

**CARACTERIZACIÓN MOLECULAR DE LA VARIACIÓN GENÉTICA EN  
CUATRO ETNIAS INDÍGENAS (PIJAO, PAEZ, EMBERA Y ZENU) Y DOS  
POBLACIONES MESTIZAS DE COLOMBIA (TOLIMA Y CÓRDOBA) MEDIANTE  
MARCADORES DEL mDNA, NRY Y AIMs**



**ANGEL ALEXANDRO CRIOLLO RAYO**

**UNIVERSIDAD DEL TOLIMA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE BIOLOGÍA  
MAESTRIA EN CIENCIAS BIOLÓGICAS  
IBAGUE, DICIEMBRE  
2012**

**CARACTERIZACIÓN MOLECULAR DE LA VARIACIÓN GENÉTICA EN  
CUATRO ETNIAS INDÍGENAS (PIJAO, PAEZ, EMBERA Y ZENU) Y DOS  
POBLACIONES MESTIZAS DE COLOMBIA (TOLIMA Y CÓRDOBA) MEDIANTE  
MARCADORES DEL mDNA, NRY Y AIMs**

**ANGEL ALEXANDRO CRIOLLO RAYO**

**Trabajo de grado para optar al título de Máster en Ciencias Biológicas**

**Directora:**

**María Magdalena Echeverry De Polanco, Doctora en Genética**

**Codirector:**

**Luis Guillermo Carvajal Carmona, Doctor en Genética**

**UNIVERSIDAD DEL TOLIMA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE BIOLOGÍA  
MAESTRIA EN CIENCIAS BIOLÓGICAS  
IBAGUE, DICIEMBRE  
2012**

Nota de aceptación

---

Presidente Jurado

---

Jurado

---

Jurado

Ibagué, diciembre de 2012

## **DEDICATORIA**

Dedico este trabajo al Maestro, hacedor de la vida, por la afortunada oportunidad al permitirme recorrer un camino, en el que he escuchado voces ancestrales, a redescubrir los senderos de antiguas aventuras humanas y permitirme recrear en mi corazón toda una historia fascinante llena de ancestrales visiones, para comprender y dimensionar en alguna medida los orígenes de una parte querida y adorada de Colombia. También porque me ha concedido toda entereza, entusiasmo y la dedicación constante, para así, hoy poder contar una nueva parte de la historia, aquella que mis abuelos Elías, María Emma, Ismael y Cristina, me enseñaron...”cuando apenas decía mama”, mientras al son de bambucos y trino de las guitarras, aprendía la esencia de la tierra amada, aquella de las pampas ardientes, de los nevados, de San Pedro en el Espinal y de la sombrerera, y, por supuesto de mis entrañables padres, hermanos, y amigos que con sus convicciones y fe inquebrantables me han ayudado a formar un espíritu de compromiso con Dios en el respeto y amor por la vida, en la sinceridad, cariño y hermandad.

## AGRADECIMIENTOS

Muchas gracias a Dios porque en este camino he conocido a una persona que me ha guiado y ha puesto en mí su confianza. Alguien quien con mucho cariño y dedicación me ha brindado fortaleza para afrontar retos académicos, científicos y personales, en quien he visto el ejemplo de una mujer, madre y científica, que con saber, paciencia y emprendimiento ha motivado un gran proyecto de vida, es la profesora María Magdalena Echeverry, quien ha sido también mi mentora desde que realizaba los estudios de pregrado en Biología, hace 11 años, y de quien he recibido una afortunada herencia de conocimientos en genética y en muchos otros aspectos invaluable de la vida que nos enriquecen como seres integrales, y quien junto con Luis Guillermo Carvajal, me han dado la oportunidad de participar en su proyecto de vida universitaria, el Grupo de Citogenética de la Universidad del Tolima, en el que han podido emprender varios retos, uno de los cuales culmina con los resultados de este trabajo y del cual me siento dichoso por haber participado.

Para Luis Guillermo mis más sinceros agradecimientos por su apoyo constante durante todo el proyecto, porque ha sido una persona que me ha inspirado a construirlo de manera ferviente, con mucha dedicación y disciplina. A quien también le debo muchas enseñanzas prácticas, especialmente durante mi estancia en Inglaterra, por orientarme en todo lo relacionado con la genotipificación de los marcadores del mDNA y del cromosoma Y, entre muchos otros aspectos del trabajo investigativo, de la convivencia y estancia en Oxford, sin lo cual no habría sido posible el presente estudio. También agradezco de manera especial a Rodrigo Prieto y Mabel Bohórquez, por su incondicional, invaluable y magnífico apoyo personal y académico, también por su entusiasmo, espíritu de compañerismo y acompañamiento durante los muestreos a las comunidades indígenas y por brindarme orientación en muchas fases del trabajo.

Este proyecto no se habría podido llevar a cabo sin la voluntad de servicio y colaboración de la Oficina de Investigaciones y Desarrollo Científico, especialmente por parte de quien fuera su director en momentos claves durante el inicio y conclusión del proyecto: John Jairo Méndez Arteaga, pues fue quien nos acompañó en la primera reunión con el Consejo Regional Indígena del Tolima – CRIT-; nos prestó su invaluable apoyo durante los muestreos y finalmente me colaboró con el trámite institucional de la pasantía en España e Inglaterra. Muchos agradecimientos a Ulises Tique y Cesar Culma entre otros miembros de la Junta directiva del CRIT por que de manera amable nos abrió las puertas a las comunidades indígenas del Tolima, así como también a una maravillosa mujer como lo fue Mergida del Carmen Fuentes, Q.E.P.D, por permitirnos llegar a varias comunidades Zenú y Embera en Córdoba, y en general, a todas las personas que hicieron posible que el proyecto llegara a distintas comunidades indígenas: Ovidio Paya, Patricia Quilcue, Alberto Moreno, Felipe Montealegre, Luis Alver Tovar,

Oscar Neira, Patricia Alvis, Álvaro Yara, Olier Basilio, Alirio Niaza, Yaqui Yagan, Isidro Mendez, Aurelio Jumi, resguardo Indígena Nasa-Paez de Gaitania, resguardo Guatavita Tua de Ortega-Tolima, resguardo Lomas de Hilarco en Coyaima- Tolima, Cabildo indígena Zenú, Porvenir La Fe, en Montelíbano-Córdoba, Comunidad indígena Embera Katío del Pindo y de Dopavarandó en Puerto Libertador-Cordoba, Cabildo indígena del Kilómetro 35, en Sahagun-Córdoba y Comunidad Embera de Mariquita-Tolima.

Agradezco de manera muy especial a:

Angel Carracedo, director del Centro Nacional de Genotipado –CEGEN- y del Instituto de Medicina legal de la Universidad Santiago de Compostela, Galicia-España, por permitirme realizar la genotipificación de los 100 SNPs en el CEGEN y también por su amable colaboración y orientación durante mi estancia allí, así como también por facilitar los datos genotípicos para las poblaciones de referencia del presente trabajo. También muchas gracias a María Torres, Juan Ansede, Karla Santos, Montserrat Santori, Jorge Cárdenas, Alex Blanco, Lorena Salazar y Luisa Fernanda Leon, quienes amablemente me colaboraron durante mi trabajo en el CEGEN. Integrantes del Grupo de Genética Poblacional Molecular del *Wellcome Trust Centre for Human Genetics* de la Universidad de Oxford-Inglaterra, especialmente a Ian Tomlinson, Luis Carvajal y Kimberley Howarth por su colaboración en la tipificación de los marcadores del mtDNA y del cromosoma Y. También a muchas gracias a las personas que hicieron de mi estancia en Oxford agradable y amable: Lupe Polanco, Rob Lawrance, Jess Reynolds, Alex Lerner y Tom Brennan. Finalmente, muchas gracias a la familia de la profesora Magdalena, como Don Raúl Polanco y Laura Polanco, por el cariño, valioso apoyo y acompañamiento en varios aspectos de este proyecto. Así como también a Aura María González del Centro Asistencial de la Universidad del Tolima, por su valiosa ayuda en la logística de la toma de muestras, y en general, a todos tanto en la Universidad del Tolima como en los municipios de Planadas y Ortega, quienes amablemente donaron una muestra de sangre para el presente proyecto. A todos y todas muchas gracias, porque su colaboración, ha permitido que este trabajo sea una realidad.

## ABREVIATURAS

La mayoría de las abreviaturas que aparecen en el presente proyecto son las registradas en la lengua inglesa. No obstante en la presente lista se adiciona su respectiva traducción al castellano.

**A:** *Adenine*. (Adenina).

**Afr:** africano.

**AIMs:** *Ancestry informative markers* (marcadores informativos de ancestro).

**AM:** *Admixture Mapping* (mapeo de la mezcla).

**Ami:** amerindio.

**C:** *Cytosine*. (Citosina).

**CEGEN:** Centro Nacional de Genotipado.

**CNVs:** *Copy number variations*. (Variaciones en el número de copias).

**DNA:** *Deoxyribonucleic acid* (Acido desoxiribonucleico).

**DYS199, SRY-2627, DYS271:** marcadores bialélicos del cromosoma Y.

**Eur:** europeo.

**G:** *Guanine*. (Guanina).

**Hg:** Hg.

**Hgs:** Hgs.

**HapMap:** *Haplotype Map*. (Mapa de haplotipos).

**HGDP:** *Human Genome Diversity Project* (Proyecto de la Diversidad Genómica Humana).

**HVR-I:** *High variable region- I*. (Región altamente variable-I).

**LA:** Latino América.

**LD:** Linkage Disequilibrium. (Desequilibrio de Ligamiento).

**MAF:** *Minor allele frequency* (Frecuencia alélica del alelo menor).

**MALDI-TOF:** *Matrix Assisted Laser Desorption/ionization – Time of Flight* (Desorción/ionización laser asistida por matriz- tiempo de vuelo).

**MCMC:** *Monte Carlo Markov Chain* (Cadenas Markov de Monte Carlo).

**mDNA:** *Mitochondrial DNA* (ADN mitocondrial).

**MDS:** *Multidimensional Scalling* (Análisis escalonado multidimensional).

**Nam:** nativo Americano.

**NCBI:** *National Center for Biotechnology Information*. (Centro Nacional para la Información Biotecnológica).

**NRY:** *No recombinant región of Y-chromosome* (Región no recombinante del cromosoma Y).

**PCR:** *Polymerase chain reaction*. (Reacción en cadena de la polimerasa).

**RFLPs:** *Restriction fragment length polymorphisms* (Polimorfismos de longitud en los fragmentos de restricción).

**RNA:** *Ribonucleic Acid*. (Ácido ribonucleico).

**SAP:** *Shrimp Alkaline Phosphatase* (fosfatasa alcalina de camarón).

**SNPs:** *Single nucleotide polymorphism* (polimorfismos de nucleótido único).

**STRs:** *Short tandem repeats* (Repeticiones cortas en tándem).

**T:** *Thymine* (Timina).



## RESUMEN

Las poblaciones colombianas humanas modernas poseen una estructura genética compleja debido a la mezcla amerindia, europea y africana, ocurrida en diferentes proporciones a lo largo del país. En el Tolima se desarrolló un estudio genético para caracterizar las tres ancestrías mencionadas, comparándolas con otra población localizada en la Costa Caribe (Córdoba). Se obtuvo ADN a partir de muestras de sangre tomadas en personas indígenas y no-indígenas procedentes de ambas regiones, para tipificar 6 haplogrupos del ADN mitocondrial (mDNA), 17 del cromosoma Y (NRY) y 100 SNPs autosómicos (AIMs) con el fin de estimar las proporciones de ancestría genética. Los indígenas del Tolima Nasa y Pijao fueron predominantemente de origen amerindio (AIMs: 91%, mDNA: 96%, NRY: 71%), mientras que los individuos mezclados no nativos presentaron principalmente un origen paterno europeo (NRY) del 70%, materno amerindio (mDNA) del 93%, y proporciones de ancestría autosómica europea del 49% y amerindia del 45% en promedio (origen biparental). Por otra parte, los indígenas Embera de Córdoba presentaron proporciones similares a los del Tolima (origen principalmente uniparental), en tanto que los Zenu mostraron proporciones similares con los individuos mezclados de la misma región: ancestría europea (AIMs: 42%, NRY: 67%), amerindia (AIMs: 35%, NRY: 10%, mDNA: 60%) and africana (AIMs: 23%, NRY: 24%, mDNA: 13%) (triparental). Los resultados sugieren que la muestra poblacional global está genéticamente estructurada, lo cual puede deberse a diferencias locales en la historia demográfica como la influencia africana mayor en algunas poblaciones de Córdoba, entre otros eventos socioculturales recientes.

### **Palabras de recuperación:**

Estructura genética poblacional  
Cromosoma Y  
DNA mitocondrial  
Mezcla genética  
Marcadores informativos de ancestría  
Nativos americanos

## ABSTRACT

Modern human Colombian populations possess a complex genetic structure because of the admixture between Amerinds, Europeans and Africans, occurred in different proportions across regions of the country. We developed a genetic study aimed to characterize Amerindian, European and African ancestry in a Colombian Andean population (Tolima), which was compared with another one located at the Caribbean coast region (Córdoba). Genetic data were obtained from indigenous and admixed individuals DNA samples, typed for 6 mitochondrial (mDNA) and 16 Y chromosome non-recombinant region (NRY) continental lineages and 100 autosomal ancestry informative markers (AIMs) to estimate genetic ancestry and the significance was evaluated by non-parametric test. In the region of Tolima, Nasa and Pijao indigenous were predominantly of Native American origin 91% (AIMs), 96% (mDNA) y 71% (NRY), whereas admixed populations were paternally and maternally respectively: 70% of European origin and 93 % Amerindian, and for the autosomal markers: 49% European and 45% Amerindian and African ancestry were <5%. On the other hand, Embera indigenous population from the Caribbean coast presented similar proportions to Tolima's Amerinds, while Zenu indigenous were similar to admixed individual from the same region, who shown European (AIMs: 42%, NRY: 67%), Amerind (AIMs: 35%, NRY: 10%, mDNA: 60%) and African ancestry (AIMs: 23%, NRY: 24%, mDNA: 13%). The results suggest that the whole sample is genetically structured, in this way the last two subpopulations have a triparental origin, while Tolima's admixed individuals are biparental, which could be related to differences in local demographic histories, such a higher African influence at Caribbean among other recent sociocultural events.

Key words:

Population genetic structure  
Y chromosome  
Mitochondrial DNA  
Genetic admixture  
Ancestry informative markers  
Native Americans

## CONTENIDO

	Pg
<b>INTRODUCCIÓN</b>	19
<b>1. PLANTEAMIENTO DEL PROBLEMA</b>	21
<b>2. JUSTIFICACIÓN</b>	22
<b>3. OBJETIVOS</b>	23
3.1 OBJETIVO GENERAL	23
3.2 OBJETIVOS ESPECÍFICOS	23
<b>4. HIPÓTESIS</b>	24
<b>5. MARCO TEÓRICO Y ANTECEDENTES</b>	25
5.1 GENÉTICA MOLECULAR Y EL ESTUDIO DE LAS POBLACIONES HUMANAS	25
5.1.1 Estudios genómicos amplios	25
5.1.2 Patrones globales de diversidad	25
5.1.3 Estructura genética de la población humana	26
5.1.4 Estructura genética de las poblaciones latinoamericanas	27
5.1.5 Estructura y diversidad genética de las poblaciones amerindias y mestizas colombianas	31
<b>6. METODOLOGÍA</b>	35
6.1 POBLACIONES Y MUESTRAS	35
6.1.1 Permiso ético y consentimiento informado	35
6.1.2 Contactos con las poblaciones Indígenas y mezcladas	35
6.1.3 Recolección de muestras y entrevistas	36

6.1.4 Poblaciones	37
6.2 EXTRACCIÓN, CUANTIFICACIÓN Y DILUCIÓN DEL DNA	40
6.3 MARCADORES GENÉTICOS Y GENOTIPIFICACIÓN	40
6.3.1 DNA autosómico (AIMs)	40
6.3.2 DNA mitocondrial (mDNA)	45
6.3.3 Región no recombinante del cromosoma Y (NRY)	48
6.4 ANÁLISIS ESTADÍSTICOS	53
<b>7. RESULTADOS</b>	<b>56</b>
7.1 ANCESTRÍA MATERNA DE LAS POBLACIONES ANALIZADAS	56
7.1.1 Comparación inter-regional	60
7.1.2 Comparación intra-regional	60
7.1.3 Análisis de distancias Fst	61
7.2 ANCESTRÍA PATERNA DE LAS POBLACIONES ANALIZADAS	64
7.2.1 Comparaciones inter-regionales	66
7.2.2 Comparaciones intra-regional	69
7.2.3 Análisis de distancias Fst	69
7.3 ANCESTRÍA A NIVEL AUTOSÓMICO	71
7.3.1 Análisis de la diversidad genética en los tolimenses	71
7.3.2 Ancestría poblacional.	73
7.3.3 Ancestría individual	80
<b>8. DISCUSIÓN</b>	<b>84</b>
8.1 LAS POBLACIONES DE ESTUDIO	84

8.2 DIVERSIDAD GENÉTICA Y DIFERENCIACIÓN POBLACIONAL EN EL TOLIMA Y CÓRDOBA	87
8.2.1 Origen materno de la diversidad	88
8.2.2 Origen paterno de la diversidad	90
8.2.3 Diversidad autosómica (AIMs)	91
8.3 ORÍGENES ANCESTRALES, MEZCLA Y ESTRUCTURA GENÉTICA DE LAS POBLACIONES ANALIZADAS	93
8.3.1 Linajes maternos y paternos en el Tolima	94
8.3.2 Estructura genética global de las poblaciones analizadas	98
<b>9. CONCLUSIONES</b>	108
<b>10. RECOMENDACIONES</b>	110
<b>REFERENCIAS</b>	111
<b>ANEXOS</b>	126

## LISTA DE FIGURAS

	Pg
<b>Figura 1.</b> Patron de diversidad global.....	26
<b>Figura 2.</b> Estructura genética poblacional global.....	27
<b>Figura 3.</b> Árbol correspondiente al clado q del cromosoma y. ....	28
<b>Figura 4.</b> Jornadas de socialización y presentación del proyecto. ....	36
<b>Figura 5.</b> Jornadas de muestreo. ....	37
<b>Figura 6.</b> Localización geográfica de los sitios de muestreo. ....	37
<b>Figura 7.</b> Fragmento del árbol filogenético del cromosoma y.....	50
<b>Figura 8.</b> Tipificación de las muestras indígenas y mestizas para los hgs mitocondriales mediante RFLPs. ....	56
<b>Figura 9.</b> Proporciones de los haplogrupos nativos americanos (A, B, C y D), africano (I) y europeo (j) en cada población analizada. ....	59
<b>Figura 10.</b> Análisis de escalamiento multidimensional (mds) para el mDNA. ....	62
<b>Figura 11.</b> Análisis de escalamiento multidimensional (mds) mDNA-2. ....	63
<b>Figura 12.</b> Agrupamiento de los genotipos del cromosoma y generado durante la tipificación de los hgs Q y R1.....	65
<b>Figura 13.</b> Agrupamiento de los genotipos del cromosoma y generado durante la tipificación de los hgs Q y R1.....	65
<b>Figura 14.</b> Proporción ancestral de los diferentes hgs de la región nry en cada población analizada. ....	68
<b>Figura 15.</b> Análisis de escalamiento multidimensional (MDS) en el cromosoma Y. .....	69
<b>Figura 16.</b> Heterocigocidad media (het) con base en los AIMs. ....	71
<b>Figura 17.</b> Análisis mds con base en los AIMs.....	72

<b>Figura 18.</b> Comportamiento del valor alfa a lo largo de las parejas de valores iteraciones- MCMC. ....	74
<b>Figura 19.</b> Estructura genética de las poblaciones estudiadas. ....	75
<b>Figura 20.</b> Proporciones (%) de ancestría nativa, africana, y europea con base en los AIMS. ....	78
<b>Figura 21.</b> Diagramas <i>box-plot</i> para la representación de la dispersión de las proporciones ancestrales. ....	81
<b>Figura 22.</b> Desviación estándar de las proporciones ancestrales analizadas. ....	82
<b>Figura 23.</b> Análisis mds mediante la matriz de distancia genética pareada individuo-individuo. ....	83
<b>Figura 24.</b> Distribución de los grupos indígenas en el Valle Alto del Magdalena durante el periodo de la conquista. ....	84
<b>Figura 25.</b> Relación de la diversidad haplogrupal y la proporción de mezcla. ....	91
<b>Figura 26.</b> Tendencia global de la heterocigocidad. ....	93
<b>Figura 27.</b> Correlación entre la heterocigocidad y la proporción de mezcla no nativa. ....	103
<b>Figura 28.</b> Estructuración genética poblacional observada en las poblaciones del estudio. ....	105

## LISTA DE TABLAS

	<b>Pg</b>
<b>Tabla 1.</b> Relación de número total de individuos muestreados. ....	38
<b>Tabla 2.</b> Individuos mestizos colectados. ....	39
<b>Tabla 3.</b> Mezcla de reacción para la pcr de amplificación. ....	42
<b>Tabla 4.</b> Mezcla para la reacción sap. ....	42
<b>Tabla 5.</b> Volúmenes para la mezcla de la reacción iPLEX. ....	43
<b>Tabla 6.</b> Tipificación de los hgs mitocondriales. ....	46
<b>Tabla 7.</b> Mezcla de reacción de la amplificación de cada hg mitocondrial. ....	47
<b>Tabla 8.</b> Mezcla de reacción de la digestión de los amplicones para cada hg mitocondrial. ....	48
<b>Tabla 9.</b> Marcadores de la región no recombinante del cromosoma y tipificados. ....	49
<b>Tabla 10.</b> Frecuencias (%) de los hgs mitocondriales fundadores amerindios A, B, C y D, el hg L africano y el J europeo. ....	58
<b>Tabla 11.</b> Matriz de los valores Fst pareados entre las poblaciones del estudio. ....	61
<b>Tabla 12.</b> Información para cada uno de los marcadores de la región NRY. ....	64
<b>Tabla 13.</b> Frecuencias (%) de los hgs de la región nry en las poblaciones estudiadas. ....	67
<b>Tabla 14.</b> Proporciones de membrecía calculadas en structure en cada uno de los tres grupos (k=3). ....	76



## LISTA DE ANEXOS

**Anexo A:** Consentimientos informados.

**Anexo B.** Detalles para cada uno de los 100 AIMs seleccionados, en orden aparece: IDrs (código del SNP), el cromosoma y la posición, el alelo 1 (A1) y el alelo 2 (A2), las frecuencias alélicas reportadas para amerindios (FA-Ami), europeos (FA-Eur), y africanos (FA-Afr), la población para la cual es informativo, los valores LSBL (Fst) y LSBL (In).

**Anexo C.** Secuencias de los cebadores para la reacción de amplificación, para las regiones de los 100 SNPs.

**Anexo D.** Secuencias de los cebadores para la reacción de extensión MassEXTEND, durante la tipificación de los 100 SNPs mediante la metodología de Sequenom MassARRAY iPLEX.

**Anexo E.** Cebadores empleados para la tipificación de los Hgs de la región NRY.

**Anexo F.** Valores P del Fst calculado a partir de las frecuencias de los Hgs mitocondriales en las poblaciones del presente estudio.

**Anexo G.** Matriz de datos para los valores valore del Fst pareados entre poblaciones, calculado a partir de las frecuencias de los Hgs mitocondriales en las poblaciones del estudio y de otras regiones de Colombia y Latinoamérica. Para la fuente de los datos en las poblaciones de referencia, ver pág.124.

**Anexo H.** (a) Matriz de valores pareados Fst con base en las frecuencias de los marcadores del cromosoma Y. Para la fuente de los datos en las poblaciones de referencia, ver pág.132. (b) Valores p de significancia de los Fst calculados.

**Anexo I.** Frecuencias génicas de los 100 AIMs. MAF: frecuencia alélica del alelo menor.

**Anexo J.** Prueba de *Mann-Whitney* para analizar estadísticamente las similitudes o diferencias inter (a) o intrapoblacionales (b) en razón de las proporciones ancestrales nativa americana, europea y africana epresentadas por los grupos de poblacionales del presente estudio. El programa estadístico empleado fue la versión libre de Infostat, v.2008 (las diferencias significativas están resaltadas en rojo).

**Anexo K.** Medidas estadísticas de resúmen, calculadas en el programa *Infostat* v.8, a partir de las proporciones ancestrales que se obtuvieron mediante el análisis de los datos genotípicos de 100 Alms en STRUCTURE v.2. Las poblaciones

analizadas fueron las indígenas (Nasa, Pijao, Embera, Zenu) y mestizas (Ibagué, Ortega, Planadas y Córdoba).

**Anexo L.** Análisis MDS con base en la matriz pareada de distancias  $F_{st}$  obtenido de las frecuencias haplogrupales del mDNA en las poblaciones objeto de estudio, entre otras poblaciones de Colombia y Latino América. Solo se representa la primera y tercera coordenada. Para la fuente de los datos en las poblaciones de referencia, ver pág.124.

**Anexo M.** Análisis multidimensional con base en la matriz de distancias  $F_{st}$  pareadas en el (a) mDNA, (b) cromosoma Y y (c) SNPs autosómicos solo en las poblaciones indígenas del estudio.

**Anexo N.** Tabla basada en un diseño de cuadro *Punnet* para el cálculo de las frecuencias de ancestría autosómica esperada, con base en el conocimiento previo de las proporciones ancestrales en el mDNA y el cromosoma Y, en una población mezclada a partir de tres orígenes ancestrales y sesgo en la mezcla y sin aporte materno africano. (a) Las proporciones ancestrales uniparentales asumidas, (b) la matriz de *Punnet* con los aporte de cada sistema genético en cada cruce con su respectiva probabilidad, y (c) suma total de las probabilidades de los aportes de cada cruce a las proporciones ancestrales con los cuales de hace la grafica.

**Anexo O.** Tabla basada en un diseño de cuadro *Punnet* para el cálculo de las frecuencias de ancestría autosómica esperada, con base en el conocimiento previo de las proporciones ancestrales en el mDNA y el cromosoma Y, en una población de tres orígenes ancestrales con sesgo relativo en la mezcla y todos los aportes posibles. (a) Las proporciones ancestrales uniparentales asumidas, (b) la matriz de *Punnet* con los aporte de cada sistema genético en cada cruce con su respectiva probabilidad, y (c) suma total de las probabilidades de los aportes de cada cruce a las proporciones ancestrales con los cuales de hace la gráfica.

**Anexo P.** Ruta de los conquistadores en la región del Tolima hacia 1550. Fuente: Triana, 1992.

**Anexo Q.** Patrón de agrupamiento de los indígenas Nasa (A: verde), Pijao (B: azul), Embera (C: amarillo) y Zenu (D: fucsia), con relación a las poblaciones de referencia (en gris): africanos (superior), europeos (izquierda) y nativos de LA (derecha). La grafica se obtuvo mediante un análisis MDS a partir de los genotipos de cada individuo.

**Anexo R.** Patrón de agrupamiento de los mestizos de Ibagué (A), Ortega (B), Planadas (C) y Córdoba (D), con relación a las poblaciones de referencia (en gris): africanos (superior), europeos (izquierda) y nativos de LA (derecha).

