diferenciación entre y dentro de las poblaciones, estos índices utilizan las frecuencias alélicas de cada marcador dentro de una población para estimar un grado de similitud genética entre poblaciones. El estadístico F es uno de

los parámetros principalmente utilizados en la estadística descriptiva para el estudio de la genética de poblaciones, el cual es una medida de la correlación entre los genes en una subpoblación con relación a la población general. (Slatkin, 1994)

Al ser un estimador, este se ve afectado por el tipo de marcador informativo utilizado, de esta manera, su interpretación es dependiente del tipo de marcador, ya que no siempre guarda relación directa con el equilibrio de Hardy-Weinberg (Holsinger & Weir, 2009). En genética de poblaciones, se puede utilizar el índice de fijación F para evaluar diferentes niveles de subestructura, dentro de estos existen: Población total (t) respecto a una subpoblación (F_{st}), Población total respecto a una región (R) y una región (R) respecto a una subpoblación (F_{sr}). Se determina que una población presenta algún nivel de subestructura cuando se evidencia una discrepancia de las frecuencias alélicas dentro de la población; como resultado, se evidenciará una variación en la distribución de frecuencias alélicas, y una posible reducción de la proporción de heterocigotos esperados entre las subpoblaciones (Hartl & Grant, 2007).

Evaluar la subestructura poblacional es una característica que se puede apreciar fácilmente en grupos poblacionales grandes, a través del uso de diferentes marcadores genéticos. Inicialmente, se usaron marcadores como el sistema HLA y grupos sanguíneos, aprovechando su patrón de herencia mendeliana y su distribución poblacional en sus diversas formas alélicas (Sandoval et al., 1993; Yunis et al., 1994, 2001). Las nuevas tecnologías trajeron consigo el avance de marcadores moleculares que ofrecen mayores ventajas (Tabla 1), que, por sus diferentes características y patrones de herencia, podrán utilizarse en función del objetivo del tipo de estudio a llevar a cabo. Dentro de estos nuevos marcadores, los más utilizados se pueden dividir entre marcadores uniparentales (el ADN mitocondrial y STRs de Cromosoma Y) y marcadores Autosómicos (STRs Autosómicos, Análisis de SNVs y Análisis de InDel). (Pereira, 2011; Vali, Brandstrom, Johansson, & Ellegren, 2008; Zaumsegel, 2013)

Tabla 1. Ventajas y Desventajas Según Tipo de Marcador de Ancestralidad.

Tipo de Marcador	Ventajas	Desventajas
mt-DNA	Linaje Específico, Resistencia a la degradación, Múltiples copias por célula.	No Información Completa Mapa Genético-Poblacional
Y-STR	Linaje Específico, Reconstrucción de Pedigrees.	Presentes solo en Hombres, No Información Completa Mapa Genético-Poblacional, Ligamiento entre Marcadores
STR Autosómico	Alta Variabilidad, Fácil Detección, Deconvolución de las Mezclas	Inestabilidad Genética, Susceptibles a Degradación del ADN
SNV	Tamaño de Fragmento Pequeño, Estabilidad Genética, Alto Potencial para Automatización	Baja Variabilidad, Mayor Dificultad y Consumo de Tiempo para la Detección
InDel	Tamaño de Fragmento Pequeño, Estabilidad Genética, Fácil Detección, Alto Potencial para Automatización	Baja Variabilidad

Tomada y Modificada de (Zaumsegel, 2013)

En la actualidad, el uso de marcadores uniparentales es utilizado frecuentemente para obtener el análisis de la ancestralidad de origen materno y paterno, de manera independiente. Una desventaja es que, al solo definir el análisis ancestral parental, su estudio representa una pequeña parte del mapa genético-poblacional que puede aportar el individuo. (Saloum de Neves Manta et al., 2013)

Los marcadores autosómicos, inicialmente fueron considerados dentro del estudio de la genética forense, pero su uso se ha expandió a la genética de poblaciones, los principalmente empleados son los STR y los SNV. No obstante, nuevos tipos de marcadores como los polimorfismos de inserción/delección (InDel) han surgido como potenciales informadores de ancestralidad, que permiten combinar de una manera deseable las características particulares de los STR y SNV. (Pereira, 2011; Pereira et al., 2009; Zaumsegel, 2013)

Se resalta la importancia que a pesar de que existe gran variedad de marcadores genéticos, su selección va en razón de la información genética que deseamos obtener de cada población, por lo tanto, su utilidad e interpretación varía en función del cumplimiento del objetivo del estudio que nos planteemos. Por ejemplo, los STRs no siempre se comportan como marcadores de diferenciación poblacional en el ámbito forense, ya que dependerá de las frecuencias alélicas de cada población, donde puede verse afectada la distribución de alelos y tomar un comportamiento hacia la normalidad. (Hartl & Grant, 2007)

Los marcadores InDel son polimorfismos largos utilizados actualmente dentro del campo de la identificación humana y la genética de poblaciones, debido a que poseen características que los definen como marcadores genéticos atractivos; dentro de sus ventajas están: Ampliamente extendidos en el genoma; presentes tanto en cromosomas autosómicos como sexuales, el polimorfismo es resultado de un único reordenamiento (Inserción o Delección), permiten determinar de manera significativa las diferencias en las frecuencias alélicas entre diferentes grupos poblacionales separados geográficamente y pueden analizarse en amplicones cortos, que favorecen la amplificación en muestras degradadas.

Por todo lo anterior, los InDel pueden ser utilizados en genética de poblaciones como marcadores informativos de ancestralidad (AIMs). Adicionalmente, el análisis de InDel puede llegar a ser superior a nivel práctico en comparación a los demás marcadores autosómicos debido a la combinación de las siguientes propiedades: una menor tasa mutacional, una fácil interpretación, la posibilidad de analizar de manera simultánea varios marcadores, obtener resultados rápidamente y la capacidad de adaptación para la amplificación de muestras con ADN en una concentración baja o degradado (Pereira et al., 2009; Santos et al., 2010; Da Costa et al., 2012).

Por otro lado, si conocemos el contexto histórico y político-social de cada país, podemos inferir el grado de mezcla de su población, motivo por el cual, si hablamos de países latinoamericanos donde existieron diferentes procesos que favorecieron el intercambio génico, como los períodos de conquista y esclavitud, así como procesos de guerras, violencia interna y migraciones forzadas, podemos concluir que su población es producto de mezcla entre diferentes poblaciones ancestrales. Para la mayoría de estos países sus poblaciones son heterogéneas, con diferentes proporciones de ancestralidad de origen europeo, africano y nativo americano; y para algunos una cuarta población ancestral del este de Asia como en Brasil, donde hubo un proceso de migración japonesa a principios del siglo XX. (Andrade et al., 2018)

En el estudio de LaRue et al, 2014 se observa cómo al aplicar los criterios de Equilibrio de Hardy Weinberg, distribución de frecuencias alélicas, subestructura poblacional y estudio de desequilibrio de ligamiento permite identificar de una manera robusta que los polimorfismos InDel son los más precisos para ser incorporados dentro de un set de panel AIM-InDel. Estas apreciaciones las podemos identificar en el panel de 46 AIM-InDel configurado por Pereira et al, 2012 en donde se determinó que los patrones de variabilidad de los InDel utilizados para las poblaciones Europea, Africana, Nativo americana y del Este asiático presentaban diferencias en la distribución de frecuencias alélicas y distanciamiento genético que permitían diferenciar entre al menos dos de los grupos poblacionales comparados, con valores del diferencial de frecuencia (δ) ≥ 0.3 o ≥ 0.4 . Adicionalmente, en los marcadores InDel no se evidenciaron desviaciones del equilibrio de Hardy Weinberg, ni presentaban asociaciones de ligamiento entre ellos.

El uso de AIM-InDel ha demostrado que pueden identificar una distribución donde alelos específicos están fijados en una población ancestral determinada en comparación a las demás, o su frecuencia es extremadamente rara, mientras en otra población es frecuente. De esta manera, la divergencia de las frecuencias alélicas en los InDel para cada población nos permite inferir el probable origen ancestral de un individuo o estimar la proporción de los

componentes de ancestralidad en individuos o poblaciones mezcladas. (LaRue et al., 2014; Manta et al., 2013; Pereira et al., 2012; N. P. Santos et al., 2010)

1.2 Estudio del poblamiento americano y Colombia

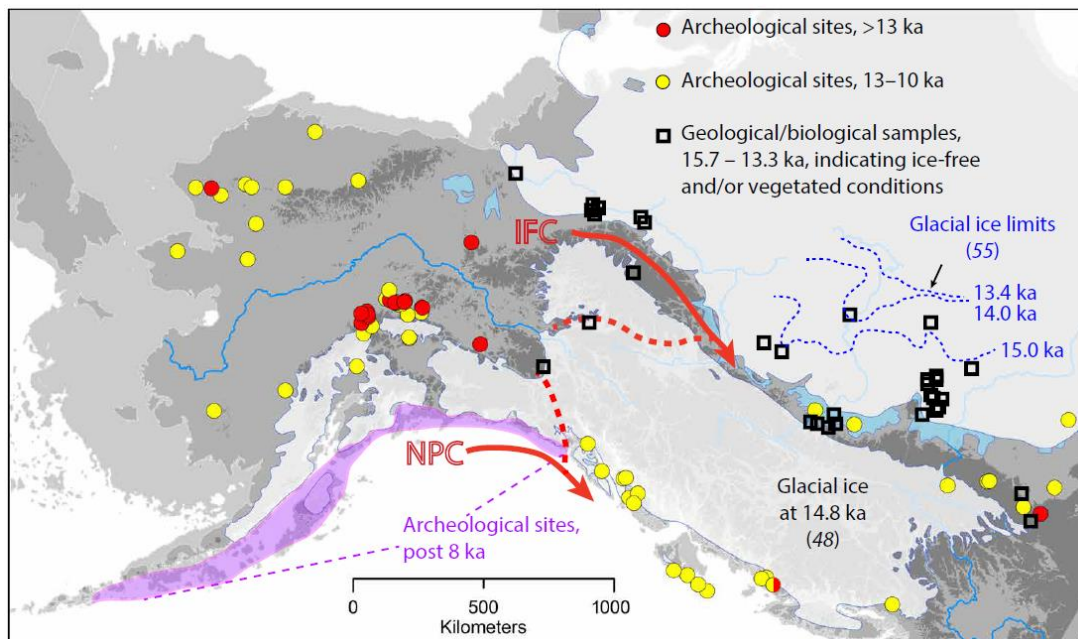
El poblamiento del continente americano es todavía tema de estudio e investigación, existen diversas teorías que aún no logran colocarse de acuerdo a los diferentes arqueólogos y antropólogos del mundo. La ocupación americana por parte de los pueblos originarios del norte del continente fue promovida por la finalización de la última era de glaciación, que permitió que se formara la masa terrestre llamada Beringia, la cual unió la región noreste de Asia con el noroeste de Norteamérica. Las múltiples teorías del poblamiento sugieren que el hombre americano provino por 3 posibles rutas: desde el noreste asiático/Siberia, desde la zona más oriental de Alaska o por el océano pacífico. (Pedersen et al., 2016; Potter, 2018).

El análisis más reciente y exhaustivo del ADN mitocondrial indica que los ancestros nativos americanos divergieron de las poblaciones de Siberia entre 25.000 a 18.000 años Ap (Llamas, 2016). Estos datos se correlacionan con los diversos hallazgos arqueológicos y antropológicos que se asocian con las migraciones de los antecesores a través de Beringia; En la actualidad, se consideran dos potenciales rutas de migración: La costa Norte del pacífico (NPC) y el Corredor Libre de Hielo (IFC) (Figura 1) (Potter, 2018).

Con respecto a los primeros pobladores suramericanos, en 1964 Bennett y Bird mediante evidencia arqueológica y geográfica de la época, propusieron un modelo donde se reconstruyen los movimientos poblacionales de los paleoindios. Según estos autores, los cazadores nómadas, pescadores y recolectores cruzaron el istmo de Panamá y entraron en la región norte de los Andes a través de los ríos Cauca y Magdalena territorio que hoy es Colombia (Rothhammer & Silva, 1989). Sin embargo, este modelo no ha podido ser confirmado en su totalidad por falta de evidencia arqueológica; a la fecha dicho

modelo se ha sustentado principalmente en argumentos geográficos, así como en datos lingüísticos y genéticos moleculares a partir de poblaciones actuales (Roewer, Nothnagel, Gusmao, Gomes, & Gonzalez, 2013; Rothhammer & Silva, 1989; Alonso Morales, 2013). En la historia del poblamiento de Colombia se ha sugerido 3 rutas: una siguiendo la Costa Pacífica, la segunda dentro de las montañas colombianas a través de los valles y del río Magdalena, y, por último, la ruta de la Costa Atlántica. (Alonso Morales et al., 2018).

Figura 1. Potenciales Rutas Migratorias de Colonización del Noroeste Americano.



Tomada de (Potter, 2018)

Colombia reconocida como nación independiente desde 1819, se encuentra ubicada en el extremo sur del puente que conecta los dos subcontinentes de las Américas, siendo el país más noroccidental de América del Sur. Limita al norte con el Océano Atlántico; al sur, con Ecuador y Perú; al oriente, con Venezuela y Brasil, y al occidente, con Panamá y el Océano Pacífico (ACEP, 1974). Alcanza una extensión de 1.138.914 Km², donde se definen cinco regiones geográficas naturales (IGAC, 2010), cada una de ellas presenta gran variabilidad a nivel topográfico y climático, a su vez, cada región exhibe una alta diversidad étnica entre los grupos que la habitan y una notable población multicultural a nivel

interregional. Actualmente se conoce la existencia de 87 etnias indígenas, 4 grupos diferenciados de población afrocolombiana, los que se ubican en el corredor del pacífico colombiano, los raizales del Archipiélago de San Andrés Providencia y Santa Catalina, la comunidad de San Basilio de Palenque y la población que reside en las cabeceras municipales o en las grandes ciudades. También se hablan 64 lenguas amerindias, además de la lengua de los raizales del Archipiélago de San Andrés, Providencia y Santa Catalina, y la lengua criolla de las comunidades de San Basilio de Palenque (Rodríguez et al., 2007).

1.3 Estudios previos de análisis de ancestralidad en Colombia

En el ámbito genético y cultural Colombia recibe el aporte de tres principales grupos poblacionales ancestrales: Nativos americanos, Europeos y Africanos, dado por la interacción de los individuos durante un primer intercambio génico en la conquista española y portuguesa alrededor de 1500 y el comercio de esclavos. (Bortolini, Salzano, Bau, Layrisse, & Petzl-Erler, 2002; W. Rojas et al., 2010; Salzano et al., 2014)

Posteriormente, la afluencia de miles de colonizadores europeos y árabes en el siglo XIX y las migraciones actuales ligadas a la situación sociopolítica nacional, han contribuido al desarrollo de la gran diversidad poblacional y cultural actual del país. Debido a esto, se cree que algunas poblaciones colombianas, a pesar de tener diferentes orígenes étnicos y geográficos, pueden presentar un nivel de subestructura genética considerablemente bajo. Además, algunos grupos aparentemente considerados como comunidades conservadas, en realidad presentan gradientes de mezcla con otros grupos étnicos presentes en el territorio colombiano. (Alonso Morales et al., 2018; M. Y. Rojas, Alonso, Sarmiento, Eljach, & Usaquen, 2013; W. Rojas et al., 2010)

Debido a la gran diversidad humana presente en la población Colombiana y la importancia de conocer nuestros orígenes y dilucidar nuestra estructura poblacional para a futuro continuar estudios que permitieran buscar

asociaciones con enfermedades de alto impacto en salud pública y generar normativas académicas y políticas que mejoren la actuación frente a grupos de riesgo, desde el Grupo de Identificación Humana e Inmunogenética del Instituto de Genética de la Universidad Nacional de Colombia, se creó el proyecto “Estructura Genética de la población Colombiana” Liderado por el Dr. Emilio Yunis entre 1988-1998.

Dentro de los primeros estudios realizados por Sandoval C y Emilio E.J (Sandoval et al., 1993) se analizaron 8 sistemas genéticos de los glóbulos rojos, para la tipificación de grupos sanguíneos en 30.259 individuos de todas las regiones del país, en el cual se estimó el índice de mestizaje de la población, calculando el aporte indígena, caucásico y afrodescendiente, siendo el primer estudio en estimar la estructura poblacional Colombiana. En este estudio, se tipificaron los grupos sanguíneos mayores ABO y Rhesus, y los sistemas menores Kell-Cellano, Duffy, Kidd, Diego, MNSs y Xg. Como resultados, se observó que para el grupo ABO, el alelo más frecuente fue el O en más del 60% de la población estudiada, seguido del A (26,9%) y B (9,3%) y el menos frecuente AB (1.9%).

Tabla 2. Grupos Sanguíneos y sus alelos.

Grupo Sanguíneo	Alelos
ABO	A, B, O,
Rhesus (Rh)	Ro (eDe), R I (Cde), R2 (cDE), Rz (CDE), Ry (CdE), r (ede), r' (Cde), r'' (edE).
Kell	K, k,
Duffy	Fya, Fyb, fy.
Kidd	Jka, Jkb, Jk.
Diego	Dia, Dib.
MNSs	MS, Ms, NS, Ns.
Xg	Xga, Xg.

Tomada de (Sandoval et al., 1993)

Con respecto a la proporción de indígena, negro y blanco para la población global, los resultados evidenciaron una población predominantemente blanca un 61-63%, seguida de indígenas 26-29% y afrodescendiente en un 10-11%. En ese mismo orden de ideas, el aporte indígena fue de predominio para las regiones de la Orinoquia y amazónica. Con relación al 10-11% de proporción afrodescendiente, mayoritariamente fue de la región pacífica y perteneciente a esta última, principalmente del departamento del Chocó.

Dentro del proyecto de “Estructura Genética de la población Colombiana” también se analizó el mestizaje de poblaciones amerindias a través de la clasificación de grupos sanguíneos y el sistema HLA. Estudios realizados por Yunis JJ. Et al, 1994 evidenciaron que utilizando el sistema HLA y tipificación de grupos sanguíneos en 4 poblaciones Amerindias Colombianas, poblaciones Wayuu y Arahucos presentaban un grado importante de mestizaje con alta variabilidad en los haplotipos del sistema HLA no asociados con poblaciones nativo americanas en un 25% y grupo sanguíneos A y B, entre un 2 a 7%. De igual manera, se realizó un estudio por Yunis JJ et al, 2001 evaluando la relación entre 3 etnias (Guambiano, Ingano y Páez) del suroccidente colombiano, donde se evidenció una relación entre las etnias guambiano e Ingano los cuales presentaban un porcentaje de mezcla asociado a proporción africana, por otro lado derivándose de un clado diferente la etnia Páez, se alejaba de estas poblaciones amerindias y se evidenciaba con un bajo grado de mezcla caracterizado por una proporción de ancestría caucásica.

Por otra parte, con el avance de las tecnologías moleculares, aparecieron nuevos marcadores de ancestralidad que trajeron consigo diferentes ventajas. (Zaumsegel, 2013). Dentro de estos, los marcadores uniparentales como el ADN mitocondrial (mtDNA) y STRs del cromosoma Y (Y-Chro), estos al no presentar recombinación, permiten determinar la ascendencia matrilineal y patrilineal, definir genealogías y patrones de proporción del sexo, respecto a eventos históricos asociados a grupos de colonizadores o migrantes. (Jobling, Hurles, & Tyler, 2004; Wilder, Kingan, Mobasher, Pilkington, & Hammer, 2004)

Diversos estudios utilizando marcadores uniparentales han demostrado su utilidad para confirmar la concordancia entre la evidencia histórica del proceso de emparejamiento asimétrico en relación con el sexo y el origen étnico que predominó en la formación de las poblaciones de América del Norte y de América Latina (Mesa et al., 2000; W. Rojas et al., 2010; Salzano et al., 2014; Xavier et al., 2015; Yunis, 2013). Uno de los principales hallazgos a partir del estudio de marcadores uniparentales, es la introducción de genes de origen europeo en la población americana a través de los hombres; ya que históricamente existió un número desigual de uniones entre hombres europeos y mujeres nativas y africanas. Así, el DNA mitocondrial muestra una contribución americana predominantemente amerindia mientras que el cromosoma Y indica una contribución principalmente europea. (Bortolini et al., 2002; Rodas, Gelvez, & Keyeux, 2003; W. Rojas et al., 2010; Tarazona-Santos, Carvalho-Silva, Pettener, Luiselli, & De Stefano, 2001; Xavier et al., 2015; Yunis, 2013; Yunis et al., 2005, 2013).

Estudios como el de Bedoya et al, 2006 en donde se analizaron marcadores uniparentales en población antioqueña, y Rojas et al, 2010, Alonso et al, 2013, Mesa et al 2000 y Rodas et al 2003, quienes analizaron poblaciones indígenas y afrocolombianas; evidenciaron el significativo grado de mezcla de las diferentes etnias y poblaciones analizadas. En el caso de Rojas et al, se observa como las tribus Ingano, Ticuna y Zenú utilizadas presentan un grado considerable de Haplogrupos mitocondriales de origen europeo, por otro lado, la población del Chocó estudiada evidenció alto porcentaje de Haplogrupos de Cromosoma Y de origen europeo.

De igual manera, Mesa et al, 2000, en donde se estimó la proporción de ancestría de una muestra de 5 etnias colombianas (Emberá, Ingano, Ticuna, Wayuu y Zenú) demostró que la poblaciones previamente mencionadas presentaban diferentes grados de mezcla de ancestría africana y europea, para la poblaciones Wayuu y Zenú se encontró un porcentaje de ancestría europea entre 46-62% por STRs de Cromosoma Y, estos datos reflejan lo posteriormente encontrado por Rojas et al, 2013, en donde se analizaron 290

individuos mayores de 18 años y que vivieran más de un año en el departamento de la guajira, en donde se observó la presencia de un gradiente de mezcla en la población residente de la guajira, con diferentes grados de ancestría amerindia. Adicionalmente, definen un grupo genealógico intermedio, Wayuu-Guajiro refutando la idea de una población Wayuu homogénea. (M. Y. Rojas et al., 2013). Estudios como el de Rodas et al, 2003 y Alonso et al, 2018 también han evidenciado un grado de mezcla considerable en individuos provenientes de etnias ubicadas en los departamentos del Tolima, Huila y Cauca como los Pijaos y Nasa.

A pesar de que los marcadores uniparentales han sido útiles en revelar las diferencias en las poblaciones colombianas con respecto al origen del perfil de ancestría paterno y materno, una descripción más exhaustiva del perfil genético de una población en términos de ancestría puede ser lograda con estudios de marcadores Autosómicos. Como se ve reflejado en la investigación realizada por Salzano & Sans, 2014, la estructura genética de la población colombiana ha sido estudiada en múltiples ocasiones para sus 6 regiones biogeográficas, observándose una amplia gama de frecuencias de poblaciones ancestrales entre sus individuos, donde no es posible definir una población colombiana homogénea general.

Dentro de los estudios autosómicos relevantes de la última década que han utilizado marcadores de tipo InDel en Colombia, observamos las investigaciones de Ibarra et al, 2014, Xavier et al, 2015, Ossa et al, 2016 y Mogollón et al, 2020. Inicialmente, Ibarra et al, 2014 analizó la utilidad de un set de 32 marcadores InDel específicos para cromosoma X vs una panel de 52 SNPs en 869 individuos proveniente de 12 poblaciones colombianas, dentro de sus resultados se evidenció un adecuado rendimiento para diferenciar las contribuciones genéticas ancestrales por parte de los dos paneles con resultados similares a pesar del uso de marcadores InDel solo para cromosoma X, esto nos permite evidenciar que la utilidad de los marcadores InDel para definir la contribución ancestral en población colombiana es altamente reproducible a los paneles utilizados con otros marcadores autosómicos. Por

otro lado, los resultados selectivos de cromosoma X, evidenciaron que la población del Chocó presentaba un menor distanciamiento genético de las poblaciones africanas y presentaba mejor precisión de estimación con el panel de SNP, pero esto puede deberse a la selectividad de InDel a cromosoma X y posibilidad de un mayor aporte de ancestría africana por parte del pool genético masculino. Por el contrario, con la población de Nariño, se observó una estimación de ancestría similar con el uso de ambos paneles.