## INTRODUCCIÓN

El conocimiento de la ancestría biogeográfica, el grado de la mezcla genética y la estructuración poblacional es relevante para los estudios de asociación genética (Winkler, Nelson, & Smith, 2010), así como también el análisis de migraciones y poblamiento (Martinez-Cortes et al., 2012; Reich et al., 2012; Salas et al., 2004; Sandoval et al., 2012), principalmente en regiones como la latinoamericana, donde ha ocurrido un proceso reciente de mezcla intercontinental. Por esta razón se han realizado varios estudios, unos enfocados principalmente en los patrones de distribución de la diversidad amerindia y su relación con el poblamiento de América (Kidd et al., 2011; Reich et al., 2012), y otros en el efecto de la mezcla genética ocurrida desde los periodos coloniales. Estos estudios en conjunto han demostrado que la variación regional en la historia demográfica y la biogeografía juegan un papel importante en la estructura genética de las poblaciones latinoamericanas. Por ejemplo, el alto componente nativo en México, el europeo en Argentina o africano en Brasil, se relaciona con las diferencias en la densidad amerindia inicial y la tasa de migración desde Europa o África respectivamente (Corach et al., 2010; Giolo et al., 2012; Silva-Zolezzi et al., 2009). En Colombia los estudios sugieren patrones de mezcla complejos (Rojas et al., 2010), que podrían explicarse por las diferencias geográficas amplias, que van desde zonas costeras, valles y regiones montañosas hasta sitios aislados como las selvas. Este hecho ha podido influenciar el patrón inicial de distribución poblacional, aislamiento y limitación de migraciones internas en periodos iniciales del crecimiento demográfico colombiano. En consecuencia, hoy en día Colombia presenta zonas biogeográficas, caracterizados por una alta densidad afrodescendiente (entre un 21 y 74%), como las costas Pacífica y Atlántica o regiones de predominio nativo como en el amazonas, la región de la Orinoquia (22 y 61%), un poco menor en la costa Caribe y el pacífico (6-21%). Mientras que la región de los Andes colombianos presenta las menores densidades indígenas (<1%), a excepción del Tolima, Risaralda y Caldas (3-5%) (DANE 2005). La región Andina es habitada por una gran parte de la población que no tiene una afiliación étnica particular, que de acuerdo con un estudio reciente, muestra una ancestría nativa mitocondrial y europea a nivel del cromosoma Y, con variaciones regionales, lo cual indica su alta diversidad (Rojas et al., 2010). No obstante, es necesario ampliar el conocimiento acerca de las proporciones de ancestría autosómica en Colombia y particularmente en el Tolima. En el presente proyecto se estudió la diversidad y estructura genética en una muestra de la población tolimese, comparándose con una población externa, localizada en una región de influencia afrodescendiente y nativa a la vez, en el departamento de Córdoba, para poder analizar la relación de las historias demográficas en los patrones de mezcla genética a nivel del cromosoma Y, el mDNA y los autosomas.

Este estudio confirma la baja diversidad en las poblaciones nativas y su alta ancestría amerindia; el sesgo en la mezcla nativo-europeo de las poblaciones

mestizas, y una variación regional en el componente africano, dependiendo principalmente de la cercanía a centros de alta densidad afrocolombiana. Adicionalmente, los resultados obtenidos podrán considerarse para el direccionamiento de futuros estudios de asociaciones genómicas en la región, como el mapeo de la mezcla, que permitan detectar variantes asociadas con la susceptibilidad de determinados síndromes como el cáncer de colorectal.

## 1. PLANTEAMIENTO DEL PROBLEMA

La población colombiana posee una diversidad étnica reconocida y distribuida diferencialmente en todo su territorio (Rodríguez et al., 2007). Se ha encontrado que estas diferencias regionales pueden estar relacionadas con la distribución de los principales componentes genéticos ancestrales que hacen parte de los colombianos, como el amerindio, el europeo y el africano (Bryc et al., 2010; Rojas et al., 2010). Particularmente la región del Tolima históricamente ha estado habitada por diversos grupos amerindios, posteriormente colonizados por inmigrantes europeos (V. Bedoya, 1950; Triana, 1992). Como resultado de este proceso demográfico, en la actualidad la región de interés posee diferentes tipos de poblaciones, unas autodenominadas de origen nativo-amerindio, que habitan principalmente en el sur del departamento como los integrantes de la etnia Nasa-Paez y los Pijaos con una proporción baja (4.3%), mientras que una gran parte de la población (94.5%) no se identifica con un grupo étnico particular, esta divergencia permite hipotetizar una correspondiente estructuración genética aun no conocida, vinculada con procesos migratorios y demográficos particulares. Por otro lado, el Tolima es una de las regiones de menor influencia afrocolombiana (1.2%) si se le compara con otros departamentos como Córdoba (13%), este hecho podría indicar una relación entre la densidad afrocolombiana y el respectivo componente en la población. Considerando lo anterior, el presente trabajo se basó en un análisis genético mediante marcadores uniparentales del cromosoma Y, el mtDNA y SNPs autosómicos, con el fin de abordar el siguiente interrogante con relación al Tolima:

¿Cuál es la estructura genética de la región, en qué medida los aportes ancestrales han afectado su diversidad génica actual y como estos dos factores se comparan con los de otras poblaciones colombianas?

## 2. JUSTIFICACIÓN

Los análisis del mDNA, NRY (Bisso-Machado, Bortolini, & Salzano, 2012; Mesa et al., 2000) y AIMS, en poblaciones Amerindias y Mestizas han sido de utilidad para dilucidar la historia del poblamiento de América, su estructura y diversidad genética (Bryc et al., 2010; O'Rourke & Raff, 2010; Perego et al., 2010; Reich et al., 2012). La importancia del uso de los tres sistemas genéticos radica en que es posible ver la interacción de la historia materna y paterna en la composición y distribución de la diversidad y las proporciones ancestrales en una población dada (Corach et al., 2010; Rojas et al., 2010). En Colombia ya se han realizado estudios en los cuales se analizan estos tres sistemas polimórficos simultáneamente, y en el caso específico del Tolima, hay dos reportes para los Pijaos (Coyaima y Natagaima), en los cuales se registran las frecuencias de los haplotipos mitocondriales y los microsatélites autosómicos (Rondon, Osorio, & Peña, 2008; Rondon, Vallejo, Osorio, & Barreto, 1999). Sin embargo, es necesario ampliar el número de marcadores para dilucidar la estructura, diversidad genética y proporciones ancestrales en una muestra de la población tolimense, tanto nativa como la mezclada. Este tipo de trabajo es necesario, por un lado para el conocimiento de la historia genética de la región, y por otro, para poder avanzar en el desarrollo de estudios de asociaciones genéticas, como el mapeo de la mezcla, que buscan genes vinculados con un estado ancestral particular, y que a la vez, aportan una pequeña fracción al riesgo total de ciertas enfermedades complejas como el cáncer de colon o de glándula mamaria. En este sentido, el presente estudio consistió en un análisis genético integral de los linajes maternos y paternos para encontrar los orígenes ancestrales de la población tolimense, la distribución de la diversidad y la extensión de la mezcla mediante el uso de SNPs autosómicos. Esto en conjunto, permitió realizar inferencias acerca de la estructuración genética en la región y compararla con otras que poseen una historia demográfica diferente como Córdoba, en la costa Caribe.

### 3. OBJETIVOS

#### 3.1 OBJETIVO GENERAL

Caracterizar molecularmente una muestra indígena proveniente de tres etnias – Pijao, Paez, Embera y Zenu- y otra compuesta por individuos mestizos no relacionados de las regiones del Tolima y Córdoba; con la finalidad de identificar grados de diferenciación genética y patrones de mezcla genética mediante el análisis de análisis de 6 Hgs del DNA mitocondrial, 100 marcadores informativos de ancestría (SNPs autosómicos) y 17 marcadores bialélicos del cromosoma Y.

#### 3.2 OBJETIVOS ESPECÍFICOS

- Determinar, a través de métodos moleculares y estadísticos, la frecuencia de los Hgs mitocondriales (A, B, C, D, L y J), marcadores bialélicos de la región no recombinante del cromosoma Y empelados para definir los linajes paternos (Q, R, J, I, G, K-R, L o NO) y 100 marcadores SNPs autosómicos - informativos de ancestro-, para establecer los orígenes ancestrales de la población de estudio.
- Comparar las frecuencias encontradas con aquellas previamente reportadas en poblaciones parentales amerindias, europeas y africanas, sin mezcla para los marcadores informativos de ancestro (AIMs), a través de lo cual se pueda estimar del grado de mezcla genética.
- Establecer la estructura, asimetría en la mezcla y la subestructuración poblacional (índices de consanguinidad F) con relación al origen ancestra, mediante los análisis moleculares y comparaciones en el mDNA y NRY.
- Generar parte de la base de datos del programa de investigación: “Análisis genético de enfermedades humanas” del grupo de investigación “Citogenética, filogenia y Evolución de Poblaciones”, que permita establecer posteriores comparaciones de frecuencias de los marcadores moleculares utilizados, con muestras de otras poblaciones europeas y amerindias.

#### 4. HIPÓTESIS

**H<sub>01</sub>** La población mestiza es predominantemente indígena para los marcadores mitocondriales y europea para los del cromosoma Y, mientras que estos dos aportes están balanceados para los marcadores autosómicos.

**H<sub>1</sub>**: No existen diferencias en el aporte de las poblaciones fundadoras, entre las muestras indígena y mestiza analizadas con relación a los marcadores utilizados.

**H<sub>02</sub>**: La muestra indígena presenta una estructura poblacional propia, posiblemente debida al contacto con las poblaciones mestizas vecinas, consecuentemente el grado de mestizaje a nivel de los marcadores de la región NRY es mayor que el de los marcadores mDNA o AIMs.

**H<sub>2</sub>**: No existen diferencias en la estructura poblacional entre las poblaciones mestizas y las vecinas nativas con relación a los marcadores utilizados tipo NRY.



## 5. MARCO TEÓRICO Y ANTECEDENTES

### 5.1 GENÉTICA MOLECULAR Y EL ESTUDIO DE LAS POBLACIONES HUMANAS

**5.1.1 Estudios genómicos amplios.** Once años antes de la obtención del código genético humano, ya existía la propuesta acerca de cómo estudiar la variación individual mediante análisis genómicos extensos (Cavalli-Sforza, 1990). Sin embargo, sólo posteriormente a la publicación de la secuencia del genoma humano (Venter et al., 2001) se empezaron a generar nuevas iniciativas y proyectos a gran escala con el propósito de entender la diversidad genética de las poblaciones humanas, posibilitando así el acceso a datos genotípicos densos, provenientes de diversos sistemas de polimorfismos (SNPs autosómicos, microsatélites, CNVs) (Cavalli-Sforza, 2005; Galanter et al., 2012; The International HapMap consortium, 2007).

Toda la información generada en estos proyectos, hoy en día está disponible en diferentes bases de datos de libre acceso y puede ser empleada para análisis comparativos intra e interpoblacionales (Hinds et al., 2005; Rosenberg et al., 2002), estudios de asociaciones amplias en el genoma y, recientemente, asociaciones basadas en la ancestría genética, para facilitar las aplicaciones médicas, entre otras (Cogent Study, 2008; Kosoy et al., 2009).

Estas bases de datos (HapMap, 1000 genomas, CEPH, SPSmart, etc.) cuentan con información amplia acerca de la variabilidad genética humana, cuyo análisis puede brindar importantes señales acerca de la evolución temprana del humano anatómicamente moderno, y de los procesos demográficos recientes que acompañaron la radiación global de *Homo sapiens*.

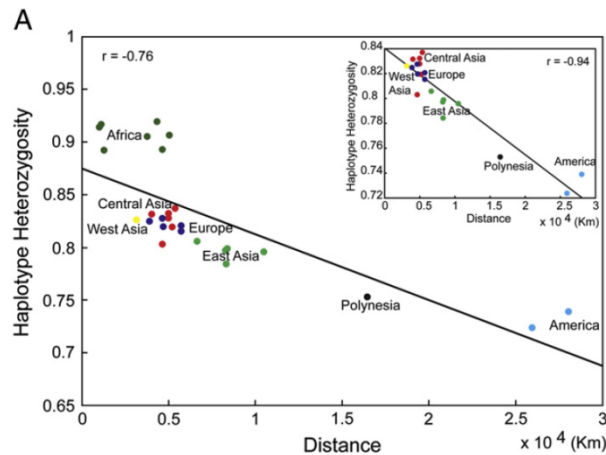
**5.1.2 Patrones globales de diversidad.** La baja diversidad genética poblacional humana actual, en comparación con otras especies, es consecuencia del origen reciente y el tamaño pequeño de la población ancestral que emigró desde el continente africano, hace aproximadamente 100 mil años (M. Campbell & Tishkoff, 2008).

Como consecuencia de este primer cuello de botella, el resto de las poblaciones del mundo originadas (no-africanas) parecen tener un subconjunto de la diversidad genética presente en África subsahariana (Liu, Prugnolle, Manica, & Balloux, 2006). Existe una correlación positiva entre la distancia geográfica y los valores de heterocigocidades esperados, para la cual se observa que los valores mayores se localizan en África y van decreciendo gradualmente a lo largo de las poblaciones localizadas en Asia, Oceanía y América (Ramachandran et al., 2005) (Figura 1).

Las poblaciones africanas, a nivel global, cuando han sido comparadas con las no-africanas, en la mayoría de los estudios de variación genética y fenotípica, en los

que se han empleado diversos tipos de marcadores para las regiones autosómicas, el cromosoma Y y el mDNA, han mostrado altos niveles de variación, debidos en gran medida a su compleja historia poblacional, variación extrema en climas, dieta y exposición a una amplia gama de enfermedades infecciosas (Jorde, Watkins, Bamshad, Dixon, & Ricker, 2000; Prugnolle, Manica, & Balloux, 2005).

**Figura 1.** Patron de diversidad global.



Fuente: Xing, et al 2010.

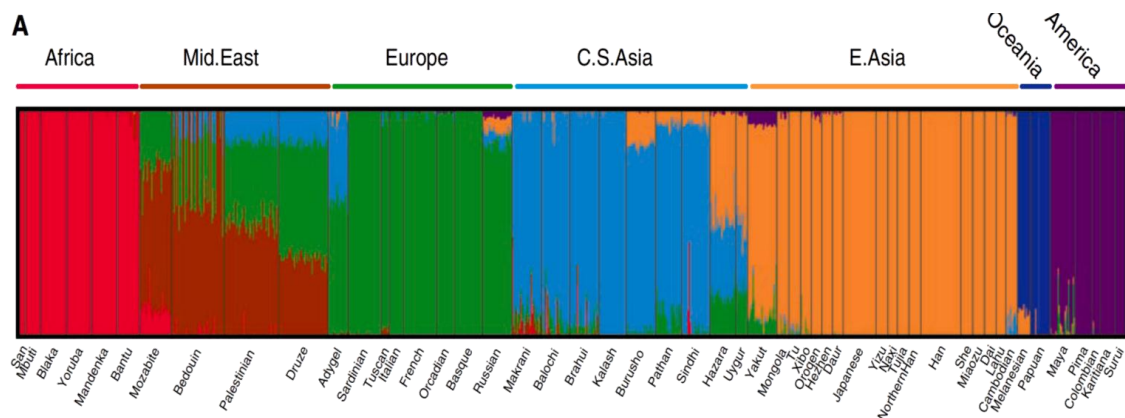
\* Diversidad haplotípica en función de la distancia geográfica de las poblaciones con relación al este de África. La correlación se incrementa cuando las poblaciones africanas se excluyen ( $r = -0.94$ ).

**5.1.3 Estructura genética de la población humana.** Mientras los estudios basados en análisis de patrones en las frecuencias de SNPs autosómicos, pueden detectar estructuración poblacional a diferentes escalas, los sistemas genéticos uniparentales como el mDNA de herencia materna y el cromosoma Y de origen paterno (específicamente la región no recombinante NRY) al carecer de recombinación, son determinantes al momento de definir genealogías, establecer estructuración patri o matrilocal (M. Campbell & Tishkoff, 2008; Wilder, Kingan, Mobasher, Pilkington, & Hammer, 2004) y patrones de simetría en la historia demográfica de las poblaciones, con relación a las proporciones del sexo en los grupos humanos colonizadores o migrantes (Jobling, Hurles, & Tyler, 2004).

Los estudios de estructura genética humana a nivel global se han abordado mediante micosatélites distribuidos en todo el genoma (Rosenberg et al., 2002) y de SNPs autosómicos que difieren en frecuencia entre poblaciones continentales

(Li et al., 2008). Estos estudios han encontrado una correlación genética entre la afiliación geográfica de las personas y su patrón de frecuencias alélicas, en varios niveles, por ejemplo a lo largo de grandes regiones geográficas o dentro de las subpoblaciones que las constituyen (figura ) (Li et al., 2008; Rosenberg et al., 2002).

**Figura 2.** Estructura genética poblacional global.



Fuente: Li, et al 2008.

\* S emplearon datos genotípicos de SNPs autosómicos en 934 individuos de 51 poblaciones alrededor del mundo. La gráfica fue obtenida mediante el análisis en el programa *frappe*, que aplica los métodos de máxima verosimilitud para asignar, a nivel individual un coeficiente de afiliación en K grupos o “clusters” predefinidos. Las etiquetas poblacionales se adicionaron posteriormente al cálculo de dicho coeficiente y se emplearon para ordenar los individuos.

**5.1.4 Estructura genética de las poblaciones latinoamericanas.** La estructura genética del continente americano esta determinada por la heterogeneidad en sus poblaciones, debido a la historia demográfica. El contacto europeo, causó casi el colapso total de la población nativa, muchos grupos enteros desaparecieron, ocurrió fusión de etnias y mezcla genética con grupos no-indígenas (Mulligan, Hunley, Cole, & Long, 2004), y emergieron las poblaciones latinoamericanas modernas, con tres orígenes ancestrales, el nativo, el europeo y el africano, (Salzano & Bortolini, 2002). Los estudios han empleado los haplogrupos mitocondriales, del cromosoma Y y los SNPs autosómicos (AIMs).

**5.1.4.1 DNA mitocondrial.** Actualmente, el mDNA de las poblaciones nativas de Latinoamérica poseen los cuatro Hgs fundadores: A, B, C y D (Schurr, 2004a), también han sido observados en muestras indígenas antiguas, constituyendo por

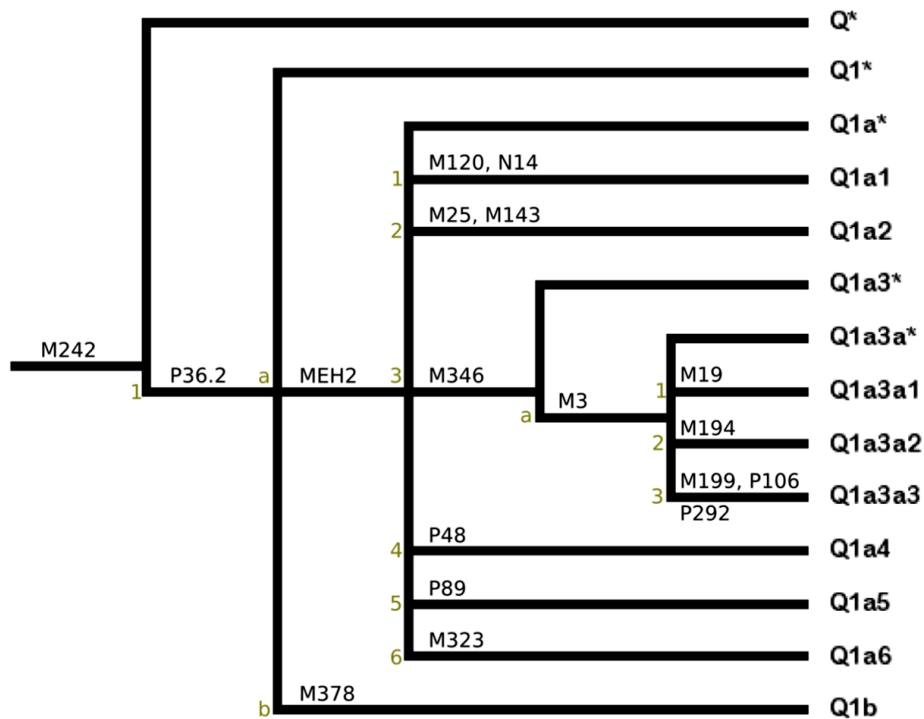
lo tanto los linajes mitocondriales fundadores en las poblaciones nativas americanas (Torroni et al., 1993). Cada uno se distingue por una única combinación de RFLPs y los polimorfismos de la región hipervariable I (HVR-I). En conjunto, estos polimorfismos comprenden del 96-100% de los Hgs mitocondriales en las poblaciones indígenas modernas del Nuevo Mundo (Bisso-Machado et al., 2012) lo cual puede ser un reflejo de los patrones de frecuencias en los momentos previos al contacto con los europeos (O'Rourke, Hayes, & Carlyle, 2000; Raff, Bolnick, Tackney, & O'Rourke, 2011).

Con las migraciones durante los periodos coloniales y postcoloniales, surgieron nuevas poblaciones no-indígenas o mezcladas, y se incorporaron nuevos Hgs en la región. En países como México, El Salvador, Panama, Venezuela y Colombia, predominan los Hgs nativos (80-93%) (Gómez-Carballeda et al., 2012; Guardado-Estrada et al., 2009; Perego et al., 2012; Salas et al., 2009), por su lado, en Brasil o Cuba este componente ancestral es menor (~33%), y aparecen los linajes africanos (28-45%) (Por ejemplo el Hg L) o europeo (22-39%) (Por ejemplo el Hg H) (Alves-Silva et al., 2000; Mendizabal et al., 2008). En la población argentina posee un perfil distinto, con un componente nativo del 66%, europeo del 39% y africano del 2% (Catelli et al., 2011).

**5.1.4.2** Región no recombinante del cromosoma Y. Contiene 20 Hgs principales (identificados desde el A, hasta el T), construidos sobre la base de 600 mutaciones (SNPs) que dan lugar a los 311 subHgs de la región NRY (Karafet et al., 2008; The Y Chromosome Consortium, 2002).

Los cromosomas Y nativos suramericanos derivan de submuestras de aquellos presentes en Siberia (Bortolini et al., 2003; Mesa et al., 2000; Underhill, Jin, Zemans, Oefner, & Cavalli, 1996), que incluyen los 4 subHgs Q1a3a y los paragrupos Q1a3\*, Q1a3a\* y Q1a\* (Figura 3). Constituyen más del 80% de los cromosomas Y indígenas de América, con altas frecuencias en poblaciones suramericanas con escasa mezcla genética; dentro de los cuales Q1a3a, posee la distribución más amplia (Bortolini et al., 2003; Jota et al., 2011; Sandoval et al., 2012; Santos, Bonatto, & Bortolini, 2007; Underhill et al., 1997; Underhill et al., 1996).

**Figura 3.** Árbol correspondiente al clado Q del cromosoma Y.



Fuente: Karafet, *et al* 2008.

\* Los nombres de las mutaciones se indican sobre las ramas y los subclados se señalan con números o letras al principio de la rama que los define, de acuerdo con la nomenclatura establecida por el Consorcio para el Cromosoma Y. Los nombres de cada subHg se indican a la derecha, al finalizar las ramas del árbol.

La mayoría de los individuos que pertenecen al Hg Q -portadores de la mutación derivada M242-, pero no a Q1a3a -alelo ancestral en M3-, actualmente pertenecen al paragrupo Q1a3\*, ya que además presentan la mutación derivada en M346 (Figura 3) (Bailliet et al., 2009; Karafet et al., 2008; Sengupta et al., 2006). Es posible encontrar otros linajes, como los europeos (Hgs: R1b1b2, I, J, G, K) o el africano (Hg E), sin embargo su presencia se puede explicar por la mezcla reciente (Bailliet et al., 2009; Blanco-Verea, Jaime, Brión, & Carracedo, 2010; González-José, Bortolini, Santos, & Bonatto, 2008; Schurr & Sherry, 2004b; Toscanini et al., 2011). En las poblaciones mestizas, de países como México, Nicaragua, Colombia o Argentina, el resultado del contacto europeo en periodos pre-y post coloniales, explica una alta frecuencia de los cromosomas Y de origen europeo, principalmente los del subclado R1b1 (G. Bedoya et al., 2006; Carvajal-Carmona et al., 2003; Corach et al., 2010; Martínez-Cortés et al., 2012; Núñez, Geppert, Baeta, Roewer, & Martínez-Jarreta, 2012), común en América Latina, definido por el estado derivado en la mutación P25 (Jobling et al., 2004; Karafet et al., 2008).

**5.1.4.3** Marcadores uniparentales y asimetría sexual en la mezcla. En las poblaciones mezcladas, la ancestría europea es encontrada en frecuencias más altas entre los cromosomas Y que en el mDNA, y por otro lado, las ancestrías nativa amerindia y africana en menor medida, se encuentran en frecuencias más altas a nivel del DNA mitocondrial que en el cromosoma Y. Este es el patrón de asimetría sexual que muestra la mayoría de las poblaciones en Sur y centro América, por ejemplo en Colombia, Costa Rica (G. Bedoya et al., 2006; Carvajal-Carmona et al., 2000) Ecuador (González-Andrade, Sánchez, González-Solórzano, Gascón, & Martínez-Jarreta, 2007), el cual indica un emparejamiento entre mujeres Nativas y hombres inmigrantes que establecieron la población neo-Americana. Concomitante con el colapso demográfico de los amerindios al primer siglo de contacto, este proceso parece haber conducido a la absorción (por la vía materna) de las poblaciones Nativas dentro de la población Criolla (Jobling et al., 2004).

En conjunto, los estudios del mDNA y del cromosoma Y indican que las poblaciones nativas americanas actuales se originaron a partir de una pequeña población protoamericana establecida por migraciones humanas procedentes del centro de Siberia, que se habrían ubicado en la región de Bering, posteriormente distribuyéndose a lo largo del continente americano, siguiendo diferentes rutas hace 15 o 20 mil años (Dulik et al., 2012; Kitchen, Miyamoto, & Mulligan, 2008; Long & Bortolini, 2011; O'Rourke & Raff, 2010; Reich et al., 2012). Por otro lado se ha encontrado que los hombres migran menos que las mujeres (Seielstad, Minch, & Cavalli, 1998), lo cual no siempre es evidente a primera vista, ya que en muchas sociedades un hombre realiza más viajes que una mujer; por ejemplo entre los pigmeos africanos, el rango de exploración de un hombre es 1.8 veces mayor que el de una mujer (Hewlett, van de Koppel, & Cavalli-Sforza, 1982).

**5.1.4.4** SNPs autosómicos (AIMs). Los AIMs son SNPs bialélicos que muestran grandes diferencias en las frecuencias entre las diferentes poblaciones humanas, y por lo tanto tienen valores  $F_{st}$  muy altos (Mao et al., 2007) debido al aislamiento ocurrido por periodos de tiempo prolongados (Clark et al., 2003; Shriver et al., 2003). Estos pueden ser empleados para estimar la ancestría biogeográfica al nivel de población, subgrupo (casos-control) e individual (Gayden, Regueiro, Martinez, Cadenas, & Herrera, 2008) así como también detectar estructuración poblacional (Galanter et al., 2012). Los estudios realizados en las poblaciones nativas de Norteamérica (Cheyenne y Pima-Arizona), Centroamérica (Maya, Pima-Mex) y Suramérica (Surui, Karitiana, Guhibia, Ticuna y Quechua), se han corroborado los patrones de diversidad observados con los marcadores uniparentales, una reducción en la diversidad en dirección norte-sur, a lo largo del continente americano, lo cual explica los patrones de migración y poblamiento. No se encontró estructuración regional (Kidd et al., 2011; Reich et al., 2012).

En la actualidad las poblaciones de LA no son homogéneas y por tanto existe una notable estructuración poblacional, con grupos en los que se puede rastrear la

ancestría a diferentes poblaciones fuente, cada una de las cuales con diferentes grados de desarrollo (Bryc et al., 2010; Sans, 2000; S. Wang et al., 2008). Por ejemplo la población Brasileña actual se formó a partir de 2.4 millones de nativos, de varias olas migratorias europeas (6.5 millones) y la traída de Africanos, ambos procesos migratorios en diferentes períodos (Jobling et al., 2004; Sans, 2000).

La densidad de nativos amerindios está relacionada con las proporciones actuales de mezcla, por ejemplo las poblaciones Mexicanas tienen cerca del 50% de ancestría amerindia, dada la alta densidad inicial, mientras que poblaciones como Río Grande del Sur en Brasil, el Valle Central de Costa Rica y Argentina poseen alta ancestría europea dada la baja densidad amerindia inicial y migraciones europeas adicionales a las del periodo colonial (Corach et al., 2010; Sans, 2000; S. Wang et al., 2008). Varias poblaciones latinoamericanas, como en México, Bolivia, Argentina, Puerto Rico, Colombia Chile, entre otras, han sido abordadas empleando los AIMs para aclarar su estado ancestral y grado de mezcla (Corach et al., 2010; Galanter et al., 2012; Martínez-Fierro et al., 2009; Watkins et al., 2012). Por ejemplo, para la población mexicana la contribución ancestral amerindia está entre un 50 - 55% y la africana es menor del 5% (Collins et al., 2004; Silva-Zolezzi et al., 2009), en contraste con Puerto Rico, donde el componente Amerindio es cerca del 12%, por su parte, el africano es >20% (Tang et al., 2007; Yang et al., 2005). En el caso de la población argentina, la ancestría amerindia está alrededor del 17 al 20% y la africana no supera el 5% (Corach et al., 2010; Seldin et al., 2007). Las poblaciones andinas en Suramérica poseen un patrón similar a las mexicanas. Estos estudios por lo tanto han mostrado diferencias en las proporciones de mezcla a lo largo de América Latina, lo cual está relacionado con las condiciones iniciales que han tenido una profunda incidencia en la composición genética actual de la región, tales como la densidad indígena inicial, la densidad de inmigrantes europeos en diferentes periodos y el traslado de Africanos desde sus colonias en África occidental hacia puertos principalmente localizados en las costas del Atlántico.

**5.1.5 Estructura y diversidad genética de las poblaciones amerindias y mestizas colombianas.** Uno de los primeros intentos por dilucidar la ancestría de la población mestiza analizó el componente africano caracterizando los principales haplotipos mitocondriales amerindios nativos, europeos ibéricos y africanos del Occidente (Salas et al., 2008). Un estudio adicional analizó la variación de los RFLPs mitocondriales solo en las poblaciones amerindias (Rodas, Gelvez, & Keyeux, 2003). Se ha estudiado la composición genética de una muestra descrita como hispanos de Antioquia, caracterizándose la asimetría sexual en la mezcla amerindio/europeo, tema referenciado en apartados anteriores (G. Bedoya et al., 2006). En estas investigaciones se compararon las frecuencias de los Hgs del mDNA de antioqueños con 5 poblaciones nativas: la Ticuna, la Ingano, la Wayuu, la Zenu y la Embera (Carvajal-Carmona et al., 2003; Carvajal-Carmona et al., 2000). Tomando los datos totales, el linaje predominante fue el Hg A y se puede observar que Antioquia muestra una mezcla alta al nivel del mDNA.

Recientemente se analizó la ancestría al nivel del mDNA en poblaciones mestizas del Sur Occidente colombiano (en los departamentos del Cauca y Valle del Cauca). El 97% de la muestra portó Hgs nativos americanos, en donde el 40% pertenecían al Hg A, 13% al B, 34% al C y un 7% al D. Dos haplotipos fueron de origen europeo, mientras que sólo uno perteneció al Hg que caracteriza a los africanos subsaharianos L2. Los afrocolombianos fueron predominantemente del Hg L (72.6%), pero el componente Amerindio es significativamente alto (23.2%), representado por el Hg C principalmente (11%). De este estudio se deduce que el componente nativo amerindio de la población colombiana analizada está más estrechamente relacionado con Centro América y el Norte de Sur América (Mesa et al., 2000).

Otros estudios han reportado estimados de mezcla mediante el uso de los haplotipos del mDNA, Y-STRs y micro satélites autosómicos 5 poblaciones amerindias colombianas -Ticuna, Ingano, Embera, Wayuu y Zenu-, los resultados de este trabajo sugieren que ha ocurrido una diferencia importante en las tasas de migración de hombres y mujeres a lo largo de la historia de las poblaciones amerindias de Sur América, indicando que no ha ocurrido dominancia matri o patrilocal (Salas et al., 2008).

De hecho, un estudio en 34 poblaciones amerindias de Sur América encontró que la mitad de estas eran predominantemente patrilocales y la otra mitad matrilocales, lo cual indica que hay una concordancia entre el tipo de organización social y el nivel de estructuración genética detectada con diferentes tipos de marcadores (Mesa et al., 2000). Los análisis de mezcla en las 5 poblaciones antes mencionadas sugieren un patrón direccional de emparejamiento que involucran principalmente hombres inmigrantes y mujeres nativas, no obstante sería necesario ampliar el tamaño de la muestra para estimados más precisos. Este patrón ha sido ya descrito también para otras poblaciones mexicanas y colombianas (Seielstad et al., 1998), dado el contexto histórico de la formación de las poblaciones modernas de LA, ya discutido.

Uno de los estudios más completos acerca de la diversidad haplogrupal mitocondrial en poblaciones Amerindias en Colombia, analizó 681 individuos pertenecientes a 25 poblaciones diferentes. Este estudio las distribuyó en tres grupos, con base en la frecuencia del linaje principal. Al calcular las frecuencias globales, los amerindios colombianos muestran esta distribución de Hgs: A (31.3%), B (30.5%), y C (28.5%), y una baja frecuencia del D (6.6%) (Keyeux, Rodas, Gelvez, & Carter, 2002). El anterior estudio planteó que los amerindios colombianos se pueden distribuir en dos grandes grupos, separados por el límite geográfico que constituye la cordillera oriental. En este sentido, se pueden encontrar aquellos que son geográfica, lingüística y genéticamente más afines a los grupos amerindios Centro Americanos y aquellos cercanos a los de Sur América, lo cual también se ha corroborado al nivel de marcadores autosómicos.



En Colombia, el conocimiento de la estructura genética de las poblaciones del centro occidente sigue siendo escaso, a pesar de que se han realizado trabajos encaminados a establecer las frecuencias alélicas y haplotípicas de RFLPs del mDNA en comunidades Indígenas como los Pijao (Rondon et al., 2008; Rondon et al., 1999) y afrodescendientes (Melton et al., 2007). En los reportes recientes para los Pijaos, se evaluó su estructura y diversidad genética caracterizando los haplotipos mitocondriales y 12 marcadores STRs autosómicos. La muestra se distribuyó entre los Hgs mitocondriales A, B y C (Rodas et al., 2003).

En cuanto al cromosoma Y, algunos pocos estudios analizan la mezcla amerindia, europeo y africano empleando marcadores bialélicos de NRY que pueden distinguir entre linajes paternos amerindios, europeos y africanos. Se realizó un estudio de mezcla en una población mezclada de Antioquia, a nivel de marcadores bialélicos de NRY. De los 80 cromosomas Y analizados, tres son probablemente de origen Africano (Hg E); y un individuo de origen amerindio (Hg Q). El polimorfismo en SRY2627, actualmente M167, que define el Hg R característico de la península Ibérica, se encontró en una mayor frecuencia entre los cromosomas Y analizados, indicando el origen ibérico de los europeos inmigrantes hacia Antioquia (Carvajal-Carmona et al., 2000). En el mismo estudio, las poblaciones indígenas Kuna, Wounana, Embera, y Wayuu fueron tipificadas para el SNP M3 (DYS199), encontrándose que toda la muestra estuvo caracterizada por el alelo T (mutación derivada), ausente en Australasia, África y Europa, por lo tanto característico de las poblaciones nativas americanas (Carvajal-Carmona et al., 2000). Otros estudios en poblaciones Wayuu, Ingano, Ticuna, Yagua y Zenu; analizados mediante los SNPs que determinan los Hgs nativos americanos, europeos y africanos. Se encontró un grado de mezcla africana y europea en baja frecuencia. El Hg predominante en los amerindios analizados fue el Q, determinado por la presencia del alelo T en M3. Otros haplotipos intermediarios (Ruiz-linares, 1999).

Con relación a la mezcla autosómica, existen dos reportes importantes. Un estudio realizado con microsatélites indicó que poblaciones de Cundinamarca y Antioquia poseen un aporte europeo entre el 55 y 65%, y en pasto fue mayor el componente indígena (>50%) (S. Wang et al., 2008). Con SNPs autosómicos en Antioquia, los componentes nativo y europeo están en similar proporción (Galanter et al., 2012). Empleando SNPs autosómicos, se ha analizado la composición ancestral en diferentes grupos étnicos como los Kogui, Arhuaco, Waunana, Embera, Zenu e Inga, pero sólo en los arhuacos se ha encontrado un aporte europeo apreciable (S. Wang et al., 2007). Otro estudio, más reciente ha detectado un componente europeo en los wayuu, pastos, awa y pijaos (Coyaima) (Galanter et al., 2012).

El segundo estudio fue realizado con 24 poblaciones distribuidas en la región Andina, la costa Caribe, la Pacífica y algunas de la Orinoquia. Para este estudio se analizaron marcadores del mDNA, de la región NRY, del cromosoma X y algunos genes autosómicos para evaluar la mezcla. Se encontró que la población

colombiana tienen predominantemente Hgs fundadores nativos (>90%) y desde el punto de vista del cromosoma Y, es principalmente no-nativo (africano o euroasiático). Es de destacar que en toda la muestra predominaron los Hgs mitocondriales A y B, por su lado, el Hg L (africano) en Chocó (68%). En cuanto al cromosoma Y, en general predominaron los Hgs europeos, seguido de los nativos, y, nuevamente en las costas se observó una alta frecuencia del Hg E africano. Las regiones Pacífico y Caribe poseen un enriquecimiento de ancestría autosómica africana alta (Chocó: 68%, Valle: 22% y Cauca: 24%, Bolívar: 44% y Magdalena del 28%). Por otra parte, en regiones como Nariño, Cauca, Huila, Casanare y Norte de Santander el aporte nativo es el principal; en otras como Caldas, Antioquia, Santander, Quindío y Magdalena el europeo está en mayor frecuencia. En conjunto, estos resultados demuestran que la mezcla entre nativos, europeos y africanos ha ocurrido en Colombia en diferentes grados y proporciones de acuerdo con la región geográfica analizada y se proponen tres tipos de patrones de mezcla: en el primero ocurrió básicamente una mezcla entre mujeres nativas y hombres europeos en una proporción 1:1, sin posteriores contribuciones. En el segundo modelo de mezcla ocurre una incorporación posterior de componente paterno nativo, que en el tercer modelo corresponde al europeo (Rojas et al., 2010).

El anterior estudio revela que Colombia posee una de las estructuras genéticas más complejas de LA, que puede estar asociada con la variación regional en las historias demográficas, posiblemente influenciada por la distribución diferencial de las agrupaciones étnicas a lo largo del país. La región del Tolima posee una composición étnica reportada por el DANE, en la cual el 94.5% de su población no se identifica con un origen ancestral o étnico particular, mientras que el 4.3% se autodenomina como indígena y el 1.2% se define como afrocolombiano. No se conoce cuál pueda ser la relación de esta distribución con la estructura genética, aunque ya se han realizado estudios en comunidades indígenas de Coyaima (Rondon et al., 2008; Rondon et al., 1999), no se ha vinculado previamente a la población mezclada, por lo que el presente estudio ha sido una oportunidad para analizar una población con una historia precolombina y postcolombina interesante que se ha comparado con otras poblaciones nativas y mestizas colombianas localizadas en Córdoba, que poseen una influencia africana mayor a la del Tolima. Finalmente, un aspecto importante es la posibilidad de emplear los datos generados en estudios posteriores de asociaciones genómicas tales como el mapeo de la mezcla, con el fin de realizar la búsqueda de nuevos genes que se encuentren aportando al riesgo de diversas enfermedades complejas como el cáncer de colon, entre otros.

## 6. METODOLOGÍA

### 6.1 POBLACIONES Y MUESTRAS

**6.1.1** Permiso ético y consentimiento informado. Para obtener las muestras de sangre, necesarias para la realización de los estudios genético-moleculares, se requirió primero la aprobación del proyecto y de los consentimientos informados, ante el comité de bioética de la universidad del Tolima (anexo 1), para garantizar que:

- Tanto las muestras obtenidas, como los resultados derivados de ellas, fueran correctamente manejados, no se comercializarán y se protegerán la privacidad y la seguridad de los participantes en la investigación.
- Los estudios derivados de las muestras colectadas acogen la declaración de Helsinki sobre *Principios éticos para las investigaciones médicas en seres humanos*; la *Declaración de las Naciones Unidas sobre los derechos de los pueblos Indígenas*, y también a la ley 21 de 1991 de la constitución Nacional de Colombia, sobre la ratificación del convenio 169 acerca de los pueblos indígenas y tribales en países independientes.

Se diseñaron dos tipos de consentimientos informados (CI) para las muestras indígenas: el CI individual, firmado por la persona donante de la muestra y el CI comunitario, firmado por el líder de la comunidad indígena. Para los casos de la población general o no-indígena, se empleó un tercer formato de CI. Adicionalmente se diseñó una entrevista para indagar acerca de la historia familiar del donante (anexo 1).

**6.1.2** Contactos con las poblaciones Indígenas y mezcladas. Con el fin de obtener la autorización para poder llevar a cabo el muestreo, se establecieron diálogos con los líderes de comunidades que se auto-definieron como grupos indígenas afiliados a las etnias Nasa, Embera, Pijao o Zenu, en los departamentos del Tolima y Córdoba.

El proceso consistió en dos socializaciones del proyecto y documentación ética, la inicial ante los gobernadores indígenas, y posteriormente una ante las comunidades interesadas. En todos los casos se informó claramente de los objetivos del proyecto y del carácter voluntario de la participación, mediante una exposición y la entrega de un resumen técnico (figura 4).

**Figura 4.** Jornadas de socialización y presentación del proyecto.



Fuente: El autor

Cuando las personas no se afiliaron o auto-identificaron como pertenecientes a un grupo étnico particular, fueron denominados en términos de “población general” para efectos de este estudio. Se consideraron aquellos que habitaban en municipios cercanos a los resguardos o cabildos indígenas participantes. No se requirió una socialización previa y fueron contactados en sitios públicos.

**6.1.3** Recolección de muestras y entrevistas. En el caso de las comunidades indígenas, una vez obtenido los permisos, se procedió a realizar el muestreo siempre en el marco de una reunión en la sede del resguardo o cabildo y con la presencia del líder indígena.

Primero se obtuvo el CI comunitario por parte del Gobernador indígena y posteriormente se tomó la muestra de sangre a cada donante sano y mayor de edad, previa lectura y/o traducción del CI individual. También se aplicó una breve entrevista para indagar acerca de la historia familiar de cada participante. En los mestizos se tomó la muestra en el domicilio del participante o en el centro de salud del municipio, previa lectura y firma del CI, y diligenciamiento de la entrevista.

En todos los casos, la toma de muestras fue realizada por personal capacitado y autorizado, siguiendo las normas de bioseguridad y protocolos establecidos por el grupo. Se tomaron tres muestras de sangre por participante en tubos con K-EDTA de 4ml debidamente rotulados, siempre en personas sanas y mayores de edad.

Una vez extraídas, se homogenizaron 4 veces y se colocaron en hielo hasta su almacenamiento en el congelador. En el caso de participantes menores, solo se obtuvo la muestra cuando una persona responsable del menor lo solicitó y autorizó con la firma del CI (figura 5).

**Figura 5.** Jornadas de muestreo.

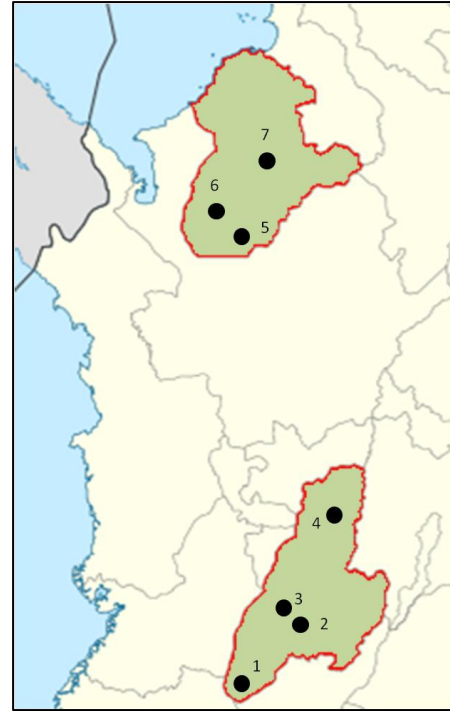


Fuente: el autor

\*La toma de muestras de sangre se llevó a cabo en los diferentes cabildos y resguardos indígenas visitados.

**6.1.4 Poblaciones.** Se tomaron muestras de sangre en siete comunidades indígenas, el resguardo de la etnia Nasa-Paez de Gaitania en Planadas-Tolima, el resguardo Pijao Lomas de Hilarco en Coyaima-Tolima, el resguardo Pijao Guatavita Tua en Ortega-Tolima, la comunidad Embera de Mariquita-Tolima, el cabildo Zenú Porvenir La Fe en Montelíbano-Córdoba, el cabildo Zenu km-35 en Sahagun-Córdooba y la comunidad Embera en Puerto Libertador-Córdoba (figura 6). De igual manera, se tomaron muestras de sangre a personas de la población general, en municipios o localidades cercanas a los resguardos o cabildos indígenas, con el objeto de analizar el flujo genético y efectos de continuidad genética (tabla 1). Todas las muestras se almacenaron en el laboratorio de Citogenética, Filogenia y Evolución de Poblaciones de la Universidad del Tolima.

**Figura 6.** Localización geográfica de los sitios de muestreo.



1. Planadas-Tolima: Indígenas de la Etnia Nasa-Paez en el Resguardo de Gaitania-Tolima. Mestizos del municipio.
2. Coyaima-Tolima: Indígenas de la etnia Pijao-Resguardo Lomas de Hilarco.
3. Ortega-Tolima: Indígenas de la etnia Pijao – Resguardo Guatavita Tua. Mestizos del municipio.
4. Mariquita-Tolima: Indígenas de la etnia Embera.
5. Montelíbano-Córdoba: Indígenas de la etnia Embera y Zenu. Mestizos del Municipio.
6. Puerto Libertador-Córdoba: indígenas de la etnia Embera.
7. Sahagun-Córdoba: Indígenas de la etnia Zenú.

Fuente: el autor.

\*Se visitaron comunidades indígenas y mestizas en los departamentos del Tolima y de Córdoba.

**Tabla 1.** Relación de número total de individuos muestreados.

Etnia	población	Total individuos	Numero efectivo
<b>Naza-Paez</b>	Planadas-Tolima (Gaitania)	178	96
<b>Embera-Katio</b>	Montelíbano y Pindo - Córdoba	34	17
<b>Pijao</b>	Ortega-Tolima (Guatavita)	94	57
<b>Pijao</b>	Coyaima-Tolima (Lomas de Hilarco)	65	25
<b>Zenu</b>	Montelíbano-Cordoba (Porvenir La Fe)	116	82
<b>Zenu</b>	Sahagun-Cordoba (km35)	139	42
<b>Embera</b>	Rio Verde-Cordoba (Puerto Libertador)	26	5
<b>Embera</b>	Mariquita-Tolima	41	7
<b>4</b>	<b>7</b>	<b>749</b>	<b>371</b>

Fuente: El autor.

\* Se aprecia la etnia, y número efectivo poblacional (es el grupo de individuos no emparentados, basado en la información familiar).

**Tabla 2.** Individuos mestizos colectados.

Población	Poblaciones Indígenas cercanas	número de total de muestras	# efectivo mestizos
Ibagué-Tolima	Pijaos	116	116
Ortega-Tolima	Pijaos	44	44
Planadas-Tolima	Nasa-Paez	101	86
Montelibano-Cordoba	Zenu	84	78
<b>TOTAL</b>		<b>397</b>	<b>376</b>

Fuente: El autor

\* Número total de muestras de sangre de la población general recolectadas en cercanías a los resguardos o cabildos indígenas.

**6.1.4.1** Selección de Individuos no emparentados y grupos familiares. Con base en las entrevistas realizadas en campo, se elaboró una base de datos con la información personal y familiar. De esta manera se establecieron los grupos familiares que componían la muestra de cada población. Posteriormente, se seleccionó un grupo de individuos no emparentados por población, los cuales fueron incluidos para los análisis genético-moleculares.

**6.1.4.2** Poblaciones de referencia. Para los análisis de ancestría genética mediante AIMS, fue necesario contar con datos genotípicos de poblaciones de referencia indígena, europea, africana y asiática, para obtener la estructura genética más probable, en vista de que algunas se contemplan como posibles poblaciones parentales de aquellas analizadas. Los genotipos fueron obtenidos de las bases de datos del proyecto HapMap y SPSmart.

Se emplearon datos genotípicos amerindios (Ami) de 21 individuos zapotecas (México), 22 tepehuanos (México), 14 nahuas (México), 25 mayas (México), 24 Quechua (Perú), 25 Aymara (Bolivia), 22 Indígenas del Norte (Colombia), 19 Indígenas del Centro (Colombia), 36 Indígenas del Sur (Colombia), 20 Panare (Amazonas-Venezuela), 20 Indígenas Pemón (Amazonas- venezuela), 20 Indígenas Warao (Amazonas-Venezuela), 20 indígenas (Venezuela) y 14 indígenas Wichi (Argentina) (datos proporcionados mediante comunicación personal del Doctor Angel Carracedo). Los datos para los europeos (Eur) provinieron de 619 españoles, 44 italianos y 56 individuos ibéricos del Centro de Estudios de Polimorfismos Humanos (CEPH) (Rosenberg et al., 2005). Los africanos (Afr) de 147 individuos Yoruba de Ibadan-Nigeria (YRI) y 110 Luhya (LWK) de Webuye-Kenya. Los asiáticos (As) fueron 100 individuos chinos Han de

Beijin (CHB) y 150 japoneses de Tokio (JPT) (The International HapMap Consortium, 2003). Para comparaciones adicionales, se emplearon 19 mestizos colombianos y 63 afrocolombianos del Valle del Cauca y Choco (datos proporcionados mediante comunicación personal del Doctor Angel Carracedo).

## 6.2 EXTRACCIÓN, CUANTIFICACIÓN Y DILUCIÓN DEL DNA

Una vez se definieron los grupos familiares y el conjunto de individuos no emparentados, se procedió a realizar la extracción y cuantificación de ADN en el laboratorio de Citogenética, Filogenia y Evolución de Poblaciones de la Universidad del Tolima. Para los Nasa (n=96), Pijao (n=52), Embera (n=29), Zenú (n=51), población general de Ibagué (n=116), Ortega (n=44), Planadas (n=47), y Córdoba (n=31). A partir de la muestra de sangre, se extrajo DNA genómico total mediante el equipo automatizado Maxwell® 16 -*Human Whole Blood Genomic DNA* (Promega tested)- empleando el kit Maxwell -16 *blood DNA purification*-(cat. AS1010). La cuantificación se hizo mediante nanodrop ND-2000, se verificó que la concentración de cada muestra estuviera por encima de 10 ng/μl, al igual que la calidad (relación 260/280 entre 1.7 y 1.9). Para la genotipificación de los marcadores, las muestras de DNA se prepararon en platos de 96 pozos a concentración de 20 ng/ μl en el *Wellcome Trust Centre for Human Genetics* de la Universidad de Oxford-UK. Todas las muestras se organizaron en 6 platos (96-Well PCR Plate, THERMO- AB-2396), distribuidos de la siguiente manera: el primer plato contuvo el DNA de los hombres indígenas, el segundo las mujeres, el tercero los grupos familiares indígenas, el cuarto los hombres de la población general, el quinto a las mujeres y el sexto contenía algunas muestras mixtas indígenas y de la población general. Posteriormente el DNA se conservo a -20°C hasta su uso.

## 6.3 MARCADORES GENÉTICOS Y GENOTIPIFICACIÓN

Se emplearon tres sistemas de polimorfismos, 100 SNPs autosómicos o AIMS bialélicos, y dos tipos de marcadores uniparentales: 6 Hgs del DNA mitocondrial y 16 SNPs bialélicos del cromosoma Y (región NRY).

**6.3.1 DNA autosómico (AIMs).** Se seleccionaron 100 SNPs autosómicos bialélicos a partir de un conjunto de 446 que hacen parte del panel LACE, diseñado por la *Fundación Pública Galega de Medicina Xenómica (SERGAS)-CIBERER*, Universidad Santiago de Compostela-España. Este panel proporciona un conjunto de marcadores con capacidad para distinguir entre las tres poblaciones ancestrales de los latinoamericanos, es decir los Ami, Afr y Eur, debido a que presentan amplias diferencias en las frecuencias interpoblacionales (Galanter et al., 2012). Las características del subpanel seleccionado para los análisis del presente proyecto aparecen en la el anexo 2.



**6.3.1.1** Genotipificación de los AIMS. La genotipificación de los AIMS se realizó mediante la plataforma de Sequenom MassARRAY iPLEX (Karafet et al., 2008; Serre & Pääbo, 2004) estandarizada en la *Fundación Pública Galega de Medicina Xenómica (SERGAS)-CIBERER*, Universidad Santiago de Compostela - Galicia, España. Esta plataforma permite dos niveles de especificidad, primero un PCR locus-específico (PCR de amplificación), y segundo, una reacción de extensión de cebadores SNP-específica (ensayo iPLEX), en la cual el primer oligonucleotídico anilla inmediatamente corriente arriba del SNP que se va a genotipificar. En el primer PCR se amplificó la región corta que contiene el SNP (amplicón entre 80 y 100 pb) y en la reacción iPLEX, el amplicon y el primer de extensión fueron incubados con didesoxinucleótidos terminadores, cuya masa estaba modificada (A: 313.0576 amu, G: 329.0526 amu; C: 289.0464 amu; T: 304.0461 amu). Una ventaja de este método es que se pueden amplificar y extender en un mismo volumen de reacción por muestra desde 1 a 40 SNPs simultáneamente.

**6.3.1.2** Diseño de cebadores y ensayos multiplex. Se diseñaron los ensayos multiplex MassEXTENDED *in-silico*, a partir del conjunto de secuencias flanqueantes para cada uno de los 100 SNPs, de manera que se pudiera acoplar la amplificación simultánea de varios SNPs en una misma reacción. Para esto se empleó el software *MassARRAY® Assay Design 3.1*, el cual está disponible en la página web de Sequenom (*My Sequenom-online-tools*: <https://www.mysequenom.com/Logon.aspx?ReturnUrl=%2fTools>

Se agruparon los cebadores en 4 ensayos multiplex. El primer ensayo con 22 SNPs, el segundo con 30 SNPs, el tercer con 29 SNPs y el cuarto con 19 SNPs. Las secuencias de los cebadores de PCR de amplificación y de extensión aparecen en el anexo 3 y 4.

El programa diseñó tres secuencias de cebadores, dos (Forward y Reverse) para la PCR de amplificación de la región flanqueante de cada SNP, y el tercero que se usó para la reacción de extensión alelo específico, mediante la hibridización en la secuencia adyacente al sitio del SNP. Los cebadores de amplificación se solicitaron y llegaron dispensados en platos de 96 pozos a una concentración de 240  $\mu\text{M}$ , posteriormente se diluyeron a 1  $\mu\text{M}$ . Los cebadores de extensión para la reacción iPLEX se ordenaron en platos de 96 pozos a concentraciones entre 250 y 450  $\mu\text{M}$ .

Una vez diseñados los ensayos y cebadores, se transfirieron 2  $\mu\text{l}$  del DNA desde los platos de 96 pozos a platos para PCR de 384 pozos mediante el uso del robot Tecan 8-tip (*freedom Evo*). Se emplearon 6 platos de 384, en cada uno se dispensó por cuadruplicado cada plato de 96 con el fin de llevar a cabo los 4 ensayos para las reacciones de amplificación y extensión.

**6.3.1.3** PCR de amplificación. Una vez listos los 6 platos con el ADN, se mezclaron los *cebadores* de acuerdo con los 4 ensayos diseñados para el PCR de amplificación. Se transfirió, la mezcla de reacción por cada pozo (Tabla 3).

**Tabla 3.** Mezcla de reacción para la PCR de amplificación.

Reactivo	Concentración en 5 $\mu$ l	Volumen/pozo	Volumen para plato de 384
H <sub>2</sub> O ultrapura	NA	1.85 $\mu$ l	888 $\mu$ l
Buffer de PCR 10x	1.25 x	0.625 $\mu$ l	300 $\mu$ l
MgCl <sub>2</sub> 25 mM	1.625 mM	0.325 $\mu$ l	156 $\mu$ l
Mix de dNTPs	500 $\mu$ M	0.100 $\mu$ l	48 $\mu$ l
Mix de Cebadores Forward y Reverse 1 $\mu$ M	100 nM	1.00 $\mu$ l	480 $\mu$ l
DNA polimerasa HotStar Taq plus 5U/ $\mu$ l	0.5 U/pozo	0.100 $\mu$ l	48 $\mu$ l
<b>Total</b>		<b>4 <math>\mu</math>l</b>	<b>1920 <math>\mu</math>l</b>

Fuente: CEGEN.

A Cada plato se le dio un golpe de vortex para mezclar bien y posteriormente una corta centrifugada a 7300 rpm durante 2 minutos para agrupar el volumen total.

La PCR de amplificación se siguió el siguiente programa en un termociclador estándar por cada plato:

- 94° C por 15 minutos
  - 94° C por 20 segundos
  - 56° C por 30 segundos
  - 72° C por 1 minuto
  - 72° C por 3 minutos
  - 4° C final
- } 45 ciclos

- Tratamiento con SAP. A los productos de PCR se les aplicó una reacción de defosforilación de los dNTPs no incorporados, empleando la mezcla de reacción SAP (Tabla 4)

**Tabla 4.** Mezcla para la reacción SAP.

Reactivo	Volumen/pozo	Volumen para plato de 384
H <sub>2</sub> O ultrapura	1.330 $\mu$ l	638.4 $\mu$ l
Buffer de SAP 10x	0.170 $\mu$ l	81.6 $\mu$ l
Enzima SAP 1U/ $\mu$ l	0.5 $\mu$ l	240 $\mu$ l
<b>Total</b>	<b>2.00 <math>\mu</math>l</b>	<b>960.0 <math>\mu</math>l</b>

Fuente: CEGEN

Se adicionaron 2  $\mu$ l de la mezcla de SAP por cada 5  $\mu$ l de reacción de PCR, luego se dió un vortex breve y se incubaron los platos en un termociclador convencional siguiendo el siguiente programa:

- 37° C por 40 minutos
- 85° C por 5 minutos
- 4° C al final

**6.3.1.4 Reacción iPLEX de Extensión.** Debido a la relación inversa entre el pico de intensidad y la masa del analito, se ajustó la concentración de los cebadores de extensión de los ensayos iPLEX, para garantizar que la intensidad de la señal detectada fuera similar para todos. Se aplicó un método general para ajustar la concentración de los cebadores de extensión. Los cebadores se dividieron en dos grupos de dilución basada en la masa, el grupo de mayor peso se diluyó a una concentración final para reacción de 1.625  $\mu\text{M}$ , y los de masa baja a 0.625  $\mu\text{M}$ .