Xavier et al, 2015 y Ossa et al en 2015 realizaron análisis en 5 poblaciones amerindias colombianas con el objetivo de caracterizar el componente de mezcla europeo y africano utilizando marcadores de ancestralidad de tipo InDel. Evidenciando que las poblaciones amerindias ubicadas en el departamento de Antioquia (etnia Emberá-Chamí), Norte de Santander (etnia Motilón-Barí) y Guainía (etnias Desano, Curripaco, Puinave, Cubeo, Guaunano y Tucano) se encontraban con aportes de más del 90% del componente de ancestralidad nativoamericano.

En el estudio de Ossa et al, 2016 se obtuvieron 761 individuos provenientes de 28 de los 32 departamentos de Colombia. Los participantes fueron obtenidos de áreas urbanas y considerados individuos con ancestría mixta. Así mismo, los investigadores incluyeron 121 individuos de tribus nativo americanas colombianas como los Guámbianos, Nasa, Emberá-Chami, Motilón-Bari, Pijaos, entre otros. Los resultados obtenidos por Ossa et al, muestran la prevalencia de ancestría europea en las regiones Caribe, Andina y Orinoquia, mientras la región Pacífica y Amazónica presentan un predominio de contribución genética Africana y Amerindia, respectivamente. La mayor variación interindividual se presentó en la Región Pacífica, de la Orinoquia y Amazónica. Con respecto al distanciamiento genético de las poblaciones estudiadas, se evidencia un distanciamiento significativo comparando las poblaciones no nativas vs nativas, pero si observamos los resultados del análisis de componentes principales, se observa como los pijaos se acercan a la región amazónica y a las subregiones andinas, lo cual explica el posible grado de mezcla que presenta esta etnia indígena en el país. Los autores

resaltan que poblaciones nativo americanas actuales pueden haber mantenido o perdido diferentes grados de pool genético ancestral, lo cual puede dificultar la inferencia de estos resultados.

Con respecto al uso de marcadores InDel en Población Afrocolombiana, Martínez B et al, 2017 analizaron una población de 187 individuos pertenecientes a San basilio de Palenque, una población fundada por esclavos negros localizada cerca de Cartagena. Esta comunidad ha sido declarada Patrimonio Cultural Inmaterial de la Humanidad por la UNESCO. Como era de esperarse se evidenció que el principal aporte de ancestralidad de esta población es africano en un 81.2%, seguido del 10,6% europeo, y un 8.2% de contribución nativo americana. Para los autores es llamativo como el componente de ancestralidad africana en San basilio de Palenque es cuatro veces mayor en comparación a los datos de la población del departamento de Bolívar, adicionalmente el segundo componente de ancestralidad en San Basilio es europeo mientras que en el departamento de Bolívar es el Nativo americano.

La publicación más reciente en donde se analiza la estructura genética de la población colombiana utilizando marcadores tipo InDel, es la aportada por Mogollón et al, 2020, en donde se refuerza que el aporte de ancestralidad colombiana este derivado principalmente de 3 subpoblaciones, nativoamericano, europea y africana. Adicionalmente, sugiere que las muestras de poblaciones afroamericanas y Nativoamericanas utilizadas presentan un considerable porcentaje de otros componentes de ancestralidad, por esta razón, no debería de hablarse de poblaciones puras, debido a que diferentes fuerzas de cambio evolutivo como deriva genética y migración puede generar cambios en las frecuencias alélicas dentro de estas poblaciones y sugieren que el termino más correcto debería ser poblaciones conservadas.

Los estudios previamente mencionados son importantes porque evidencian que Colombia no contiene un pool genético uniforme, sino una alta heterogeneidad del componente Afrodescendiente, Caucásico-mestizo y

Amerindio, dependiendo de la región estudiada. A pesar de que algunos estudios han tratado de caracterizar la poblaciones amerindias en Colombia, podemos considerar que las muestras utilizadas son pequeñas y restringidas a departamentos específicos, por lo tanto, no son suficientes para obtener una estimación precisa de la ancestralidad de la población colombiana, ya que las muestra utilizadas, especialmente las de muestras de mayor tamaño son individuos nativos americanos que pueden considerarse poco representativas debido al grado de mezcla presente en estas etnias.

En conclusión, a la fecha se han realizado varios estudios genéticos de muestras de diferentes departamentos del país; estos estudios han estado enfocados en evaluar la diversidad y estructura genética de la población colombiana. Si bien, estos estudios han realizado aportes importantes al conocimiento de la historia biológica de las poblaciones humanas colombianas, presentan algunas limitaciones en cuanto a criterios de inclusión, tamaño y procedencia de las muestras.

1.4 Importancia del estudio de la subestructura poblacional a través de la ancestralidad genética en la población colombiana.

Múltiples investigaciones se han realizado en búsqueda de determinar la ancestría genética en Colombia, obteniendo resultados limitados y discrepantes; debido a la diversidad de metodologías de muestreo planteadas, así como las diferencias en el número y tipo de marcadores utilizados. En la actualidad, continúa siendo de gran interés el conocimiento y análisis de la composición del origen ancestral de la población colombiana. (Alonso Morales et al., 2018; Ibarra, Restrepo, et al., 2014; Moncada J, 2018; Ossa et al., 2016; Xavier et al., 2015). En los últimos años, los estudios poblacionales han documentado la considerable heterogeneidad y diversidad genética de la población de América Latina, siendo Colombia una de las poblaciones más diversas con un intenso flujo de carga genética y cultural dado sus procesos históricos y ubicación geográfica que ha favorecido la presencia de esta

pluralidad genética. (Carvajal-Carmona et al., 2000; Price et al., 2007; Rodas et al., 2003; W. Rojas et al., 2010; Sans, 2000).

El estudio de la variación genética mediante el uso de polimorfismos de ADN distribuidos a lo largo del genoma, ha permitido una mejor comprensión de la historia de la diversidad de poblaciones humanas (Pereira et al., 2009). El uso de estos marcadores también permite una mayor aproximación a la estimación de la ancestría, información muy útil para los estudios poblacionales en virtud de los datos relevantes que pueden proporcionar. Esta información esclarecería las dinámicas de la migración entre diferentes poblaciones, patrones de mezcla y la influencia de fuerzas de cambio evolutivo en la composición genética de las comunidades en cuestión (Manta et al., 2013). En adición, el mapeo de la mezcla genética de las poblaciones puede resultar en un enfoque económico importante para la búsqueda de enfermedades genéticas, tal es el caso, evidenciado en las poblaciones afroamericanas en cáncer de próstata, siendo una herramienta prometedora en poblaciones latinas dado su mestizaje proveniente de diferentes continentes lo que proporciona un grado diferente de riesgo genético. (Price et al., 2007).

Por lo tanto, se ha visto un creciente interés en estudiar y comprender la ancestralidad de las poblaciones y usar esta información como una herramienta útil para estudios de asociación, genético-poblacionales, socio-económicos, economía y política en salud y forenses. El resultado esperado es la disminución de los posibles efectos negativos o sesgados que puedan generarse debido a una estratificación genética en la población que no ha sido considerada o ha sido subestimada. (Freedman et al., 2004; Pritchard & Rosenberg, 1999; Tang et al., 2005)

Los estudios genético-poblacionales en enfermedades oncológicas han sido un foco de especial interés en la última década, donde definir la relación y el riesgo de presentación, comorbilidades, mortalidad y supervivencia con respecto al origen étnico ancestral, son tópicos de estudio de gran relevancia en distintos grupos de investigación internacional.(Carvalho et al., 2015; da Silva et al.,

2017; Gimbrone et al., 2017; Hernandez-Suarez et al., 2014; Hines et al., 2017; Karol et al., 2017; Perez-Mayoral et al., 2019; Torres et al., 2019). Así mismo, como lo plantean Chande et al, 2021 la importancia de conocer la proporción del componente de ancestralidad nos permite determinar el impacto de este en el riesgo y prevalencia de diferentes enfermedades, lo cual no solo nos ayuda a determinar políticas en salud pública, sino que nos permite dar un paso más cerca hacia la medicina de precisión.

La leucemia mieloide aguda (LMA) un tipo de neoplasia hematológica que se define como la transformación maligna y producción incontrolada de células hematopoyéticas inmaduras. Constituye un 20% de las leucemias agudas diagnosticadas en la edad pediátrica, con una tasa de recurrencia del 40% y una supervivencia general de los pacientes con recaída a los 3 años menor del 60%. (Amankwah & Hale, 2018; Estey, 2018; McNeer et al., 2019; Walsh et al., 2014). Debido al alto impacto en la salud infantil y salud pública de esta enfermedad oncológica, nuestro grupo de investigación inició el macroproyecto *“Caracterización molecular y farmacogenética en una muestra de pacientes con Leucemia Mieloide Aguda y su correlación con la estratificación del riesgo y respuesta al tratamiento. Una aproximación hacia los patrones moleculares de la Leucemia Mieloide Aguda Pediátrica en población colombiana”*; del cual se deriva el presente estudio, donde se espera una vez obtenidos los resultados utilizarlos como base de referencia poblacional para la comparación de los perfiles de ancestría de pacientes con leucemia mieloide aguda.

Pereira et al, 2020 Decidieron evaluar la precisión de la inferencia de ancestralidad de diferentes sets de marcadores utilizados actualmente en una muestra de poblaciones suramericanas, en donde consideraron que independientemente de la capacidad de estos marcadores para separar poblaciones de diferentes continentes o regiones, rara vez se había investigado la incertidumbre asociada con las estimaciones proporcionadas por estos paneles y su capacidad para informar con precisión las diferentes contribuciones ancestrales en individuos de poblaciones mezcladas. Dentro de sus resultados resaltaron la importancia de tener un adecuado equilibrio y

compromiso entre no solo el número de marcadores utilizados y la capacidad de estos para diferenciar poblaciones ancestrales, sino también una diferenciación balanceada entre grupos de referencia, para obtener valores más precisos de ancestría genética. Adicionalmente, concluyeron que un aspecto a considerar al inferir la ancestralidad de los individuos es el impacto de la selección de poblaciones de referencia apropiadas.

Considerando la complejidad de la dinámica de mezcla de América del Sur y que los paneles de marcadores actualmente utilizados fueron diseñados para maximizar las diferencias entre continentes, estos se usan comúnmente para determinar las principales contribuciones de ancestralidad continental en individuos o poblaciones mezcladas. Sin embargo, los patrones de mezcla en América del Sur presentan contribuciones diferenciales de varias poblaciones africanas y europeas de diferentes regiones del continente, y estudios recientes han dilucidado que la presencia de europeos del norte está más restringida al sur, mientras que los eventos de mezcla de europeos occidentales son más generalizados. (Gouveia et al, 2020) Adicionalmente, estudios anteriores informaron la ausencia de una determinación precisa dentro de los grupos de África subsahariana, Europa y Asia oriental por parte de diferentes paneles de marcadores (Al-Asfi et al., 2018; Nakanishi et al., 2018; Mogensen et al., 2020). Por lo tanto, esto nos hace pensar que no solo se podría estar afectando la resolución y precisión de determinación de ancestralidad africana o europea, sino adicionalmente dado las diferencias y grandes comunidades amerindias a través del continente americano, sea importante determinar poblaciones de referencia nativo americanas altamente conservadas y apropiadas de cada población como en el caso de Colombia.

Tradicionalmente los estudios genético-poblacionales en Colombia se han limitado a un número restringido de muestras, tomadas en su mayoría de las comunidades urbanas y en su selección se han basado principalmente en criterios de autodeterminación. Si bien, en este tipo de estudios los muestreos son por conveniencia, se ha caído en la tendencia de realizarlos sin el mayor rigor metodológico, ni matemático (Alonso, L.A, Usaquen, W 2013).

Adicionalmente, estudios como el de Price et al, 2007 en donde se busca determinar un mapeo de la estructura nativo americana en Suramérica utilizan bases de datos de poblaciones colombianas de estudios como los mencionados previamente, en donde se ha evidenciado que las poblaciones amerindias y afrocolombianas presentan un grado considerable de mezcla.

Por lo tanto, se pretende con el presente estudio determinar una población nativo americana y afrocolombiana lo más conservada posible que permita ampliar el conocimiento sobre la estructura genética de la población colombiana y adicionalmente que la comunidad científica pueda apropiarse de estos resultados como una fuente de población de referencia apropiada para estudios del componente ancestral o biogeográfico a futuro.

2. Objetivos

2.1 Objetivo general:

Caracterizar el perfil de ancestría y analizar el nivel de sub-estructura genética de los 3 principales grupos étnicos Amerindios, Afrodescendientes y Caucásico-mestizos en una muestra de población Colombiana mediante un panel de 46 marcadores tipo AIM InDel.

2.2 Objetivos Específicos:

- Caracterizar el perfil de ancestría mediante un panel de 46 AIM-InDel en una muestra de población Amerindia Colombiana.
- Caracterizar el perfil de ancestría mediante un panel de 46 AIM-InDel en una muestra de población Afrodescendiente Colombiana.
- Caracterizar el perfil de ancestría mediante un panel de 46 AIM-InDel en una muestra de población Caucásico-Mestiza Colombiana.
- Analizar los patrones de mezcla y su variación interpoblacional e interindividual dentro de la muestra de los 3 grupos poblacionales (Amerindios, Afrodescendientes y Caucásico-mestizos).
- Comparar los resultados de subestructura poblacional con las diferentes poblaciones colombianas utilizadas como referencia en las diferentes bases de datos a nivel internacional.

3. Metodología

3.1 Población de Estudio

Las muestras fueron seleccionadas del banco de muestras del Grupo de Identificación Humana e Inmunogenética de la Universidad Nacional de Colombia, las cuales fueron recolectadas como parte del proyecto “Estructura Genética de la población Colombiana” entre 1988-1998. Las muestras cuentan con el consentimiento informado correspondiente y algunas han sido parte de investigaciones previas, las cuales fueron procesadas y tipificadas para marcadores STRs, Cromosoma Y, ADN mitocondrial, y Sistema HLA. Adicionalmente, se seleccionaron algunas muestras existentes en el banco del Instituto de Genética de Servicios Médicos Yunis Turbay y Cia SAS. las cuales cuentan con el Consentimiento informado para su utilización en estudios de caracterización genética. Se estimó un tamaño de muestra de 450 individuos por conveniencia, 150 individuos para cada subgrupo poblacional (Amerindios, Afrocolombianos y con Ancestralidad Mixta).

El procesamiento de las muestras se llevó a cabo en el Laboratorio del Instituto de Genética de la Universidad Nacional de Colombia y el Laboratorio de Servicios Médicos Yunis Turbay.

3.2 Extracción de ADN genómico

Las muestras de ADN refrigerado inicialmente fueron evaluadas en una electroforesis en gel de agarosa al 0.5% en Buffer SB 1X para verificar la calidad del ADN. Aquellas en donde se evidenció fragmentación y mala calidad del ADN, se descartaron y se realizó la extracción del ADN a partir de muestra de sangre periférica congelada dentro del banco de muestras del Grupo de Identificación Humana e Inmunogenética de la Universidad Nacional de Colombia, las cuales se mantenían en congelación a -80°C. Para la extracción de ADN, se realizaron alícuotas de 200uL y se siguió el protocolo del *QIAamp DNA Mini Blood Kit* (QIAGEN).

3.3 Amplificación PCR Multiplex

Se realizó la amplificación por PCR Multiplex. Los primers, las condiciones de reacción de la mezcla maestra y las condiciones de termociclado, fueron diseñados según lo reportado por Pereira et al, 2012

3.3.1 Primers

Los primers que se utilizaron para la amplificación de los marcadores InDel por PCR Multiplex se obtuvieron del set de 46 marcadores AIM-InDel validado por Pereira et al.,2012. El diseño de los primers se ajustó a los criterios definidos por los autores de la siguiente manera:

- Tamaño Promedio del Amplicon 60-160bp (Algunos 230bp)
- Tm Optimo 60°C (Mínimo 58°C)
- Contenido de GC optimo 50% (mínimo 45%)
- El conjunto de primers fue evaluado para no presentar hibridaciones con otras regiones del genoma utilizando la Herramienta BLAST de NCBI - <http://blast.ncbi.nlm.nih.gov>

Tabla 3. Listado de los pares de primers utilizados para los 46 marcadores.

MID	rs number	Labeled primer	Unlabeled primer*
MID-1470	rs2307666	6FAM-GAGTCTGACCCTTCATAAGC	<i>g</i> CCATGGTGATATTACGTCCC
MID-777	rs1610863	6FAM-TGGAAGACACGTCTAAGAG	<i>g</i> TATTCCTCCAGGCTCTTTGC
MID-196	rs16635	6FAM-CCAAGTTCTAGCCATATGGA	<i>gtttctt</i> GACTATCTTCTCTGACCATC
MID-881	rs1610965	6FAM-TTGGCTCCCTATGATAATCC	<i>gtttc</i> TTGTGTTCCCAAAGTTCTCC
MID-3122	rs35451359	6FAM-TCACAAGTCCGGAATACCAG	<i>g</i> AGTTATGGGATGGGAAGGAG
MID-548	rs140837	6FAM-AGTCAGGACTGAAGAAACCC	<i>gttt</i> CAGTAAACAAGAGCCCGTG
MID-659	rs1160893	6FAM-CACTGCATCAGACTGACTTC	<i>gtttctt</i> GGTGCTTTTGCTTTGAATTG
MID-2011	rs2308203	6FAM-TGAGAACTAGGAGCTCTGG	<i>gf</i> TTCCCTAAGAGCCACTGACAT
MID-2929	rs33974167	6FAM-TGTGATGTGGATAGGCAAGG	<i>g</i> AGGCTCCATTGTGTTAAGAGG
MID-593	rs1160852	6FAM-TGCTCACTTGTAGTGAGACC	<i>gf</i> TTGCGTTTAGGTCCCTTCTG
MID-798	rs1610884	6FAM-ACGACAGTGTTCACAAAGAG	<i>g</i> CTGTTGTCTGACCTGTGAAG
MID-1193	rs2067280	6FAM-GCTGGTGTAGTTTTTCTCTCC	<i>gf</i> TCCACCATCTACCTTCTATG
MID-1871	rs2308067	6FAM-TTGTAGTCAGAGAGTGTGCC	<i>g</i> AGCCTTTTCCCTAACGTCAC
MID-17	rs4183	6FAM-AGAACTGCAACCCCTCAAG	<i>G</i> ATCCCAGACACTGAAGATG
MID-2538	rs3054057	6FAM-CTCGCAAAGTAGGCAAGTTC	<i>g</i> ACACCAACAATCTTGGCACC
MID-1644	rs2307840	6FAM-ACACCCTTAAGATCTGACC	<i>GGT</i> CTAAAGTCACTGACACAG
MID-3854	rs60612424	VIC-TCACCTTATTCAGGTTGC	<i>GCC</i> AGGGATTTAGTGTAGAG
MID-2275	rs3033053	VIC-CTACCTGACTACCACCTATG	<i>g</i> ACCCAGCCTATCTGACTTTG
MID-94	rs16384	VIC-TGGTGGCTCATGACTTTTG	<i>gttt</i> ACAGGGTCTCGCTATGATGC
MID-3072	rs34611875	VIC-AGCTTTTTCGGCAACTCTC	<i>gtt</i> TGGATGTGTCTGAGCTCAAC
MID-772	rs1610859	VIC-GTCTCRTTTTCTGCACTAG	<i>gtttctt</i> ATCCTTCTGCTCACTCTACC
MID-2313	rs3045215	VIC-GCACACATGCAGAAATGCAG	<i>G</i> TTGTAACATCTGTGAGGTC
MID-397	rs25621	VIC-TGGGCTTCTTCTGGGAAAAC	<i>g</i> CCACTTCAGGCGTTTTGTGTC
MID-1636	rs2307832	VIC-TTAGGAAGAGGTGCTATGGG	<i>g</i> CCTCCTTTGAAGACACACAG
MID-51	rs16343	VIC-AAGATTGGAGGGAAAAGTGC	<i>g</i> CGTCCTCCACCTTCTTTTTTC
MID-2431	rs3031979	VIC-AGGAGGAGCTGATAGACTTC	<i>G</i> CAGTGTGCAACTGATACG
MID-2264	rs34122827	VIC-CTTTGGCTATCTGTCTCAC	<i>G</i> TAGGAGACCACTCACATTC
MID-2256	rs133052	NED-ATCGAACCCTTCTAAGGAC	<i>G</i> CAAGAAAAGGAATCCAGGC
MID-128	rs6490	NED-ATCAGGAGACAATCCAGCAG	<i>g</i> TCCAGCCATTCAGACAAAGG
MID-15	rs4181	NED-GGGTTATTTGCCTCATCTCC	<i>gtttctt</i> AGGTATTCTCTGTTCCACG
MID-2241	rs3030826	NED-ACATACACGTGGAAGACTGC	<i>gtt</i> ACTGTGCGACTGATCCAATAG
MID-419	rs140708	NED-CAGGAAAGTATGGCCATTC	<i>G</i> TCCATGTTTTCTTTGAGCATC
MID-943	rs1611026	NED-TCTTCTACCCCTGTTAGTG	<i>G</i> ACAAGATCACTAGCTTGAC
MID-159	rs16438	NED-ACCAGAGCACTACAGCCTTT	<i>g</i> CAAGGYAGTAACAATGAGGG
MID-2005	rs2308161	NED-TGTAGCGCAATATAGGCAG	<i>G</i> AAAGTTGTGGCTTAACTGG
MID-250	rs16687	NED-ATGGAGCAGTAAAGCAGCAC	<i>G</i> TCACTTTGGTTTTTGCAGG
MID-1802	rs2307998	NED-ACGGTCAACTTTGTAGCTCC	<i>g</i> CCAGTTGAGAATCACTGCAC
MID-1607	rs2307803	NED-TGTTGCAGAAGAACAACC	<i>G</i> ATAAGCACCTAACTCCAG
MID-1734	rs2307930	PET-TTCGTGTTCTCACACTGTCC	<i>G</i> TGCATCCCATACAACAGTAC
MID-406	rs25630	PET-TGGCTGCTGTAGATTGTAGG	<i>g</i> ACAAATGGACAACGGCCAAG
MID-1386	rs2307582	PET-AGAGGATCATGGAGACCAAC	<i>gttt</i> ATGTTCCAAGTCAGCAGCAC
MID-1726	rs2307922	PET-GGTCCAAATGCACCACAATC	<i>G</i> CTCTGCTATTTTGGTTTGC
MID-3626	rs11267926	PET-TGTTGGTCTCTCCTTTTCC	<i>G</i> GTGACCCCTCTTTATCTC
MID-360	rs25584	PET-AGATCAACTGCCAATCTGGG	<i>g</i> CTCAAGTGACCAACCCACCT
MID-1603	rs2307799	PET-TTACAATTTCAAGCCTCCGC	<i>G</i> GAGCTGTAGTCTGAGTAG
MID-2719	rs34541393	PET-GTCAGGAGTCTAGAAAATTC	<i>G</i> GGTGATGAAATGTTCCGAA

* Lowercase italic letters represent nucleotide tails added to the primers.

Tomada de (Pereira et al.,2012)

3.3.2 Condiciones de Reacción para la PCR

Para la preparación del pool de primers, se resuspendió cada uno de los primers con 100 uL de agua de biología molecular libre de RNAsas para una concentración final de 0,1 uM de cada primer. Una vez resuspendidos se dejaron agitando a temperatura ambiente por 40 min y luego se centrifugaron por 5 segundos. Una vez preparado cada primer, se procedió a realizar la mezcla de primers que se utilizó en la PCR Multiplex, para ello, se tomó 10 uL de cada primer para un volumen final de 920 uL, contando forward y reverse. Se revuelve la mezcla de primers y se centrifugan. Posteriormente, se separaron en alícuotas de 50 uL en tubos Eppendorf de 1.5 uL. Cada tubo, se forró con papel aluminio para proteger los primers fluoromarcados. De igual manera, para mayor seguridad los tubos con la mezcla de primers se guardaron en una caja forrada en aluminio y se mantuvieron en refrigeración a -20°C. Los primers MID-196, MID-881, MID-772, MID-51, MID-1607, MID-3626, MID-1603 y MID-2719 se les realizó un spike con 10 uL de primers adicional en la mezcla de primers para obtener una mejor calidad en la genotipificación.

El procedimiento del protocolo de la PCR Multiplex se llevó a cabo acorde con las condiciones presentadas en la tabla a continuación:

Tabla 4. Condiciones de reacción de mezcla maestra.

Reactivo	Concentración	Volumen
QIAGEN Multiplex Master Mix	2x	5.0 uL
Mezcla de Primers	0.1uM (Cada Primer)*	1.0 uL
ADN	1 ng	4.0 uL
Volumen Final Reacción		10.0 uL

*Primers MID-196, MID-881, MID-772, MID-51, MID-1607, MID-3626, MID-1603 y MID-2719 concentración 0.2uM.

3.3.3 Condiciones de Termociclado

Tabla 5. Condiciones de termociclado.

Condiciones de Termociclado		
Paso	T ° C – Tiempo	Ciclos
HotStart	95°C – 15:00 Min	1x
Denaturación	94°C – 0:30 Min	29x
Hibridación	60°C – 1:30 Min	
Extensión	72°C – 0:45 Min	
Final	72°C – 60:00 Min	1x

3.4 Electroforesis Capilar

Una vez se obtuvo los productos del amplificado de PCR, estos fueron preparados para la Electroforesis Capilar, este proceso se realizó en un analizador genético ABI 3500 (Life Technologies) con protocolo estandarizado en el Laboratorio de Servicios Médicos Yunis Turbay. La recolección y análisis de datos se hizo mediante el software Genemapper v.4.1 (Life Technologies).

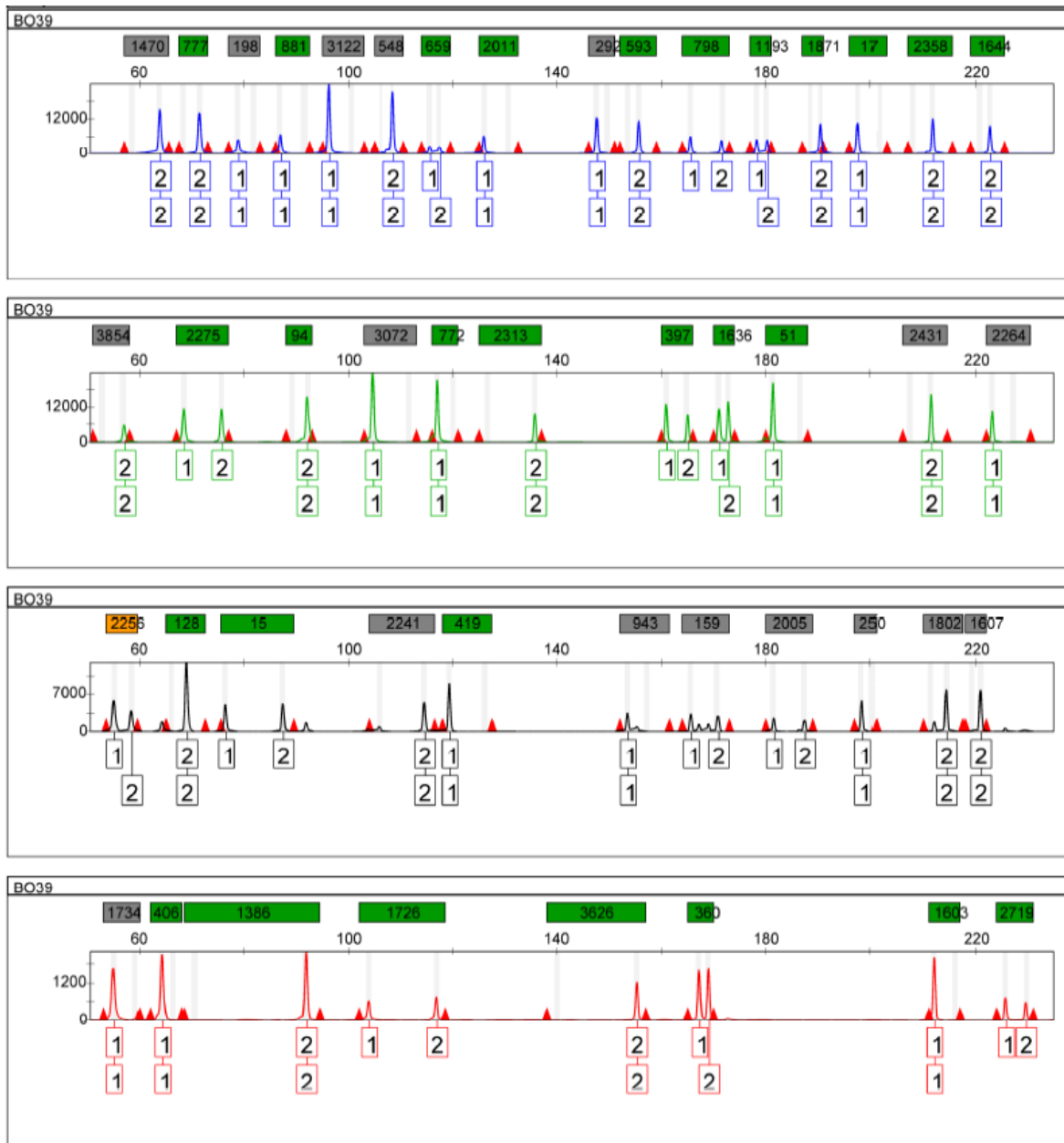
3.5 Análisis Estadístico

Las frecuencias alélicas, el equilibrio de Hardy Weinberg, la diversidad genética, el análisis molecular de la varianza (AMOVA) y las distancias genéticas F_{st} fueron calculadas mediante el software Arlequín v. 3.5.2.2 (Excoffier et al., 2010) y Genepop v4.6 (Rousset F, 2017).

La distribución de los componentes de ancestralidad genética Nativo americana (NAM), Europea (EUR) y Africana (AFR) en cada individuo se estimó utilizando el software STRUCTURE v.2.3.4.21. (Pritchard et al., 2010; Falush et al., 2003), El análisis se llevó a cabo utilizando un periodo de corrido de 100.000 iteraciones, seguido de 100.000 repeticiones de Método de Montecarlo basado en cadenas de Markov (MCMC). En los análisis se asumió una contribución trihíbrida de nativos americanos, europeos y

africanos (es decir, $K = 3$). Se utilizó el modelo de mezcla y se correlacionaron y actualizaron las frecuencias alélicas utilizando únicamente individuos con POPFLAG = 1, se utilizaron como referencias las muestras de HGDP-CEPH (Human Genome Diversity Cell Line Panel) y de muestras de referencia de poblaciones africanas, europeas, del este de Asia y Nativo americanas (Pereira et al., 2012). Adicionalmente, para algunos análisis se incluyeron datos de estudios previos en muestras de población colombiana (Ossa et al, 2016) y (Mogollon et al, 2020). Con esta misma información y para facilitar la interpretación de estos resultados, se realizará un análisis de componentes principales mediante el software MVSP (Multivariate Statistical Package) v.3.22 (Kovach, 2007).

Figura 2. Electroferograma de los 46 marcadores InDel amplificados.



4. RESULTADOS

4.1 Genotipificación y Descripción de la Población:

Se obtuvo en total 474 muestras de individuos para genotipificación, en la mayoría de ellas se logró el perfil de marcadores de ancestría completo, sin embargo, fueron excluidas 7 muestras; ya que no fue posible identificar algún marcador en 5 de ellas y en las otras 2 muestras no se pudo evidenciar más de 5 marcadores del perfil de ancestría utilizado. En definitiva, la población de estudio para análisis fue de 153 amerindios (AMR), 164 afrocolombianos (AFC) y 150 individuos de ancestralidad mixta (MIX).

Para el grupo Amerindio se lograron muestras de 28 etnias indígenas colombianas con mayor predominio de comunidades localizadas en la región sur del país. Los pueblos indígenas participantes y su tamaño poblacional se describen en la figura 3. Aproximadamente el 42% de la muestra poblacional amerindia se obtuvo de la región amazónica colombiana. El departamento con mayor representación por individuos indígenas fue el departamento del Magdalena con aproximadamente un 33% del total de la muestra, sin embargo, el departamento de Vaupés fue el departamento con mayor aporte en número de comunidades indígenas.

Los individuos de la población afrocolombiana en un 100% se localizaron en el departamento del Chocó. Los municipios en donde se obtuvieron las muestras fueron Quibdó, Bahía Solano, Condoto, Istmina y Tadó. El municipio con mayor número de muestras incluidas en el estudio fue el municipio de Tadó con una representación de aproximadamente el 24% de la muestra afrocolombiana.

La muestra poblacional de individuos con ancestralidad mixta se obtuvo de los departamentos de Cundinamarca, Boyacá y Santander, y de la ciudad de Bogotá con una representación del 42%, 24%, 15.3% y 18.7%, respectivamente.



Comunidad Indígena	N° Individuos
Arhuacos	29
Barasano	2
Bari	17
Carapana	1
Chimila	12
Curripako	6
Desano	4
Guayabero	8
Inga	4
Kawiyari	1
Kogui	9
Kubeo	8
Makuna	1
Misak	1
Nasa	4
Nukak	7
Piapoco	1
Piratapuyo	2
Puinave	8
Siriano	1
Tikuna	4
Tukano	2
Tunebo	7
Wanano	2
Wayuu	4
Waunano	3
Yeral	2
Yukpa	3

Figura 3. Mapa de las comunidades indígenas incluidas en el estudio.

4.2 Análisis Genético Descriptivo y Diversidad Poblacional:

Se calcularon las frecuencias alélicas para los 46 marcadores InDel en cada uno de los grupos poblacionales, el análisis fue realizado mediante el software Genepop v4.6. Los datos de las frecuencias alélicas por locus poblacionales están reportados en la Tabla 6. En este estudio se detectaron individuos con un tercer alelo en los marcadores MID-548, MID-1636, MID-2264 y MID-360. El tercer alelo en el marcador MID-548 fue hallado únicamente en la muestra de población afrocolombiana con una frecuencia del 3.66%, siendo esta la mayor prevalencia identificada en todos los marcadores que presentaron un tercer alelo (Figura 4). Por otro lado, el marcador MID-1636 con una frecuencia del 0.33% estuvo presente en la población amerindia. Un tercer alelo en el marcador MID-360 fue detectado con una baja frecuencia (<0.004) en la población mezclada y con una mayor frecuencia (>0.02) en la muestra afrocolombiana. En las tres muestras poblacionales se logró evidenciar individuos con un tercer alelo en el marcador MID-2264, con una mayor frecuencia en la muestra afrocolombiana (>0.02).

Los análisis del equilibrio de Hardy Weinberg, diversidad genética, el análisis de la varianza molecular y estadístico F_{st} se realizaron en el software Arlequín v. 3.5.2. En el análisis de la diversidad genética se evidenció que los valores más bajos de Heterocigosidad fueron encontrados en la población amerindia con una H_o de 0.24120 en comparación a las poblaciones afrodescendientes y mestizas donde se evidenció una H_o de >0.36 . (Tabla 7). Con respecto a las poblaciones amerindias, la gran mayoría de las etnias se encontraron con heterocigosidades menores de 0.3, sin embargo, la muestra de población Inga, Puinave y Siriano fueron las menos conservadas con H_o mayor de 0.30. Dentro de las etnias más conservadas encontramos a los ticunas, los yukpas, los piratapuyos y los chimila con H_o menores o igual a 0.20. En la población afrodescendiente la de menor Heterocigosidad fue la muestra de población del municipio de Istmina con una H_o de 0.34 y las poblaciones con mayor Heterocigosidad fueron la de Quibdó y Bahía Solano. (Tablas Suplementarias).

Figura 4. Electroferograma evidencia tercer alelo en marcador MID-548

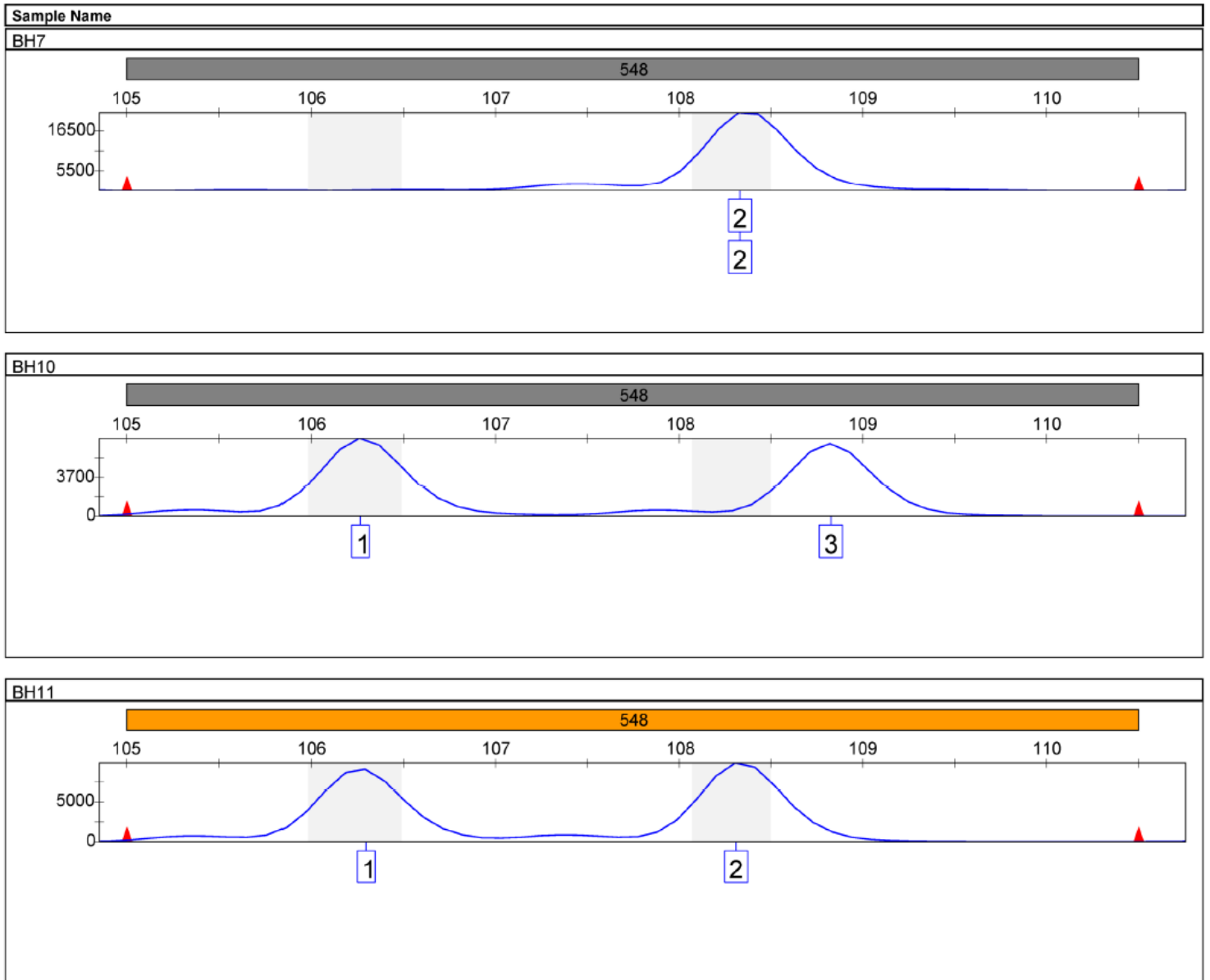


Tabla 6. Frecuencias alélicas por locus para los subgrupos poblaciones.

Frecuencias Alelicas por Locus para 3 poblaciones colombianas									
Locus	Caucasicos			Amerindios			Afrocolombianos		
	Alelo 1	Alelo 2	Alelo 3	Alelo 1	Alelo 2	Alelo 3	Alelo 1	Alelo 2	Alelo 3
1470	0.2967	0.7033		0.0229	0.9771		0.2073	0.7927	
777	0.5233	0.4767		0.7843	0.2157		0.4146	0.5854	
198	0.2967	0.7033		0.0621	0.9379		0.4329	0.5671	
881	0.9167	0.0833		0.9739	0.0261		0.5671	0.4329	
3122	0.9700	0.0300		0.9967	0.0033		0.7744	0.2256	
548	0.1800	0.8200		0.0882	0.9118		0.3506	0.6128	0.0366
659	0.2248	0.7752		0.1961	0.8039		0.2409	0.7591	
2011	0.8100	0.1900		0.9248	0.0752		0.3963	0.6037	
2929	0.7633	0.2367		1,000	0,000		0.8537	0.1463	
593	0.0333	0.9667		0.0163	0.9837		0.5000	0.5000	
798	0.3167	0.6833		0.1013	0.8987		0.3720	0.6280	
1193	0.1000	0.9000		0.0980	0.9020		0.1463	0.8537	
1871	0.2667	0.7333		0.4216	0.5784		0.0732	0.9268	
17	0.5833	0.4167		0.9216	0.0784		0.6433	0.3567	
2358	0.2300	0.7700		0.0621	0.9379		0.0549	0.9451	
1644	0.6342	0.3658		0.3725	0.6275		0.3262	0.6738	
3854	0.0900	0.9100		0.0294	0.9706		0.5183	0.4817	
2275	0.2233	0.7767		0.4346	0.5654		0.3598	0.6402	
94	0.4533	0.5467		0.6046	0.3954		0.1189	0.8811	
3072	0.9667	0.0333		0.9869	0.0131		0.4116	0.5884	
772	0.8867	0.1133		0.7876	0.2124		0.9360	0.0640	
2313	0.4833	0.5167		0.8366	0.1634		0.2835	0.7165	
397	0.5733	0.4267		0.3627	0.6373		0.7348	0.2652	
1636	0.8267	0.1733		0.9869	0.0098	0.0033	0.4482	0.5518	
51	0.7333	0.2667		0.8725	0.1275		0.2622	0.7378	
2431	0.1900	0.8100		0.2026	0.7974		0.1280	0.8720	
2264	0.3500	0.6000	0.0500	0.1373	0.8431	0.0196	0.3293	0.6494	0.0213
2256	0.4967	0.5033		0.6013	0.3987		0.1921	0.8079	
128	0.2933	0.7067		0.2418	0.7582		0.1067	0.8933	
15	0.6067	0.3933		0.8529	0.1471		0.3963	0.6037	
2241	0.2433	0.7567		0.0556	0.9444		0.5732	0.4268	
419	0.7833	0.2167		0.6667	0.3333		0.9116	0.0884	
943	0.6259	0.3741		0.2664	0.7336		0.3628	0.6372	
159	0.6367	0.3633		0.7908	0.2092		0.7012	0.2988	
2005	0.3967	0.6033		0.2712	0.7288		0.1677	0.8323	
250	0.6333	0.3667		0.4216	0.5784		0.7073	0.2927	
1802	0.0467	0.9533		0.0131	0.9869		0.3537	0.6463	
1607	0.4233	0.5767		0.5458	0.4542		0.3598	0.6402	
1734	0.7400	0.2600		0.5196	0.4804		0.8384	0.1616	
406	0.7167	0.2833		0.6961	0.3039		0.3018	0.6982	
1386	0.3967	0.6033		0.5752	0.4248		0.2226	0.7774	
1726	0.7167	0.2833		0.7124	0.2876		0.4329	0.5671	
3626	0.4033	0.5967		0.0395	0.9605		0.2409	0.7591	
360	0.7500	0.2467	0.0033	0.6569	0.3431		0.5549	0.4238	0.0213
1603	0.5700	0.4300		0.7680	0.2320		0.2073	0.7927	
2719	0.4867	0.5133		0.5294	0.4706		0.3720	0.6280	

Tabla 7. Equilibrio de Hardy-Weinberg y Heterocigosidad observada y esperada por locus y por subgrupos poblaciones.