Posteriormente se preparó la mezcla para la reacción iPLEX, como se indica en la tabla 5.

**Tabla 5.** Volúmenes para la mezcla de la reacción iPLEX.

Reactivo	Conc. en 9 $\mu\text{l}$	Volumen/pozo	Volumen para plato de 384
H <sub>2</sub> O ultrapura	NA	0.755 $\mu\text{l}$	362.40 $\mu\text{l}$
Buffer iPLEX plus 10x	0.222x	0.200 $\mu\text{l}$	96.0 $\mu\text{l}$
Mix de terminación iPLEX (incluye los ddNTPs terminadores)	1x	0.200 $\mu\text{l}$	96.0 $\mu\text{l}$
Mix de cebadores (7 $\mu\text{M}$ y 14 $\mu\text{M}$ )*	1.625 $\mu\text{M}$ y 0.625 $\mu\text{M}$	0.804 $\mu\text{l}$	385.92 $\mu\text{l}$
Enzima iPLEX (termosequenasa)	1x	0.041 $\mu\text{l}$	19.68 $\mu\text{l}$
<b>Total</b>		<b>2.00 <math>\mu\text{l}</math></b>	<b>960 <math>\mu\text{l}</math></b>

Fuente: CEGEN.

\*7  $\mu\text{M}$  y 14  $\mu\text{M}$  son las concentraciones de los cebadores de menor y mayor masa.

Se mezcló bien mediante vortex y llevó a centrifuga a 7300 rpm por 2 minutos, y, posteriormente se llevó a termociclador convencional para desarrollar los ciclos de la reacción iPLEX empleando un programa con dos pasos para un total de 200 ciclos cortos, de la siguiente manera:

- 94° C por 30 segundos
  - 94° C por 5 segundos
  - 52° C por 5 segundos
  - 80° C por 5 segundos
  - 72° C por 3 minutos
  - 4° C final
- \* 5 ciclos; \*\* 40 ciclos
- 

El programa se puede guardar en la memoria de un termociclador de la siguiente manera:

- I. 94° C por 30 segundos
- II. 94° C por 5 segundos
- III. 52° C por 5 segundos
- IV. 80° C por 5 segundos

- V. GOTO III, 4 veces más
- VI. GOTO II, 39 veces más
- VII. 72° C por 3 minutos
- VII. 4° C final

El programa empleó dos bucles de ciclado, uno de 5 ciclos que va inmerso en el de 40 ciclos, por esta razón el programa aplicado fue de 200 ciclos.

Las muestras se denaturaron a 94° C, el anillamiento a 52° C por 5 segundos. El ciclo de anillamiento y extensión se repitió 4 veces más para un total de 5 ciclos y luego retornó a la denaturación a 94° C por 5 segundos y nuevamente entró al ciclo de anillamiento-extensión. Los 5 ciclos de anillamiento-extensión con un solo paso de denaturación fueron repetidos 39 veces para un total de 40. Los 40 ciclos con los 5 internos de anillamiento-extensión sumaron en total 200 ciclos (5x40). Un paso final de extensión se realizó a 72° C por 3 minutos, y, finalmente la muestra se conservó a 4° C hasta la limpieza con resina.

- Tratamiento con resina. Posteriormente se tomaron los productos de la reacción iPLEX y se desalinizaron (remoción de iones Na<sup>+</sup>, K<sup>+</sup>, Mg<sup>++</sup>, los cuales pueden generar ruido de fondo en el espectro de masas) para optimizar el análisis de espectrometría de masas. Cada muestra se diluyó con 16 µl de agua y se emplearon 6.0 mg de *espectro clean resina/pozo* (Sequenom: resina catiónica). Después se centrifugó a 4000 rpm para precipitar toda la resina en el fondo de cada pozo.
- Dispensación en los bioarreglos SpectroCHIP®. Una vez se tuvieron los analitos desalinizados, se empleó un nanodispensador robótico para dispensar 25 nl de producto de reacción iPLEX en los bioarreglos de sequenom *SpectroCHIP*, cada uno de 384 puntos (para depositar cada muestra) los cuales poseen una matriz de ácido 3-hidroxipicolínico.
- Análisis de espectrometría de masas MALDI-TOF: Este método se denomina MALDI por las siglas en inglés *Matrix-Assisted Laser Desorption/Ionization* (desorción/ionización láser asistida por matriz) y TOF por el detector de iones que se acopla al MALDI, del inglés *Time-Of-Flight*. Para el análisis se empleó un espectrómetro compacto de masas (Sequenom, manufacturado por Bruker Ins.) y el programa *MassARRAY Workstation* version 3.4, para procesar y analizar el bioarreglo *SpectroCHIP*. El chip se colocó en el compartimento del espectrómetro de masas donde fue expuesto al vacío y cada punto recibió el impacto de un láser para ionizar (ionización suave), evaporar la muestra y transferir las moléculas electrostáticamente al espectrómetro de masas “tiempo de vuelo” donde las muestras fueron separadas de la matriz iónica e individualmente detectadas y analizadas basados en las proporciones masa/carga. Cada chip tardó en ser analizado entre 30 y 60 minutos. El producto final es un conjunto de espectros por cada muestra.

- Análisis estadístico. Los espectros resultantes fueron analizados por medio del programa *Spectro Typer* 4.0.20. Cada muestra debió tener asociados tres picos: los dos homocigotos (alelo1 de masa mayor y el alelo 2 de masa menor) y el heterocigoto (pico de masa inesperada). El programa analizó la tasa de intensidad de los picos en relación con el ruido de fondo para generar los genotipos y respectivos *clusters* empleando los modelos de mezcla Gausiana. Por cada SNP se obtuvo una gráfica que agrupó las muestras con el alelo 1, el alelo 2 y los heterocigotos en un sistema cartesiano, donde los homocigotos aparecen a 0° y 90°, mientras los heterocigotos a 45°.

Al analizar una muestra para un SNP determinado, solo se asignó el genotipo cuando este se agrupaba en los límites de los núcleos definidos por el modelo de *cluster*. No se asignó el genotipo cuando las muestras agruparon en sitios intermedios entre los núcleos, debido a que pudieron estar contaminadas. Igualmente tampoco se asignó cuando estuvieron en la región de baja señal, pues posiblemente no se pudieron amplificar correctamente o bien la reacción de extensión no funcionó. Al final, se produjo una matriz de datos genotípicos con la cual se realizaron los análisis genéticos de ancestría y mezcla.

**6.3.2 DNA mitocondrial (mDNA).** Se tipificaron los Hgs mitocondriales fundadores A, B, C, D, de las poblaciones Amerindias, el Hg L de poblaciones Africanas y el J para las poblaciones europeas de acuerdo con la metodología previamente reportada (Bailliet, Rothhammer, Carnese, Bravi, & Bianchi, 1994; Carvajal-Carmona et al., 2003; Moraga et al., 2000) (Tabla 6). Esta parte de los experimentos se desarrollaron en el *Wellcome Trust Centre for Human Genetics* - Universidad de Oxford- Reino Unido.

En el caso de los 4 Hgs mitocondriales amerindios, el procedimiento se hizo usando la PCR para amplificar en el mDNA, los sitios asociados a las enzimas de restricción *HaeIII*, *HincII*, y *AluI* (*New England Biolabs*), y a la presencia o ausencia de la delección de 9 pb en COII-tRNA<sup>Lys</sup>. Para los Hgs L (Lewontin, 1972) y J (Prugnolle et al., 2005) se emplearon las enzimas *HPAI* y *BstNI* respectivamente (*New England Biolabs*). Los cebadores y variaciones específicas para la PCR están indicados en la tabla 6.

**Tabla 6.** Tipificación de los Hgs mitocondriales.

Hg	secuencia	Enzima de restricción	MgCL <sub>2</sub> (mM)	Tamaño del fragmento amplificado (pb)	Tamaño de los fragmentos de restricción
A	MIL582 5' TGTAGCTTACCTCCTCAAAGC 3' MIH745 5' TTGATGGTGGTGATTAGAGG 3'	+Hae III	1.0	163	82/105
D	MIL5150 5' CCTACTACTATCTCGCACCTG 3' MIH5281 5' GTGAATTCTTCGATAATGGCC 3'	-Alu I	1.5	131	26/105
B	MIL8209 5' CATCGTCCTAGAAATTAATTCC 3' MIH8304 5' CTTTACAGTGGGCTCTAGAGG 3'	Repetición en región V	1.5	95 pb (dos copias repetidas) 86 pb (una copia repetida)	N.A.
C	MIL13209 5' GCCCTTACACAAAATGACATCAA 3' MIH13437 5' GGAGCACATAAATAGTATGGC3'	-Hinc II	1.5	228	100/50
L	MIL3522 5' CTACCATCACCCCTCTACATC 3' MIH3650 5' CACCCTGATCAGAGGATTG 3'	+HapI	1.5	190	90/100
J	MIL13626 5' CCTAACAGGTCAACCTCG 3' MIH 13729 5' CTGCGAATAGGCTTCCGGCT 3'	+BstNI	1.5	110	62/42

Fuente: Bailliet, et al., 1994; Carvajal-Carmona, et al., 2003; Moraga et al., 2000

\* Secuencias de los cebadores, condiciones de la PCR y tamaños de los productos para la amplificación de los Hgs del mDNA A, B, C, D, L y J. La pérdida o ganancia del sitio para la enzima de restricción se denota con un + o -, respectivamente y determina el Hg

Los cebadores se llevaron a 100  $\mu\text{M}$  y almacenados a  $-20^{\circ}\text{C}$  hasta su uso. Para la mezcla de la PCR, se empleó una solución de cebadores a 10  $\mu\text{M}$  cada uno. En cada uno de los 6 platos iniciales se tipificaron los Hgs A, B y D. Se transfirió 1  $\mu\text{l}$  de DNA a platos nuevos de PCR de 96 pozos (96-Well PCR Plate, THERMO- AB-0800) y se le adicionó la mezcla de reacción para el Hg a tipificar (volumen total=25  $\mu\text{l}$ ) (tabla 7). La tipificación de los Hgs C, L y J se realizó solo en aquellos individuos que no expresaron el A, B o D.

**Tabla 7.** Mezcla de reacción de la amplificación de cada Hg mitocondrial.

reactivo	Hgs		
	A	L, J, C	B, D
Buffer	2.5 $\mu\text{l}$	2.5 $\mu\text{l}$	2.5 $\mu\text{l}$
Primer Reverse	2.0 $\mu\text{l}$	0.5 $\mu\text{l}$	2.0 $\mu\text{l}$
Primer Forward	2.0 $\mu\text{l}$	0.5 $\mu\text{l}$	2.0 $\mu\text{l}$
dNTPs	1.0 $\mu\text{l}$	1.0 $\mu\text{l}$	1.0 $\mu\text{l}$
Mg <sup>++</sup>	1.5 $\mu\text{l}$	1.5 $\mu\text{l}$	1.5 $\mu\text{l}$
Taq polimerasa	0.25 $\mu\text{l}$	0.25 $\mu\text{l}$	0.25 $\mu\text{l}$
Solución Q	5.0 $\mu\text{l}$	N.A	N.A
H <sub>2</sub> O	14.75 $\mu\text{l}$	17.75 $\mu\text{l}$	14.75 $\mu\text{l}$

Fuente: Carvajal-Carmona, et al 2003.

Posteriormente, el plato se rotuló y selló con una cinta adhesiva (96-well PCR film AB-0558, THERMO) para PCR. Se llevó al termociclador (96-well MJ-PTC 225) con el siguiente programa (2 horas y 45 minutos aprox):

- 1: 95°C por 5 minutos
- 2: 94 °C por 1 minuto
- 3: 55 °C por 1 minuto
- 4: 72 °C por 1 minuto
- 5: GOTO 2, 35 veces
- 6: 72°C por 10 minutos
- 7: 8°C al finalizar.

Posteriormente se verificó la presencia de las bandas del producto de amplificación, corriendo 5  $\mu\text{l}$  de producto de amplificación con 3  $\mu\text{l}$  de buffer de carga (Orange G) en un gel de agarosa al 1% preparado con bromuro de etidio (10  $\mu\text{l}$  de bromuro a 10 mg/ml por cada 500 ml de gel preparado). El corrido se realizó durante 20 minutos a 120 voltios y 400 mA en cámara de electroforesis horizontal (BIORAD *Mini-Sub Cell GT Cell Tank and Lid #166-4000EDU*). El revelado se realizó en un sistema fotodocumentador (SYNGENE). Para el Hg B, se realizó una electroforesis directa, corriendo 5  $\mu\text{l}$  de producto de amplificación con 3  $\mu\text{l}$  de buffer de carga (Orange G) en un gel de agarosa al 3 % preparado con bromuro de etidio (10  $\mu\text{l}$  de bromuro a 10 mg/  $\mu\text{l}$  por cada 500 ml de gel preparado). El corrido se realizó en 40 minutos, 120 voltios y 400 mA en cámara de electroforesis horizontal (BIORAD *Mini-Sub Cell GT Cell Tank and Lid #166-4000EDU*). El

revelado se realizó en un sistema fotodocumentador (SYNGENE). El Hg se reveló por la ausencia de la repetición. Para los demás Hgs, se realizó el proceso de digestión con las respectivas enzimas. Se transfirieron 15 µl de producto de amplificación en un nuevo plato de 96 pozos y se adicionó por cada pozo 10 µl de mezcla para digestión (volumen total=25 µl) (tabla 8), posteriormente se selló el plato y llevó a incubación durante 3 horas a 37°C.

**Tabla 8.** Mezcla de reacción de la digestión de los amplicones para cada Hg mitocondrial.

reactivo	Hgs		
	A, D	J, C	L
Buffer*	2.5 µl	2.5 µl	2.5 µl
Enzima	0.5 µl	0.5 µl	1.0 µl
BSA	N.A	2.5 µl	N.A
H <sub>2</sub> O	7.0 µl	4.5 µl	6.5 µl
<b>TOTAL</b>	<b>10 µl</b>	<b>10 µl</b>	<b>10 µl</b>

Fuente: Carvajal-Carmona, et al 2003

\*El buffer varió de acuerdo a la enzima de restricción empleada. *Hapl*: Nebbuffer4; *Bstnl*: Nebbuffer 2; *Hincll*: Nebbuffer 3; *HaeIII*: Nebbuffer 2.

Finalmente se revelaron los productos de digestión para establecer el Hg de cada muestra. Se corrieron 5 µl de producto de amplificación con 3 µl de buffer de carga (Orange G) en un gel de agarosa al 3% preparado con bromuro de etidio (10 µl de bromuro a 10 mg/ml por cada 500 ml de gel preparado). El corrido se realizó durante 30 minutos a 120 voltios y 400 mA en cámara de electroforesis horizontal (BIORAD *Mini-Sub Cell GT Cell Tank and Lid* #166-4000EDU).

El revelado se realizó en un sistema foto-documentador (SYNGENE) y los genotipos se asignaron manualmente mediante inspección visual de las fotos del gel. El Hg A se reveló por la presencia del sitio de digestión para *HaeIII*, el C y el D por la ausencia del sitio de corte de las enzimas *Hincll* y *Alul* respectivamente, el L y J por la presencia de los sitios de corte de las enzimas *Hapl* y *Bstnl* respectivamente. Los datos genotípicos para los Hgs mitocondriales se depositaron en una matriz de datos para su posterior análisis.

**6.3.3** Región no recombinante del cromosoma Y (NRY). Se tipificaron 16 SNPs bialélicos que definen algunos linajes del cromosoma Y para agrupar los individuos del estudio de acuerdo con el origen ancestral más probable: amerindio, africano y europeo asumido por la evidencia existente (tabla 9) (Griffiths & Murillo Araujo, 2000; Harris & Hopkinson, 1972; Ramachandran et al., 2005; Relethford & Harding, 2001; The Y Chromosome Consortium, 2002; Xing et al., 2010).



La tipificación se hizo a través de PCR alelo específico - *Kaspar-Bioscience* (<http://www.kbioscience.co.uk/>)- y detectados con un sistema de TaqMan/ABI en RT-PCR. Los experimentos se desarrollaron en el *Wellcome Trust Centre for Human Genetics* - Universidad de Oxford- UK. El proceso siguió una secuencia, primero se identificaron los Hgs o linajes más probables, es decir el Q y el R1 a los cuales pertenecen los subHgs más frecuentes de las poblaciones indígenas (Q1a3a, Q1a3, Q1a3a2\* y Q1a3a3) y europeas (R1b1b2, R1b) respectivamente (figura 7).

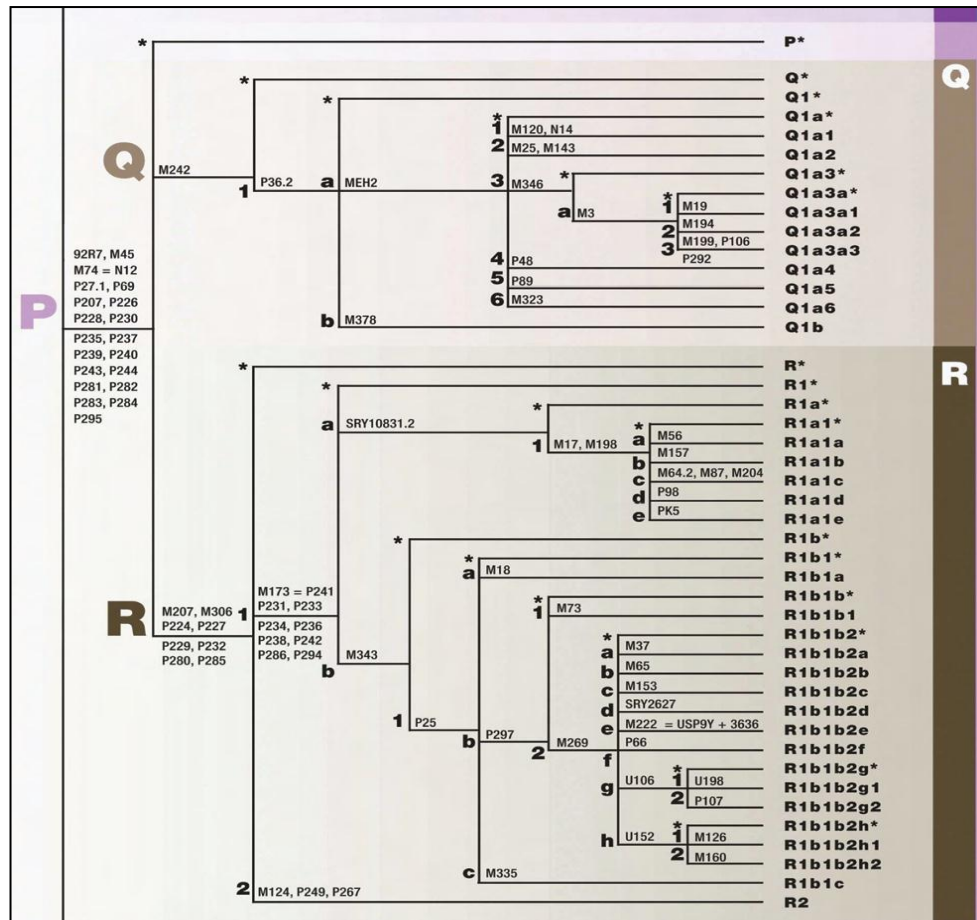
**Tabla 9.** Marcadores de la región no recombinante del cromosoma Y tipificados.

Nombre del Hg	Nombre de la mutación	Código del SNP
E	M96	rs9306841
L	M22	rs3913
K-R	M9	rs3900
P	M45	rs2032631
Q	M242	rs8179021
Q1a3*	M346	NP
Q1a3a*	M3	rs3894
Q1a3a1	M19	rs3910
Q1a3a2	M194	rs2032677
R1	M173	rs2032624
R1b	M343	rs9786184
R1b1	P25	rs150173
R1b1b2	M269	rs9786153
G	M201	rs2032636
I	M170	rs2032597
J	M304	rs13447352
NO	M231	rs9341278

Fuente: Karafet, et al 2008.

\* Se tipificaron Hgs nativos (Q, Q1a3\*, Q1a3a\*, Q1a3a1, Q1a3a2), europeos (R1, R1b, R1b1, R1b1b2, I), euroasiáticos (G, J, K-R), africano (E), indú (L) y polinesio (N-O), en las muestras indígenas y mestizas colombianas. Cada Hg se caracteriza por medio de una mutación (M) y esta a su vez se puede rastrear en las bases de datos mediante el código rs (reference SNP).

**Figura 7.** Fragmento del árbol filogenético del cromosoma Y.



Fuente: Karafet, et al 2008.

\* Se representan las ramas que conducen a los subHgs de los linajes principales Q y R. Cada Hg principal puede ser determinado por una o varias mutaciones y estan indicadas sobre la bifurcación, mediante una nomenclatura que emplea los prefijos M o P. Las mutaciones de los sublinajes estan sobre la rama.

Una vez identificadas todas las muestras con Hgs Q o R1, se separaron en dos platos diferentes para poder tipificar los subHgs de manera eficiente. Aquellas muestras negativas para Q y R1 se colocaron en otro plato para tipificar los Hgs E (africano), I (europeos), G y J (Medio Oriente). Aquellas muestras que durante el proceso no se pudieron asignar a un Hg particular, se dejaron solo hasta el nivel de clado (por ejemplo aquellas que estuvieron en el clado K-R).

Para los análisis de cromosoma Y, se empleó el DNA a una concentración de 20ng/μl. Como control negativo se emplearon muestras de mujeres o la mezcla de

reacción sin ADN. Los controles positivos consistieron en muestras de origen geográfico conocido. El proceso fue el siguiente:

Diseño de Cebadores: Una vez definidos los SNPs del cromosoma Y, se procedió a diseñar los cebadores para emplear en la PCR alelo específico. Para esto se obtuvieron las secuencias de las regiones flanqueantes de cada SNP desde la página del dbSNP y se cargaron en el programa *PrimerPicker Lite* –Kbioscience.

Debido a que cada SNP posee solo dos formas alélicas, el programa generó dos cebadores, el A1 y el A2 para la identificación de cada uno (en posición forward o reverse); de igual forma también generó los cebadores C1 y C2 para escoger el de mejor desempeño (Común 1 y 2, en posición forward o reverse) (anexo 5) por ejemplo así:

ID= rs2032597

SNP1=A

SNP2=C

A1=

GAAGGTGACCAAGTTCATGCTTACTATTTTATTTACTTAAAAATCATTGTTCA

A2=

GAAGGTCGGAGTCAACGGATTACTATTTTATTTACTTAAAAATCATTGTTCC

C1=AGGTCCTCATTTTACAGTGAGACACAA

C2=CAGTGAGACACAACCCACACTGAAA

Dirección=Forward

Secuencia=AAGTTATGTTTTTCATATTCTGTGCATTATACAAATTAATTTTATTT  
ACTTAAAAATCATTGTTCC[**A/C**]TTTTTTTCAGTGTGGGTTGTGTCTCACTGTAAA  
ATGAGGACCTGTTTTTGTGTGGTCTTAAATGTTGAAAGTAATTGG

Los cebadores A1 y A2 se diseñaron con una secuencia que puede interactuar con los fluoróforos *Fam* y *Vic* respectivamente. Los cebadores, se llevaron a una concentración de 100 µM, posteriormente se preparó la mezcla de cebadores a tipificar para cada SNP por separado (A1 +A2+ C1 o C2) con los siguientes volúmenes:

A1= 12 µl

A2= 12 µl

C1= 30 µl

H<sub>2</sub>O= 46 µl

Posteriormente se colocaron 2.0  $\mu$ l de DNA/muestra a una concentración de 20 ng/ $\mu$ l en plates ópticos de 96 pozos (*MicroAmp® Optical 96-Well Reaction Plate with Barcode* cat. # 4306737). Como control negativo se emplearon muestras de mujeres y pozos sin DNA y como control positivo se emplearon muestras de referencia con Hg conocido. Luego se colocó el plate con el DNA en hielo, mientras se preparó la mezcla para la PCR de la siguiente manera:

Mezcla de reacción para la PCR/muestra

Kaspar 2x = 5  $\mu$ l

Mezcla de cebadores= 0.14  $\mu$ l

H<sub>2</sub>O= 3.86  $\mu$ l

La Hotstart Taq, el buffer, los dNTPs, el cloruro de magnesio (50mM), y los fluoroforos Vic y Fam, vienen incluidos en la preparación comercial Kaspar 2X (cat # 1536, KBioscience), la cual se conservó siempre a -20°C o en hielo durante la preparación de la mezcla. Se dispensaron 9  $\mu$ l de mezcla de reacción por muestra (pozo) (volumen total: 11  $\mu$ l) de manera automatizada o manual y posteriormente se cubrió el plate con una película adhesiva óptica (*MicroAmp® Optical Adhesive Film, cat# 4313663*) y se llevo a un termociclador G-Storm GS4, aplicando el programa:

Temperatura de denaturación: 94 °C

Temperatura máxima de anillamiento: 65°C

Temperatura mínima de anillamiento: 57°C

Elongación: 60°C

ciclos: 10

**Hot star:**

Denaturación: 94°C

*hold for:* 15 min

Lid T<sup>o</sup>: 112°C

**star cycle:**

Número de ciclos: 26

Temperatura *step1*

Temperatura procesado: 94°C

sostener por: 20seg

T<sup>o</sup> ramp: max; set max

T<sup>o</sup> increment: 0

time: 0

Temperatura *step2*

Temperatura procesado: 57°C

sostener por: 1 min.

Posteriormente, se realizó la discriminación alélica mediante un sistema de TaqMan/ABI (*Fast Real Time PCR System 7900 HT, Applied Biosystems*) y los resultados de la fluorescencia se interpretaron en el programa SDS 3.3 ABS, a través del cual se asignaron los genotipos manualmente a partir de la visualización de los *cluster* formados. Se generaron los genotipos y exportaron a una matriz de datos para su análisis.

#### 6.4 ANÁLISIS ESTADÍSTICOS

Los genotipos se organizaron en tres matrices independientes, que correspondieron a los datos del cromosoma Y, mDNA, y los AIMs. Posteriormente cada matriz fue organizada en los formatos requeridos por los programas de análisis.

Con relación a los marcadores uniparentales del mDNA y del cromosoma Y, los datos genotípicos se analizaron con el software ARLEQUIN 3.5.1.2. Se calcularon las frecuencias alélicas y haplotípicas con la fórmula de diversidad genética no sesgada de Nei ( $H$ ), que emplea el estimador de la heterocigocidad virtual como medida de la diversidad genética total. Este parámetro es definido como:

$$H = \frac{n}{n-1} \times \left( 1 - \sum_{i=1}^K P_i^2 \right)$$

Donde  $n$  es el tamaño de la muestra analizada,  $K$  es el número de Hgs y  $P_i$  es la frecuencia del  $i_{ésimo}$  Hg en la población.  $H$  genera la probabilidad de que un par de alelos independientes, obtenidos al azar en la población sean diferentes (proporción de heterocigotos entre cigotos imaginarios; en sistemas diploides; es la probabilidad de heterocigosis para un locus dado) y va desde 0 a 1 (Kimura, 1983). El cálculo se basa en las frecuencias alélicas.

Se calcularon los valores de  $F_{st}$  pareados entre poblaciones (incluyendo la significancia), empleando el análisis de la varianza en el programa ARLEQUIN 3.5.1.2. (Excoffier, Laval, & Schneider, 2005). Este  $F_{st}$  maneja algoritmos que tienen en cuenta diferencias en el tamaño de la muestra, así corrige sesgos al comparar poblaciones de diferente tamaño y pondera el tamaño del denominador en los estimativos para cada marcador (Weir & Cockerham, 1984).