Amerindios					Afrodescendientes					Caucasico-Mestizos				
MD	Ho	He	P-value	SE	MD	Ho	He	P-value	SE	MD	Ho	He	P-value	SE
1470	0,02143	0,02127	1,00000	0,00000	1470	0,31690	0,33850	0,45965	0,00050	1470	0,39333	0,41871	0,55431	0,00052
777	0,27143	0,31244	0,17043	0,00035	777	0,49296	0,48226	0,86196	0,00035	777	0,47333	0,50058	0,51652	0,00050
198	0,10000	0,09534	1,00000	0,00000	198	0,50704	0,49818	0,86580	0,00034	198	0,38000	0,41871	0,32461	0,00047
881	0,03571	0,03520	1,00000	0,00000	881	0,47887	0,49928	0,73545	0,00041	881	0,16667	0,15329	0,59975	0,00050
3122	0,00714	0,00714	1,00000	0,00000	3122	0,35211	0,38670	0,28494	0,00044	3122	0,06000	0,05839	1,00000	0,00000
548	0,12857	0,14531	0,19468	0,00039	548	0,42254	0,51909	0,00000	0,00000	548	0,33333	0,29619	0,16585	0,00038
659	0,21429	0,29442	0,00257	0,00005	659	0,40141	0,38329	0,66218	0,00047	659	0,35570	0,34974	1,00000	0,00000
2011	0,09286	0,12696	0,01556	0,00013	2011	0,47887	0,46195	0,71624	0,00043	2011	0,28667	0,30883	0,42457	0,00048
2929	0,00000	Locus is monomorphic: no test done			2929	0,23944	0,25292	0,51017	0,00046	2929	0,35333	0,36252	0,82001	0,00036
593	0,00714	0,00714	1,00000	0,00000	593	0,43662	0,49689	0,17364	0,00035	593	0,06667	0,06466	1,00000	0,00000
798	0,10000	0,14531	0,00444	0,00007	798	0,50000	0,46392	0,37474	0,00048	798	0,51333	0,43423	0,03718	0,00020
1193	0,18571	0,19201	0,65626	0,00047	1193	0,24648	0,25788	0,52817	0,00051	1193	0,16000	0,18060	0,16468	0,00037
1871	0,44286	0,49350	0,23490	0,00043	1871	0,10563	0,10041	1,00000	0,00000	1871	0,36000	0,39242	0,30424	0,00045
17	0,10000	0,12074	0,09471	0,00029	17	0,54225	0,45994	0,05247	0,00022	17	0,46000	0,48774	0,50584	0,00048
2358	0,08571	0,08233	1,00000	0,00000	2358	0,05634	0,05494	1,00000	0,00000	2358	0,35333	0,35538	1,00000	0,00000
1644	0,50000	0,45663	0,27132	0,00044	1644	0,32394	0,40611	0,05129	0,00021	1644	0,51678	0,46553	0,21482	0,00043
3854	0,05000	0,04892	1,00000	0,00000	3854	0,50704	0,49181	0,73467	0,00044	3854	0,14000	0,16435	0,09661	0,00029
2275	0,51429	0,49677	0,73433	0,00042	2275	0,43662	0,47629	0,73646	0,00048	2275	0,36667	0,34807	0,63850	0,00048
94	0,45000	0,48026	0,48226	0,00049	94	0,20423	0,19517	1,00000	0,00000	94	0,49333	0,49730	1,00000	0,00000
3072	0,00000	Locus is monomorphic: no test done			3072	0,51408	0,46392	0,21058	0,00045	3072	0,06667	0,06466	1,00000	0,00000
772	0,25714	0,34603	0,00529	0,00007	772	0,09859	0,11915	0,09158	0,00029	772	0,21333	0,20165	0,69492	0,00047
2313	0,27857	0,26091	0,53007	0,00051	2313	0,34507	0,40300	0,09520	0,00033	2313	0,50000	0,50111	1,00000	0,00000
397	0,47857	0,45876	0,71410	0,00048	397	0,33099	0,37633	0,17830	0,00038	397	0,48000	0,49088	0,86778	0,00034
1636	0,01429	0,01426	1,00000	0,00000	1636	0,47887	0,47937	1,00000	0,00000	1636	0,32000	0,28754	0,25260	0,00046
51	0,18571	0,19201	0,65639	0,00047	51	0,33099	0,34646	0,62803	0,00046	51	0,37333	0,39242	0,53748	0,00049
2431	0,30714	0,33382	0,44408	0,00051	2431	0,21127	0,21152	1,00000	0,00000	2431	0,27333	0,30883	0,18278	0,00039
2264	0,21429	0,24946	0,04996	0,00021	2264	0,52113	0,47191	0,43556	0,00049	2264	0,46000	0,51672	0,24866	0,00045
2256	0,46429	0,47719	0,85774	0,00035	2256	0,32394	0,28189	0,12777	0,00034	2256	0,46000	0,50165	0,32909	0,00044
128	0,37857	0,36536	0,81659	0,00039	128	0,19718	0,17837	0,36217	0,00046	128	0,38667	0,41596	0,43036	0,00051
15	0,20714	0,24063	0,14478	0,00036	15	0,50704	0,46952	0,37391	0,00050	15	0,46667	0,47884	0,86335	0,00036
2241	0,07857	0,10177	0,04656	0,00021	2241	0,42958	0,48363	0,22286	0,00046	2241	0,38000	0,36948	0,82609	0,00038
419	0,35714	0,45223	0,01429	0,00012	419	0,12676	0,14343	0,19029	0,00039	419	0,34000	0,34058	1,00000	0,00000
943	0,32374	0,39542	0,05043	0,00022	943	0,46479	0,45787	1,00000	0,00000	943	0,38095	0,46992	0,02301	0,00015
159	0,29286	0,33382	0,19966	0,00040	159	0,41549	0,41514	1,00000	0,00000	159	0,47333	0,46419	0,86091	0,00034
2005	0,32857	0,39693	0,05222	0,00019	2005	0,28169	0,26278	0,52839	0,00051	2005	0,51333	0,48025	0,49364	0,00048
250	0,42143	0,48313	0,16180	0,00039	250	0,35211	0,41218	0,10126	0,00029	250	0,49333	0,46600	0,48832	0,00050
1802	0,02143	0,02127	1,00000	0,00000	1802	0,45070	0,47629	0,59462	0,00048	1802	0,09333	0,08928	1,00000	0,00000
1607	0,37857	0,49603	0,00625	0,00007	1607	0,40845	0,46195	0,20094	0,00039	1607	0,46000	0,48988	0,50316	0,00047
1734	0,47857	0,50156	0,61653	0,00048	1734	0,22535	0,23262	0,71651	0,00045	1734	0,36000	0,38609	0,40455	0,00051
406	0,31429	0,41567	0,00655	0,00009	406	0,37324	0,36179	0,81679	0,00038	406	0,42000	0,40747	0,83993	0,00036
1386	0,42857	0,48705	0,16476	0,00035	1386	0,32394	0,37277	0,17100	0,00036	1386	0,47333	0,48025	0,86583	0,00034
1726	0,37857	0,41267	0,40948	0,00054	1726	0,54225	0,48621	0,22645	0,00042	1726	0,40667	0,40747	1,00000	0,00000
3626	0,05714	0,05571	1,00000	0,00000	3626	0,32394	0,34251	0,62382	0,00049	3626	0,46000	0,48292	0,60876	0,00050
360	0,45000	0,44524	1,00000	0,00000	360	0,52113	0,52230	0,74385	0,00041	360	0,36000	0,37790	0,25536	0,00041
1603	0,27857	0,34201	0,04251	0,00020	1603	0,33099	0,30466	0,40851	0,00050	1603	0,51333	0,49184	0,61807	0,00047
2719	0,41429	0,50138	0,04353	0,00020	2719	0,45775	0,46770	0,85766	0,00034	2719	0,40000	0,50132	0,01485	0,00013
Ho =	0,24120				Ho =	0,36819				Ho =	0,36435			

*Lo resaltado en verde evidencia un valor P menor a 0.05

La Tabla 7, refleja los resultados del análisis del equilibrio de Hardy-Weinberg, mostrando que la gran mayoría de los 46 loci para las poblaciones de afrodescendientes y del grupo de ancestralidad mixta no presentaban desviación del equilibrio. Es llamativo para la población afrodescendiente únicamente la presencia de desequilibrio en el marcador MID-548. Adicionalmente, para la población amerindia se evidencio desviación del equilibrio de Hardy-Weinberg para algunos loci, sin embargo, al evaluar cada etnia solo un marcador en Arhuacos (MID-2264), Curripakos (MID-2431) y Tunebos (MID-94) se encontraban en desequilibrio.

Los resultados del cálculo de AMOVA demuestran que en promedio se obtuvo un porcentaje de la varianza entre poblaciones de 16.67% y entre los individuos de 79.11%. Adicionalmente, el porcentaje de la varianza para los individuos dentro de su población fue de 1.93%, lo que indica que los individuos de cada población tienden a ser homogéneos. (Tabla 8).

Tabla 8. Cálculo de AMOVA para los subgrupos poblacionales.

Fuente de Variación	DF	Suma de Cuadrados	Componentes de Varianza	% de Variación
Entre Grupos	2	1.023,943	1,59184	16,67
Entre Poblaciones dentro de los grupos	35	443,346	0,21855	2,29
Entre Individuos dentro de las poblaciones	426	3.373,609	0,18409	1,93
Entre individuos	464	3.504,000	7,55454	79,11
Total	927	8.344,898	9,54902	

4.3 Análisis de la Subestructura y Distanciamiento Genético:

Para evaluar el distanciamiento genético y el nivel de subestructura de las poblaciones estudiadas en este proyecto, se calculó el estadístico F_{st} (Tabla 9). Los resultados evidencian una menor diferenciación entre las poblaciones de Bogotá, Boyacá, Cundinamarca y Santander. Así mismo, los valores de diferenciación fueron más altos entre las poblaciones amerindias en comparación a las poblaciones afrodescendientes. Adicionalmente, entre las poblaciones afrodescendientes llama la atención como la población de Quibdó y Bahía Solano a pesar de que el grado de los valores de distanciamiento F_{st} son bajos, el valor de P es <0.05 indicando que existe una diferenciación estadísticamente significativa con respecto a las poblaciones de Condoto, Istmina y Tadó.

En el caso de las poblaciones amerindias, las etnias Arhuacos, Bari, Chimila, Guayaberos y Kogui son las poblaciones que más se alejan genéticamente del resto de etnias. Por otro lado, en las poblaciones Inga, Waunano, Yeral, Tukano y Barasano no se encontraron valores de P estadísticamente significativos ($P>0.05$) frente a las poblaciones de Boyacá, Cundinamarca y Santander, es decir, no se presenta un distanciamiento entre estas poblaciones.

Lo anteriormente mencionado, puede correlacionarse con los datos evidenciados en los dendogramas realizados a partir de la diferenciación de locus, en donde se refleja como las etnias Arhuacos, Bari, Chimila, Guayaberos y Kogui se localizan más cercanas a las poblaciones de ancestralidad mixta. (Figura 5).

Tabla 9. Distanciamiento genético. Valores Fst se muestran debajo de la diagonal. Los P-Valores se muestran por arriba de la diagonal. Los P-Valor >0.05 están resaltados en Verde. Lo P-Valor <0.05 entre poblaciones que pertenecen al mismo grupo ancestral están resaltados en amarillo.

	Santander	Cundinamarca	Boyaca	Bogotá	Tadó	Istmina	Condoto	Quibdo 1	Quibdo 2	Bahía Solano	Wounaan	Wanano	Yeral	Tikuna	Siriano	Makuna
Santander	*	0.67568+-0.0455	0.75676+-0.0579	0.55856+-0.0522	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.09910+-0.0286	0.05405+-0.0201	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Cundinamarca	0.00191	*	0.49550+-0.0595	0.68468+-0.0433	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.18919+-0.0394	0.12613+-0.0309	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Boyaca	0.00236	0.00006	*	0.39640+-0.0594	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Bogotá	0.00038	0.00107	0.00168	*	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.04505+-0.0203	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Tadó	0.13878	0.15936	0.15992	0.16987	*	0.11712+-0.0237	0.00901+-0.0091	0.00000+-0.0000	0.31532+-0.0412	0.00000+-0.0000	0.00000+-0.0000	0.01802+-0.0121	0.00000+-0.0000	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Istmina	0.16255	0.18750	0.19027	0.20140	0.00395	*	0.48649+-0.0309	0.00000+-0.0000	0.08108+-0.0252	0.00000+-0.0000	0.00000+-0.0000	0.00901+-0.0091	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Condoto	0.17639	0.20101	0.19716	0.21633	0.00872	0.00001	*	0.00000+-0.0000	0.02703+-0.0139	0.00000+-0.0000	0.00000+-0.0000	0.01802+-0.0121	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Quibdo 1	0.06672	0.09065	0.09060	0.09547	0.01950	0.02470	0.02882	*	0.16216+-0.0424	0.83784+-0.0451	0.00000+-0.0000	0.00901+-0.0091	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Quibdo 2	0.11296	0.13827	0.13403	0.14797	0.00191	0.00629	0.00984	0.00396	*	0.16216+-0.0326	0.00000+-0.0000	0.00901+-0.0091	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Bahía Solano	0.06242	0.08618	0.08352	0.09530	0.02079	0.03025	0.03315	0.00436	0.00589	*	0.00000+-0.0000	0.01802+-0.0121	0.00901+-0.0091	0.00000+-0.0000	0.99099+-0.0030	0.99099+-0.0030
Wounaan	0.10020	0.09378	0.10760	0.13103	0.21510	0.27900	0.27989	0.18167	0.21982	0.15198	*	0.55856+-0.0618	0.56757+-0.0360	0.09009+-0.0303	0.99099+-0.0030	0.99099+-0.0030
Wanano	0.07399	0.05169	0.06477	0.07551	0.23000	0.27484	0.28424	0.17003	0.20641	0.15030	0.08642	*	0.99099+-0.0030	0.30631+-0.0388	0.99099+-0.0030	0.99099+-0.0030
Yeral	0.05728	0.04969	0.07468	0.09654	0.21053	0.22901	0.24284	0.13572	0.16773	0.10809	0.05286	0.00595	*	0.14141+-0.0309	0.99099+-0.0030	0.99099+-0.0030
Tikuna	0.17255	0.15042	0.16521	0.18245	0.30639	0.35319	0.35855	0.26108	0.29525	0.23389	0.24952	0.13312	0.16814	*	0.99099+-0.0030	0.99099+-0.0030
Siriano	0.05937	0.02065	0.04950	0.05064	0.20830	0.26351	0.27619	0.14806	0.18125	0.10534	0.05764	0.06667	0.08711	0.27779	*	0.99099+-0.0030
Makuna	0.00137	0.01035	0.02022	0.03767	0.19256	0.22203	0.23316	0.09795	0.15381	0.09624	0.06380	0.03291	0.03507	0.27362	0.13793	*
Kubeo	0.09827	0.08573	0.08986	0.11198	0.25588	0.29889	0.30795	0.20387	0.23418	0.16848	0.06126	0.00891	0.04280	0.09881	0.04542	0.02766
Kawiyari	0.06340	0.06601	0.07645	0.07432	0.18724	0.23639	0.25638	0.12870	0.17964	0.09715	0.26671	0.09973	0.07259	0.29727	0.03846	0.20000
Desano	0.10213	0.08517	0.08578	0.11515	0.26578	0.30571	0.31439	0.21010	0.23942	0.18009	0.01142	0.06122	0.02417	0.07183	0.01961	0.01554
Curripako	0.09445	0.07422	0.08089	0.08965	0.24479	0.28727	0.29985	0.19382	0.22176	0.16895	0.05915	0.00832	0.01062	0.15030	0.04955	0.00005
Carapana	0.01742	0.01504	0.00868	0.05499	0.20625	0.23769	0.25320	0.13931	0.16679	0.10099	0.11269	0.02663	0.01024	0.18114	0.06897	0.10000
Barasano	0.06132	0.05127	0.04798	0.07645	0.19995	0.25756	0.25094	0.15386	0.18061	0.11668	0.10843	0.06149	0.03704	0.16615	0.00901	0.12634
Yukpa	0.12970	0.12754	0.11990	0.15305	0.29079	0.32475	0.33814	0.23693	0.27627	0.21834	0.22203	0.23123	0.18024	0.32245	0.29272	0.23534
Wayuu	0.08691	0.08188	0.07475	0.11148	0.24661	0.29594	0.29283	0.19683	0.22368	0.15519	0.05036	0.03898	0.07792	0.12016	0.02554	0.12118
Tunebo	0.12256	0.11076	0.11466	0.13899	0.29527	0.34059	0.33945	0.24250	0.27237	0.21264	0.15707	0.03180	0.01820	0.14089	0.13446	0.15732
Tukano	0.05353	0.03944	0.04458	0.06240	0.23161	0.26455	0.28056	0.16550	0.20737	0.14551	0.01050	0.02913	0.01502	0.19989	0.04518	0.03121
Puinave	0.06696	0.05192	0.05170	0.07395	0.23924	0.27034	0.27885	0.17493	0.20526	0.14289	0.04076	0.00163	0.05536	0.11051	0.00302	0.02259
Piratapuyo	0.08232	0.07345	0.08745	0.09739	0.25409	0.30113	0.31353	0.19923	0.22967	0.16167	0.05380	0.05674	0.08000	0.11798	0.09116	0.18273
Piapoco	0.08377	0.04275	0.04649	0.07742	0.26842	0.33032	0.31606	0.19891	0.23421	0.16714	0.05358	0.06801	0.02897	0.07951	0.17391	0.15385
Nasa	0.11110	0.08522	0.07915	0.11844	0.26221	0.31353	0.31427	0.22338	0.24289	0.19558	0.22131	0.06193	0.11111	0.23671	0.13573	0.22427
Nukak	0.09206	0.07669	0.08054	0.09070	0.26273	0.31453	0.31891	0.21022	0.24545	0.17732	0.07467	0.05035	0.07462	0.09577	0.00800	0.07886
Kogui	0.16566	0.14432	0.15932	0.18853	0.31463	0.34973	0.35342	0.26339	0.28461	0.23258	0.14889	0.12496	0.01579	0.23709	0.16255	0.13957
Guayabero	0.13727	0.10718	0.12028	0.14038	0.27840	0.32348	0.32235	0.22442	0.25806	0.19659	0.05840	0.02588	0.09263	0.17806	0.01451	0.12421
Misak	0.05352	0.05513	0.04373	0.05481	0.25050	0.29364	0.31161	0.18949	0.23359	0.13796	0.27294	0.13333	0.09637	0.32790	0.07143	0.32258
Inga	0.04938	0.03176	0.04470	0.06289	0.22135	0.25660	0.26912	0.15933	0.18809	0.11477	0.00936	0.02038	0.05436	0.06877	0.06412	0.01143
Chimila	0.17354	0.13613	0.14548	0.16886	0.32567	0.36158	0.37669	0.27110	0.30486	0.24109	0.20490	0.12406	0.12066	0.20081	0.14395	0.19786
Bari	0.17996	0.15282	0.15852	0.19766	0.33092	0.37112	0.37598	0.28634	0.31668	0.25497	0.17387	0.13664	0.12037	0.24776	0.21317	0.17874
Arhuacos	0.10543	0.08703	0.09727	0.11655	0.27458	0.31547	0.32100	0.21719	0.25125	0.18870	0.14438	0.06237	0.04770	0.20670	0.14314	0.12355

*Las casillas resaltadas en azul evidencian un P-Valor >0.05 en poblaciones con tamaño de muestra muy bajo (n=1).

Continuación Tabla 9. Distanciamiento genético. Valores Fst se muestran debajo de la diagonal. Los P-Valores se muestran por arriba de la diagonal. Los P-Valor >0.05 están resaltados en Verde. Lo P-Valor <0.05 entre poblaciones que pertenecen al mismo grupo ancestral están resaltados en amarillo.

Kubeo	Kawiari	Desano	Curripako	Carapana	Barasano	Yukpa	Wayuu	Tunebo	Tukano	Puinave	Piratapuyo	Piapoco	Nasa	Nukak	Kogui	Guayabero
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.02703+-0.0139	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.04505+-0.0152	0.0000+-0.0000	0.03604+-0.0201	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.04505+-0.0244	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.03604+-0.0278	0.0000+-0.0000	0.01802+-0.0121	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.07207+-0.0227	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.10811+-0.0326	0.0000+-0.0000	0.03604+-0.0201	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.01802+-0.0121	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.04505+-0.0152	0.0000+-0.0000	0.00901+-0.0091	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.03604+-0.0148	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.03604+-0.0201	0.0000+-0.0000	0.00901+-0.0091	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.00901+-0.0091	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.00901+-0.0091	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.03604+-0.0148	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.00901+-0.0091	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.00901+-0.0091	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.00901+-0.0091	0.0000+-0.0000	0.01802+-0.0121	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.02703+-0.0139	0.0000+-0.0000	0.0000+-0.0000	0.9909+-0.0030	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000	0.0000+-0.0000
0.0909+-0.0235	0.9909+-0.0030	0.50450+-0.0493	0.19820+-0.0572	0.9909+-0.0030	0.03604+-0.0201	0.07207+-0.0378	0.16216+-0.0264	0.00901+-0.0091	0.55856+-0.0466	0.03604+-0.0148	0.19820+-0.0353	0.9909+-0.0030	0.01802+-0.0121	0.03604+-0.0148	0.0000+-0.0000	0.04505+-0.0152
0.52252+-0.0354	0.9909+-0.0030	0.84685+-0.0310	0.65766+-0.0334	0.9909+-0.0030	0.9909+-0.0030	0.06306+-0.0194	0.28829+-0.0539	0.35135+-0.0550	0.9909+-0.0030	0.39640+-0.0388	0.61261+-0.0379	0.9909+-0.0030	0.31532+-0.0204	0.92793+-0.0196	0.00901+-0.0091	0.20721+-0.0385
0.18919+-0.0316	0.9909+-0.0030	0.51351+-0.0731	0.66667+-0.0562	0.9909+-0.0030	0.63964+-0.0394	0.13514+-0.0244	0.13514+-0.0389	0.26126+-0.0438	0.62162+-0.0345	0.9909+-0.0235	0.25225+-0.0588	0.9909+-0.0030	0.08108+-0.0212	0.16216+-0.0379	0.37838+-0.0485	0.09910+-0.0286
0.01802+-0.0121	0.9909+-0.0030	0.16216+-0.0264	0.02703+-0.0194	0.9909+-0.0030	0.08108+-0.0252	0.05405+-0.0242	0.12613+-0.0201	0.00000+-0.0000	0.14414+-0.0201	0.00000+-0.0000	0.14414+-0.0242	0.9909+-0.0030	0.02703+-0.0139	0.00000+-0.0000	0.00000+-0.0000	0.00901+-0.0091
0.40541+-0.0412	0.9909+-0.0030	0.62162+-0.0317	0.87387+-0.0238	0.9909+-0.0030	0.61261+-0.0354	0.27928+-0.0497	0.11712+-0.0305	0.10811+-0.0353	0.63964+-0.0417	0.41441+-0.0388	0.35135+-0.0344	0.9909+-0.0030	0.12613+-0.0337	0.74775+-0.0385	0.10811+-0.0326	0.28829+-0.0485
0.30631+-0.0243	0.9909+-0.0030	0.72072+-0.0450	0.76577+-0.0366	0.9909+-0.0030	0.37838+-0.0504	0.20721+-0.0194	0.21622+-0.0278	0.12613+-0.0337	0.39640+-0.0411	0.35135+-0.0438	0.36937+-0.0438	0.9909+-0.0030	0.27928+-0.0417	0.20721+-0.0385	0.19820+-0.0424	0.10811+-0.0297
*	0.9909+-0.0030	0.89189+-0.0371	0.45946+-0.0252	0.9909+-0.0030	0.18919+-0.0370	0.00000+-0.0000	0.63063+-0.0272	0.00000+-0.0000	0.33333+-0.0451	0.76577+-0.0311	0.52252+-0.0466	0.9909+-0.0030	0.00901+-0.0091	0.66667+-0.0310	0.00000+-0.0000	0.07207+-0.0264
0.08831	*	0.36036+-0.0359	0.48649+-0.0454	0.9909+-0.0030	0.70270+-0.0327	0.24324+-0.0333	0.19820+-0.0182	0.12613+-0.0309	0.37838+-0.0556	0.14414+-0.0364	0.36937+-0.0394	0.9909+-0.0030	0.23423+-0.0492	0.45045+-0.0575	0.09910+-0.0286	0.11712+-0.0273
0.03660	0.08808	*	0.93694+-0.0151	0.9909+-0.0030	0.60360+-0.0360	0.05405+-0.0201	0.87387+-0.0238	0.01802+-0.0121	0.83784+-0.0334	0.93694+-0.0203	0.74775+-0.0451	0.9909+-0.0030	0.06306+-0.0194	0.97297+-0.0125	0.00000+-0.0000	0.04505+-0.0152
0.00137	0.10540	0.03010	*	0.9909+-0.0030	0.71171+-0.0479	0.08108+-0.0252	0.27027+-0.0489	0.00901+-0.0091	0.87387+-0.0306	0.40541+-0.0245	0.64865+-0.0485	0.9909+-0.0030	0.05405+-0.0201	0.56757+-0.0471	0.00901+-0.0091	0.00000+-0.0000
0.07630	0.18519	0.16175	0.06393	*	0.65766+-0.0238	0.20721+-0.0408	0.78378+-0.0274	0.09090+-0.0192	0.9909+-0.0030	0.9909+-0.0030	0.9909+-0.0030	0.9909+-0.0030	0.14414+-0.0337	0.9909+-0.0030	0.25225+-0.0379	0.18018+-0.0469
0.01158	0.04450	0.00000	0.00754	0.00713	*	0.17117+-0.0252	0.44144+-0.0628	0.10811+-0.0297	0.27027+-0.0359	0.23423+-0.0278	0.32432+-0.0454	0.9909+-0.0030	0.39640+-0.0364	0.80180+-0.0214	0.02703+-0.0194	0.17117+-0.0370
0.14584	0.31272	0.12168	0.09675	0.06716	0.12628	*	0.01802+-0.0121	0.00000+-0.0000	0.10811+-0.0264	0.00000+-0.0000	0.17117+-0.0316	0.9909+-0.0030	0.03604+-0.0201	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.01562	0.16705	0.02418	0.02494	0.07032	0.00883	0.14681	*	0.01802+-0.0121	0.40541+-0.0339	0.44144+-0.0286	0.61261+-0.0297	0.9909+-0.0030	0.10811+-0.0353	0.86486+-0.0389	0.00000+-0.0000	0.00901+-0.0091
0.07056	0.13299	0.06988	0.10161	0.07004	0.06877	0.25762	0.06292	*	0.11712+-0.0360	0.00000+-0.0000	0.20721+-0.0408	0.9909+-0.0030	0.00000+-0.0000	0.01802+-0.0121	0.00000+-0.0000	0.00000+-0.0000
0.01359	0.16469	0.04215	0.03030	0.11448	0.06195	0.20496	0.02918	0.11560	*	0.50450+-0.0309	0.35135+-0.0497	0.9909+-0.0030	0.20721+-0.0360	0.39640+-0.0411	0.01802+-0.0121	0.06306+-0.0194
0.01611	0.09144	0.03824	0.00652	0.14127	0.00633	0.12550	0.00822	0.07426	0.02358	*	0.61261+-0.0466	0.9909+-0.0030	0.01802+-0.0121	0.67568+-0.0455	0.00000+-0.0000	0.00901+-0.0091
0.02537	0.23553	0.03591	0.02227	0.08108	0.05674	0.28422	0.02372	0.04613	0.05185	0.02072	*	0.9909+-0.0030	0.07207+-0.0227	0.96396+-0.0142	0.00901+-0.0091	0.03604+-0.0201
0.00325	0.29032	0.08675	0.08081	0.04348	0.13673	0.24272	0.10667	0.03464	0.02408	0.02860	0.05512	*	0.43243+-0.0522	0.9909+-0.0030	0.27928+-0.0370	0.77477+-0.0511
0.11532	0.23065	0.07008	0.09852	0.08966	0.00574	0.20220	0.10657	0.14473	0.08400	0.04785	0.18199	0.05358	*	0.02703+-0.0139	0.00000+-0.0000	0.00000+-0.0000
0.00881	0.05318	0.03822	0.00145	0.01196	0.03488	0.16672	0.02593	0.05375	0.02144	0.00861	0.03259	0.08204	0.09330	*	0.00000+-0.0000	0.04505+-0.0203
0.14155	0.30683	0.10428	0.09283	0.06118	0.11144	0.20521	0.13567	0.15063	0.19011	0.12806	0.13706	0.04739	0.16966	0.14014	*	0.00000+-0.0000
0.02979	0.22466	0.04673	0.06021	0.09718	0.02588	0.23661	0.03581	0.13690	0.06433	0.04697	0.10566	0.08636	0.15899	0.09410	0.19299	*
0.14415	0.08696	0.12473	0.07635	0.08000	0.01961	0.16443	0.10945	0.16212	0.17068	0.09389	0.19001	0.17857	0.17264	0.09414	0.19019	0.29661
0.03817	0.08865	0.05827	0.00087	0.14699	0.03285	0.16813	0.02564	0.06810	0.02381	0.03531	0.06024	0.08736	0.05815	0.01873	0.07189	0.03047
0.09451	0.20770	0.09970	0.10869	0.10616	0.16007	0.24075	0.15753	0.19529	0.17388	0.10563	0.14707	0.06340	0.21476	0.11592	0.15322	0.12100
0.13658	0.31292	0.12844	0.13014	0.12198	0.18051	0.23878	0.07722	0.15972	0.11182	0.12613	0.17301	0.08109	0.21812	0.11693	0.15760	0.17129
0.08261	0.16425	0.11843	0.11166	0.06358	0.10711	0.18182	0.09485	0.09377	0.10561	0.06933	0.12569	0.10499	0.09548	0.10258	0.10574	0.13473

*Las casillas resaltadas en azul evidencian un P-Valor >0.05 en poblaciones con tamaño de muestra muy bajo (n=1).

Continuación Tabla 9. Distanciamiento genético. Valores Fst se muestran debajo de la diagonal. Los P-Valores se muestran por arriba de la diagonal. Los P-Valor >0.05 están resaltados en Verde. Lo P-Valor <0.05 entre poblaciones que pertenecen al mismo grupo ancestral están resaltados en amarillo.

Misak	Inga	Chimila	Bari	Arhuacos
0.99099+-0.0030	0.03604+-0.0148	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.05405+-0.0148	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.07207+-0.0297	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00901+-0.0091	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.58559+-0.0365	0.00901+-0.0091	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.88288+-0.0184	0.03604+-0.0201	0.00901+-0.0091	0.12613+-0.0278
0.99099+-0.0030	0.75676+-0.0364	0.04505+-0.0203	0.01802+-0.0121	0.11712+-0.0490
0.99099+-0.0030	0.10811+-0.0264	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.99099+-0.0030	0.09910+-0.0212	0.04505+-0.0244	0.04505+-0.0203
0.99099+-0.0030	0.54054+-0.0470	0.08108+-0.0212	0.17117+-0.0438	0.12613+-0.0278
0.99099+-0.0030	0.82883+-0.0425	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.16216+-0.0424	0.13514+-0.0244	0.03604+-0.0148	0.01802+-0.0121
0.99099+-0.0030	0.97297+-0.0125	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.60360+-0.0334	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.99099+-0.0030	0.09910+-0.0212	0.19820+-0.0264	0.27027+-0.0359
0.99099+-0.0030	0.65766+-0.0543	0.00901+-0.0091	0.00000+-0.0000	0.00901+-0.0091
0.99099+-0.0030	0.13514+-0.0339	0.01802+-0.0121	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.88288+-0.0327	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.69369+-0.0385	0.00000+-0.0000	0.00901+-0.0091	0.00000+-0.0000
0.99099+-0.0030	0.86486+-0.0244	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.91892+-0.0298	0.01802+-0.0121	0.02703+-0.0194	0.00000+-0.0000
0.99099+-0.0030	0.99099+-0.0030	0.31532+-0.0339	0.12613+-0.0364	0.09910+-0.0212
0.99099+-0.0030	0.24324+-0.0385	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.73874+-0.0539	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
0.99099+-0.0030	0.11712+-0.0273	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
*	0.52252+-0.0354	0.10811+-0.0402	0.05405+-0.0201	0.03604+-0.0148
0.01326	*	0.05405+-0.0201	0.00000+-0.0000	0.03604+-0.0201
0.23104	0.07365	*	0.00000+-0.0000	0.00000+-0.0000
0.26712	0.12272	0.18988	*	0.00000+-0.0000
0.15262	0.07026	0.15195	0.13175	*

*Las casillas resaltadas en azul evidencian un P-Valor >0.05 en poblaciones con tamaño de muestra muy bajo (n=1).

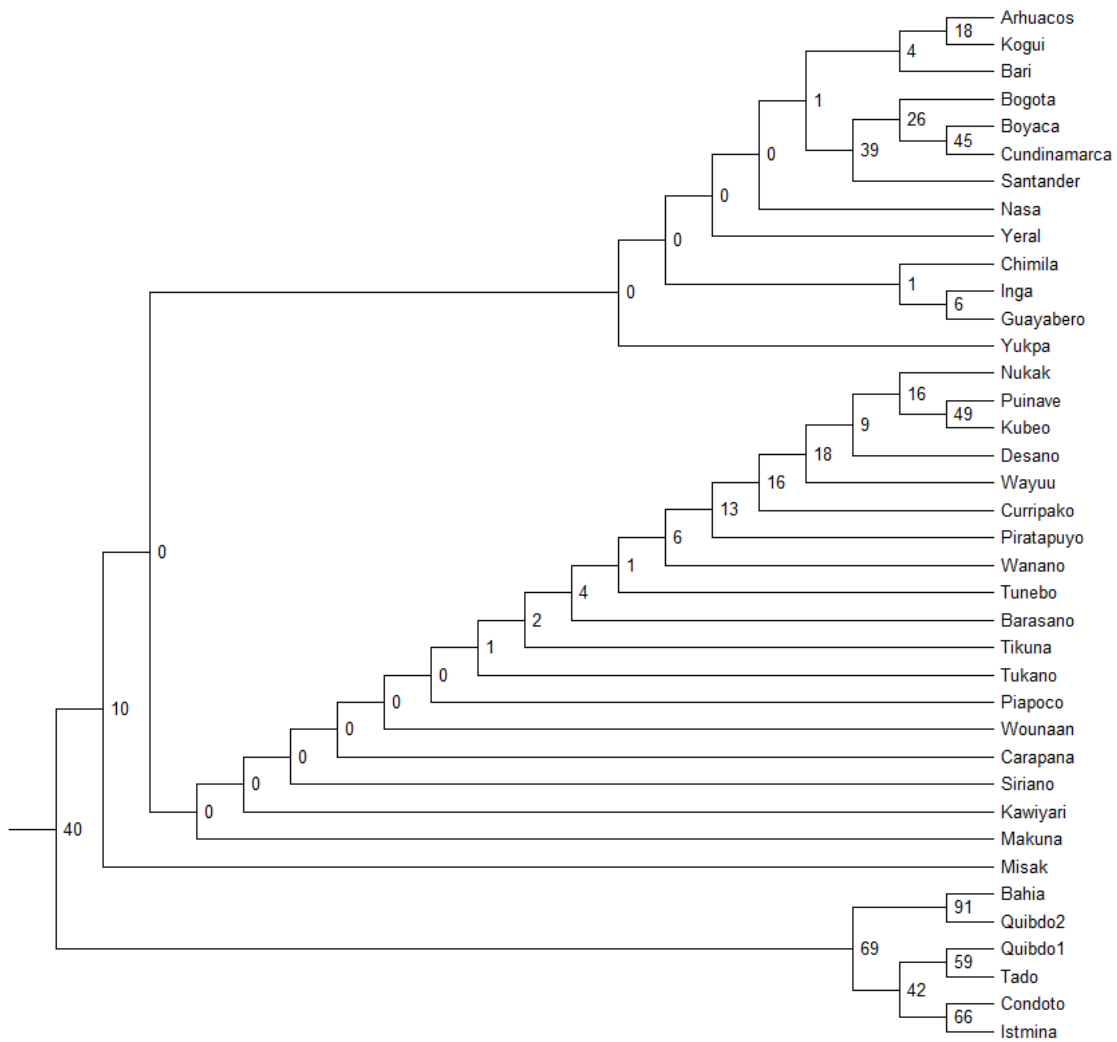


Figura 5. Dendrograma construido a partir de los valores Fst de 38 poblaciones colombianas.

Analizando en conjunto los datos previamente mencionados, se sugiere que las muestras de las 38 poblaciones incluidas en este estudio pueden agruparse en 3 grupos, debido a la diferenciación genética estadísticamente significativa entre ellas (P -Valor < 0.05), esta asociación se comprueba al realizar el cálculo del estimado Fst y Dendrograma por análisis de distancia genética de Cavalli-Sforza agrupando las poblaciones en los grupos Amerindios, Afrodescendiente y Ancestralidad Mixta. (Tabla 10 y Figura 6)

Tabla 10. Valores Fst para los 3 subgrupos poblacionales.

Tabla. Valores Fst para los 3 subgrupos poblacionales ancestrales			
	Afrocolombianos	Amerindios	Caucasico- Mestizos
Afrocolombianos	*	0.00000+-0.0000	0.00000+-0.0000
Amerindios	0.25723	*	0.00000+-0.0000
Ancestralidad Mixta	0.14028	0.09347	*

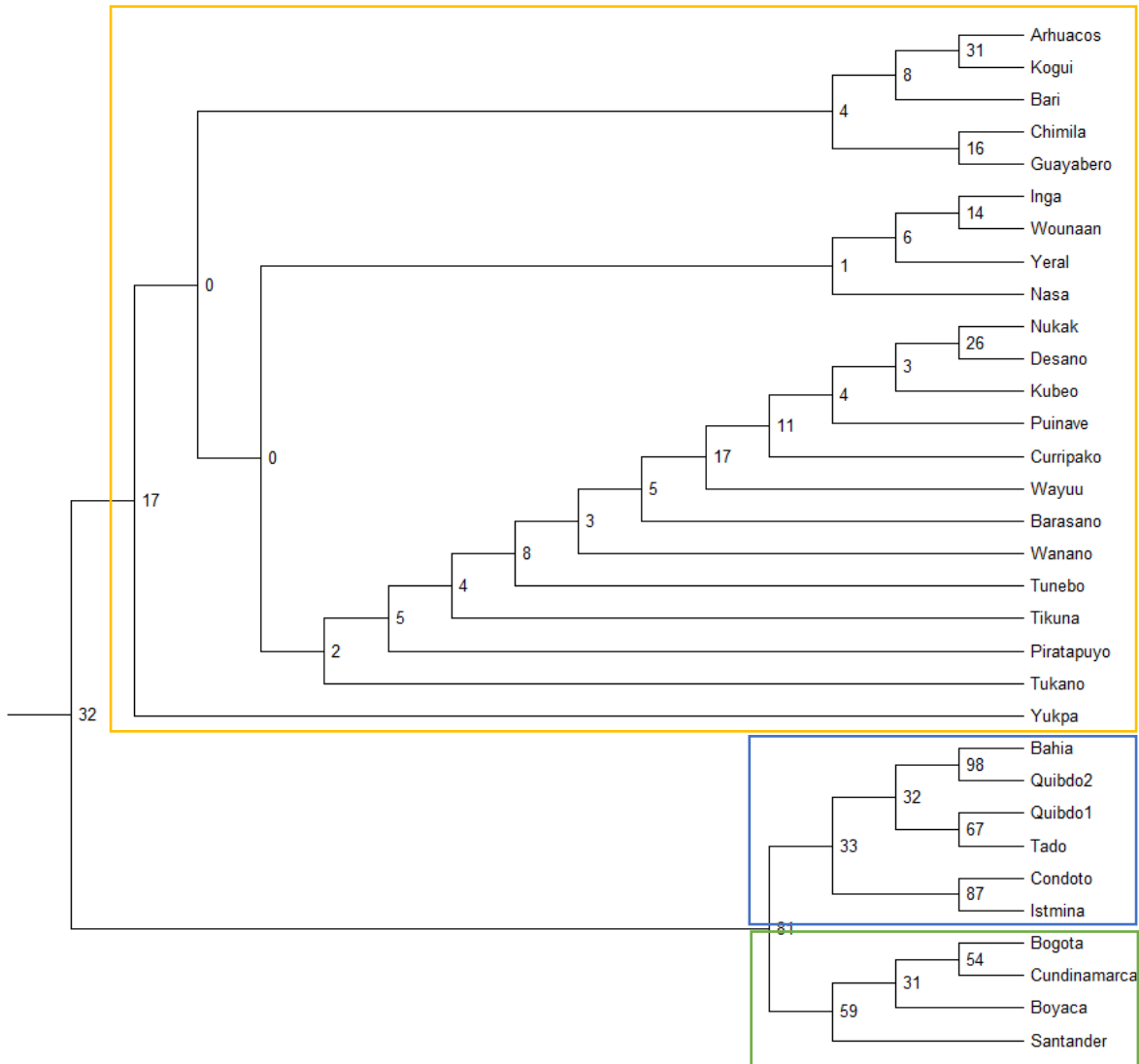


Figura 6. Dendrograma UPGMA construido a partir del análisis de distancia genética de Cavalli-Sforza en 33 grupos de poblaciones de Colombia. Se excluyeron las poblaciones amerindias con un n=1. Se resalta en cuadro de colores las agrupaciones sugeridas por el Dendrograma.

4.4. Estimación de Componentes Ancestrales:

Las tablas 11 y 12 representan los porcentajes de ancestralidad obtenidos para cada grupo poblacional incluyendo las muestras de referencia para África (AFR), Europa (EUR), Asia del este (EAS) y Nativos americanos (NAM). Los resultados obtenidos evidencian una diferencia en los porcentajes del componente ancestral cuando se utilizan como base un K=4 vs K=3 para la estimación de la inferencia de ancestría en la muestra poblacional estudiada.

En el grupo de amerindios los porcentajes de ancestralidad nativo-americana varían de 88,2% vs 94%, obteniendo un cambio de alrededor de 6% en la inferencia del componente ancestral. A su vez, la estimación de la proporción del componente europeo en el grupo de ancestralidad mixta se ajustó de un 58,1% a 55,8%. No hubo diferencia en la estimación del componente africano para el grupo de muestra de afrodescendientes colombianos.

Tabla 11. Porcentaje de proporciones utilizando K=4.

Población	1	2	3	4
AFR	0,959	0,013	0,017	0,011
EUR	0,014	0,944	0,023	0,019
ASN	0,014	0,035	0,910	0,041
NAM	0,013	0,042	0,043	0,901
AMR	0,022	0,058	0,038	0,882
AFC	0,682	0,175	0,049	0,094
MIX	0,047	0,581	0,068	0,304

Tabla 12. Porcentaje de proporciones utilizando un K=3.

Población	1	2	3
AFR	0,955	0,022	0,022
EUR	0,025	0,938	0,037
NAM	0,021	0,053	0,927
AMR	0,012	0,048	0,940
AFC	0,680	0,200	0,120
MIX	0,065	0,558	0,376

En contraste con los resultados previos, se realizaron diferentes simulaciones en STRUCTURE con un $K=2-7$. Se seleccionó $K=3$ como el k más probable dado que es el valor más bajo de k en el que captura el grado máximo de estructura detectado en los datos y dado que también muestra el valor máximo de Δk , según el método de Evanno et al. (2005), (Figura 7).

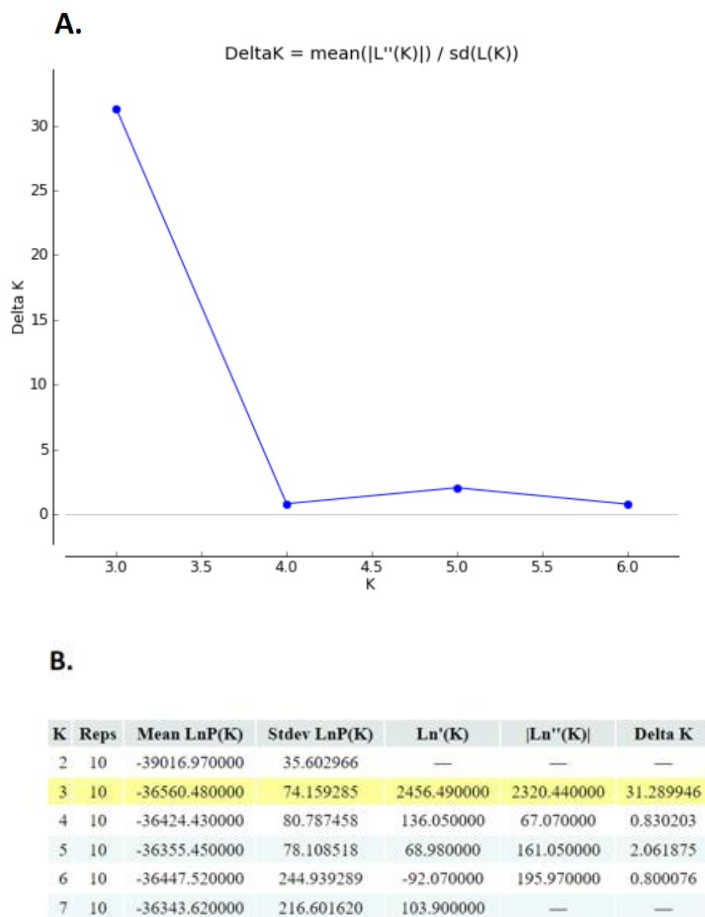
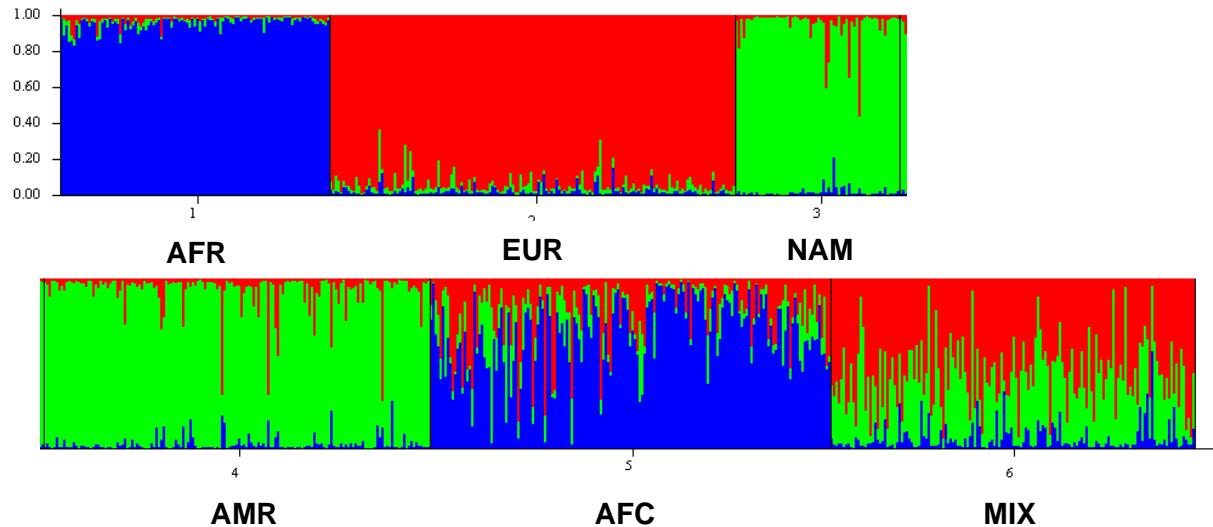


Figura 7. Resultados del método de Evanno utilizados para estimar el valor de k más adecuado para el análisis de STRUCTURE. A. Valores modales de Δ -K calculados por el método de Evanno, Regnaut y Goudet (2005). B. Tabla de resumen de estadísticas donde $k=3$ está resaltado.

En la figura 8, se evidencia del análisis de STRUCTURE donde muestra la composición ancestral para todos los individuos agrupándolos en los 3 subgrupos poblaciones analizadas amerindios (AMR), afrodescendientes (AFC) y Ancestralidad Mixta (MIX), incluyendo las muestras de referencia de África (AFR), Europa (EUR) y Nativos americanos (NAM).

Figura 8. Bar Plot de STRUCTURE utilizando K=3.



Estos resultados se correlacionan con lo reportado previamente a nivel de valores F_{st} y dendograma a partir de distancias genéticas, en donde se pueden distinguir la diferenciación de las poblaciones incluidas en este estudio en los 3 subgrupos poblacionales ancestrales.

En las muestras de poblaciones amerindias y afrodescendientes se observan individuos con diferentes grados de aporte de otros componentes ancestrales, lo que refleja un grado de mezcla significativo en estos individuos.

Se realizó una alineación por grado de porcentaje del componente nativo americano a través de STRUCTURE en donde se pudo observar que algunos individuos de poblaciones amerindias presentaban en su estimación una mayor ancestralidad europea o un grado de otros componentes ancestrales superior al nativo americano. No obstante, algunas de las muestras de referencia utilizadas por Pereira et al también presentaban hallazgos similares (Figura 9 y 10).

Figura 9. Bar plot STRUCTURE BY SORTING de población amerindia (AMR) utilizando K=3.