Dicho parámetro se calculó así:

$$F_{st} = \frac{\alpha_1}{\alpha_1 + \beta_1}$$

Donde  $\alpha_1$  y  $\beta_1$  son los componentes de la varianza para el  $i_{ésimo}$  locus dentro y entre las poblaciones respectivamente. Estos dos parámetros se calculan así:

$$\alpha_1 = \frac{[(2 \sum n_i \sum (p_{ilu} - p_{lu})^2) - (r - 1)bl]}{2(r - 1)nc}$$

Y

$$\beta_1 = \frac{2 \sum n_i \alpha_{il}}{r(2n - 1)}$$

Donde  $n_i$  es el tamaño de la muestra en cada población,  $P_{ilu}$  y  $P_{lu}$  son las frecuencias del  $i$ ésimo alelo en el  $l$ ésimo locus y su valor esperado a través de las poblaciones,  $r$  es el número de poblaciones y  $\alpha_{il}$  son las diversidades genéticas.

Con la matriz pareada de valores  $F_{st}$  se realizaron análisis de escalamiento multidimensional en el programa *GenALEx* ver. 6.3 (Peakall & Smouse, 2006, 2012) como medio para una representación visual de las distancias genéticas. En conjunto, estos análisis permitieron definir los ancestros paternos y maternos de las poblaciones analizadas, evaluar la presencia asimetría sexual en la mezcla y la diferenciación poblacional.

Los datos genotípicos de los AIMS se organizaron en dos archivos formato de texto plano para el programa Plink, el .ped (ID familia, ID individuo, ID padre, ID madre, sexo y fenotipo en filas y en columnas los genotipos) y el archivo .map (código de los SNPs, cromosoma y posición). En Plink se realizó el control de calidad de los datos genotípicos excluyendo aquellos individuos con menos del 5% de datos y posteriormente se excluyeron los SNPs con menos del 5% de datos.

También se calculó la heterocigocidad esperada y la observada y las frecuencias genotípicas. En *GenALEx* ver. 6.3 se calcularon los valores para la heterocigocidad, las matrices pareadas de valores  $F_{st}$  interpopulacionales (algoritmos de Weir y Cockerham) y las matrices de distancias individuales para los análisis de escalamiento multidimensional (MDS).

Las proporciones de ancestría individual y poblacional se calcularon mediante el uso de los métodos aplicando los algoritmos de agrupación Bayesiana aplicados en el programa *STRUCTURE* v 2.3.3 (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000). Cada análisis de *STRUCTURE* se desarrolló sin asignar previamente una determinada población a cada individuo, con 5 repeticiones por cada K asumido (k=2 hasta k=7) y empleando 40.000 periodos *burn-in* y 80.000 cadenas de Markov. Se empleó el modelo de mezcla y frecuencias correlacionadas mediante la aplicación de la opción  $\alpha$  (un valor  $\alpha$  para cada población), donde  $\alpha$  es el parámetro Dirichlet para el grado de mezcla. Para cada prueba se empleó la opción  $\lambda=1$ , donde  $\lambda$  parametrizó con antelación las frecuencias alélicas y se basó en la distribución de Dirichlet para estas. Se detectó el número de *clusters* más probables empleando el método de Evanno (Evanno, Regnaut, & Goudet, 2005).

Las gráficas fueron desarrolladas en los programas Clumpp y Distruct (Ewens, 1972; Kimura, 1969). Se calcularon también el logaritmo de la verosimilitud (LOD score) en cada análisis para varios grupos poblacionales (k) mediante STRUCTURE.

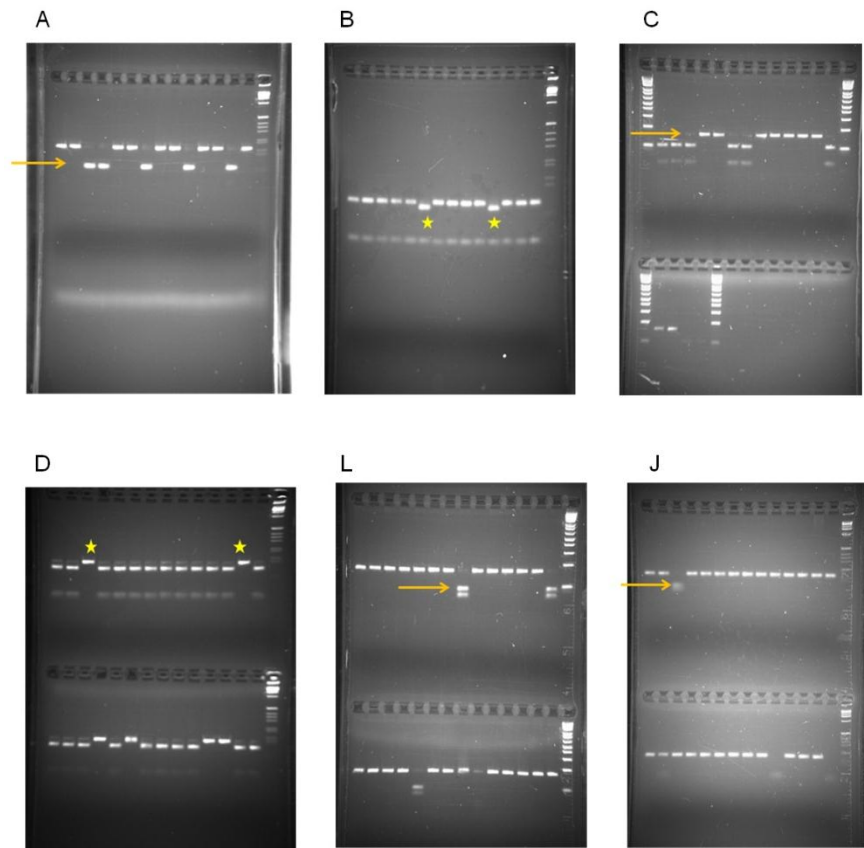
Los datos de las proporciones de ancestría individual fueron analizados aplicando la prueba de *Mann-Whitney* en el programa estadístico INFOSTAT v.2008, con el objetivo de determinar la significancia de las diferencias entre las poblaciones del presente estudio. Adicionalmente se calculó la desviación estándar y se realizó un Box-plot (INFOSTAT v.2008) de las proporciones obtenidas en STRUCTURE por cada población, y de esta manera, establecer el grado de variación de los datos y relacionarlo con la diversidad y la estructuración poblacional.

## 7. RESULTADOS

### 7.1 ANCESTRÍA MATERNA DE LAS POBLACIONES ANALIZADAS

Con el fin de analizar la ancestría materna, se tipificaron los Hgs mitocondriales A, B, C, D L y J en muestras poblacionales de los grupos indígenas Nasa-Paez, Pijao, Embera y Zenu, igualmente, en 4 poblaciones mestizas: Ibagué, Planadas, Ortega (Tolima) y Montelíbano (Córdoba). Los criterios para asignar cada Hg fueron: 1. Presencia o ausencia de los sitios de restricción en el mDNA para las enzimas empleadas, lo cual se refleja en el patrón de bandas. 2. Presencia o ausencia de la delección de 9 pb en COII-tRNA<sup>Lys</sup> que determina el Hg B (figura 8).

**Figura 8.** Tipificación de las muestras indígenas y mestizas para los Hgs mitocondriales mediante RFLPs.



Fuente: el autor.

\* La definición de los Hgs A, C, D, L y J se realizó por medio de patrones electroforéticos de los amplicones previamente digeridos con las enzimas de



restricción *HaeIII+*, *HincIII-*, *AluI-*, *HapI+* y *BstNI+* respectivamente. En los casos anteriores, el Hg está dado por la presencia (+) o no (-) del sitio de digestión, mientras que para el Hg B la determinación fue directamente en el amplicón. En cada caso, los geles de agarosa se prepararon al 3% con bromuro de etido, el corrido fue de 40 minutos a 120 voltios y 400 mA empleando el marcador de peso molecular 1kb, y se visualizaron en un sistema fotodocumentador. La asignación del Hg en cada muestra se realizó mediante inspección visual de la foto del gel, revisando el tamaño de banda esperada en pares de bases (pb): 82/105 (A); 86 pb -una copia- (B); 228 (C); 131 (D); 90/100 (L) y 62/42 (J). Las muestras positivas para cada Hg aparecen señaladas con una flecha o estrella amarilla.

Las poblaciones indígenas del estudio pueden ser distribuidas en dos grupos: aquellas que hacen parte de la región del Tolima (Nasa y piajo), objeto de estudio y dos poblaciones -Embera y Zenu-, localizadas en Córdoba; las cuales permiten realizar comparaciones concernientes a la localización geográfica e historia demográfica. En el caso particular de los Embera, aunque la muestra poblacional está constituida por dos submuestras, una que proviene de Córdoba (Katio) y otra del Norte del Tolima (Chamí), estos últimos son de reciente establecimiento allí, pues como consecuencia del desplazamiento, proceden de los departamentos de Antioquia y Córdoba. Ambas submuestras son equivalentes por provenir de una misma etnia y pertenecer al mismo rango de dispersión geográfica y familia lingüística.

De acuerdo con la tabla 10 y figura 9, tanto en las poblaciones indígenas como en las mestizas están representados en mayor proporción los Hgs nativos americanos típicos (60-98%). Otros Hgs como el africano y europeo se reflejan más en las poblaciones mestizas o en los zenues. En cuanto a la diversidad, se pueden notar tres grupos, los indígenas Embera y Nasa con las diversidades más bajas (0.55 y 0.67 respectivamente), seguido por los mestizos del Tolima que presentaron una diversidad intermedia y los zenues y mestizos de Córdoba con las diversidades más altas (0.81 y 0.80 respectivamente).

En cuanto a la distribución de los Hgs, el D fue el menos frecuente en todas las poblaciones, oscilando entre un 6 y 13%, en tanto que el A fue uno de los más frecuentes y constante (entre 20 y 40%). El Hg B fue el segundo más frecuente y constante entre los mestizos (26 - 34%), por su parte, el C lo fue más en los indígenas (20-32%) (Mestizos: 11-29%). En los Embera, por un lado es notable la ausencia de los Hgs C y D, y por otro el Hg B es el más alto (58.6%), y a su vez esta fue la población con la diversidad más baja.

El componente haplogrupal desconocido es similar entre las poblaciones Zenu y mestizos de Córdoba, incluso son los valores más altos (20 y 26% respectivamente), seguido por los mestizos de Ibagué (8.7%).

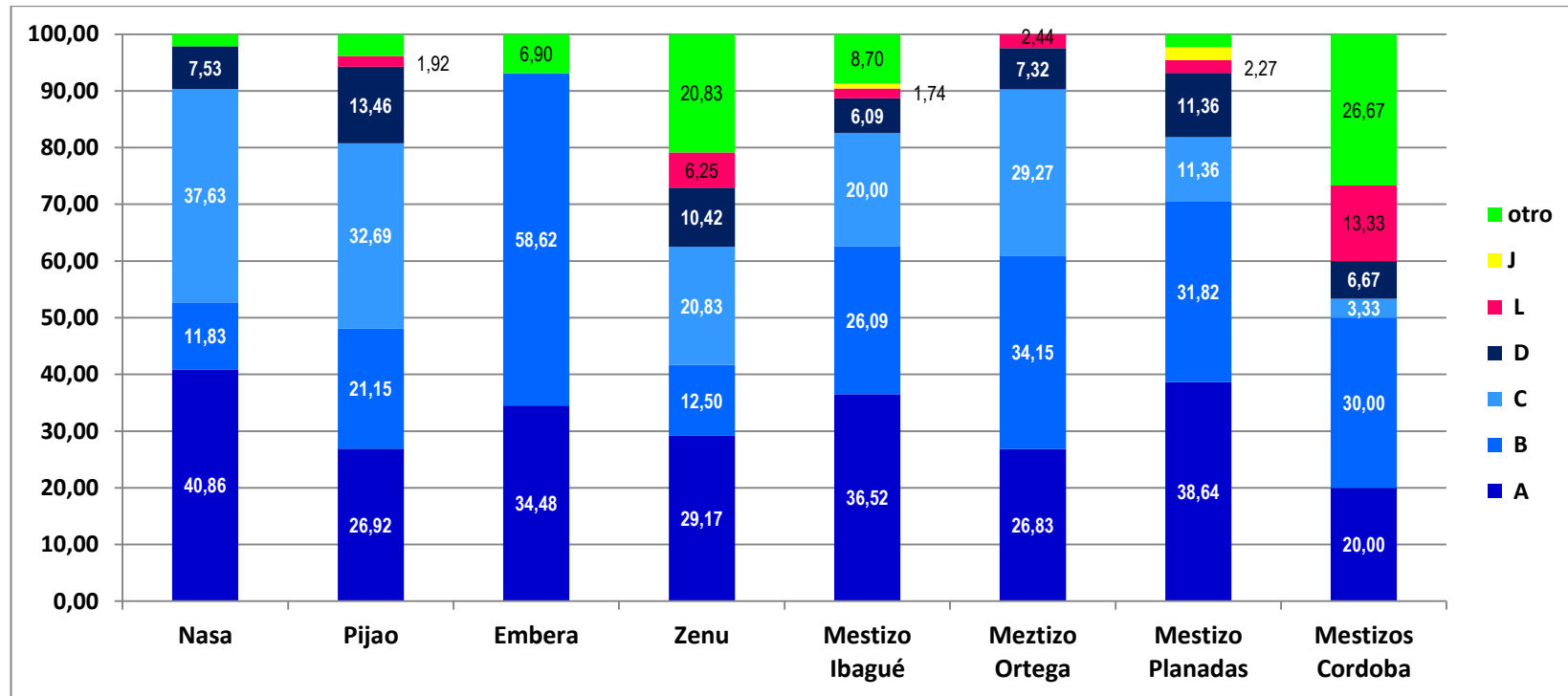
**Tabla 10.** Frecuencias (%) de los Hgs mitocondriales fundadores amerindios A, B, C y D, el Hg L africano y el J europeo.

Población	n	Hg							<i>h</i>
		A	B	C	D	L	J	otro	
<b>Nasa</b>	93	40,86	11,83	37,63	7,53	0,00	0,00	2,15	<b>0.6786</b> ± 0.0267
<b>Pijao</b>	52	26,92	21,15	32,69	13,46	1,92	0,00	3,85	0.7707 ± 0.0256
<b>Embera</b>	29	34,48	58,62	0,00	0,00	0,00	0,00	6,90	<b>0.5517</b> ± 0.0621
<b>Zenú</b>	48	29,17	12,50	20,83	10,42	6,25	0,00	20,83	0.8147 ± 0.0240
<b>Ibagué</b>	115	36,52	26,09	20,00	6,09	1,74	0,87	8,70	0.7535 ± 0.0205
<b>Ortega</b>	41	26,83	34,15	29,27	7,32	2,44	0,00	0,00	0.7378 ± 0.0287
<b>Planadas</b>	44	38,64	31,82	11,36	11,36	2,27	2,27	2,27	0.7389 ± 0.0400
<b>Córdoba</b>	30	20,00	30,00	3,33	6,67	13,33	0,00	26,67	0.8023 ± 0.0348
<b>Total</b>	452	31,68	28,27	19,39	7,85	3,49	0,39	8,92	

Fuente: el autor.

\* Las poblaciones analizadas fueron: indígenas Nasa, Pijao, Embera y Zenu; las mestizas del Tolima (Ibagué, Ortega y Planadas) y de Córdoba. *h* corresponde a la diversidad genética no sesgada de Nei-1987, basada en las frecuencias calculadas en Arlequin. En las poblaciones indígenas Nasa, Pijao y Embera y mestizas del Tolima los Hgs nativos americanos fundadores, están por encima del 90%, mientras que en los mestizos de Córdoba y los Zenues están entre el 60 y 72% y en los cuales aparece la participación de Hgs como el africano (>5%) u otros no identificados en el presente trabajo (>20%).

**Figura 9.** Proporciones de los haplogrupos nativos americanos (A, B, C y D), africano (L) y europeo (J) en cada población analizada.



Fuente: el autor.

\* En todas las poblaciones, la ancestría materna a nivel mitocondrial esta determinada por la mayor frecuencia del componente nativo americano, el cual es menor en las poblaciones mestizas de Córdoba y en los indígenas Zenu, dado que en estos aparecen los Hgs africano entre otros no identificados en este trabajo. El Hg europeo aparece sólo en dos poblaciones mestizas del Tolima (Ibagué y Planadas).

**7.1.1 Comparación inter-regional.** Las poblaciones mestizas del Tolima poseen en promedio un componente nativo mitocondrial mayor (93.2%) que la población mestiza de Córdoba (60%), incluso que los zenues (73%). Se puede notar que el aporte africano es más alto en las poblaciones cercanas a la costa Caribe (Zenú y mestizos de Córdoba), aunque también se presenta en los Pijao (1.9%) y los mestizos del Tolima (2.2%) pero en menor frecuencia.

Los indígenas del Tolima –Nasa y Pijao- poseen los 4 Hgs fundadores en un perfil de frecuencias similares entre sí. Difiere de los zenues debido a que estos últimos poseen un alto componente que no se pudo determinar. Con relación a los Embera, difieren en el perfil de las frecuencias, ya que éstos solo presentaron dos Hgs fundadores predominantes, el A y el B.

**7.1.2 Comparación intra-regional.** Las poblaciones indígenas del Tolima, en promedio -96%-, poseen una alta frecuencia de los Hgs nativos fundadores (Nasa: 98% y Pijao: 94.2%). En los mestizos, la población de Ibagué fue la más baja (88%), con relación a los de Ortega (97.5%) y Planadas (93%). En promedio, la población mestiza del Tolima posee más del 90% de componente nativo mitocondrial. El componente africano es bajo en ambos grupos poblacionales (<5%), no obstante es casi tres veces mayor en los mestizos, en los que también aparece un componente europeo (Hg J) bajo (1%), ausente en los indígenas. Por otro lado, la fracción de otros Hgs fue menor al 10%, siendo mayor en la población mestiza que en los indígenas. En cuanto al perfil de frecuencias de los Hgs nativos, el A es el más frecuente y similar entre mestizos e indígenas (~35%), el C es más frecuente en indígenas (35.86%) y el B lo es en los mestizos (29%).

En Previos estudios para la población Pijao, el Hg C fue determinado empleando dos mutaciones simultáneas: pérdida del sitio Hinc II en la posición 13259 *AluI* 13262. Encontrándose presente solo la primera de ellas, razón por la cual no se consideró la presencia del Hg C en los Pijao (Rondon et al., 1999). Actualmente se contempla que para definir el Hg C es suficiente mediante la pérdida del sitio *HincII* en la posición 13259 (Derenko et al., 2007; Moraga et al., 2000; Moraga, Santoro, Standen, Carvallo, & Rothhammer, 2005). Por esta razón en el presente estudio se reportó dicho Hg.

La población indígena Embera muestra una mayor prevalencia de Hgs nativos fundadores (93%), con relación a los zenues (73%). Por otro lado, en los mestizos de Córdoba solo el 60% de los Hgs son los fundadores A, B, C o D. El componente africano es del 13% en los mestizos y 6% en los indígenas Zenu, el cual no fue encontrado la muestra Embera. No se detectó el Hg europeo J en ninguna de las poblaciones de Córdoba. En cuanto al perfil de frecuencias de los Hgs nativos, el A y el B son los más frecuentes en los zenues; mientras que el C es más frecuente en los mestizos, seguido por el A.

**7.1.3 Análisis de distancias  $F_{st}$ .** Se realizó un análisis de distancias  $F_{st}$ , interpretado como la fracción de heterocigocidad perdida debido a la subdivisión poblacional ( $1-F_{st}$ ) mediante el programa *Arlequin*. Por tratarse de un sistema polimórfico haploide, se partió del cálculo de la heterocigocidad virtual obtenida a partir de las frecuencias haplogrupales. En la Tabla 11 se reportan los valores  $F_{st}$ .

**Tabla 11.** Matriz de los valores  $F_{st}$  pareados entre las poblaciones del estudio.

	Nasa	Pijao	Embera	Zenú	Ibagué	Ortega	Planadas	Córdoba
Nasa	0							
Pijao	0.00934	0						
Embera	<b>0.20913</b>	<b>0.14504</b>	0					
Zenú	<b>0.03893</b>	0.01369	<b>0.1503</b>	0				
Ibagué	<b>0.02976</b>	0.00942	<b>0.07746</b>	0.01272	0			
Ortega	<b>0.0367</b>	-0.00642	<b>0.08523</b>	<b>0.04008</b>	0.00499	0		
Planadas	<b>0.05971</b>	0.02581	0.04625	<b>0.0359</b>	-0.0035	0.00967	0	
Córdoba	<b>0.14536</b>	<b>0.07531</b>	<b>0.0807</b>	0.02151	<b>0.04406</b>	<b>0.06851</b>	<b>0.04564</b>	0

Fuente: el autor.

Alta Moderada baja

\* Los valores significativos ( $P < 0.05$ ) (ver Anexo 6) están en negrilla y la escala del grado de la diferenciación poblacional se da de acuerdo a la literatura: baja diferenciación (0 - 0.05), moderada (0.05 – 0.15), alta (0.15 – 0.25) y muy alta ( $>0.25$ ) (Hartl & Clark, 2007).

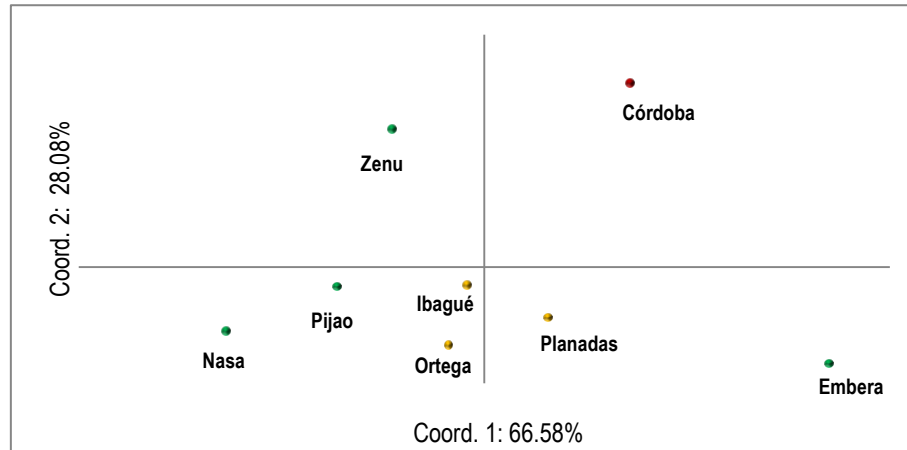
Con la matriz de distancias  $F_{st}$  se hizo un análisis de escalamiento multidimensional en el programa *GenALEX*. Este método multivariado define coordenadas para cada población, de manera que las distancias entre ellas son el reflejo de la distancia  $F_{st}$  original. En esencia, es un proceso por el cual un conjunto de datos multidimensionales se localizan en los ejes mayores de variación. Cada eje sucesivo explica proporcionalmente menos variación de la total, de manera que cuando se tienen distintos grupos, los primeros dos o tres ejes revelan la separación entre los grupos (Peakall & Smouse, 2006, 2012).

De acuerdo con figura 10, los ejes 1 y 2 explican la mayor parte de la variación (94.66%), así, las poblaciones mestizas del Tolima desde el punto de vista mitocondrial son genéticamente afín a los Pijao. La población mestiza de Córdoba está en un nivel de distancia baja-moderada de las mestizas del Tolima ( $F_{st}$  entre 0.04 y 0.14) y moderada con relación a los Embera, Nasa y Pijao. Las distancias Embera-Nasa y Embera-Zenu fueron las más altas, y moderadas con los mestizos de Córdoba, Ibagué y Ortega.

Cuando se hace un análisis multidimensional escalonado, basado en la matriz de distancias genéticas (anexo 7), tomando en consideración varias poblaciones indígenas y mestizas de Colombia y Latino América, el 69.4% de la variación se explica por las coordenadas 1 y 2. Aparece un perfil en el cual la

mayoría de las poblaciones mestizas están agrupadas en el centro, también se pueden notar otros dos grupos, uno que asocia ligeramente las poblaciones mestizas de Córdoba y Quibdó (Chocó-Colombia) y una población indígena del occidente colombiano –Waunana-. El otro grupo asocia las poblaciones mestizas de Cali y el Cauca, los indígenas Nasa (Colombia) y Maya de Centro América. Las poblaciones indígenas de Colombia Nukak, Kogi y Embera son las más externas y distantes genéticamente (Figura 11).

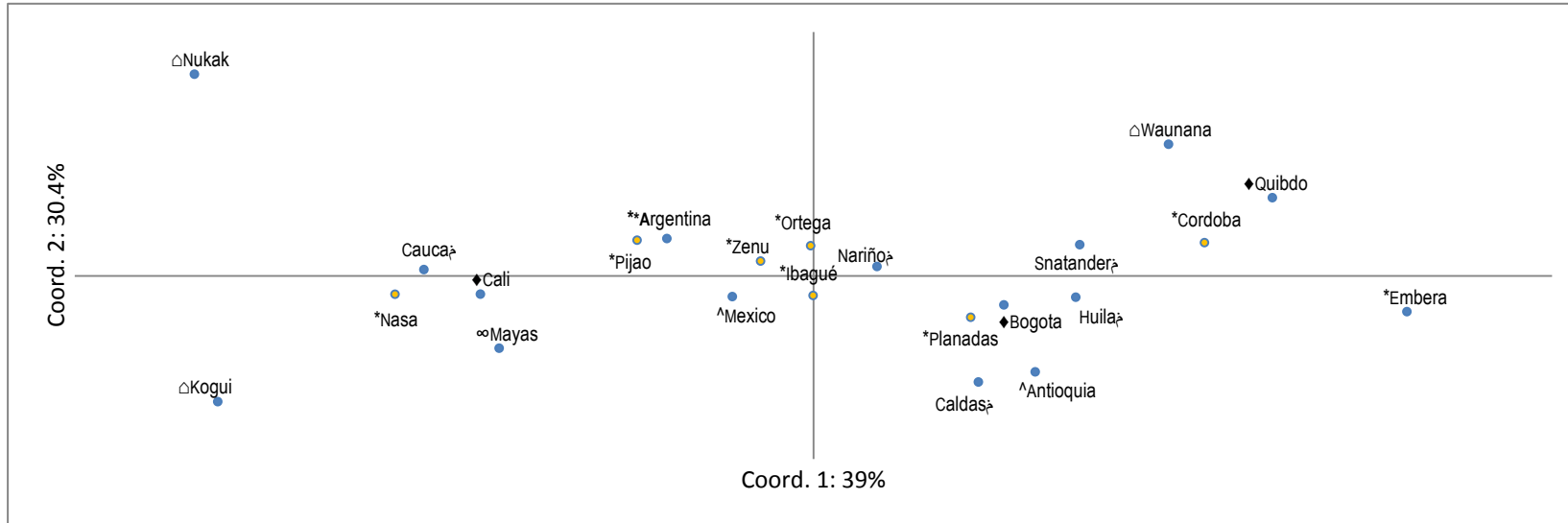
**Figura 10.** Análisis de escalamiento multidimensional (MDS) para el mDNA.



Fuente: el autor.

\* Cálculos con base en la matriz de datos pareados para los valores  $F_{st}$  entre las poblaciones del estudio. La matriz de valores  $F_{st}$  fue obtenida mediante el análisis de las frecuencias de los Hgs mitocondriales. Las poblaciones más distantes entre sí, desde el punto de vista mitocondrial son la embera, Zenu, Nasa y los mestizos de Córdoba.

**Figura 11.** Análisis de escalamiento multidimensional (MDS) mDNA-2.



\*Fuente: el autor.  
 ∞Fuente: Bailliet, et al 1994.  
 △Fuente: Keyeux, et al 2002  
 ◆Fuente: Keyeux, et al 2003  
 ^Fuente: Carvajal-Carmona, et al 2003  
 \*\*Fuente: Catelli, et al 2011  
 ↗Fuente: Rojas, et al 2010

\*. Calculos obtenidos mediante la matriz de distancias Fst calculada a partir de las frecuencias de los Hgs mitocondriales en las poblaciones del Tolima, Córdoba, y otras de Colombia y Latinoamérica. Se observan poblaciones nativas distantes, especialmente aquellas que han permanecido en regiones apartadas (Nukak, Kogui y Embera). La mayoría de las poblaciones mestizas aparecen agrupadas en el centro de la gráfica.

## 7.2 ANCESTRÍA PATERNA DE LAS POBLACIONES ANALIZADAS

El análisis de la ancestría paterna se llevó a cabo mediante la tipificación de 17 marcadores bialélicos de la región no recombinante del cromosoma Y (NRY) en 215 hombres para asignar cada individuo a un Hg ancestral, lo cual es posible debido a que cada uno de éstos presenta diferencias interpoblacionales de frecuencias (Karafet et al., 2008). Cada Hg se define por la presencia del estado derivado de la mutación. En la tabla 12 se reportan los estados alélicos para cada uno de los SNPs empleados.

**Tabla 12.** Información para cada uno de los marcadores de la región NRY.

Nombre del Hg	Nombre de la mutación	Código del SNP	Alelo ancestral	Alelo derivado
E	M96	rs9306841	C	G
L	M22	rs3913	A	G
K-R	M9	rs3900	C	G
P	M45	rs2032631	G	A
Q	M242	rs8179021	C	T
Q1a3*	M346	NP	C	G
Q1a3a*	M3	rs3894	C	T
Q1a3a1	M19	rs3910	T	A
Q1a3a2	M194	rs2032677	T	C
R1	M173	rs2032624	A	C
R1b	M343	rs9786184	C	A
R1b1	P25	rs150173	C	A
R1b1b2	M269	rs9786153	T	C
G	M201	rs2032636	G	T
I	M170	rs2032597	A	C
J	M304	rs13447352	A	C
N	M231	rs9341278	G	A

Fuente: Karafet, et al 2008.

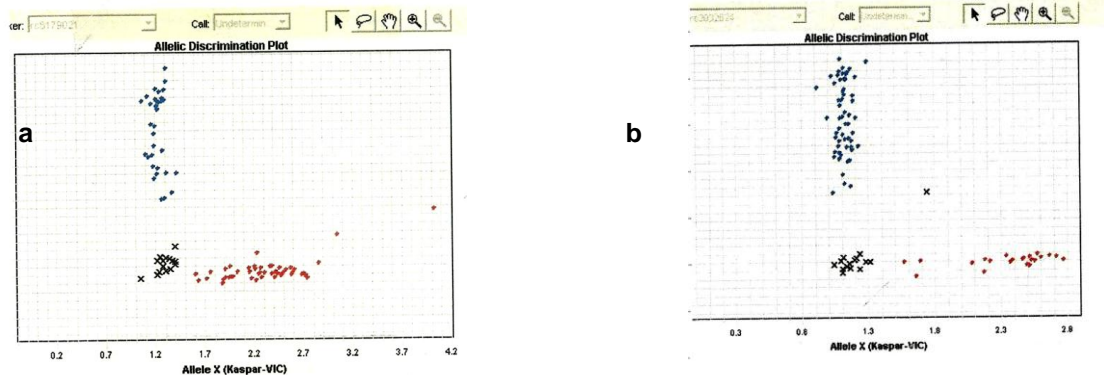
\* NP: no presenta. Se emplearon 17 SNPs en la determinación de la ancestría paterna de las poblaciones del Tolima y de Córdoba en Colombia. La primera columna identifica el nombre del haplogrupo, la siguiente corresponde a la nomenclatura de la mutación, luego el código *reference SNP* (rs) por medio del cual se accede en las bases de datos, luego aparece el alelo ancestral y derivado de acuerdo con lo reportado (Karafet et al., 2008).

Debido a que las poblaciones analizadas eran indígenas y mestizas colombianas, inicialmente se tipificaron los Hgs más probables, previendo la distribución



geográfica de los mismos y los procesos migracionales recientes en Latino América. En este sentido, se tipificó en todas las muestras de hombres el Hg Q, de origen nativo amerindio, y el R1 de origen europeo (Figuras 12 y 13).

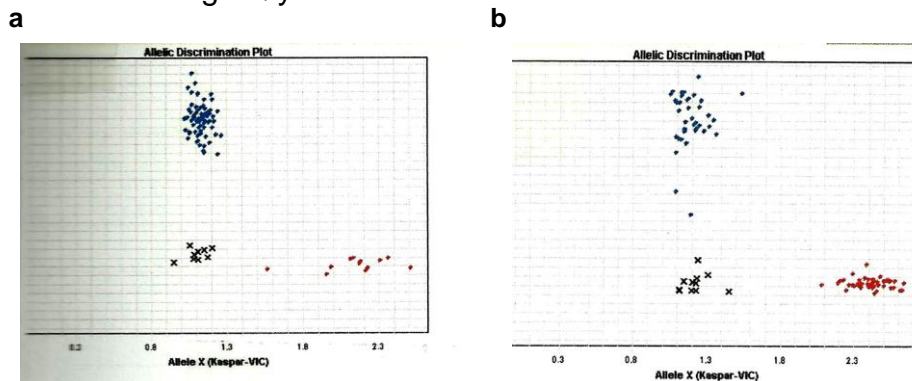
**Figura 12.** Agrupamiento de los genotipos del cromosoma Y generado durante la tipificación de los Hgs Q y R1.



Fuente: el autor.

\* Tipificación en poblaciones Nasa, Pijao, Embera y Zenú. (a) Para el Q (rs 8179021), los individuos con el alelo derivado están sobre el eje X (fluorescencia de la sonda Vic). (b) En la tipificación del R1 el alelo derivado esta sobre el eje X (señal de la sonda VIC). La señal de los alelos ancestrales están en el eje Y (sonda FAM) y la de los controles negativos o no amplificados, en el centro.

**Figura 13.** Agrupamiento de los genotipos del cromosoma Y generado durante la tipificación de los Hgs Q y R1.



Fuente: el autor.

\*Tipificación en poblaciones mestizas del Tolima y de Córdoba. (a) Para el Q (rs 8179021), los individuos con el alelo derivado están sobre el eje X (fluorescencia de la sonda Vic). En la tipificación del R1 (B), el alelo derivado está sobre el eje X (señal de la sonda VIC). La señal de los alelos ancestrales están en el eje Y (sonda FAM) y la de los controles negativos o no amplificados, en el centro.

Aquellos individuos que presentaron el alelo ancestral para ambos Hgs, fueron evaluados para otros probables como el E (africano), L (subcontinente indú), G, J (euroasiáticos), I (europeo) N (polinesio). Cuando no presentaron alguno de los anteriores (ancestral en todos los marcadores empleados) se confirmaron los clados principales K-R y P. Posteriormente, en aquellos individuos con Hg Q o R1 se les determinó el subHg correspondiente. Las frecuencias de los Hgs encontrados están dados en la tabla 13 y figura 14.

**7.2.1 Comparaciones inter-regionales.** De acuerdo con la figura 14 se puede observar que las poblaciones indígenas Nasa, Embera y Pijao poseen la mayor frecuencia del componente nativo (Hgs Q1a3a\* y Q1a3\*) americano paterno, oscilando entre el 66 y 71%, mientras que en los mestizos y los indígenas Zenu, dicho componente se encuentra en menor frecuencia (9.5-29%). En términos generales, se observa que la ancestría nativa posee distribución clinal, mayor en las poblaciones indígenas y decreciendo en las mestizas.

El componente europeo estuvo dado por los subHgs R1b1b2, R1 y el I, mientras que el euroasiático por los Hgs G, J y el clado K-R. Ambos componentes se manejarán en conjunto como ancestría europea dado que comparten el mismo rango geográfico. Este componente presentó una distribución opuesta al nativo americano, es decir bajo en las poblaciones indígenas Nasa, Pijao y Embera (entre 21 y 33%) e incrementando en los indígenas zenues y las poblaciones mestizas (58 - 78%).

Dentro de los Hgs no nativos descritos anteriormente, el R es el más frecuente seguido por los Hgs J, I, G y el clado K-R. Los individuos mestizos que presentaron la mutación derivada en el SNP M9 y ancestral para las mutaciones características del Hg Q, R, L, NO y P, fueron asignados al clado K-R. Posiblemente los Hgs específicos para estos individuos sean el K o T, de origen euroasiático, los cuales no fueron tipificados en el presente estudio. Estos individuos sólo se presentaron en las poblaciones mestizas del Tolima. La ancestría africana está en mayor frecuencia en la población mestiza de Córdoba (23%), y en una frecuencia menor en las poblaciones de Planadas (~12%) o Ibagué (~11%).

En otras poblaciones se le puede encontrar pero en frecuencias menores al 5% (indígenas Nasa, Pijao y mestizos de Ortega). En cuanto a la diversidad (Tabla 13), los indígenas Nasa, Pijao y Embera presentaron las menores diversidades. Por su parte, los mestizos de Córdoba y Planadas presentaron la mayor diversidad.

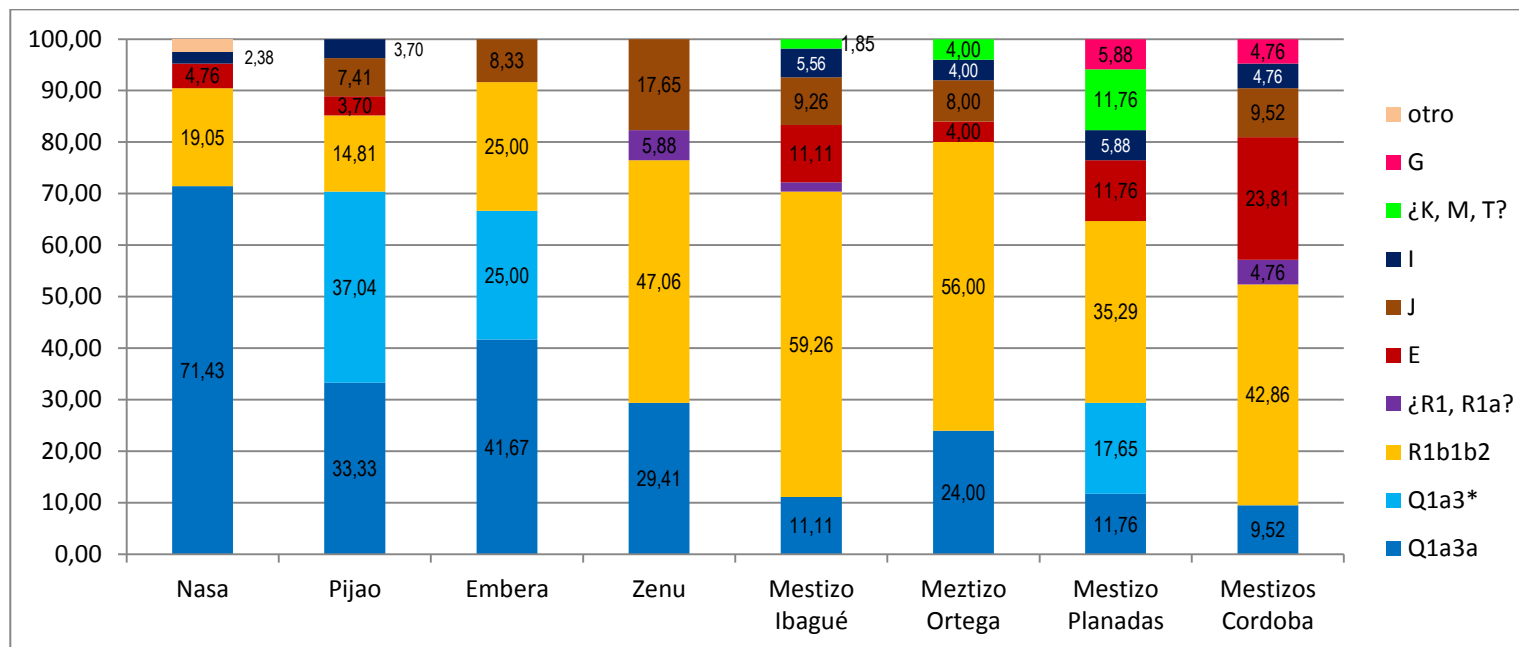
**Tabla 13.** Frecuencias (%) de los Hgs de la región NRY en las poblaciones estudiadas.

Población	Hg											
	Q1a3a*	Q1a3*	R1b1b2	¿R1, R1a?	E	J	I	¿K, M, T?	G	otro	<i>h</i>	
<b>Nasa</b> (n=42)	71,43	0,00	19,05	0,00	4,76	0,00	2,38	0,00	0,00	2,38	0.4611 ± 0.0819	
<b>Pijao</b> (n=27)	33,33	37,04	14,81	0,00	3,70	7,41	3,70	0,00	0,00	0,00	0.4929 ± 0.1075	
<b>Embera</b> (n=12)	41,67	25,00	25,00	0,00	0,00	8,33	0,00	0,00	0,00	0,00	0.53 ± 0.1359	
<b>Zenú</b> (n=17)	29,41	0,00	47,06	5,88	0,00	17,65	0,00	0,00	0,00	0,00	0.6985 ± 0.075	
<b>Ibagué</b> (n=54)	11,11	0,00	59,26	1,85	11,11	9,26	5,56	1,85	0,00	0,00	0.6155 ± 0.0682	
<b>Ortega</b> (n=25)	24,00	0,00	56,00	0,00	4,00	8,00	4,00	4,00	0,00	0,00	0.6277 ± 0.0883	
<b>Planadas</b> (n=17)	11,76	17,65	35,29	0,00	11,76	0,00	5,88	11,76	5,88	0,00	0.8015 ± 0.0646	
<b>Córdoba</b> (n=21)	9,52	0,00	42,86	4,76	23,81	9,52	4,76	0,00	4,76	0,00	0.7714 ± 0.0725	

Fuente: el autor.

\* Se muestran los Hgs típicos de los nativos americanos (Q1a3a\* y Q1a3\*), europeos ( R1b1b2, R1, I), euroasiáticos (G, J), y africanos (E). Las poblaciones analizadas fueron: indígenas Nasa, Pijao, Embera y Zenu; las mestizas del Tolima (Ibagué, Ortega y Planadas) y de Córdoba. *h* corresponde a la diversidad genética no sesgada de Nei-1987, basada en las frecuencias calculadas en Arlequin. Las poblaciones más diversas fueron la Nasa, Pijao y Embera, en tanto que las de mayor diversidad fueron las mestizas de planadas y Córdoba.

**Figura 14.** Proporción ancestral de los diferentes Hgs de la región NRY en cada población analizada.



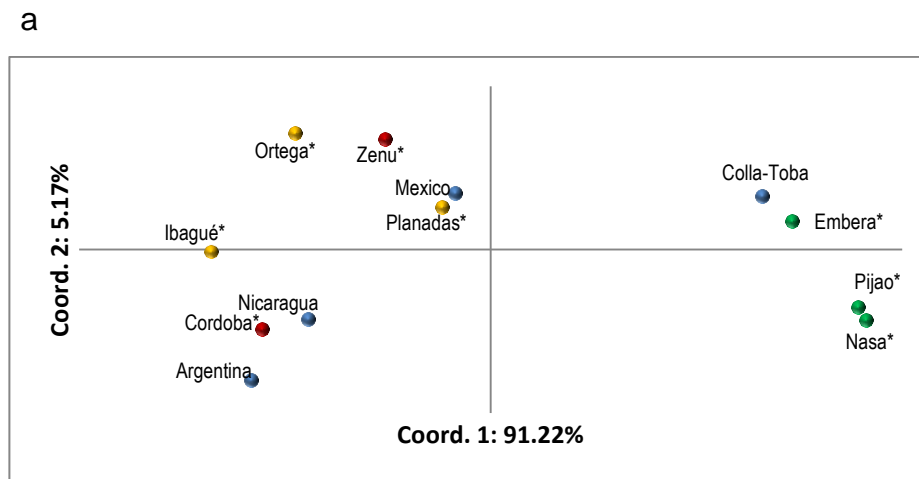
Fuente: el autor.

\* Los Hgs nativos americanos se presentaron en una frecuencia mayor en los indígenas Nasa, Pijao y Embera (66-71%) que en los mestizos y zenues (9.5-29%). El componente ancestral europeo o euroasiático es más alto en mestizos y zenues (58 - 78%) que en los indígenas Pijao, Embera y Nasa (21 - 33%). La ancestría africana es mayor al 20% en los mestizos de Córdoba, alrededor del 11% en los mestizos de Ibagué y Planadas y menor al 5% o ausente en las demás poblaciones.

**7.2.2 Comparaciones intra-regional.** En los indígenas del Tolima, Nasa y Pijao la proporción de ancestría nativa promedio fue del 71%. En los Pijao, este componente esta representado por los 2 paragrupos (Q1a3a\*, Q1a3\*), y en los Nasa tan solo por el Q1a3a\*. En ambas poblaciones indígenas existe una ancestría europea promedio de 20.3%. En los mestizos la estructura se invierte, siendo 60.42% europea y 17.71% indígena. Solo en los mestizos de Planadas están presentes los paragrupos Q1a3a\* y Q1a3\*. El componente africano (E) fue menor al 5% en los Nasa, Pijao y mestizos de Ortega, y cerca del 11% en los mestizos de Ibagué y Planadas. Las poblaciones mestizas del Tolima presentan algunas particularidades que las diferencian de las indígenas, por ejemplo la presencia del Hg G y el clado K-R. Por otro lado, se puede observar que las poblaciones mestizas más cercanas a las comunidades indígenas Nasa o Pijao (Planadas y Ortega), poseen un mayor componente nativo. En cuanto a las poblaciones de Córdoba, los indígenas Embera presentaron principalmente ancestría nativa americana (67%), mientras que en los los Zenu predominó el componente europeo (80.3%). En los mestizos de la misma región el componente ancestral europeo fue el más frecuente (67%) seguido por el africano (24%).

**7.2.3 Análisis de distancias  $F_{st}$ .** De acuerdo con el análisis escalonado multidimensional (Figura 15 a y b), el 96.02% de la variación se explica por medio de las dos primeras coordenadas. Las poblaciones mestizas aparecen agrupadas aparte de las indígenas, ya que los valores de distancias  $F_{st}$  entre ambos tipos de poblaciones son significativamente altos, moderado o muy alto ( $p < 0.05$ ), oscilando desde 0.12 a 0.3. Los zenues se agrupan también con los mestizos, lo cual se refleja en valores  $F_{st}$  bajos, pero altos con los indígenas. Se emplearon otras poblaciones mestizas o indígenas de Latino América con el fin de establecer comparaciones y analizar si los patrones de diversidad se compartían a lo largo de la región latino americana, encontrándose que mediante el análisis de las distancias  $F_{st}$  se observa el mismo patrón de agrupamiento (figura 15).

**Figura 15.** Análisis de escalamiento multidimensional (MDS) en el cromosoma Y.



b

	Nasa	Pijao	Embera	Zenu	Ibagué	Ortega	Planadas	Córdoba	Nicaragua	Nam-Argentina	Argentina	México
Nasa	0											
Pijao	-0.0222	0										
Embera	0.03798	0.04792	0									
Zenu	<b>0.19113</b>	<b>0.1698</b>	0.08031	0								
Ibagué	<b>0.32359</b>	<b>0.31621</b>	<b>0.24399</b>	0.02031	0							
Ortega	<b>0.25495</b>	<b>0.24705</b>	<b>0.16044</b>	0.02393	0.01107	0						
Planadas	<b>0.13981</b>	<b>0.12908</b>	0.06286	0.00347	0.0482	0.00632	0					
Córdoba	<b>0.28592</b>	<b>0.26486</b>	<b>0.19498</b>	0.02028	0.0049	0.01924	0.00107	0				
Nicaragua	<b>0.23156</b>	<b>0.22312</b>	<b>0.16419</b>	0.00383	0.00679	0.00459	0.01547	0.01419	0			
Colla-Toba	0.00326	0.00552	0.04054	<b>0.09246</b>	<b>0.23359</b>	<b>0.15918</b>	<b>0.06518</b>	<b>0.19307</b>	<b>0.16138</b>	0		
Argentina	<b>0.28773</b>	<b>0.27809</b>	<b>0.22184</b>	0.02321	0.00555	0.01728	<b>0.04818</b>	0.00131	<b>0.00965</b>	<b>0.21599</b>	0	
México	<b>0.12234</b>	<b>0.11952</b>	<b>0.06451</b>	0.01053	<b>0.046</b>	0.01022	-0.0172	0.02237	<b>0.02193</b>	<b>0.06979</b>	<b>0.0536</b>	0

Valor P<0.05

Fuente: el autor.

Los datos para estas poblaciones (\*) son del presente trabajo

Datos mestizos Argentina, fuente: Corach, et al 2010.

Datos mestizos Nicaragua, fuente: Nuñez, et al 2012.

Datos indígenas Argentina (Colla-Toba), fuente: Toscanini, et al 2011.

Datos Mestizos México, fuente: Martínez-Cortés, et al 2012

Muy alta
  Alta
  Moderada
  baja

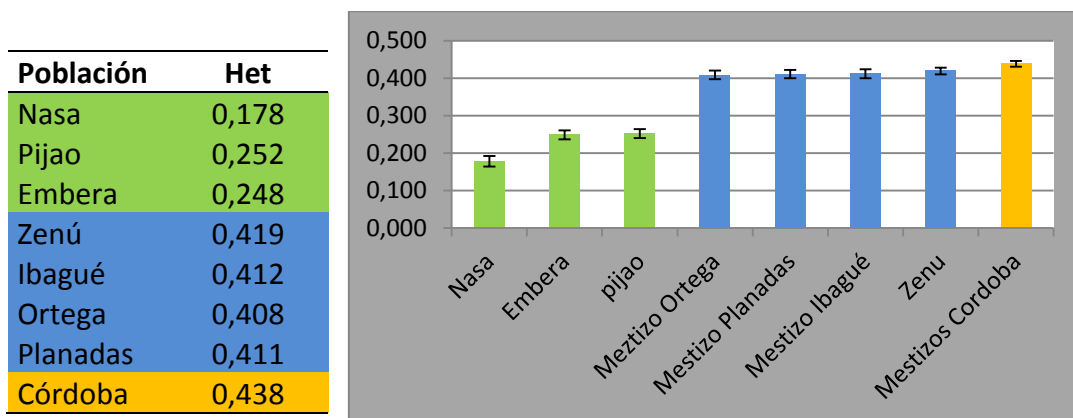
\*Cálculos obtenidos mediante el análisis de las frecuencias de los Hgs de en la región NRY en las poblaciones del Tolima y Córdoba.(b) Matriz de datos pareados para los valores Fst, con valores significativos ( $p < 0.05$ ) (ver Anexo 8) resaltados en negrilla y la escala del grado de la diferenciación poblacional se da de acuerdo a la literatura: baja diferenciación (0 - 0.05), moderada (0.05 – 0.15), alta (0.15 – 0.25) y muy alta ( $> 0.25$ ) (Hartl & Clark, 2007).

### 7.3 ANCESTRÍA A NIVEL AUTOSÓMICO

Se obtuvieron los genotipos de 446 individuos para 100 AIMs. Posteriormente los datos se acomodaron en los formatos .ped y .map para el programa PLINK, en el cual se realizó el control de calidad para extraer inicialmente aquellos individuos y SNPs con más del 10% de datos genotípicos faltantes. Después de lo cual, los análisis se realizaron con 411 individuos (se extrajeron 35 individuos= 7.8% de la muestra total). En este programa también se calcularon las frecuencias génicas (Anexo 9).

**7.3.1** Análisis de la diversidad genética en los tolimeses. El análisis de la heterocigocidad (Het) reveló que los valores más bajos se presentaron en las poblaciones indígenas Nasa (0.178), Pijao (0.252) y Embera (0.248), mientras que las mestizas y los Zenú presentaron valores altos, entre 0.408 y 0.438 (Figura 16).

**Figura 16.** Heterocigocidad media (Het) con base en los AIMs.



**Fuente:** el autor.

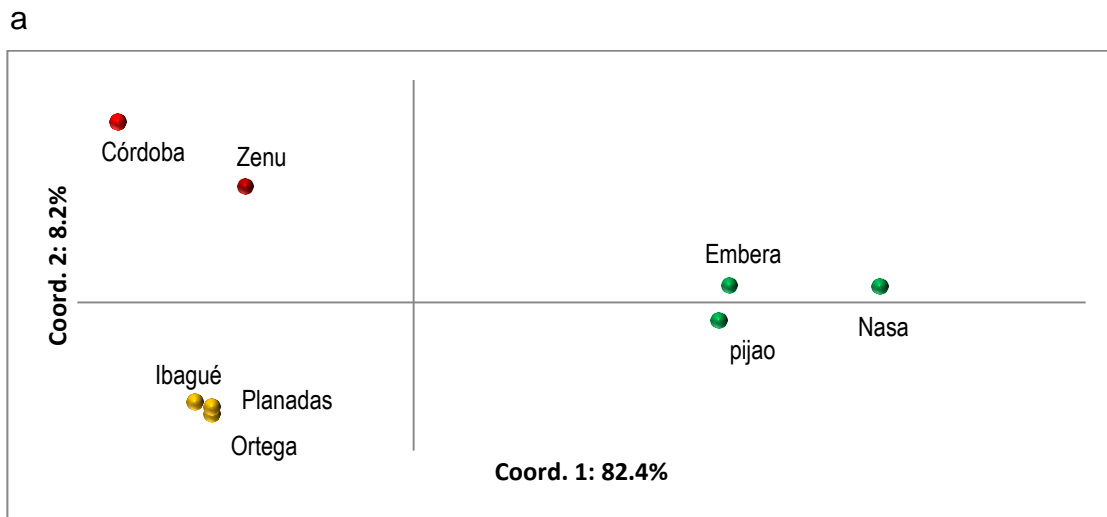
\* Cálculos obtenidos con base en los genotipos para los 100 AIMs en las poblaciones indígenas Nasa, Pijao, Embera y Zenu, y las mestizas de Ibagué, Ortega, Planadas y Córdoba.

Cuando se realizó el análisis de escalamiento multidimensional (MDS) para las 8 poblaciones de estudio, basados en los valores pareados de distancias  $F_{st}$ , se obtuvo la formación de dos grupos claramente definidos y distantes, el de los

mestizos del Tolima, y el que reúne los indígenas Nasa, Pijao y Embera, esta separación se explica por el 82.2% de la variación total entre ambos grupos.

En cuanto a los mestizos de Córdoba y los Zenú, aparecen alejados de los otros dos grupos, lo cual corresponde al 8.2% de la variación total (Figura 17).

**Figura 17.** Análisis MDS con base en los AIMs.



b

	Nasa	Pijao	Embera	Zenu	Ibague	Ortega	Planadas	Cordoba
Nasa	0							
Pijao	0,02223	0						
Embera	0,02895	0,00636	0					
Zenu	0,20916	0,11448	0,10712	0				
Ibague	0,22537	0,14263	0,14191	0,02053	0			
Ortega	0,20806	0,11005	0,10814	0,0123	0,00899	0		
Planadas	0,2219	0,12144	0,11578	0,01227	0,00633	0,0019	0	
Cordoba	0,30673	0,18674	0,17502	0,00827	0,02466	0,02821	0,02348	0

Fuente: el autor.

\*(a) Análisis MDS, obtenido mediante los valores de distancia Fst calculado mediante las frecuencias de los genotipos para los 100 AIMs en las poblaciones del Tolima y Córdoba. (b) Matriz de valores Fst.