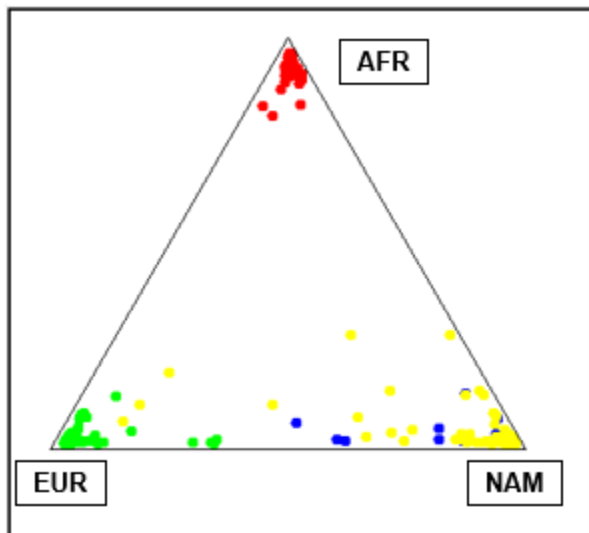
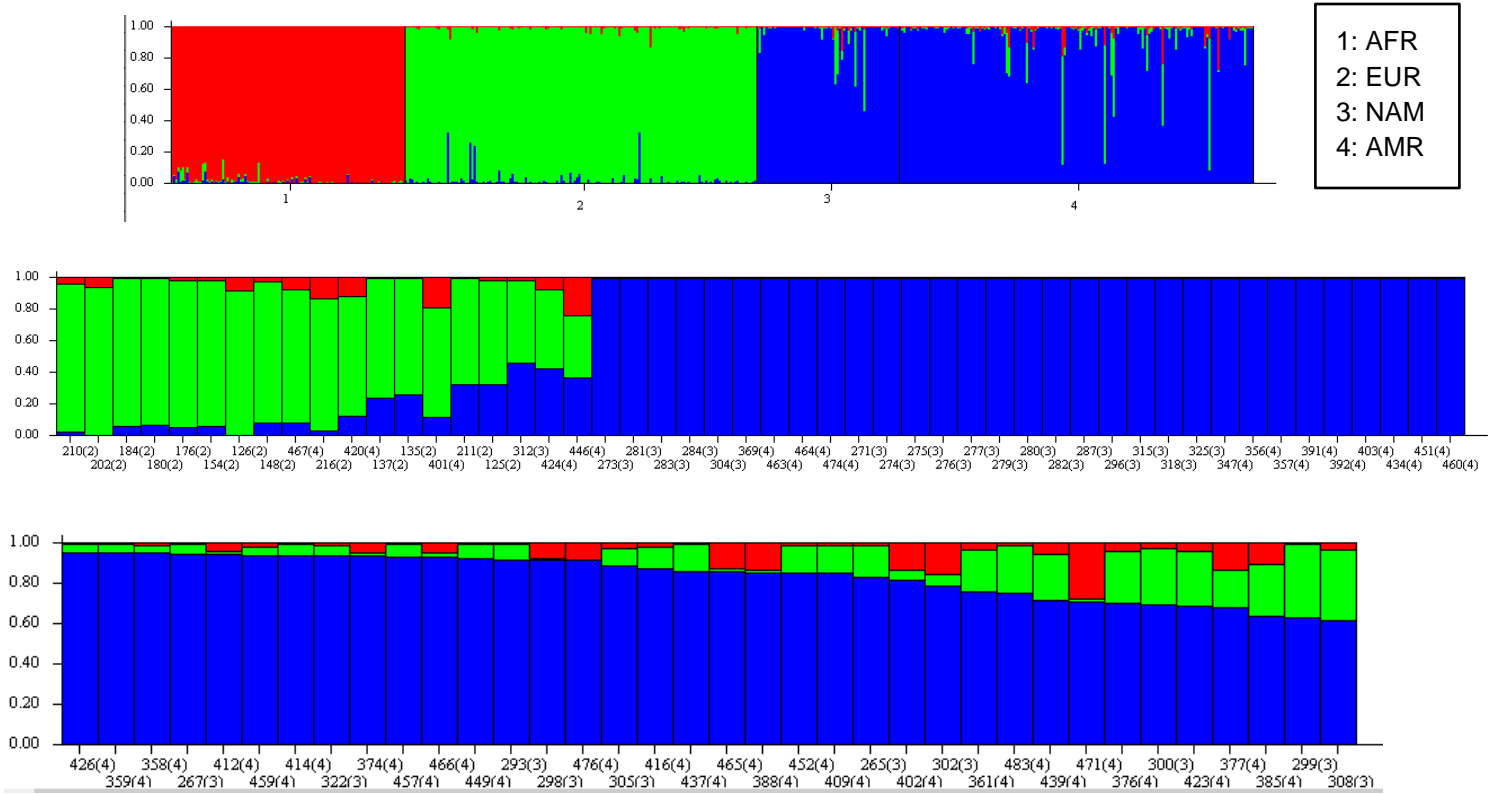


Figura 10. Plot en Triangulo de las 3 agrupaciones de poblaciones ancestrales de referencia en comparación con poblaciones amerindias. Agrupación de color rojo: Población de referencia africana. Agrupación de color verde: Población de referencia europea. Agrupación de color amarillo: Población de referencia nativo americana. Agrupación de color Azul: Población amerindia del estudio.

En consecuencia, se decidió realizar un análisis secundario en donde se excluyeron aquellos individuos amerindios menos conservado (Componente NAM <90%) y además se ajustó la población afrodescendiente, excluyendo aquellos con componente AFR <60%. Como resultado, se excluyó un 8% de la población amerindia (n=13) y un 14% de la población afrodescendientes (n=24) del total de población inicial. Los Arhuacos y Chimila fueron las etnias con mayor aporte de individuos excluidos con una proporción de 30.8% (n=4) y 15.4% (n=2), respectivamente. Por otro lado, los municipios donde se excluyeron mayor número de individuos afrodescendientes fueron Quibdó con 12 individuos (50%) y bahía solano con 6 individuos (25%). El resultado final de población para el análisis secundario fue un total de 140 individuos amerindios y 140 individuos afrodescendientes. (Figura 11 y 12).

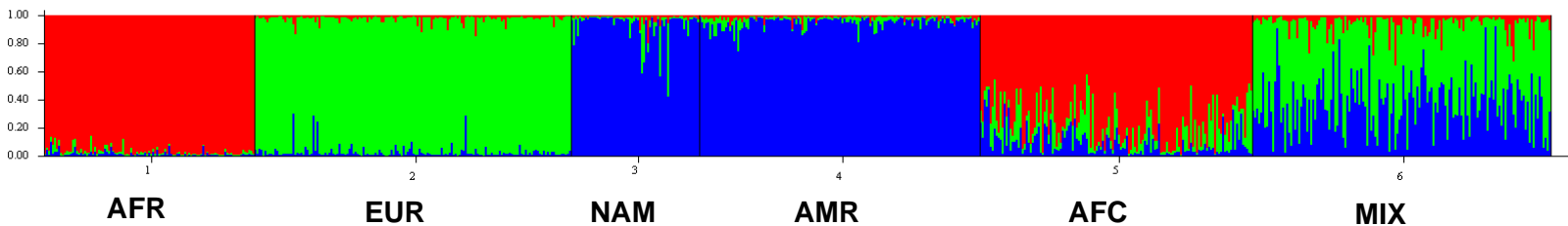


Figura 11. Estimación de componentes ancestrales (k=3) en las poblaciones de referencia y los 3 subgrupos de poblaciones colombianas.

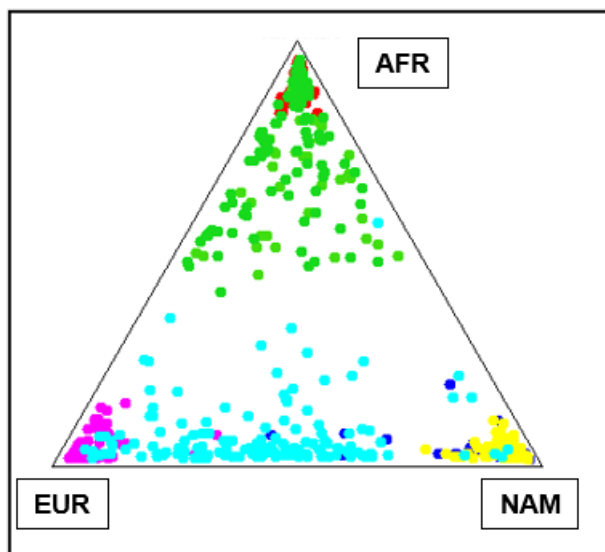


Figura 12. Plot en Triangulo utilizando un K=3 de población ancestral de referencia en comparación a las poblaciones incluidas en este estudio posterior al proceso de ajuste a individuos más conservados. Color Verde: población afrodescendiente. Color azul oscuro: población nativo americana. Color azul celeste: población de ancestralidad mixta.

La Figura 13 y la Tabla 13 muestran los porcentajes de ancestralidad obtenidos del análisis secundario para cada una de las 38 poblaciones colombianas, así como, su proporción de la estimación y agrupamiento en los 3 subgrupos poblaciones. La población de ancestralidad mixta presentó una mayor proporción de ancestralidad europea (57.8%) y nativo americana (36.3%) en comparación a la africana (5,9%). Las poblaciones amerindias incluidas en este estudio evidencian en promedio un alto grado de componente ancestral nativo americano, todas las etnias incluidas presentaron una estimación de más del 90%, la media para el grupo amerindio fue del 95%. Los Bari, Carapana, Desano, Kogui, Piapoco, Piratapuyo, Tikuna, Yeral y Yukpa fueron las poblaciones con mayor inferencia en componente nativo americano (>97%), inclusive la etnia Piapoco llegó a casi un 99%. El componente europeo fue del 3,1% en este subgrupo, siendo el segundo componente más representativo, encontrándose en las etnias Arhuacos, Inganos, Misak, Nasa, Puinave y Tukano en mayor medida, presentando un porcentaje $\geq 5\%$ en el componente Europeo. Por último, para la población amerindia el componente africano fue menor del 2% de estimación, siendo las etnias Inga, Kawayari, Barasano y Waunano en orden de frecuencia las de mayor proporción africana.

Figura 13. Diagrama circular de los componentes ancestrales para los 3 subgrupos poblaciones analizados. El componente Europeo, Nativo-americano y Africano corresponde al color Amarillo, Verde y Naranja respectivamente.

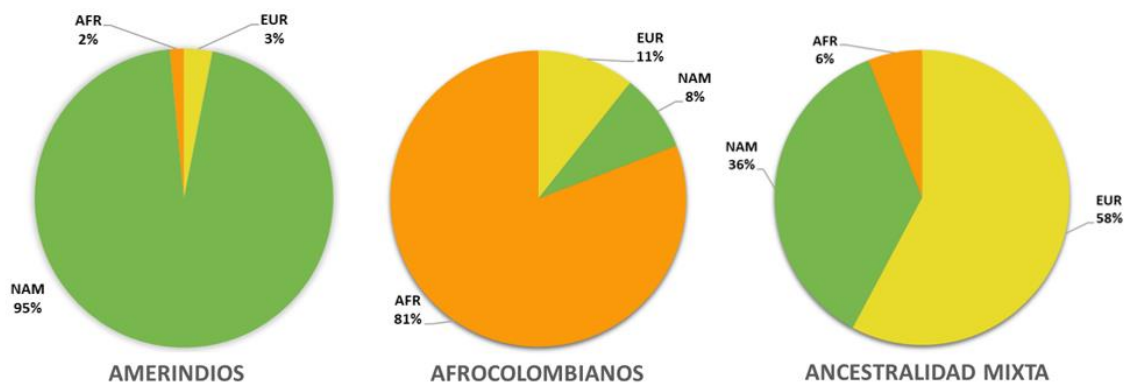


Tabla 13. Estimación de ancestría en las 38 poblaciones colombianas analizadas

Poblaciones	% Componente ancestral		
	EUR	NAM	AFR
Arhuacos	6,9	90,9	2,2
Bari	1,4	97,3	1,2
Chimila	1,7	96,4	1,9
Inga	7,0	90,2	3,8
Misak/Guambiano	9,8	90,3	0,9
Guayabero	1,7	96,9	1,4
Kogui	1,2	97,9	0,9
Nukak	2,7	96,4	0,8
Nasa/Paez	8,2	90,6	1,2
Piapoco	0,9	98,7	0,4
Piratapuyo	1,4	97,9	0,7
Puinave	5,0	93,2	1,8
Tukano	5,7	93,2	1,2
Tunebo	2,4	96,7	0,8
Wayuu	1,4	96,8	1,8
Yukpa	2,0	97,0	1,0
Barasano	1,9	95,0	3,3
Carapana	1,1	97,9	1,0
Curripako	3,2	95,5	1,3
Desano	1,6	97,5	0,9
Kawiyari	4,4	92,0	3,5
Kubeo	1,7	96,1	2,2
Makuna	2,2	96,7	1,0
Siriano	2,7	96,4	0,9
Tikuna	2,0	97,2	0,8
Yeral	0,9	97,8	1,3
Wanano	2,6	95,0	2,5
Wounaan	2,3	95,9	1,7
Amerindios (AMR)	3,1	95,5	1,5
Bahía Solano	13,2	15	71,7
Quibdo_1	11,2	9,4	79,4
Quibdo_2	14,8	11	74,2
Condoto	5,1	5	90,3
Istmina	5,7	4,2	90,1
Tadó	13,2	6,1	80,7
Afrocolombianos (AFC)	10,5	8,4	81,1
Bogotá	63,3	33,2	3,5
Boyacá	55,1	38,3	6,6
Cundinamarca	55,2	39,4	5,4
Santander	57,7	34,2	8,2
Ancestralidad Mixta (MIX)	57,8	36,3	5,9

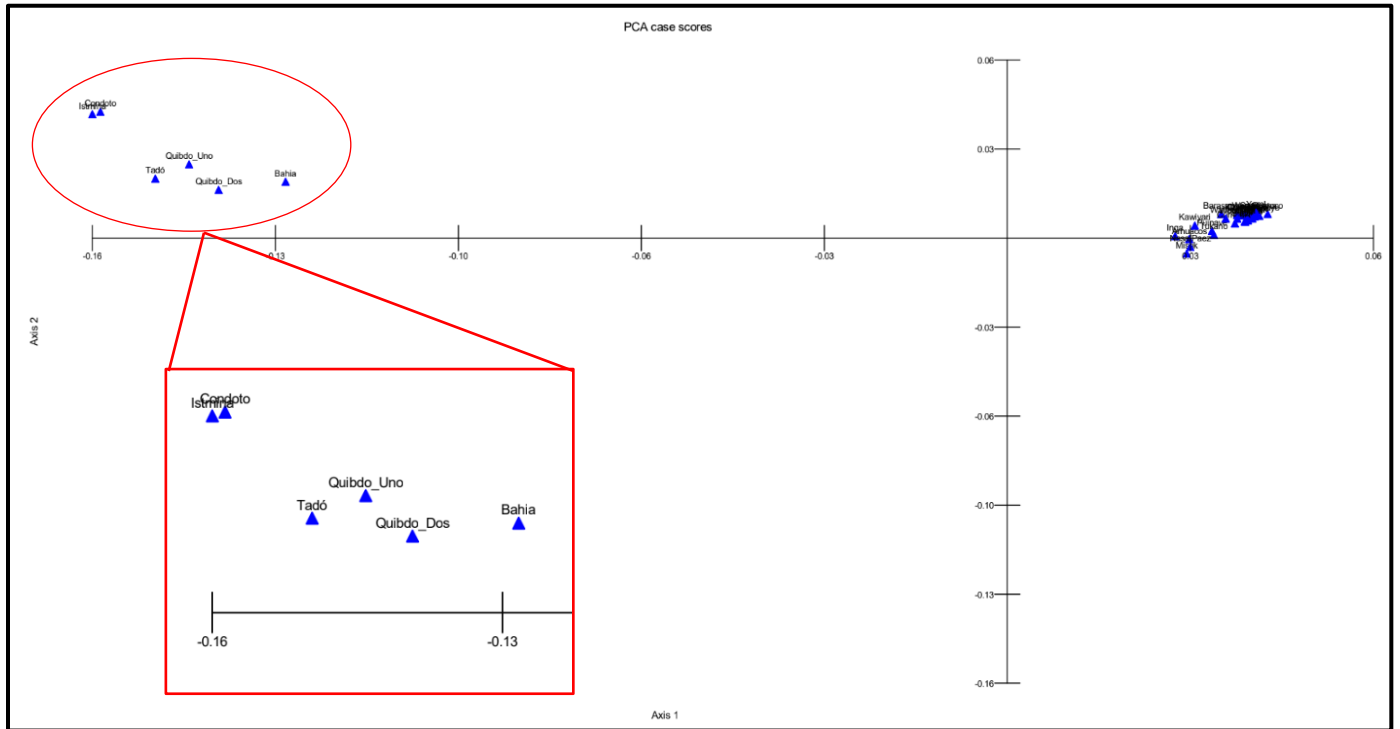
Por su parte, como era de esperarse para la población afrocolombiana su principal representación con el 81,1% fue el componente africano, en segundo grado, el componente europeo con un 10,5% y en menor proporción la ancestralidad nativo-americana en 8.4%. Los municipios de Bahía solano y Quibdó_2 fueron las poblaciones de menor ancestralidad africana, reflejado en gran medida a que fueron las poblaciones afrodescendientes con mayor proporción del componente europeo, y presentan en un grado considerable proporciones entre el 25-30% de otros componentes ancestrales.

Los resultados del análisis secundario reflejan unas poblaciones afrodescendientes y amerindias más homogéneas, como se aprecia en la figura de la estimación del componente ancestral para cada uno de los subgrupos poblacionales (Figura 11 y 12), no obstante, la inferencia de la estimación promedio del componente nativo americano no varía significativamente (94% vs 95,5%). Por el contrario, para las poblaciones afrodescendiente el cambio en la estimación del componente ancestral es notorio, sin embargo, el uso de estos análisis secundarios no estaría reflejando el grado de mezcla real de esta población.

Por otro lado, los municipios de Condoto e Istmina fueron las comunidades más representativas del componente africano en los individuos afrocolombianos estudiados, con más del 80% de ancestralidad africana. En recapitulación, la muestra de población afrodescendiente colombiana a pesar de que se localiza en mayor medida hacia el clúster de ancestralidad africana (Figura 12), continúa exhibiéndose muy dispersa, reflejando que la población afrodescendiente estudiada no es del todo homogénea.

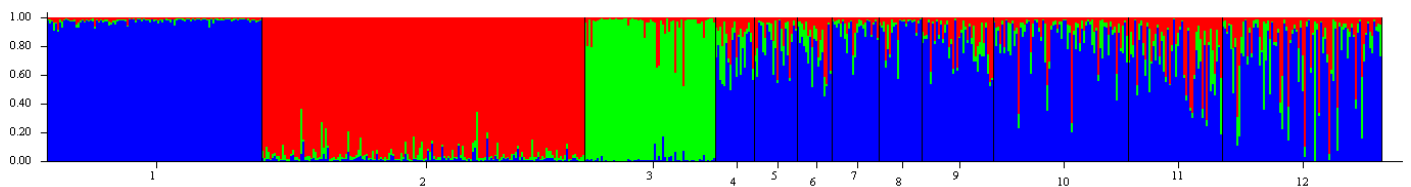
Esto último se puede visualizar como mayor detalle en la figura 14, en donde al realizar un análisis de componentes principales se percibe como las poblaciones afrodescendientes se alejan de las demás y entre ellas los municipios de Condoto e Istmina presentan un grado mayor de distanciamiento.

Figura 14. Análisis de componentes principales para los subgrupos poblacionales.



Teniendo en cuenta los resultados previamente comentados, se compararon nuestras muestras poblacionales de Chocó con otras poblaciones afrocolombianas reportadas en la literatura (Tabla 14 y Figura 15); para la figura 15 se tomaron los genotipos de las poblaciones afrodescendientes reportadas en Mogollón et al, 2020 y Ossa et al, 2015.

Figura 15. Comparación de la estimación del componente ancestral en muestras poblacionales afrodescendientes de Colombia.



**1: AFR Referencia 2: EUR Referencia 3: NAM Referencia 4: Bahía Solano 5: Quibdo_1 6: Quibdo_2
7: Condoto 8: Istmina 9: Tadó 10: Chocó Ossa et al 11: Chocó Mogollón et al 12: SAN Mogollón et al**

Tabla 14. Comparación de estimación del componente ancestral utilizando marcadores de ancestría tipo INDEL en poblaciones afrodescendientes colombianas.			
Poblaciones	% Componente Ancestral		
	EUR	NAM	AFR
Bahía Solano	21,5	22,1	56,4
Quibdo_1	31,0	12,8	56,2
Quibdo_2	17,0	12,5	70,5
Condoto	7,7	7,4	84,9
Istmina	11,6	7,3	81,0
Tadó	17,3	9,5	73,2
Chocó AFC 1er Análisis	20,0	12,0	68,0
Chocó Ossa et al*	16,3	11,7	72,0
Chocó Mogollón et al	11,0	9,3	79,7
SAN Mogollón et al	26,0	11,3	62,7

AFC: Muestras de afrocolombianos del presente estudio, SAN: San Andrés

* Calculada de población del departamento del chocó reportada en Ossa et al, 2015

Inicialmente, en la Figura 15, podemos apreciar una mayor homogeneidad en la asignación de las proporciones de ancestralidad en las poblaciones afrocolombianas utilizadas en el presente estudio en comparación a las otras muestras poblacionales reportadas previamente en la literatura, en donde se observan individuos con una mayor proporción de ancestralidad europea o nativo americana.

Nuestra población afrodescendiente del chocó, presentó una distribución de las proporciones de ancestralidad similares a las poblaciones utilizadas en previas investigaciones, con un predominio de ancestralidad africana, seguido de europea y en menor grado nativo americana. Aunque es de resaltar, una mayor similitud a lo reportado por Ossa et al, 2020, en donde su población presentó una proporción promedio africana del 72%, en segundo grado una proporción europea del 16,3% y en menor medida una proporción nativo americana del 11,7%. El comportamiento de la población afrodescendiente de San Andrés reportada por Mogollón et al, 2020 es aquella en la que se observa mayor diferenciación, mostrando un mayor grado de

ancestralidad europea con una proporción del 26%, datos que son esperados debido al componente histórico y mestizaje presentado en la isla. (Tabla 14).

Con los resultados obtenidos de las poblaciones amerindias y afrodescendientes incluidas en nuestro estudio, las cuales reflejan una mayor homogeneidad, se procedió a analizar si el uso de poblaciones propias con un menor grado de mezcla permitían identificar una mejor inferencia en la estimación de la proporción de ancestría en población mezclada utilizándolas como poblaciones de referencia.

Tabla 15. Estimación del componente ancestral en población mezclada utilizando como referencia poblaciones previas.

Poblaciones	% Componente ancestral		
	EUR	AFR	NAM
AFR REF	3,0%	94,5%	2,6%
EUR REF	95,0%	2,0%	3,0%
NAM REF	5,4%	1,5%	93,0%
MIX	58,2%	5,1%	36,7%

Tabla 16. Estimación componente ancestral en población mezclada utilizando como referencia amerindios propios + afrodescendiente y europeos de poblaciones previas.

Poblaciones	% Componente ancestral		
	EUR	AFR	NAM
AFR REF	2,4%	95,2%	2,3%
EUR REF	95,5%	1,6%	2,8%
AMR	2,1%	1,1%	96,9%
MIX	56,7%	4,7%	38,6%

Tabla 17. Estimación componente ancestral en población mezclada utilizando como referencia poblaciones previas + amerindios propios.

Poblaciones	% Componente ancestral		
	EUR	AFR	NAM
AFR REF	2,2%	95,8%	2,0%
EUR REF	96,3%	1,4%	2,2%
NAM REF + AMR	3,2%	1,2%	95,7%
MIX	59,8%	4,5%	35,8%

Tabla 18. Estimación componente ancestral en población mezclada utilizando como referencia poblaciones previas + amerindios y afrodescendientes propios.			
Poblaciones	% Componente ancestral		
	EUR	AFR	NAM
AFR REF + AFC	7,3%	88,0%	4,6%
EUR REF	95,2%	1,9%	2,9%
NAM REF + AMR	3,6%	1,5%	94,9%
MIX	57,4%	5,5%	37,2%

Tabla 19. Estimación componente ancestral en población mezclada utilizando como referencia amerindios y afrodescendientes propios + europeos de poblaciones previas.			
Poblaciones	% Componente ancestral		
	EUR	AFR	NAM
AFC	6,0%	89,8%	4,2%
EUR REF	93,8%	2,6%	3,6%
AMR	2,5%	1,6%	95,8%
MIX	53,0%	7,5%	39,5%

Tabla 20. Estimación componente ancestral en población mezclada utilizando como referencia afrodescendientes propios + nativo americanos y europeos de poblaciones previas.			
Poblaciones	% Componente Ancestral		
	EUR	AFR	NAM
AFC	7,1%	87,9%	5,0%
EUR REF	92,7%	3,3%	4,0%
NAM REF	5,8%	2,5%	91,7%
MIX	54,3%	8,1%	37,6%

Comparando los resultados de la inferencia de la estimación de la proporción de ancestralidad a partir de poblaciones de referencia usualmente utilizadas vs poblaciones propias con menor grado de mezcla, se observó que para el componente nativo americano la asignación de la proporción aumento entre 2 y 3% (Tabla 16 y 19) cuando la población utilizada como referencia nativo americana fue población autóctona más conservada. Por el contrario, la combinación de estas últimas muestras con otras poblaciones nativo americanas utilizadas como referencia, hacen que la estimación de la proporción disminuya a 35,8% (Tabla 17); el cálculo del valor P indico que estos datos no fueron estadísticamente significativos.