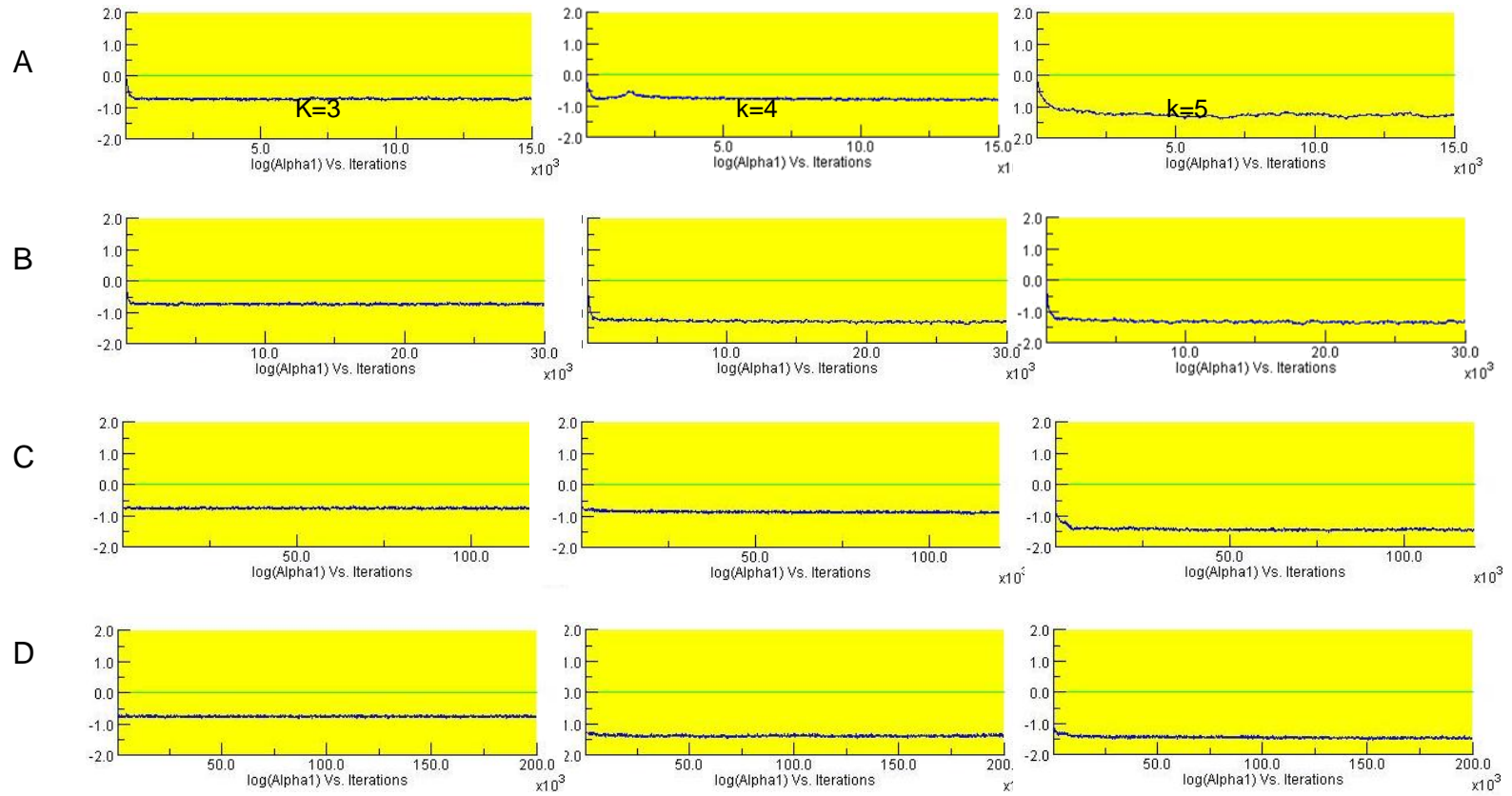


**7.3.2 Ancestría poblacional.** Para el análisis de la estructura genética poblacional y los cálculos de las proporciones de ancestría, se realizaron distintas simulaciones en el programa STRUCTURE empleando los datos genotípicos de 1889 individuos pertenecientes a las 8 poblaciones del presente estudio y las de referencia (africanos, europeos y nativos de América).

Inicialmente se determinaron las condiciones apropiadas para lograr la estabilidad del valor alfa a lo largo de las iteraciones, para que la estimación de las proporciones fuera confiable. Esto se hizo empleando distintas parejas de valores de iteraciones-MCMC de 5000-10000, 10000-20000, 40000-80000 y 100000-100000. Cada simulación se realizó bajo la asunción de diferentes grupos poblacionales -*k*- oscilando entre 2 y 7 ( $K=2, K=3 \dots K=7$ ) con 5 réplicas para cada *k* evaluado. En cada caso con o sin asignación previa de cada individuo a la población de origen predefinida por el muestreo, y, bajo los modelos de mezcla (*Admixture*) y de frecuencias génicas correlacionadas. El valor óptimo encontrado fue la pareja de 100000-100000 iteraciones-MCMC, ya que a valores inferiores, la probabilidad alfa no presentó estabilidad durante las iteraciones y las MCMC (Figura 18). El modelo: sin asignación previa de cada individuo a una población predefinida, permitió un comportamiento más estable del valor alfa en las diferentes simulaciones realizadas.

Con los datos de STRUCTURE, se calculó el número *K* más probable mediante el método de *Evanno*, disponible en el programa *STRUCTURE harvester*, de acceso libre en internet (Evanno et al., 2005; Rosenberg et al., 2005).

**Figura 18.** Comportamiento del valor alfa a lo largo de las parejas de valores iteraciones- MCMC.



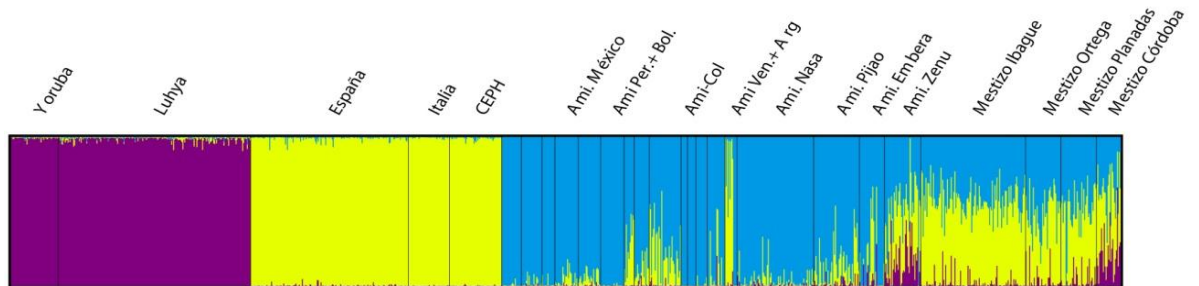
Fuente: el autor.

\*(A) 5000-10000, (B) 10000-20000, (C) 40000-80000 y (D) 100000-100000, para  $k=3$ ,  $K=4$  y  $K=5$ . Los valores más estables para los diferentes  $K$  asumidos son evidentes en las simulaciones C y D.

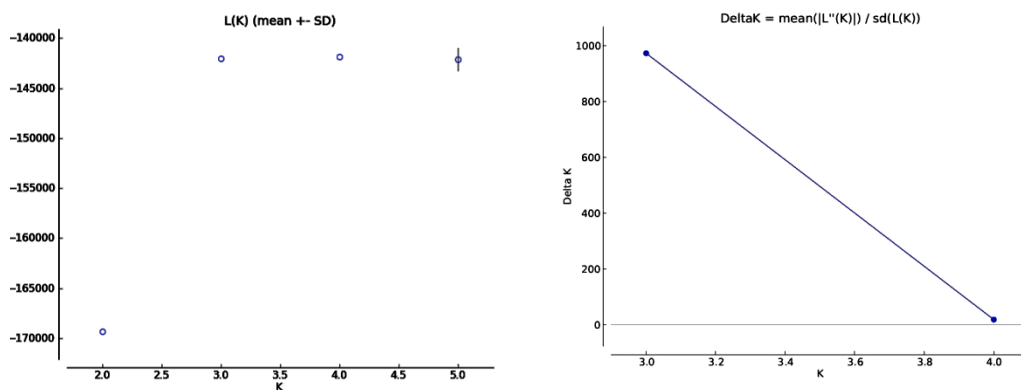
Se calculó la magnitud de cambio de probabilidades de valores K sucesivos ( $\Delta K$ ). Basado en este análisis, se obtuvo que el K=3 fue el número poblacional más probable y presente en la muestra analizada, explicando la mayor parte de la subestructura poblacional en la muestra analizada (Figura 19b). Con los datos del K óptimo se realizó la gráfica de estructuración poblacional. (Figura 19a).

**Figura 19.** Estructura genética de las poblaciones estudiadas.

a.



b.



Fuente: el autor.

\*(a) Plot de la estructura poblacional inferida con base en 1189 individuos, y datos genotípicos de 100 AIMs analizados en STRUCTURE. Cada individuo es representado por una línea vertical fraccionada en K- segmentos de diferentes colores que representan la fracción de pertenencia a K grupos. Cada gráfica se produjo en *Distruct*, y se basaron en 5 replicas para K=3, promediadas en el programa *Clumpp*. Las poblaciones de referencia africanas (morado), europeas (amarillo) y nativas americanas (azul), presentaron una estructura uniforme y claramente diferenciadas, por su parte, en las mestizas del Tolima se observa un patrón de mezcla genética europea e indígena, lo cual también sucede en los

mestizos de Córdoba y Zenu, quienes adicionalmente poseen un componente ancestral africano. (b) Tasa de cambio de la distribución de probabilidad (media  $\pm$ SD) calculada como  $L'(K) = L(K) - L(K - 1)$ , en donde se puede apreciar que después de  $K=2$ , no ocurre algún cambio o incremento significativo en la probabilidad para los  $K$  sucesivos.

En términos cualitativos, de acuerdo con la figura 19 se observa que dos de las poblaciones continentales parentales asumidas presentaron una estructura genética claramente definida y diferenciadas por patrones propios de frecuencias, identificadas por colores: África (violeta) y Europa (amarillo), mientras que en los nativos de América, aunque muestran un patrón de frecuencias principal (azul) que los separa de las dos anteriores, se evidencia una leve mezcla europea, localizada en las poblaciones indígenas de Colombia y Argentina.

Al observar las poblaciones objeto del presente estudio, los indígenas Nasa, Pijao y Embera presentan una estructura similar a los de referencia, es decir un componente nativo principal y una leve mezcla europea. En cuanto a los Zenu, su estructura obedece a un patrón de tres poblaciones ancestrales: nativo americano, europeo y el africano; el cual también es compartido con la población mestiza de Córdoba. Los Mestizo del Tolima son una población que en conjunto están constituidos por dos componentes ancestrales principales, el nativo y el europeo.

En cuanto a las proporciones ancestrales, en la tabla 16 se presentan los resultados cuantitativos de las mismas. Los africanos y europeos presentaron proporciones de membresía entre el 97 y 98 % exclusivamente por los grupos 1 y 2 respectivamente, indicando por lo tanto la ancestría que define cada uno.

En los nativos americanos, se observa variación entre las poblaciones al momento de agruparse, sin embargo hay un consenso general, por ejemplo, los de México, Perú, Bolivia y Venezuela se afiliaron ampliamente al grupo 3 - entre el 94 y 99%-, mientras que los de Argentina y Colombia entre el 42 y 88 %. Esto indica que el grupo 3 marca la ancestría nativa americana.

Es importante señalar que el patrón mostrado por las poblaciones colombianas y argentinas obedece a la mezcla, principalmente de origen europea (figura 19,  $K=3$ ) que oscila entre un 10 y 56% (Tabla 14).

Las diferencias interpoblacionales o intrapoblacionales, en razón de las proporciones ancestrales, fueron evaluadas mediante la prueba estadística de *Mann-Whitney* usando el programa Infostat versión 2008. Los valores del estadístico  $W$ , así como de la probabilidad de todas las comparaciones realizadas se pueden ver en el anexo 10.

**Tabla 14.** Proporciones de membresía calculadas en STRUCTURE en cada uno de los tres grupos ( $K=3$ ).

Origen geográfico	Población	Grupo 1	Grupo 2	Grupo 3	n
África	Africano	0.991	0.006	0.003	51
África	Africano	0.9836	0.0114	0.005	206
España	Caucásicos	0.005	0.987	0.008	619
Italia	Caucásicos	0.002	0.993	0.005	44
CEPH	Caucásicos	0.0026	0.9886	0.0088	56
México	Zapotecas	0.001	0.007	0.992	21
México	Tepehuanos	0.0036	0.0194	0.977	22
México	Nahua	0.001	0.015	0.984	14
México	Maya	0.0046	0.0464	0.949	25
Perú	Quechua	0.0056	0.054	0.9404	24
Bolivia	Aymara	0.001	0.004	0.995	25
Colombia-Norte	Nativos	0.0036	0.2761	0.7203	11
Colombia-Centro	Nativos	0.0098	0.1004	0.8898	16
Colombia-Sur	Nativos	0.0078	0.2262	0.766	34
Venezuela	Penare	0.0006	0.002	0.9974	7
Venezuela	Pemon	0.0006	0.002	0.9974	9
Venezuela	Warao	0.001	0.0148	0.9842	12
Venezuela	Indígena	0.0126	0.0888	0.8986	19
Argentina	Wichi	0.0124	0.5647	0.4229	13
Colombia	Nasa	0.0066	0.0344	0.959	82
Colombia	Pijao	0.0176	0.1281	0.8543	49
Colombia	Embera	0.0334	0.1226	0.844	27
Colombia	Zenú	0.1721	0.3491	0.4788	39
Colombia	Ibagué	0.0232	0.5496	0.4272	112
Colombia	Ortega	0.039	0.4519	0.5091	38
Colombia	Planadas	0.039	0.4709	0.4901	38
Colombia	Córdoba	0.2248	0.4219	0.3533	26

Fuente: el autor.

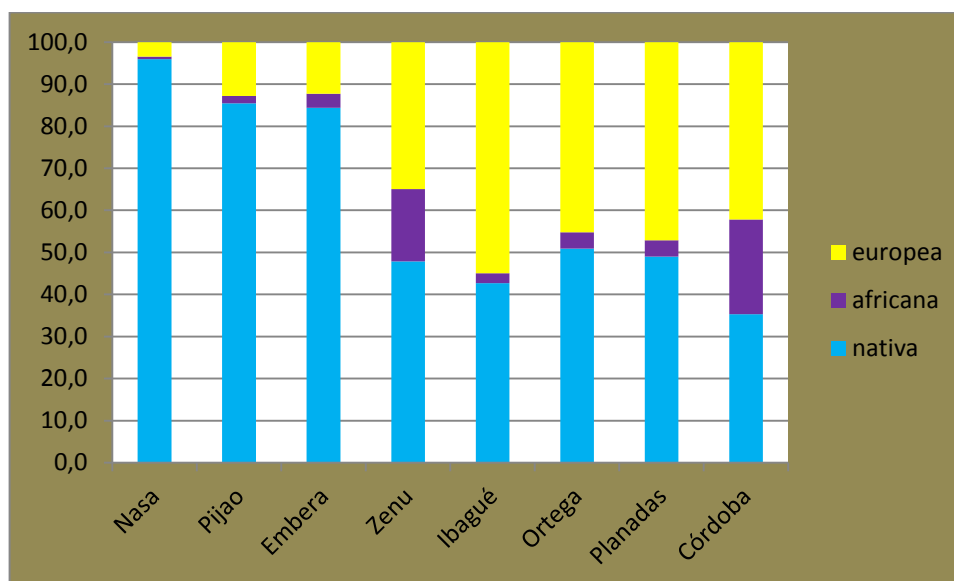
\* El análisis en STRUCTURE se realizó bajo las condiciones mencionadas en el texto. El grupo 1 define la ancestría africana, el 2 la europea y el 3 la nativa americana debido a las afiliaciones de las respectivas poblaciones parentales o de referencia. Los datos genotípicos con los cuales se obtuvieron las proporciones ancestrales de las poblaciones de referencia fueron proporcionados por el Doctor Angel Carracedo.

En promedio, la proporción de mezcla europea en las poblaciones mestizas del Tolima fue significativamente mayor ( $0.52 \pm 0.13$ ) que en las indígenas Pijao y nasa ( $0.07 \pm 0.1$ ) ( $p < 0.0001$ ), por otro lado, el componente nativo americano en estas

últimas fue significativamente mayor ( $0.92\pm 0.11$ ) que en los mestizos de la misma región ( $0.46\pm 0.13$ ) ( $p < 0.0001$ ). En cualquier caso, la ancestría africana media fue baja ( $< 5\%$ ) (mestizos:  $0.03\pm 0.05$ , indígenas:  $0.01\pm 0.02$ ) y no se diferenció significativamente entre ambos grupos poblacionales ( $p > 0.05$ ), así como tampoco al realizar comparaciones solo entre mestizos o los indígenas ( $p > 0.05$ ) (figura 20).

La población mestiza de Córdoba, y los Zenu, presentaron tres componentes ancestrales (Figura 20): nativo ( $0.35\pm 0.15$  y  $0.48\pm 0.19$ ), europeo ( $0.42\pm 0.15$  y  $0.35\pm 0.15$ ) y africano ( $0.23\pm 0.15$  y  $0.17\pm 0.13$ ), encontrándose diferencias significativas entre ambas poblaciones sólo para el nativo americano ( $p = 0.0042$ ), el cual fue mayor en los Zenu, y en quienes por otro lado, las tres proporciones

**Figura 20.** Proporciones (%) de ancestría nativa, africana, y europea con base en los AIMS.



Población	nativa	africana	europea	n
<b>Nasa</b>	95.9	0,66	3,44	82
<b>Pijao</b>	85,43	1,76	12,81	49
<b>Embera</b>	84,4	3,34	12,26	27
<b>Zenu</b>	47,88	17,21	34,91	39
<b>Ibagué</b>	42,72	2,32	54,96	112
<b>Ortega</b>	50,91	3,9	45,19	38
<b>Planadas</b>	49,01	3,9	47,09	38
<b>Córdoba</b>	35,33	22,48	42,19	26

Fuente: el autor.

\* Calculos realizados para las poblaciones indígenas Nasa, Pijao, Embera y Zenú, y poblaciones mestizas de Tolima y Córdoba analizadas. Los resultados se obtuvieron a partir del análisis de los genotipos para 100 AIMs en los programas STRUCTURE y *Clumpp*. La ancestría nativa es mayor en las poblaciones indígenas Nasa, Pijao y Embera (84-95%) que en los mestizos y los Zenu (35-50%). El componente ancestral africano fue más alto en los mestizos de Córdoba y en los Zenu (17 y 22% respectivamente) que en las demás poblaciones (<5%).

ancestrales fueron significativamente diferentes ( $p < 0.01$ ). El análisis de los mestizos no reveló diferencias entre la ancestría nativa y la europea ( $p = 0.0854$ ) pero sí en las demás comparaciones (Nam vs Afr,  $p < 0.0022$  y Eur vs Afr,  $p < 0.0001$ ).

Los Embera presentaron unas proporciones ancestrales similares a los indígenas del Tolima (Figura 20) sin diferencias estadísticas entre ambos grupos ( $p > 0.05$ ), y se observa que el componente ancestral nativo americano es significativamente mayor que los demás (europeo y africano), dentro de cada una de las tres poblaciones de nativos muestreadas: Nasa ( $0.96 \pm 0.06$ ), Pijao ( $0.85 \pm 0.14$ ) y Embera ( $0.84 \pm 0.22$ ) ( $p < 0.0001$ ). Estas tres poblaciones indígenas conservan una proporción ancestral amerindia significativamente mayor ( $p < 0.0001$ ) que cualquiera de las poblaciones mestizas analizadas incluyendo los Zenu (figura 20).

En las poblaciones mestizas del Tolima, se encontraron dos componentes ancestrales principales: el nativo y el europeo, que al compararlos entre Ortega (nam:  $0.51 \pm 0.15$ , eur:  $0.45 \pm 0.14$ ) y Planadas (nam:  $0.49 \pm 0.15$ , eur:  $0.47 \pm 0.15$ ) no se hallaron diferencias significativas ( $p > 0.05$ ) y dentro de cada una de estas poblaciones ambas ancestrías fueron estadísticamente similares ( $p = 0.1502$  y  $p = 0.7871$ ). Por su lado, Ibagué presentó una ancestría europea mayor y amerindia menor (nam:  $0.43 \pm 0.11$ , eur:  $0.55 \pm 0.10$ ) con respecto a Ortega o Planadas ( $p < 0.05$ ). Ambos componentes ancestrales en Ibagué se diferenciaron significativamente ( $p < 0.0001$ ).

Los mestizos de Córdoba y los Zenu presentaron una ancestría africana mayor (17 y 23% respectivamente) comparada con cualquiera de las poblaciones mestizas o indígenas del Tolima y Embera (<5%) ( $p < 0.0001$ ). La ancestría nativa entre Ibagué (~43%), Ortega (~51%), Planadas (49%) y los Zenu (~48%), no se diferenció significativamente ( $p > 0.05$ ); y dicho componente amerindio, fue menor en los mestizos de Córdoba (35%) que en cualquiera de las anteriores cuatro poblaciones ( $p$  para cada comparación en el rango:  $0.0002 \leq p \leq 0.0068$ ).

Con respecto a la ancestría europea presentada por los Zenu (~35%), Ortega (45%), Planadas (47%) y mestizos de Córdoba (42%) (Figura 20), no se evidenciaron diferencias significativas ( $p$  en el rango:  $0.0982 \leq p \leq 0.1224$ ); y en Ibagué, dicha proporción ancestral fue mayor que para cualquiera de las

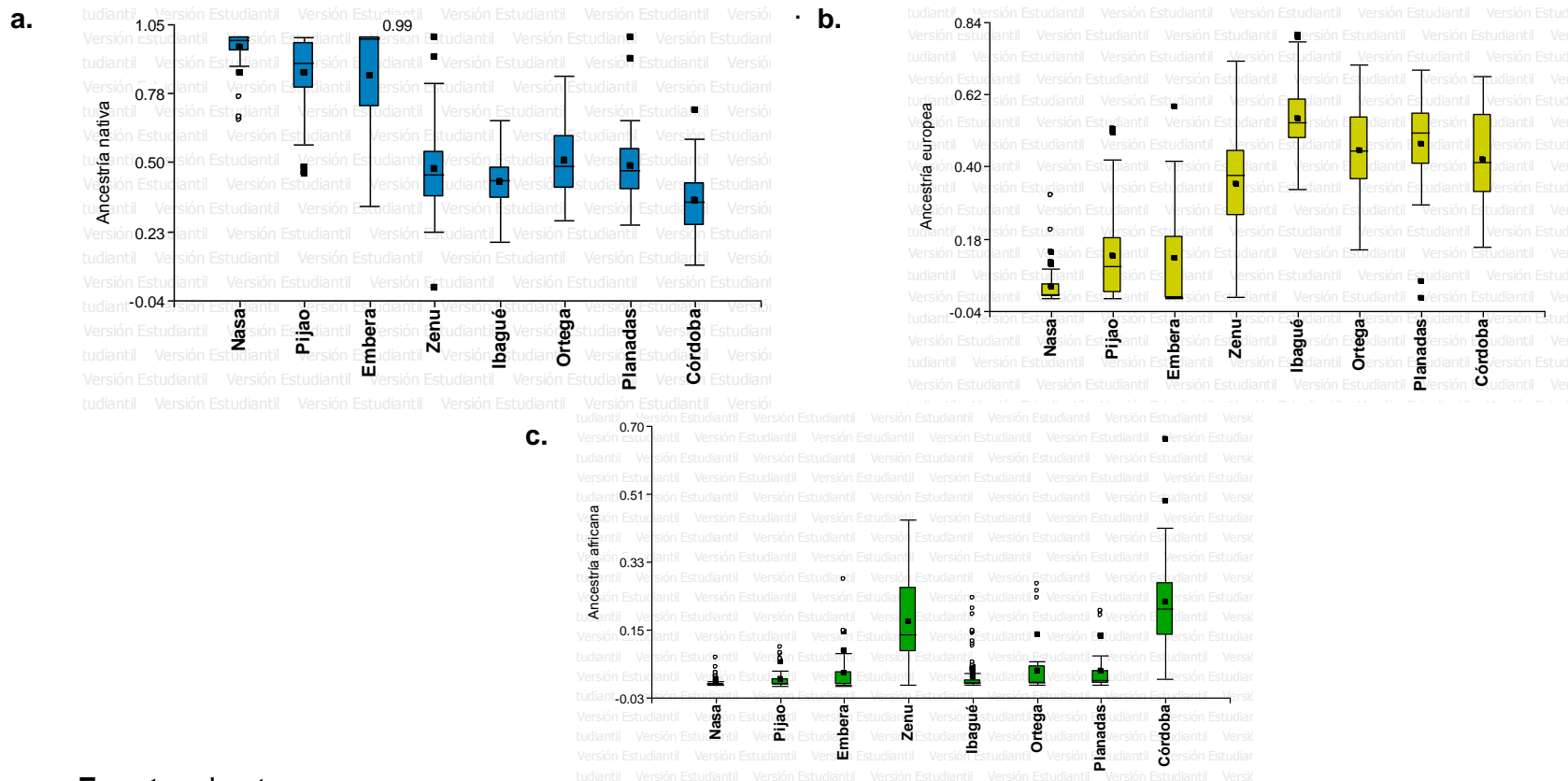
anteriores cuatro poblaciones (~55%) ( $p$  para cada comparación en el rango:  $0.0001 \leq p \leq 0.0117$ ).

**7.3.3** Ancestría individual. Se evaluó la dispersión y la desviación estándar (D.E) de las proporciones ancestrales nativa, europea y africana que fueron estimadas mediante el programa STRUCTURE a través de las frecuencias para 100AIMs, en cada una de las poblaciones del presente trabajo. Los diagramas *Box-plot* (Figura 21) permitieron observar que las poblaciones mestizas del Tolima, Córdoba y los Zenu presentan una amplia variación en las ancestrías nativa y europea hacia ambos lados de la media, como lo determina la posición de los valores máximos y mínimos. La variación en la ancestría africana tuvo efectos principalmente sobre la población Zenu y los mestizos de Córdoba (D.E afr: 0.131 y afr: 0.148) (Figura 21 y 22), lo cual estaría indicando un flujo reciente del aporte africano. En la población indígena Nasa no se observa una dispersión amplia de los datos en las proporciones nativas o europeas, como se puede apreciar, los cuartiles 1 y 3 están por encima del 0.9 y por debajo de 0.1 respectivamente, lo cual indica su homogeneidad, también evidenciado a través de los bajos valores de desviación estándar (D.E. eur: 0.058, nam: 0.062) (Figura 22 y anexo 11), cercanos a las poblaciones indígenas de Centroamérica y Suramérica.

En cuanto a los Pijao y Embera, los cuartiles 1 y 3 en la ancestría nativa están por encima de 0.78 y para la europea por debajo de 0.18, adicionalmente se nota la dispersión asimétrica hacia un solo lado de la media, lo cual fue más pronunciado en los Embera (Figura 22) y se corrobora porque fue la población con la mayor desviación estándar (D.E nam: 0.22, Eur: 0.185) (figura 22), indicando la presencia de individuos con alta ancestría indígena y otros con un componente europeo también alto.



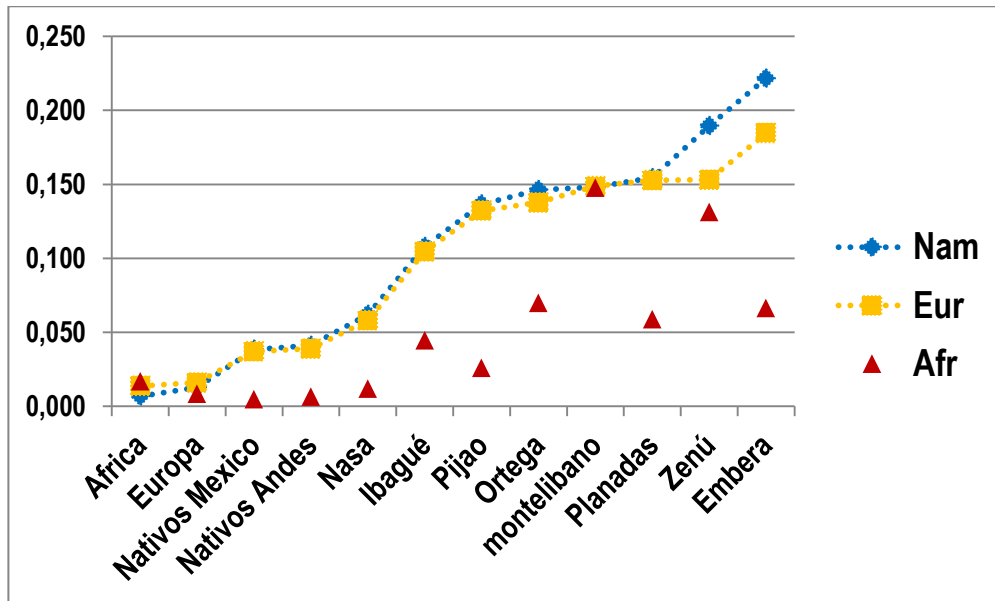
**Figura 21.** Diagramas *box-plot* para la representación de la dispersión de las proporciones ancestrales.



**Fuente:** el autor.

\* Ancestrías nativa (a), europea (b) y africana (c), estimadas en STRUCTURE mediante 100AIMs. Las gráficas se generaron en Infostat v.2008 y muestran la distribución de las proporciones ancestrales por cuantiles en las poblaciones indígenas Nasa, Pijao, Embera y Zenu; y en las poblaciones mestizas de Ibague, Ortega, Planadas y Córdoba. Aparecen también los valores mínimos y máximos (bigotes), el primer y tercer cuartil (cajas), el valor medio (línea media), y los individuos con valores extremos.

**Figura 22.** Desviación estándar de las proporciones ancestrales analizadas.



Fuente: el autor.