5. DISCUSIÓN:

Los análisis realizados en este estudio sugieren que las muestras de las 38 poblaciones pueden agruparse principalmente en 3 grupos: Amerindios, afrodescendientes y de ancestralidad mixta. Adicionalmente, los resultados descritos en esta investigación permitieron caracterizar el perfil de ancestría y analizar el nivel de sub-estructura genética de los tres grupos poblacionales mencionados, mediante un panel de 46 marcadores tipo AIM InDel, así como, realizar comparaciones y analizar la influencia que podría presentarse al utilizar poblaciones propias conservadas para una mejor estimación de la proporción de ancestría.

Respecto a la caracterización del perfil de ancestría en la población de ancestralidad mixta, nuestros resultados indicaron una mayor proporción de ancestralidad europea (57.8%), seguido del componente nativo-americano (36.3%) y africano (5.9%). Lo reportado es consistente con lo señalado por otros autores en estudios previos. Ossa et al, 2016, señala que todas las subregiones andinas están caracterizadas por una contribución predominante de ascendencia europea, siendo la región Central Oeste la que presentó mayor proporción europea (67%). Así mismo, este estudio identificó al componente nativo-americano como el segundo aporte más importante de ancestralidad en esta población.

Por su parte, Ibarra et al, 2014 encontró que en las poblaciones de las regiones de los Andes centrales y orientales (Santander-Cundinamarca-Boyacá y Huila), la proporción de contribución europea estimada era mayor que la esperada según los marcadores de ancestralidad InDel en comparación a las proporciones promediadas de polimorfismos de origen europeo en el ADNmt y el cromosoma Y. Lo que sugiere, en primera medida una fase de mayor colonización europea en nuestra población, y en segunda instancia, una mezcla genética caracterizada por la presencia del componente nativo-americano, que podría estar explicado por las prácticas de abuso sexual contra las mujeres indígenas por parte de los colonos europeos. (Quijano, A, 2000; Aristizábal, S., 2000)

La diversidad genética de la región andina comprende una mezcla compleja de múltiples factores socioculturales. Esta región ha sido históricamente una importante zona de actividad económica en el país, lo que ha contribuido a la migración de los habitantes de otras regiones de Colombia, convirtiéndole en una zona de alta densidad poblacional. Así mismo, muchos de los habitantes de las áreas rurales y selváticas del país afectados por el conflicto armado, han migrado forzosamente a las grandes urbes. Este factor migratorio, en uno de los factores que más ha favorecido el resultado de la población mezclada actualmente presente en gran parte de la región andina y se relaciona con lo mencionado por Mogollón et al, 2020, quienes observaron estimaciones para las poblaciones de Bogotá, Tolima y Huila similares a lo reportado en nuestra poblaciones de ancestralidad mixta y quienes comentan que estos hallazgos pueden deberse a los patrones demográficos y migratorios que exhiben estas poblaciones y resaltan la importancia que la ancestralidad y la diferenciación genética no debería establecerse únicamente considerando la división territorial.

A pesar de que Colombia ocupa el tercer lugar en Latinoamérica con mayor porcentaje de población afrodescendiente, los cuales representan alrededor de 5 millones de habitantes que se encuentran ubicados principalmente en la región del Pacífico (DANE, 2019), existen muy pocos estudios genéticos de la población afrocolombiana pues los estudios previos de ancestralidad genética colombiana se han ocupado en su mayoría de poblaciones mestizas o comunidades nativo-americanas.

Con relación a la población afrodescendiente en el presente estudio, se identificaron hallazgos no descritos previamente en población afrocolombiana que son de gran interés para la comunidad científica. En primera instancia, se detectó la presencia de un tercer alelo en el marcador MID-548 con una frecuencia del 3.66%, siendo un hallazgo que podría asumirse como un alelo de fijación o alelo extremadamente raro, que puede contribuir a inferir el componente ancestral africano de la población afrocolombiana. Los individuos con un tercer alelo en homocigosis para el marcador MID-548, presentaron una inferencia de la estimación del componente africano mayor del 95%. A su vez, el locus MID-548 no se encontró en equilibrio en la población afrocolombiana, siendo la explicación más probable la presencia de este tercer alelo en dicho marcador.

Los análisis realizados para la estimación de ancestría en población afrocolombiana presentados en este estudio son consistentes con lo señalado por otros autores. Medina-Rivas et al, 2016, reportaron altos niveles de proporción promedio de ancestría africana en población de Chocó (75,8%), siendo este el componente dominante, además, observaron que el componente no africano se dividía en proporciones similares entre las ancestralidades nativo americana y europea, hallazgo que nuestro estudio también evidenció. Es llamativo como la población del Chocó se diferencia de otras poblaciones afrodescendientes del nuevo mundo, quienes presentan un menor grado de ancestría nativo americano y mayor influencia de porcentaje europeo. Este patrón es indicativo de un contacto más prolongado y sostenido entre afrodescendientes y comunidades indígenas en el Chocó en comparación con lo ocurrido en el Caribe o en los Estados Unidos. (Martínez, B et al, 2020)

En un estudio realizado en el corregimiento de San Basilio de Palenque (Martínez, B et al, 2017), se encontró un porcentaje de ancestría africana del 81,2%, seguido de un 10.6% para el componente europeo, y de un 8.2% para el nativo americano. Este estudio fue realizado en una de la comunidades afrodescendiente más conservadas en el país y nuestros resultados presentaron unas proporciones similares para algunas de nuestras poblaciones (Istmina y Condoto), las cuales presentaron en promedio porcentajes mayores al 80% para ancestría africana, seguido de una proporción promedio de entre el 10% y 7% para el componente europeo y ancestralidad nativo americana.

El pueblo de San Basilio de Palenque es un pueblo que fue fundado por esclavos africanos fugados, quienes se establecieron en la región montañosa cercana a Cartagena, en el norte de Colombia. Se considera el primer territorio libre de América, símbolo de resistencia contra la conquista militar española, que se ha mantenido aislado en gran medida de la cultura predominante colombiana. Este pueblo conserva con mayor fidelidad las costumbres, ritos y formas de expresión de comunidades africanas lo que favorece el patrimonio genético africano y podría ser un factor determinante para comprender el alto porcentaje del componente de ancestría africano presente en esta

población en comparación con otras poblaciones afrocolombianas (Restrepo, E., Pérez, J.N. et al, 2005).

Por otra parte, en el análisis de la estimación del componente de ancestralidad genético, se identificó una menor proporción africana en las poblaciones de Bahía Solano (56,4%) y Quibdó (56,2%) en comparación con las poblaciones de Condoto e Istmina (>80%). Así mismo, al utilizar como una medida de distanciamiento genético el índice F_{st} , se evidenció una diferencia estadísticamente significativa ($p < 0.05$) entre las poblaciones de Quibdó y Bahía Solano, respecto a las poblaciones de Condoto, Istmina y Tadó. Estos hallazgos destacan la diversidad genética que las poblaciones afrocolombianas pueden tener entre sí y reitera la necesidad de continuar estudiándolas.

Ahora bien, continuando con los análisis presentados por Martínez, B et al., 2017 et al, quien comparó los resultados obtenidos de estimación de ancestría con valores previos en una subregión del departamento de Bolívar (Montes de María), que también incluía población de San Basilio. Los resultados indicaron que la ascendencia africana de San Basilio fue de cuatro veces mayor que la encontrada en el departamento de Bolívar, demostrando una vez más que pese a que son poblaciones afrodescendientes, estas poblaciones no son homogéneas entre sí.

Esta observación de la diversidad en comunidades afro también ha sido descrita en poblaciones afrodescendientes de otras regiones de Latinoamérica. Carvalho et al, 2018 señala una correlación existente entre el grado de aislamiento y el grado de mezcla en población afrodescendiente de Brasil (Quilombos: Kalunga, Sacutiaba y Mocambo). Sus resultados indican que para la población Kalunga, la cual es la comunidad más aislada, presenta un mayor componente africano (67,3%), en contraste con Sacutiaba, que es la población más pequeña e impactada por la migración, dado principalmente por la interacción que tienen con las zonas urbanas que se encuentran alrededor de esta región, en la cual se identificó el mayor componente europeo (46,8 %). Por último, Mocambo es vecino de una población nativa americana y presentó la mayor contribución amerindia. (12,2%).

Esta correlación mencionada previamente entre el grado de aislamiento y mestizaje de población afrodescendiente es consistente con los análisis de subestructura genética y distanciamiento genético identificados en nuestros afrocolombianos. Las poblaciones con un mayor componente de ancestría africana encontradas en nuestro estudio, se encuentran ubicadas en zonas de difícil acceso y lejanas de la capital del Departamento del Chocó, lo que promueve el aislamiento de sus habitantes con las regiones circundantes del departamento. Por el contrario, los municipios de Bahía Solano y Quibdó, además de contar con mejores vías de acceso y por ende mayor tránsito, han experimentado un incremento significativo de la demanda turística lo cual puede contribuir a una mayor mezcla de su población.

Adicionalmente, existen otros aspectos importantes que han contribuido en gran medida al flujo genético en esta región, los cuales son derivados del contexto socio-político que afecta al departamento del Chocó, como la presencia del conflicto armado en zonas urbanas y rurales, el desplazamiento forzado, la presencia de narcotráfico, el paso de contrabando y la gran oleada de migrantes hacia zonas fronterizas del litoral Pacífico. (Martínez, Y.S., 2021; Mosquera, A.Y., 2016)

De igual manera a lo descrito anteriormente en el análisis de la población de ancestralidad mixta, en donde se señalaba la posible relación entre la proporción de ancestría amerindia en esa población y las prácticas de abuso sexual contra las mujeres indígenas, ocurrió algo similar en la población afrodescendiente, especialmente dada la condición de esclavos de los africanos llegados a las tierras de las Américas, en donde este tipo de violencia sexual se daba bajo circunstancias de castigo o de dominación, entre otros. (Wade, P., 1995; Aristizábal, S., 2000; Hoffmann, O., 2007). Estudios en población afrocolombiana como el de Medina-Rivas et al, 2016 y Alonso et al, 2013, reflejan biológicamente como el haplotipo no africano del cromosoma Y esta predominantemente abarcado por el componente europeo.

Respecto a la población amerindia, se logró obtener una muestra de 28 comunidades indígenas de alrededor del país. El departamento de Vaupés aportó el mayor número de etnias. Como se describe previamente en los resultados de este estudio, el análisis de ancestría en población amerindia arrojó un alto porcentaje de componente ancestral

nativo americano, con una estimación mayor al 90%, presente en todas las etnias estudiadas, seguido del componente europeo con un 3.1%, y un porcentaje menor del 2% para el componente africano. Datos similares fueron reportado por Mogollón et al., 2020, para algunas comunidades indígenas amazónicas (Cocamas, Uitoto, Ticunas) y del norte del país (Wayúus y Zenús).

En cuanto a los análisis de diferenciación genética en nuestra población amerindia y según los resultados presentados en la Tabla 9, los grupos indígenas del norte (Arhuaco, Kogui, Bari y Chimila) se distanciaron significativamente de las demás. Lo anterior, puede estar relacionado con diversos factores como la localización geográfica de las comunidades indígenas analizadas en este estudio, dado que la gran mayoría provienen principalmente de la región Amazónica. A su vez, los orígenes ancestrales de estas comunidades pueden jugar un papel importante en la comprensión de la diferenciación genética identificada, así como las lenguas que predominan en esta población. Estas observaciones ha sido objeto de estudio en investigaciones previas realizadas en población amerindia de Colombia (Yunis et al., 1994; Callegari-Jacques et al., 2011; Wang et al., 2007; Fuselli et al., 2003; Tarazona-Santos et al., 2001).

La diferenciación entre los individuos de las comunidades indígenas de las regiones Orinoquía y Amazonas en comparación a la región Caribe y Andina ya había sido reportado (Keyeux & Usaquén, 2006). Lo cual, es concordante con nuestros resultados de distanciamiento genético y cálculos de F_{st} . Inicialmente, algunas investigaciones han sugerido teorías en donde todas las poblaciones nativo americanas presentes en América del sur emergieron de un evento único migratorio, seguido de una divergencia influenciada por elementos como barreras geográficas y climáticas, así como, patrones de asentamiento y estructura social. (Rothhammer & Silva, 1989; Reich et al., 2012; Rothhammer & Dillehay, 2009). Sin embargo, otros autores recientemente proponen que ciertas comunidades amazónicas se originaron de una población con vínculos genéticos más cercanos a las poblaciones indígenas del sur de Asia, que antes de llegar a la Amazonía se habrían mezclado con un linaje asociado a la primera población americana, explicando a través de estos modelos las posibles diferencias entre las poblaciones indígenas del caribe y amazónicas. (Skoglund et al., 2015; Harris et al., 2018).

Complementariamente, Romero et al, 2013 señala una relación entre la variable lingüística-geográfica y la distribución de ADN mitocondrial (ADNmt) en población indígena de las regiones del norte y sur de Colombia. A pesar de que se utilizó un marcador de ancestralidad sexual, es llamativo que cuando las poblaciones fueron agrupadas por localización geográfica (norte-Caribe, sur-Amazonia, Oeste-Pacífico, Este-Orinoquía), la diferencias entre la frecuencia del haplogrupo A fue mayor en la región norte de Colombia (50% de frecuencia), disminuyendo en un 20% en la región del sur del país, mientras que la frecuencia del grupo C fue menor en el norte y tuvo su valor más alto en el sur de Colombia, demostrando que las frecuencias de los haplogrupos de ADNmt difieren entre las regiones del norte y sur de Colombia, lo cual nos permite explicar otros factores que pueden favorecer las diferencias entre las etnias del norte y sur del país (Yunis, J et al., 2013).

Por otra parte, para el análisis del distanciamiento genético en nuestra población de estudio, se hizo uso del dendograma como herramienta de análisis visual, en donde es posible observar un acercamiento de los pueblos indígenas del norte con la población de ancestralidad mixta (Figura 5). Aquí es importante recordar que el segundo porcentaje de ancestralidad presente en la población de la región Andina de nuestro estudio fue la nativo americana con un 37.3%. Así mismo, otra posible explicación, recae en los grandes y atractivos entornos ecológicos y el enorme capital cultural y social que estas tribus indígenas del norte poseen, han dado también lugar al interés de quienes desean conocer un poco más sobre estos pueblos ancestrales y sus entornos, lo que ha ocasionado un aumento relevante del turismo en la región del Caribe colombiano. Según el Instituto Distrital de Turismo de Santa Marta (Indetur) en el último año la ciudad de Santa Marta recibió 2'978.376 visitantes nacionales y en los últimos cuatro años ingresaron más de 11 millones de visitantes, lo que podría favorecer al grado de mestizaje por medio del flujo genético entre los indígenas de esta región (Indetur, 2022).

Otro de los factores que afecta considerablemente a la población amerindia en Colombia es el impacto del conflicto armado en sus territorios, lo que ha provocado que muchos habitantes de estas comunidades estén desplazándose constantemente hacia

entornos urbanos de manera transitoria, para luego regresar nuevamente a sus regiones. Esta situación es evidenciada a diario en ciudades de los departamentos de Santander, Norte de Santander, Boyacá y Cundinamarca, donde los indígenas encuentran lugar en zonas concurridas de las grandes urbes para trabajar y asentarse temporalmente.

De acuerdo con los datos suministrados por la ACNUR, aproximadamente 70,000 de los desplazados internos registrados son indígenas. El pueblo Arhuaco, ubicado en las altas partes de la Sierra Nevada, cuyo resguardo principal está en la ciudad de Santa Marta, es uno de los pueblos más abatidos por el conflicto armado colombiano, provocando que los pueblos indígenas que habitan la Sierra constantemente tengan que emigrar de su lugar de residencia o vivienda (ACNUR, 2011).

Una situación similar ocurre con los resguardos Motilón-Barí, donde habitan los indígenas Barí, también conocidos como motilones, los cuales se encuentran localizados en el territorio del Catatumbo, uno de los principales corredores del conflicto armado en Colombia. En la actualidad, este territorio se encuentra bajo el mando de los grupos armados al margen de la ley, grupos organizados de narcotráfico y la presencia de la fuerza pública.

La organización de la Comisión de la Verdad señala a la violencia sexual y la prostitución forzada, como unas de las grandes afectaciones ocasionadas por las confrontaciones de los diversos actores de la guerra en estos territorios, lo que puede contribuir indiscutiblemente a un mestizaje genético entre los habitantes de estas comunidades indígenas (Comisión para el esclarecimiento de la verdad, la convivencia y la no repetición, 2022).

6. CONCLUSIONES:

- La población con ancestralidad mixta presento mayor proporción de ancestralidad europea (57.8%), seguido del componente nativo-americano (36.3%) y africano (5.9%). Resultados similares a lo reportado por otros grupos de investigación.
- La población Amerindia mostro una estimación primaria nativo-americana con 94%, en segunda medida europea con 5% y en menor proporción africana con 1%
- El presente proyecto presento el mayor número de comunidades amerindias incluidas en un estudio de análisis de ancestría genética en población colombiana.
- Recomendamos utilizar nuestra población amerindia como población de referencia para la estimación de ancestría NAM en Colombia, ya que refleja mejor la diversidad indígena colombiana.
- Para la población afrocolombiana, presento principalmente un componente AFR 68%, seguido de EUR 20% y NAM 12%.
- Los resultados indican que ocurrieron diferentes dinámicas de mezcla de la población afrocolombiana dentro del departamento del chocó, consideramos que se presenta una gran variabilidad en el grado de mezcla entre aquellas poblaciones pequeñas y alejadas frente a poblaciones urbanas vecinas, por lo tanto, los datos deben ser tratados de manera particular en estudios de asociación con enfermedades, en el análisis forense y pruebas de identificación humana.
- Los resultados expuestos amplían y destacan la diversidad genética de las poblaciones amerindias y afrocolombianas, sin embargo, continúan siendo pocos los estudios realizados en poblaciones afrodescendientes en donde todavía quedan preguntas por resolver acerca de su reciente historia demográfica y estructura poblacional.

- Al comprender la estructura genética de las poblaciones humanas, podemos reconstruir la historia y comprender las bases genéticas de las enfermedades. Por lo tanto, el trabajo presentado contribuye al esfuerzo continuo por documentar la variabilidad genética humana, al enfocarse en explorar poblaciones que han sido poco estudiadas y resaltamos la necesidad de continuar estudiando nuestra población afrodescendiente.

7. LISTADO DE TABLAS Y FIGURAS

7.1 TABLAS

Tabla 1. Ventajas y Desventajas Según Tipo de Marcador de Ancestralidad.

Tabla 2. Grupos Sanguíneos y sus alelos.

Tabla 3. Listado de los pares de primers utilizados para los 46 marcadores.

Tabla 4. Condiciones de reacción de mezcla maestra.

Tabla 5. Condiciones de termociclado.

Tabla 6. Frecuencias alélicas por locus para las 3 subpoblaciones ancestrales.

Tabla 7. Equilibrio de Hardy-Weinberg y Heterocigosidad observada y esperada por locus y por subgrupos poblacionales.

Tabla 8. Cálculo de AMOVA para los subgrupos poblacionales

Tabla 9. Distanciamiento genético. Valores F_{st} se muestran debajo de la diagonal. Los P-Valores se muestran por arriba de la diagonal.

Tabla 10. Valores F_{st} para los 3 subgrupos poblacionales.

Tabla 11. Porcentaje de proporciones utilizando un $K=4$.

Tabla 12. Porcentaje de proporciones utilizando un $K=3$.

Tabla 13. Estimación de ancestría en las 38 poblaciones colombianas analizadas.

Tabla 14. Comparación de estimación del componente ancestral utilizando marcadores de ancestría tipo INDEL en poblaciones afrodescendientes colombianas.

Tabla 15. Estimación del componente ancestral en población mestiza utilizando como referencia poblaciones previas.

Tabla 16. Estimación componente ancestral en población mestiza utilizando como referencia amerindios propios + afrodescendiente y europeos de poblaciones previas.

Tabla 17. Estimación componente ancestral en población mestiza utilizando como referencia poblaciones previas + amerindios propios.

Tabla 18. Estimación componente ancestral en población mestiza utilizando como referencia poblaciones previas + amerindios y afrodescendientes propios

Tabla 19. Estimación componente ancestral en población mestiza utilizando como referencia amerindios y afrodescendientes propios + europeos de poblaciones previas.

Tabla 20. Estimación componente ancestral en población mestiza utilizando como referencia afrodescendientes propios + nativo americanos y europeos de poblaciones previas.

Libro Excel. Tablas Suplementarias

7.2 FIGURAS

Figura 1. Potenciales Rutas Migratorias de Colonización del Noroeste Americano.

Figura 2. Electroferograma de los 46 marcadores InDel amplificados.

Figura 3. Mapa de las comunidades indígenas incluidas en el estudio.

Figura 4. Electroferograma evidencia tercer alelo en marcador MID-548.

Figura 5. Dendograma construido a partir de los valores F_{st} de 38 poblaciones colombianas.

Figura 6. Dendograma UPGMA construido a partir del análisis de distancia genética de Cavalli-Sforza en 33 grupos de poblaciones de Colombia.

Figura 7. Resultados del método de Evanno utilizados para estimar el valor de k más adecuado para el análisis de STRUCTURE.

Figura 8. Bar Plot de STRUCTURE utilizando $K=3$.

Figura 9. Bar plot STRUCTURE BY SORTING de población amerindia (AMR) utilizando $K=3$.

Figura 10. Plot en Triangulo-1 de las 3 agrupaciones de poblaciones ancestrales de referencia en comparación con poblaciones amerindias.

Figura 11. Estimación de componentes ancestrales ($k=3$) en las poblaciones de referencia y los 3 subgrupos de poblaciones colombianas.

Figura 12. Plot en Triangulo utilizando un $K=3$ de población ancestral de referencia en comparación a las poblaciones incluidas en este estudio posterior al proceso de ajuste a individuos más conservados.

Figura 12. Diagrama circular de los componentes ancestrales para los 3 subgrupos poblaciones analizados.

Figura 14. Análisis de componentes principales para los subgrupos poblaciones.

Figura 15. Comparación de la estimación del componente ancestral en muestras poblacionales afrodescendientes de Colombia.

8. Bibliografía.

- ACEP. (1974). La población de Colombia. *Publicación auspiciada con motivo del Año Mundial de Población, por el "Committee For International Coordination Of National Research In Demography". Colombia.*
- ACNUR – Agencia de la ONU para los refugiados. Comunidades Indígenas. (2011). <https://www.mipresente.org>.
- Alonso Morales, L. A., Casas-Vargas, A., Rojas Castro, M., Resque, R., Ribeiro-Dos-Santos, A. K., Santos, S., . . . Usaquen, W. (2018). Paternal portrait of populations of the middle Magdalena River region (Tolima and Huila, Colombia): New insights on the peopling of Central America and northernmost South America. *PLoS One*, *13*(11), e0207130. doi:10.1371/journal.pone.0207130
- Alonso Morales, L.A., Usaquen, W. (2013). Y-chromosome and surname analysis of the native islanders of San Andrés and Providencia (Colombia). *JCHB*, *14*. doi: 10.1016/j.jchb.2012.11.006
- Al-Asfi, M et al. (2018). Assessment of the Precision ID Ancestry panel. *Int J Legal Med*. *132*(6):1581-1594. doi: 10.1007/s00414-018-1785-9.
- Amankwah, E. K., & Hale, G. A. (2018). Hispanic Ethnicity and the Risk of Pediatric Leukemia Relapse. *Clin Pediatr (Phila)*, *57*(6), 656-659. doi:10.1177/0009922817732147
- Andrade, R. B., Amador, M. A. T., Cavalcante, G. C., Leitao, L. P. C., Fernandes, M. R., Modesto, A. A. C., . . . Santos, N. P. C. (2018). Estimating Asian Contribution to the Brazilian Population: A New Application of a Validated Set of 61 Ancestry Informative Markers. *G3 (Bethesda)*, *8*(11), 3577-3582. doi:10.1534/g3.118.200650
- Aristizábal, S. (2000) LA DIVERSIDAD ÉTNICA Y CULTURAL DE COLOMBIA: UN DESAFÍO PARA LA EDUCACIÓN. *Pedagogía y Saberes*. 15.
- Bortolini, M., Salzano, F., Bau, C., Layrisse, Z., & Petzl-Erler, M. (2002). Y chromosome biallelic polymorphisms and Native American population structure. *Annals of human genetics*.
- Callegari-Jacques, S. M., Tarazona-Santos, E. M., Gilman, R. H., Herrera, P., Cabrera, L., Dos Santos, S. E. B., ... Salzano, F. M. (2011). Autosome STRs in native South America—Testing models of association with geography and language. *American Journal of Physical Anthropology*, *145*, 371–381.
- Carvajal-Carmona, L., Soto, I., Pineda, N., Ortiz-Barrientos, D., Duque, C., Ospina-Duque, J., . . . Ruiz-Linares, A. (2000). Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *American Journal of Human Genetics*, *67*, 1287-1295.
- Carvalho, C et al. (2018). Ancestry analysis in rural Brazilian populations of African descent. *Forensic Sci Int Genet*. *36*:160-166. doi: 10.1016/j.fsigen.2018.06.018
- Carvalho, D. C., Wanderley, A. V., Amador, M. A., Fernandes, M. R., Cavalcante, G. C., Pantoja, K. B., . . . Dos Santos, N. P. (2015). Amerindian genetic ancestry and INDEL polymorphisms associated with susceptibility of childhood B-cell Leukemia in an admixed population from the Brazilian Amazon. *Leuk Res*. doi:10.1016/j.leukres.2015.08.008
- Chande, A et al. (2021). The Impact of Ethnicity and Genetic Ancestry on Disease Prevalence and Risk in Colombia. *Front Genet*. *17*:12:690366. doi: 10.3389/fgene.2021.690366
- Comisión para el esclarecimiento de la verdad, la convivencia y la no repetición. (2022). Informe Final. Julio. Colombia.

- Da Costa Francez, P. A., Rodrigues, E. M., de Velasco, A. M., & dos Santos, S. E. (2012). Insertion-deletion polymorphisms--utilization on forensic analysis. *Int J Legal Med*, 126(4), 491-496. doi:10.1007/s00414-011-0588-z
- Da Silva, E. M., Fernandes, M. R., de Carvalho, D. C., Leitao, L. P. C., Cavalcante, G. C., Pereira, E. E. B., . . . Dos Santos, N. P. C. (2017). Effect of genetic ancestry to the risk of susceptibility to gastric cancer in a mixed population of the Brazilian Amazon. *BMC Res Notes*, 10(1), 646. doi:10.1186/s13104-017-2963-4
- DANE - Departamento Administrativo Nacional de Estadística. POBLACIÓN NEGRA, AFROCOLOMBIANA, RAIZAL Y PALENQUERA RESULTADOS DEL CENSO NACIONAL DE POBLACIÓN Y VIVIENDA 2018. (2019). Colombia. <https://www.dane.gov.co › files › grupos-etnicos>
- Estey, E. H. (2018). Acute myeloid leukemia: 2019 update on risk-stratification and management. *Am J Hematol*, 93(10), 1267-1291. doi:10.1002/ajh.25214
- Excoffier, L., & Lischer, H. E. L. (2010). An integrated software package for population genetics data analysis. *Molecular Ecology Resources*, 10, 564-567.
- Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., . . . Altshuler, D. (2004). Assessing the impact of population stratification on genetic association studies. *Nat Genet*, 36(4), 388-393. doi:10.1038/ng1333
- Fuselli, S., Tarazona-Santos, E., Dupanloup, I., Soto, A., Luiselli, D., & Pettener, D. (2003). Mitochondrial DNA diversity in South America and the genetic history of andean highlanders. *Molecular Biology and Evolution*, 20, 1682-1691.
- Gimbrone, N. T., Sarcar, B., Gordian, E. R., Rivera, J. I., Lopez, C., Yoder, S. J., . . . Cress, W. D. (2017). Somatic Mutations and Ancestry Markers in Hispanic Lung Cancer Patients. *J Thorac Oncol*, 12(12), 1851-1856. doi:10.1016/j.jtho.2017.08.019
- Harris, D. N., Song, W., Shetty, A. C., Levano, K. S., Cáceres, O., Padilla, C., . . . Guio, H. (2018). Evolutionary genomic dynamics of Peruvians before, during, and after the Inca empire. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E6526-E6535.
- Hartl, D. L., & Grant, A. (2007). *Principles of Population Genetics* (4th Edition ed.): Sinauer Associates.
- Hernandez-Suarez, G., Sanabria, M. C., Serrano, M., Herran, O. F., Perez, J., Plata, J. L., . . . Tenesa, A. (2014). Genetic ancestry is associated with colorectal adenomas and adenocarcinomas in Latino populations. *Eur J Hum Genet*, 22(10), 1208-1216. doi:10.1038/ejhg.2013.310
- Hines, L. M., Sedjo, R. L., Byers, T., John, E. M., Fejerman, L., Stern, M. C., . . . Slattery, M. L. (2017). The Interaction between Genetic Ancestry and Breast Cancer Risk Factors among Hispanic Women: The Breast Cancer Health Disparities Study. *Cancer Epidemiol Biomarkers Prev*, 26(5), 692-701. doi:10.1158/1055-9965.EPI-16-0721
- Hoffmann, O. (2007). Libro Comunidades Negras en el pacífico colombiano. https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1038&context=abya_yala.
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nature Reviews*, 10, 639-650.
- Ibarra, A., Freire-Aradas, A., Martinez, M., Fondevila, M., Burgos, G., Camacho, M., . . . Gusmao, L. (2014). Comparison of the genetic background of different Colombian populations using the SNPforID 52plex identification panel. *Int J Legal Med*, 128(1), 19-25. doi:10.1007/s00414-013-0858-z
- Ibarra, A., Restrepo, T., Rojas, W., Castillo, A., Amorim, A., Martinez, B., . . . Gusmao, L. (2014). Evaluating the X chromosome-specific diversity of Colombian populations using insertion/deletion polymorphisms. *PLoS One*, 9(1), e87202. doi:10.1371/journal.pone.0087202
- IGAC, I. G. A. C. (2010). Geografía y Cartografía de Colombia. <https://www.igac.gov.co/>.

- INDETUR – Instituto Distrital de Turismo Santa Marta. (2022). <https://www.santamarta.gov.co/instituto-distrital-de-turismo-indetur>
- Jobling, M., Hurler, M., & Tyler, C. (2004). Human evolutionary genetics, origins, peoples and disease. *Cambridge-UK: Garland Sciences*.
- Karol, S. E., Larsen, E., Cheng, C., Cao, X., Yang, W., Ramsey, L. B., . . . Relling, M. V. (2017). Genetics of ancestry-specific risk for relapse in acute lymphoblastic leukemia. *Leukemia*, 31(6), 1325-1332. doi:10.1038/leu.2017.24
- Kovach WL. 2007. MVSP - A MultiVariate statistical package for windows: 1-135. Available from: <https://www.kovcomp.co.uk/mvsp/>
- Keyeux G, Usaquén W. 2006. Rutas migratorias hacia Sudamérica y poblamiento de las cuencas de los ríos Amazonas y Orinoco, deducidas a partir de estudios genéticos moleculares. In: Gaspar Morcote, Mora S, Calvo C, editors. Pueblos y paisajes antiguos de la selva amazónica. Bogotá: Universidad Nacional de Colombia, Editorial Unibiblos. p 415.
- Lamprea Bermúdez, N. (2009). CARACTERIZACIÓN GENÉTICA DE LA POBLACIÓN HUMANA DE SAN ANDRÉS Y PROVIDENCIA A PARTIR DE LOS MARCADORES MICROSATÉLITES (STR'S) EMPLEADOS POR EL COMBINED DNA INDEX SYSTEM (CODIS). Repositorio Institucional UN. Retrieved from <https://repositorio.unal.edu.co/handle/unal/8061>.
- LaRue, B. L., Lagace, R., Chang, C. W., Holt, A., Hennessy, L., Ge, J., . . . Budowle, B. (2014). Characterization of 114 insertion/deletion (INDEL) polymorphisms, and selection for a global INDEL panel for human identification. *Leg Med (Tokyo)*, 16(1), 26-32. doi:10.1016/j.legalmed.2013.10.006
- Liebel, M. (2017). Infancias latinoamericanas: Civilización racista y limpieza social. Ensayo sobre violencias coloniales y postcoloniales. *Sociedad e Infancias*. 1:19-38. <http://dx.doi.org/10.5209/SOCI.55646>.
- Llamas, B. e. a. (2016). Ancient Mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.*, Vol. 2, no. 4, e1501385. doi:10.1126/sciadv.1501385
- Manta, F. S., Pereira, R., Caiafa, A., Silva, D. A., Gusmao, L., & Carvalho, E. F. (2013). Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. *Ann Hum Biol*, 40(1), 94-98. doi:10.3109/03014460.2012.742138
- Martinez, B et al. (2017). Ancestry estimates in afrodescendant population from San Basilio de Palenque, Colombia. *FSI: Genetics Supplement Series*. e224-e225. doi: 10.1016/j.fsigss.2017.09.105
- Martinez, B et al. (2020). Searching for the roots of the first free African American community. *Sci Rep* 10, 20634. doi: 10.1038/s41598-020-77608-8
- Martínez, Y. S. (2021). DINÁMICA DEL CONFLICTO ARMADO EN LA SUBREGIÓN DEL BAJO ATRATO DEPARTAMENTO DEL CHOCÓ DESDE MEDIADOS DEL SIGLO XX HASTA LA ACTUALIDAD. Repositorio universidad de Córdoba. <https://repositorio.unicordoba.edu.co/ucordoba>.
- Mateus, G et al. (2020). Origins, Admixture Dynamics, and Homogenization of the African Gene Pool in the Americas. *Mol Biol Evol*. 37(6):1647-1656. doi: 10.1093/molbev/msaa033.
- McNeer, N. A., Philip, J., Geiger, H., Ries, R. E., Lavalley, V. P., Walsh, M., . . . Kentsis, A. (2019). Genetic mechanisms of primary chemotherapy resistance in pediatric acute myeloid leukemia. *Leukemia*, 33(8), 1934-1943. doi:10.1038/s41375-019-0402-3
- Mesa et al. (2000). Autosomal, mtDNA, and Y-Chromosome Diversity in Amerinds: Pre- and Post-Columbian Patterns of Gene Flow in South America. *Am J Hum Genet*.

- Medina-Rivas et al. (2016). Chocó, Colombia: a hotspot of human biodiversity. *Rev Biodivers Neotrop*, 6(1): 45-54.
- Mogollón et al. (2020). Contrasting the ancestry patterns of three distinct population groups from the northernmost region of South America. *Am J Phys Anthropol.*, 173(3), 437-447. doi: 10.1002/ajpa.24130
- Moncada J, U. W. (2018). *Análisis de la estructura genética de una muestra de población del departamento de Amazonas, Colombia*. (Magister en Genética Humana). Universidad Nacional de Colombia, Repositorio Digital Universidad Nacional de Colombia. Retrieved from http://bdigital.unal.edu.co/70219/1/TESIS%20FINAL_Julie%20Moncada.pdf
- Mongensen, H et al. (2020). Ancestry prediction efficiency of the software GenoGeographer using a z-score method and the ancestry informative markers in the Precision ID Ancestry Panel. *Forensic Sci Int Genet.* 44:102154. doi: 10.1016/j.fsigen.2019.102154
- Mosquera, A. Y. (2016). Impacto del conflicto armado en los medios de vida de la población afrodescendiente del corregimiento de Tutunendo municipio Quibdó, departamento del Chocó. Repositorio institucional – Universidad de Medellín. <https://repository.udem.edu.co/bitstream/handle/11407/2987/Impacto%20del%20conflicto%20armado%20en%20los%20medios%20de%20vida%20de%20la%20poblaci%C3%B3n%20afrodescendiente%20del%20corregimiento%20de%20Tutunendo%20municipio%20Quibd%C3%B3,%20departamento%20del%20Choc%C3%B3.pdf?sequence=1>.
- Nakanishi H et al. (2018). Analysis of mainland Japanese and Okinawan Japanese populations using the precision ID Ancestry Panel. *Forensic Sci Int Genet.* 33:106-109. doi: 10.1016/j.fsigen.2017.12.004
- Noguera, M. C., Schwegler, A., Gomes, V., Briceno, I., Alvarez, L., Uricoechea, D., . . . Gusmao, L. (2014). Colombia's racial crucible: Y chromosome evidence from six admixed communities in the Department of Bolivar. *Ann Hum Biol*, 41(5), 453-459. doi:10.3109/03014460.2013.852244
- Ossa, H., Aquino, J., Sierra, S. . . Gusmao, L. (2015). Analysis of admixture in Native American populations from Colombia. *Fosensic Sci. Int. Gene.* 1129. 3.
- Ossa, H., Aquino, J., Pereira, R., Ibarra, A., Ossa, R. H., Perez, L. A., . . . Gusmao, L. (2016). Outlining the Ancestry Landscape of Colombian Admixed Populations. *PLoS One*, 11(10), e0164414. doi:10.1371/journal.pone.0164414
- Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., . . . Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature*, 537(7618), 45-49. doi:10.1038/nature19085
- Pereira, R. (2011). *Bridging the gap between SNPs and STRs: insertion deletion polymorphisms in forensic genetics, principles and applications*. (p.h.D Forensic Sciences and Pathology). Universidad Santiago de Compostela, Retrieved from <https://dialnet.unirioja.es/servlet/tesis?codigo=113551>
- Pereira, R., Phillips, C., Alves, C., Amorim, A., Carracedo, A., & Gusmao, L. (2009). A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophoresis*, 30(21), 3682-3690. doi:10.1002/elps.200900274
- Pereira, R., Phillips, C., Pinto, N., Santos, C., dos Santos, S. E., Amorim, A., . . . Gusmao, L. (2012). Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PLoS One*, 7(1), e29684. doi:10.1371/journal.pone.0029684
- Perez-Mayoral, J., Soto-Salgado, M., Shah, E., Kittles, R., Stern, M. C., Olivera, M. I., . . . Cruz-Correa, M. (2019). Association of genetic ancestry with colorectal tumor location in Puerto Rican Latinos. *Hum Genomics*, 13(1), 12. doi:10.1186/s40246-019-0196-4

- Potter, B. A. (2018). Current evidence allows multiple models for the peopling of the Americas. *Science Advances*. doi:10.1126/sciadv.aat5473
- Price, A. L., Patterson, N., Yu, F., Cox, D. R., Waliszewska, A., McDonald, G. J., . . . Reich, D. (2007). A genomewide admixture map for Latino populations. *Am J Hum Genet*, *80*(6), 1024-1036. doi:10.1086/518313
- Pritchard, J., & Rosenberg, N. (1999). Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies. *Am J Hum Genet*, *65*, 220-228.
- Pritchard JK, Wen X, Falush D. 2010. Documentation for structure software: Version 2 . 3. Univ Chicago 6:321-326.
- Quijano, A. (2000). Coloniality of Power and Eurocentrism in Latin America. *International Sociology*. 15:215. doi: 10.1177/0268580900015002005
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Ruiz-Linares, A. (2012). Reconstructing native American population history. *Nature*, *488*, 370–374.
- Restrepo, E., Pérez, J.N. (2005). SAN BASILIO DE PALENQUE: CARACTERIZACIONES Y RIESGOS DEL PATRIMONIO INTANGIBLE. *Jangwa Pana. Revista del Programa de Antropología de la Universidad del Magdalena*, *4*: 58-69.
- Rodas, C., Gelvez, N., & Keyeux, G. (2003). Mitochondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and Mestizo populations. *Human Biology*, *75*, 13-30.
- Rodríguez et al. (2007). Colombia una nación multicultural, su diversidad étnica. Retrieved from https://www.dane.gov.co/files/censo2005/etnia/sys/colombia_nacion.pdf
- Roewer, L., Nothnagel, M., Gusmao, L., Gomes, V., & Gonzalez, M. (2013). Continent-Wide Decoupling of Y-Chromosomal Genetic Variation from Language and Geography in Native South Americans. *PLoS One*, *9*, e1003460.
- Rojas, M. Y., Alonso, L. A., Sarmiento, V. A., Eljach, L. Y., & Usaquen, W. (2013). Structure analysis of the La Guajira-Colombia population: a genetic, demographic and genealogical overview. *Ann Hum Biol*, *40*(2), 119-131. doi:10.3109/03014460.2012.748093
- Rojas, W., Parra, M. V., Campo, O., Caro, M. A., Lopera, J. G., Arias, W., . . . Bedoya, G. (2010). Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. *Am J Phys Anthropol*, *143*(1), 13-20. doi:10.1002/ajpa.21270
- Rondon et al. (2008). Diversidad genética en poblaciones humanas de dos regiones Colombianas. *Colombia Médica*, *39*(2), 52-60.
- Rothhammer, F., & Dillehay, T. D. (2009). The late Pleistocene colonization of South America: An interdisciplinary perspective. *Annals of Human Genetics*, *73*, 540–549.
- Rothhammer, F., & Silva, C. (1989). Peopling of Andean South America. *American Journal of physical anthropology*, *78*, 403-410.
- Rousset F. 2017. Genepop version 4.6.9.
- Saloum de Neves Manta, F., Pereira, R., Vianna, R., Rodolfo Beuttenmuller de Araujo, A., Leite Goes Gitai, D., Aparecida da Silva, D., . . . Gusmao, L. (2013). Revisiting the genetic ancestry of Brazilians using autosomal AIM-Indels. *PLoS One*, *8*(9), e75145. doi:10.1371/journal.pone.0075145
- Salzano et al. (2014). Interethnic admixture and the evolution of Latin American populations. *Genetics and Molecular Biology*.
- Sandoval, C., De la Hoz, A., & Yunis, E. J. (1993). Estructura Genética de la Población Colombiana. *Revista de la Facultad de Medicina, Universidad Nacional de Colombia*, *41*(1), 3-14.
- Sans, M. (2000). Admixture studies in Latin America: from the 20th to the 21st century. *Human Biology*, *72*, 155-177.

- Santos, C., Phillips, C., Oldoni, F., Amigo, J., Fondevila, M., Pereira, R., . . . Lareu, M. V. (2015). Completion of a worldwide reference panel of samples for an ancestry informative Indel assay. *Forensic Sci Int Genet*, *17*, 75-80. doi:10.1016/j.fsigen.2015.03.011
- Santos, N. P., Ribeiro-Rodrigues, E. M., Ribeiro-Dos-Santos, A. K., Pereira, R., Gusmao, L., Amorim, A., . . . Santos, S. E. (2010). Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat*, *31*(2), 184-190. doi:10.1002/humu.21159
- Skoglund, P., Mallick, S., Bortolini, M. C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M. L., . . . Reich, D. (2015). Genetic evidence for two founding populations of the Americas. *Nature*, *525*, 104–108.
- Slatkin, M. (1994). Gene Flow and population structure. *Ecological Genetics*, 1-11.
- Tang, H., Quertermous, T., Rodriguez, B., Kardia, S., Zhu, X., Brown, A., & Pankow, M. (2005). Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies. *Am J Hum Genet*, *76*, 268-275.
- Tarazona-Santos, E., Carvalho-Silva, D. R., Pettener, D., Luiselli, D., & De Stefano, G. F. (2001). Genetic differentiation in South Amerindians is related to environmental and cultural diversity: evidence from the Y chromosome. *American Journal of Human Genetics*, *68*, 1485-1496.
- Torres, D., Lorenzo Bermejo, J., Garcia Mesa, K., Gilbert, M., Briceno, I., Pohl-Zeidler, S., . . . Hamann, U. (2019). Interaction between genetic ancestry and common breast cancer susceptibility variants in Colombian women. *Int J Cancer*, *144*(9), 2181-2191. doi:10.1002/ijc.32023
- Vali, U., Brandstrom, M., Johansson, M., & Ellegren, H. (2008). Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genet*, *9*, 8. doi:10.1186/1471-2156-9-8
- Wade, P. (1995) Blackness and race mixture: the dynamics of racial identity in Colombia. Baltimore: JHU Press.
- Wallace D.C. (1995). Mitochondrial DNA variation in human evolution, degenerative disease, and aging. *American Journal of Human Genetics*, *57*(2), 201-223.
- Walsh, K. M., de Smith, A. J., Welch, T. C., Smirnov, I., Cunningham, M. J., Ma, X., . . . Wiemels, J. L. (2014). Genomic ancestry and somatic alterations correlate with age at diagnosis in Hispanic children with B-cell acute lymphoblastic leukemia. *Am J Hematol*, *89*(7), 721-725. doi:10.1002/ajh.23727
- Wang, S., Lewis, C. M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., . . . Ruiz-Linares, A. (2007). Genetic variation and population structure in native Americans. *PLoS Genetics*, *3*, 2049–2067.
- Wilder, J., Kingan, S., Mobasher, Z., Pilkington, M., & Hammer, M. (2004). Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nat Genet*, *36*(10), 1122-1125.
- Winkler, C., Nelson, G., & Smith, M. (2010). Admixture mapping comes of age. *Annual Review of Genomics Human Genetics*, *11*, 65-89.
- Xavier, C., Builes, J. J., Gomes, V., Ospino, J. M., Aquino, J., Parson, W., . . . Goios, A. (2015). Admixture and genetic diversity distribution patterns of non-recombining lineages of Native American ancestry in Colombian populations. *PLoS One*, *10*(3), e0120155. doi:10.1371/journal.pone.0120155
- Yunis et al. (1994). Major Histocompatibility Complex Class II alleles and Haplotypes and Blood Groups of Four Amerindian Tribes of Northern Colombia. *Human Immunology*.
- Yunis et al. (2001). Genetic Relationship of the Guambiano Paez, and Ingano Amerindians of Southwest Colombia Using Major Histocompatibility Complex Class II Haplotypes and Blood Groups. *Human Immunology*.

- Yunis, J. J. (2013). Mitochondrial DNA (mtDNA) haplogroups in 1525 unrelated individual from 11 Departamentos of Colombia. *Genetics and Molecular Biology*.
- Yunis, J. J., Acevedo, L. E., Campo, D. S., & Yunis, E. J. (2005). Population data of Y-STR minimal haplotypes in a sample of Caucasian-Mestizo and African descent individuals of Colombia. *Forensic Sci Int*, 151(2-3), 307-313. doi:10.1016/j.forsciint.2005.02.005
- Yunis, J. J., Acevedo, L. E., Campo, D. S., & Yunis, E. J. (2013). Geno-geographic origin of Y-specific STR haplotypes in a sample of Caucasian-Mestizo and African-descent male individuals from Colombia. *Biomedica*, 33(3), 459-467. doi:10.7705/biomedica.v33i3.807
- Zaumsegel, D. (2013). *Binary Polymorphisms as Ancestry Informative Markers*. (Doctor rerum medicinalium). Universitat zu Koln, Retrieved from <https://repository.publisso.de/resource/frl:5424570-1/data>