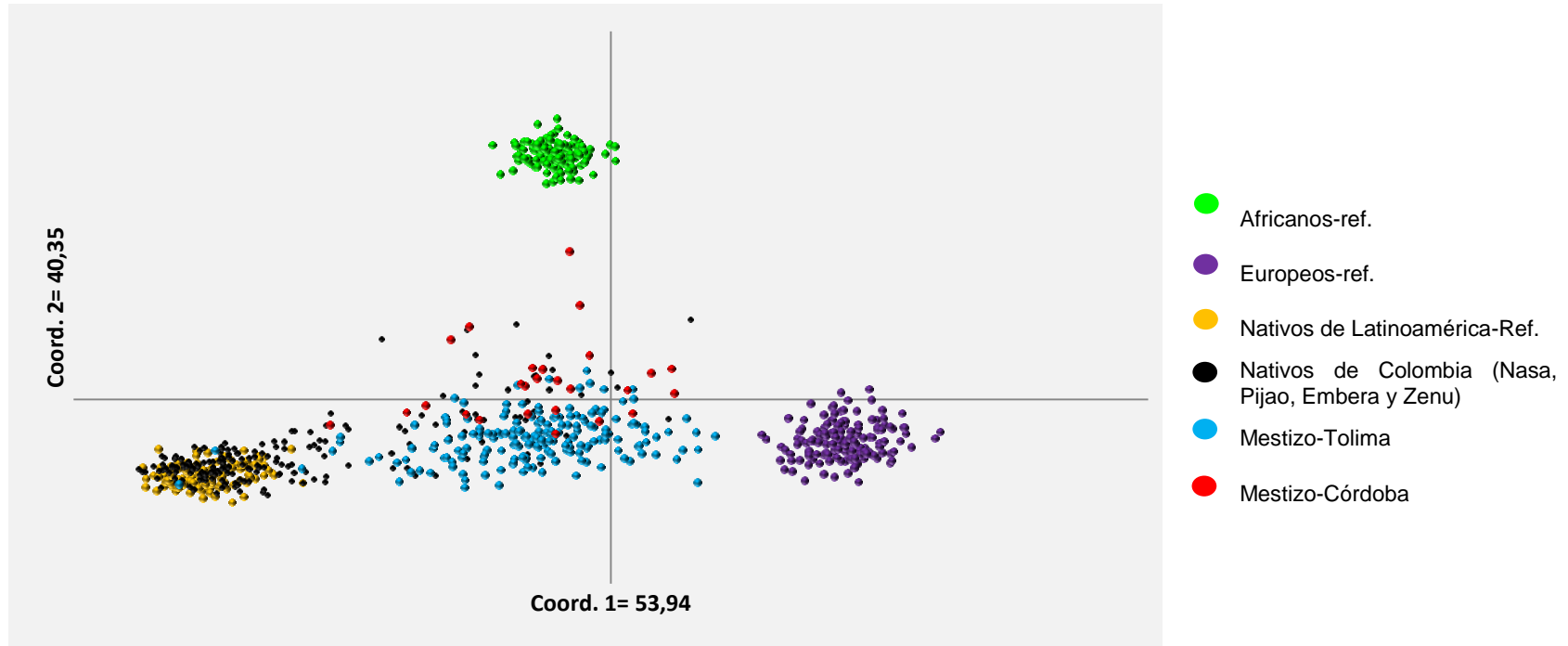
\*Se observa el comportamiento de la desviación estándar de las proporciones nativa, europea y africana en los individuos de cada población del presente estudio y de las poblaciones de referencia. (b) Tabla de datos a partir de la cual se generó la figura en a.

En Ibagué, se observó una distribución similar hacia ambos componentes ancestrales: nativo y europeo (Figura 21). Esto muestra que es una población con homogeneidad en las proporciones ancestrales (D.E. nam: 0.108 y eur: 0.105) (figura 30), que disminuyó en Ortega (D.E. nam: 0.146 eur: 0.138) y en Planadas (D.E. nam: 0.154 eur: 0.153).

Se realizó un análisis MDS, basado en la matriz de distancias genéticas pareadas individuo-individuo, el cual permitió ver la disposición relativa de cada individuo con relación a las tres poblaciones ancestrales simultáneamente. Se evidenció que las dos primeras coordenadas permiten una diferenciación clara entre los individuos de las tres poblaciones de referencia o ancestrales: africanos, europeos y nativos de LA, formando así grupos definidos (Figuras 23).

Los indígenas Nasa, Pijao y Embera, son cercanos genéticamente con los nativos de Centroamérica y los Quechuas de Perú como era de esperarse, mientras los mestizos del Tolima están distribuidos principalmente entre los europeos y nativos. En cuanto a los de Córdoba, se nota el efecto una tercera población ancestral, los africanos, como ya se ha observado. La tercera y cuarta coordenada no explican alguna diferenciación adicional entre las poblaciones del estudio (Figuras 23).

**Figura 23.** Análisis MDS mediante la matriz de distancia genética pareada individuo-individuo.



Fuente: el autor.

\* Calculos desarrollados a partir de los genotipos de 100 AIMs en los individuos de las poblaciones de referencia (Ref.) africanos, europeos y nativos de LA y las poblaciones de este estudio (Indígenas y mestizos). Se muestra los valores de Informatividad para los ejes 1 y 2.

## 8. DISCUSIÓN

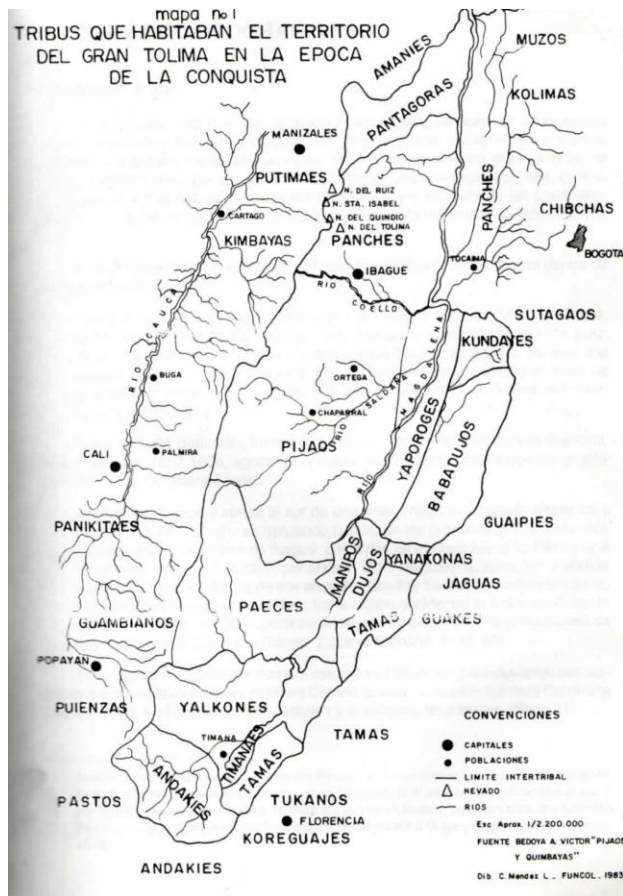
### 8.1 LAS POBLACIONES DE ESTUDIO

El territorio geográfico que hoy en día ocupa la región del Tolima, tiene una historia de 13 mil años. Se sabe por ejemplo, a través de excavaciones arqueológicas en los municipios de Chaparral, Roncesvalles y Fresno, que los primeros habitantes de la región existieron entre los años 11.000 y 3.500 a.C., también que hacia el año 1000 a.C. comenzó un proceso de transformación cultural que se caracterizó por la formación de pequeñas aldeas de agricultores y ceramistas que extrajeron oro de la cuenca del río Saldaña e iniciaron la orfebrería (Banco de la República, 2012). Durante este periodo prehispánico se desarrollaron varios grupos étnicos como los panches en el norte, pijaos centro-sur y paeces en el sur (figura 24). Dentro de los cuales, los pijao, probablemente ascendiente de la familia Karib o Caribe (Triana, 1992), han sido referenciados por los cronistas e historiadores de las indias, refiriéndose a ellos así:

*“...dueños de una agresividad que los conducían a no tolerarse ni siquiera a sí mismos, navegantes insignes, pescadores habilidísimos, guerreros afortunados, audaces e inteligentes, con un desarrollo físico superior a los demás grupos etnológicos de la América septentrional, raza llena de audacias y proyecciones militares, que asombran al investigador de nuestra protohistoria, pues mejor táctica guerrera no la tuvieron las dinastías incaicas y aztecas, modelos de conquistadores en el continente...”*

La exploración de la región –El alto Magdalena- empezó alrededor de 1550, con la fundación de Ibagué y posteriormente de otros centros urbanos, como Saldaña en 1664, Chaparral en 1769, entre otros, los cuales recibieron apoyo como asentamientos españoles en medio de un ambiente hostil y de resistencia nativa, como en ninguna otra región de Colombia. No existen datos claros acerca de la densidad poblacional indígena al momento del contacto ibérico, pero han quedado registros de los diferentes eventos que afectaron la demografía indígena, como por ejemplo las diversas plagas de viruela y sarampión que azotaron en repetidas ocasiones sus poblaciones, al igual que los suicidios colectivos y las guerras de sometimiento, o durante el trabajo en las minas y las rebeliones, en las cuales morían principalmente los hombres, mientras las mujeres sobrevivían (V. Bedoya, 1950; Triana, 1992).

**Figura 24.** Distribución de los grupos indígenas en el valle alto del Magdalena durante el periodo de la Conquista.



Fuente: Triana 1992.

Los historiadores del Tolima han estado interesados en el proceso de la mezcla ocurrido en la región, e incluso para 1950 ya se realizaban estimaciones empíricas de las proporciones de dicha mezcla:

*“La gran mayoría humana del Tolima y Huila, o llevan sangre del pijao en un 50% o la conservan pura con todas sus características psicofisiológicas, Natagaima, Coyaima, Chaparral, Ortega, Suarez, Cunday Melgar, Espinal, Coello, Venadillo, Ambalema, Mariquita, y otros centros correspondientes al Huila y el valle del Magdalena y la costa atlántica, constituyen núcleos de pálido y del indio pijao...”* (V. Bedoya, 1950).

También hay referencias de la dirección en que dicha mezcla ha ocurrido, principalmente entre mujeres nativas y hombres españoles, estableciendo un nuevo grupo, sin identidad ibérica o nativa. Después del sometimiento y el triunfo de gobierno español, el poblamiento del Tolima obedeció a políticas de segregación social, implicando la fundación de pueblos indios y la delimitación de resguardos, prohibiendo a los indígenas construir en centros urbanos recién

fundados, en los cuales solo vivían, o bien los españoles, o criollos, o los mestizos con reconocimiento paternos (Rueda, Henrik, & Orlando, 1996; Triana, 1992).

Estas circunstancias sociales han tenido un profundo efecto en la estructura de la población del Tolima en la actualidad, así hoy en día el 4.3% de la población reconoce su identidad cultural ancestral indígena y un 94.5% que no tiene alguna afinidad étnica. Por otro lado los nativos no son el único grupo étnico reconocido, también están los afrocolombianos. Durante los periodos pre y post coloniales, la región no fue permeada masivamente por la afluencia de africanos traídos como esclavos, ya que sus asentamientos principales se localizaron en las costas pacífica y atlántica, pero, posiblemente debido a oleadas migratorias recientes, hoy en día se aprecia que un 1.2% de los tolimenses son afrodescendientes (DANE, 2005).

Si bien, por un lado, la evidencia histórica aclara muchos aspectos del origen de los pobladores de la región del Tolima, y por otro, existen algunos estudios de la diversidad genética indígena en la región (Rondon et al., 2008; Rondon et al., 1999), no es claro cuáles han sido las consecuencias del contacto europeo en los tiempos pre y postcoloniales en términos de la estructura genética poblacional; por ejemplo, si ha ocurrido remplazo o continuidad de los patrones de diversidad nativos. Por esta razón mediante la presente investigación, se ha indagado desde el punto de vista genético, (1) la naturaleza de la variación en la mezcla genética, (2) el rol de las características geográficas de la región en la estructuración de su diversidad y (3) la existencia o no de diferenciación poblacional. Así mismo se ha comparado con otras regiones culturalmente diferentes para analizar la influencia de la variación en la historia demográfica.

Por las anteriores razones, el presente estudio involucró 3 poblaciones no-indígenas del Tolima, en distintas zonas geográficas -eje sur-centro- del Tolima, y los grupos étnicos cercanos: por un lado los Pijaos con una lengua extinta debido las condiciones sociales post-colombianas (V. Bedoya, 1950; Durbin & Seijas, 1973; Triana, 1992) y los Nasa-Paez o paces, con su lengua actualmente vigente, el *Naza-Yuwe* (Ecue, 1989; Gardner, 1978; Gerdel, 1989; Yule-Yatacué, 1996). Esto ha permitido por una parte discutir las relaciones y el flujo genético entre dichas poblaciones, así como también evaluar el proceso de continuidad genética y analizar la influencia de la mezcla con relación a la conservación de la lengua nativa, pues como se sabe, la estructura genética de las poblaciones humanas depende en una medida de factores socioculturales.

Adicionalmente, se realizaron comparaciones con poblaciones indígenas y mestizas de otra región geográfica (Córdoba), para observar la influencia de historias demográficas diferentes, así como también patrones similares de características e influjos culturales externos, sobre la diversidad y la estructura genética. Los grupos indígenas de Córdoba fueron: La etnia Embera, del tronco lingüístico Chibcha con una lengua nativa aun presente (Mortensen, 1999; Pardo,

1984); y los Zenúes quienes ya la perdieron (Cardona, 2001). También se escogió una población mestiza del mismo departamento (Montelíbano).

Se seleccionó el departamento de Córdoba por pertenecer a la región Caribe, con población afrocolombiana presente con una densidad del 13%, mayor que en el Tolima (2%) (DANE, 2005). Por otro lado, el departamento de Córdoba alberga un grupo indígena aislado, en términos culturales, del resto de las etnias indígenas y población mestiza, probablemente debido a políticas tribales que le han permitido sobrevivir a lo largo de los últimos 500 años; dichas circunstancias socioculturales, comparables con las presentadas por los Nasa-Paez. Adicionalmente, este departamento alberga una etnia bajo condiciones cercanas a los Pijao del Tolima: es decir con autoreconocimiento pero sin lengua nativa vigente. Las similitudes o diferencias socioculturales mencionadas entre ambas regiones permiten, por ejemplo, realizar inferencias acerca de la relación entre el grado de mezcla y aspectos antropológicos como la supervivencia de tradiciones y la lengua e, incluso, discutir el posible papel de la organización tribal en la conservación de la etnia, en vista de que ambos factores pueden estar relacionados. Por último, con este panorama demográfico y cultural descrito, entonces se pueden empezar a discutir los hallazgos genético-poblacionales.

## **8.2 DIVERSIDAD GENÉTICA Y DIFERENCIACIÓN POBLACIONAL EN EL TOLIMA Y CÓRDOBA**

Antes de abordar las poblaciones del presente estudio, es necesario hacer una introducción a los factores demográficos que han afectado las poblaciones de la región latinoamericana. El continente americano ha atravesado por 5 fases históricas distintas, las cuales son el marco para un modelo integrativo que explique la evolución morfológica y genética en las poblaciones nativas americanas y mezcladas, cada una de las cuales ha dejado su huella en la arquitectura de su diversidad genética (Long & Bortolini, 2011).

La primera fase ocurrió antes de la expansión geográfica de la especie humana y la formación de poblaciones continentales locales; como consecuencia, existen unos polimorfismos comunes a todos, incluyendo los indígenas americanos. La segunda fase fue el efecto fundador ocurrido hace 40-80 mil años, en el momento en que los humanos salieron de África para colonizar todas las regiones del planeta, lo cual afectó profundamente las poblaciones no-africanas (Ramachandran et al., 2005).

Tercero, la evidencia arqueológica, anatómica, lingüística y genética ha demostrado que los habitantes originales de América llegaron del noreste de Asia durante el Pleistoceno tardío, luego de la última gran glaciación hace 15-30 mil años (Brace et al., 2001; Goebel, 1999; Goebel, Waters, & O'Rourke, 2008; Greenberg, Turner, & Zegura, 1986; Wallace, Garrison, & Knowler, 1985), y

constituían una pequeña población fundadora, de la cual descienden todos los indígenas actuales (Kitchen et al., 2008; Tamm et al., 2007). América fue el último continente en ser colonizado por los humanos modernos y muy probablemente los ancestros asiáticos llegaron a través del estrecho de Bering, la porción de tierra que conecta Asia y América en algunos periodos estacionales, los asentamientos más antiguos de Norte y Sur América datan desde 14.500 años (Shastri, 2007a).

Cuarto, el proceso que condujo al desarrollo de nuevas culturas y lenguas, restringiendo el contacto entre diferentes grupos poblacionales indígenas, disminuyendo así el flujo genético y por lo tanto incrementando la diferenciación entre los grupos (Hunley & Long, 2005; Hunley, Spence, & Merriwether, 2008). La diversidad generada se debe en parte a que el continente se expande hacia ambas latitudes y comprende todas las zonas ecológicas, de igual manera muchas barreras naturales que impiden el libre movimiento de las personas. Por lo tanto, es probable que tal diversidad cultural y lenguas surgieran *In-Situ*, durante el curso de crecimiento poblacional y la adaptación a diversos ambientes (Fagundes, Kanitz, & Bonatto, 2008; Goebel et al., 2008).

Quinto, el contacto europeo, causó casi el colapso total de la población nativa, muchos grupos enteros desaparecieron, ocurrió fusión de etnias y mezcla genética con grupos no-indígenas (Mulligan et al., 2004). Cada una de estas fases ha tenido profundos efectos, incluso con la posibilidad de haber transformado o eliminado las huellas genéticas de sucesos antiguos.

No obstante, a pesar de la variedad de ambientes y la riqueza cultural, los amerindios portan niveles bajos de diversidad genética, atribuido a la existencia de un efecto fundador durante las etapas tempranas del poblamiento (Santos et al., 2007) a partir de las migraciones asiáticas hacia Las Américas (Mulligan et al., 2004; Torroni et al., 1992; S. Wang et al., 2007) y el ocurrido por el contacto europeo (O'Fallon & Fehren-Schmitz, 2011).

**8.2.1** Origen materno de la diversidad. Las variaciones en las frecuencias haplogrupales del mDNA de los nativos analizados en el presente estudio, fueron consistentes con los patrones continentales de distribución; por ejemplo, el A es más frecuente en la parte norte de Suramérica, y prácticamente desaparece en la Patagonia, mientras que el D es altamente frecuente en los pueblos nativos más australes. Por otro lado, los Hgs B y C no tienen un patrón latitudinal, y son frecuentes en los Andes y la Amazonía (Bisso-Machado et al., 2012). Reportes anteriores en los Pijao y Embera son consistentes con lo hallado en el presente estudio, predominando los linajes en un orden de menor a mayor frecuencia: A, B y C (Mesa et al., 2000; Rondon et al., 2008; Rondon et al., 1999).

Uno de los factores que más incrementó la diversidad haplogrupal mitocondrial en las poblaciones fue la aparición de linajes distintos a los fundadores amerindios. En diferentes estudios se ha observado que la mezcla europea y africana incrementa la diversidad mitocondrial (Salas et al., 2008). Por otra parte, las



poblaciones menos diversas como los Nasa y Embera, aparecen distantes de las demás poblaciones en el análisis MDS. Los Nasas presentaron un índice de diversidad similar a los reportados para otros grupos indígenas como los Mayas (Bailliet et al., 1994), en tanto que el de los Embera se asemeja al de comunidades como los Waunana (Análisis MDS, anexo 12), con quienes de hecho comparten la misma distribución geográfica y la familia lingüística. Sin embargo, el índice de diversidad no es tan bajo como el de los Kogui y Nukak, poblaciones aun más aisladas en la Sierra Nevada y en el Amazonas respectivamente (Keyeux et al., 2002). La razón de la baja diversidad en los Embera es consecuencia de la ausencia de los Hgs C y D entre otros, lo cual ya ha sido reportado previamente (Mesa et al., 2000).

Los mestizos del Tolima y los Pijao son más diversos que los indígenas Nasa o Embera, y están estrechamente relacionados entre sí desde el punto de vista mitocondrial. Esto es consecuencia de que por un lado, comparten proporciones cercanas de Hgs amerindios fundadores y por otro, el Hg L africano presenta frecuencias similares en ambos grupos. Es importante destacar que existe una cercanía con otras poblaciones mestizas de Colombia o Latino América, por ejemplo con los mestizos de Cali (Salas et al., 2008), Bogotá (Rodas et al., 2003), México (Carvajal-Carmona et al., 2003), Argentina (Catelli et al., 2011), lo cual se evidenció mediante el patrón de agrupamiento en análisis multidimensional.

Finalmente, los mestizos de Montelíbano y los Zenú presentaron los valores de diversidad haplogrupal mitocondrial más altos y cercanos entre sí, lo que implica también una cercanía genética de ambos grupos poblacionales en Córdoba, como se aprecia en el análisis MDS que incluye sólo las poblaciones del estudio. La alta frecuencia de Hgs no determinados, encontrada en ambas poblaciones también pudo aportar a la diversidad observada. Es posible que otros linajes, no analizados en el presente estudio, hagan parte de estas fracciones desconocidas; dentro de los cuales podría estar el Hg europeo H que puede alcanzar >40% en algunas poblaciones latinoamericanas como en Cuba o Brasil y de igual forma el Hg africano U (Alves-Silva et al., 2000; Mendizabal et al., 2008).

Se ha observado que la diversidad haplogrupal mitocondrial en las poblaciones mestizas es cercana a la presentada por las etnias que históricamente han habitado la misma región, por ejemplo la diversidad y perfil haplogrupal en la población antioqueña es similar a la presentada por los Embera, lo cual indica una continuidad en los patrones de frecuencias entre las poblaciones (G. Bedoya et al., 2006; Carvajal-Carmona et al., 2000). Por otro lado, es interesante notar que en el análisis MDS en el que se comparó con varias poblaciones colombianas, la población mestiza de la región de Córdoba y los mestizos de Quibdó mostraron una relativa cercanía, lo cual puede ser el reflejo de una misma historia materna, ya que estas poblaciones comparten la misma ecoregión del Darién y departamentos cercanos (Choco, Córdoba y Antioquia), probablemente influenciados por patrones de frecuencias de linajes mitocondriales nativos

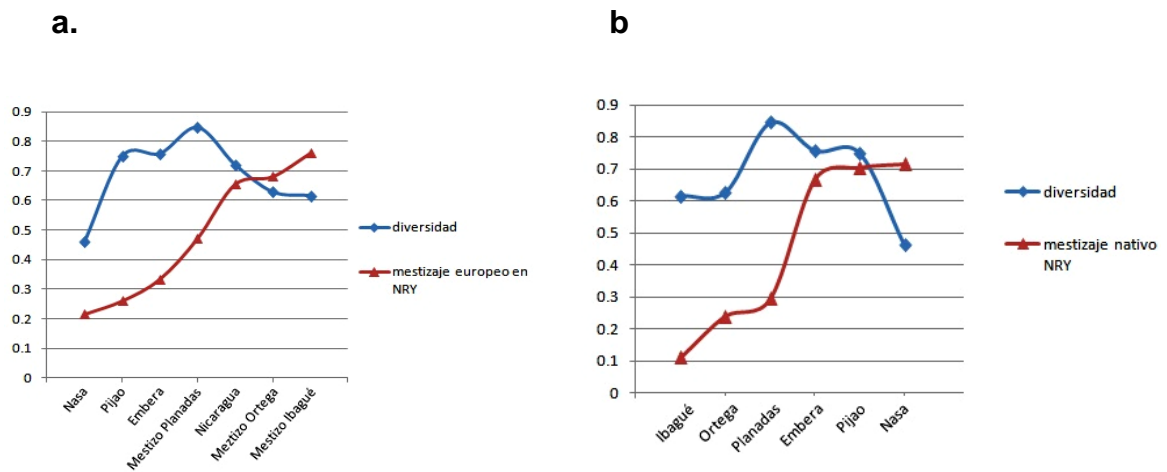
locales. Un aspecto interesante del análisis de la diversidad del componente matrilineal revela que la menor diversidad coincide con aquellos grupos donde muchos aspectos socioculturales aun tienen arraigo, por ejemplo, la conservación de la lengua nativa entre otras. Mientras que aquellas comunidades con un nivel mayor de diversidad existe menos posibilidad de encontrar dicha conservación, posiblemente por el efecto de la mezcla y el contacto con la cultura ibérica o afrodescendiente.

**8.2.2 Origen paterno de la diversidad.** La diversidad de origen paterno de las poblaciones analizadas está relacionada con el número de Hgs NRY y la distribución de las frecuencias de los mismos (Karafet et al., 2002 ). Las poblaciones más diversas fueron los mestizos de Planadas y Córdoba, dada la representación haplogrupal amplia y de varios orígenes: europeo (R1b1b2, I), medio euroasiático (J, G), africano (E) y nativo americano (Q), y, en el caso particular de Planadas, unos individuos a los cuales se les identifico como pertenecientes al clado K-R pero sin poder definir el Hg específico. Con relación a estos linajes paternos, los mestizos del Tolima, de Córdoba y los indígenas Zenu reflejan una estructura haplogrupal similar, no obstante existen diferencias como la presencia exclusiva en el Tolima de un clado K-R o el mayor componente africano en los mestizos de Córdoba.

La historia paterna de los indígenas –excepto los Zenu- tiene una menor influencia de Hgs distintos al Q, independientemente de la región que se analizó. En los nasas, Pijao y Embera, la mayor frecuencia del Hg Q hace de estas, las poblaciones menos diversas, ya que si bien existen otros Hgs no nativos como el I, J y el R1b1b2, entre otros; están en bajas frecuencias con relación a Q. Lo anterior refleja el agrupamiento cercano de las tres poblaciones indígenas mencionadas, observado en el análisis MDS, lo cual indica una completa diferenciación con respecto a las mestizas. Recientes estudios, basados en datos genéticos de sistemas haploides han mostrado que el aislamiento regional puede disminuir la diversidad y causar diferenciación poblacional como la observada entre los nativos andinos –Quechua y Aymara- y grupos amazónicos aislados como los Karitiana o Surui, donde la diversidad es baja. No obstante parece ser que la mezcla europea está aportando a la diversidad de las poblaciones andinas (Hunley & Healy, 2011). En este sentido, por ejemplo en el presente estudio se observó una relación entre la diversidad haplogrupal y la contribución de linajes europeos o euroasiáticos en las poblaciones que no tienen otros componentes importantes como el africano en los Zenues o mestizos de Córdoba (figura 25). El incremento gradual en la fracción europea o nativa en las poblaciones del Tolima (adicional a una población externa) está relacionado también con un incremento en la diversidad, sin embargo al aumentar cualquiera de las dos ancestrías mas allá de los niveles encontrados en Planadas, la diversidad empieza a decrecer, por lo tanto, se puede decir que en esta población existe un balance entre ambas ancestrías que permite los máximos niveles de diversidad encontrados en el presente estudio.

A diferencia del mtDNA, en el caso del cromosoma Y no hay una continuidad clara entre los patrones de frecuencias de linajes nativos en el Tolima y sus correspondientes poblaciones mestizas. Por ejemplo, aunque los Nasa y mestizos de Planadas se encuentran en la misma área geográfica, existe un perfil de frecuencias diferente que indica que son poblaciones distantes desde el punto de vista de la historia paterna, lo cual se evidencia en el análisis multidimensional basado en las distancias pareadas  $F_{st}$ . En otras poblaciones de Latino América también se observa una discontinuidad genética entre los indígenas y mestizos (Corach et al., 2010; Toscanini et al., 2011). En Argentina y Nicaragua, la población mestiza presentan niveles de diversidad cercanos y el mismo conjunto de linajes africanos y euroasiáticos (Corach et al., 2010; Núñez et al., 2012), similar a los mestizos del Tolima y Córdoba, por esta razón aparecen cercanos en el análisis multidimensional.

**Figura 25.** Relación de la diversidad haplogrupal y la proporción de mezcla.



Fuente: el autor.

\*Comparaciones para la mezcla europea (a) y (b) nativa observada en el cromosoma Y. A medida que la ancestría nativa o europea incrementa, también ocurre un incremento de la diversidad, sin embargo, después de un valor límite (diversidad en Planadas), la diversidad empieza a disminuir.

A pesar de las diferencias resaltadas arriba, ambos sistemas genéticos coinciden en que las poblaciones indígenas: Nasa, Embera y Pijao, tienen diversidades bajas, contrario a las mestizas, pero particularmente los de Córdoba tienen los valores mayores, todas estas observaciones tienen un factor común, el grado de mezcla genética, que se volverá a discutir más adelante.

**8.2.3** Diversidad autosómica (AIMs). Los patrones de distribución de la diversidad entre las poblaciones mestizas e indígenas del Tolima mostraron historias materna y paterna diferentes, aunque con patrones comunes como ya se vió. Estos

marcadores son esencialmente útiles para definir los linajes fundadores, sobretodo en poblaciones con una historia reciente de mezcla genética, dado que tanto el mtDNA como la región NRY, se comportan como un locus único, o un supergen, ya que al no recombinar el ligamiento es total.

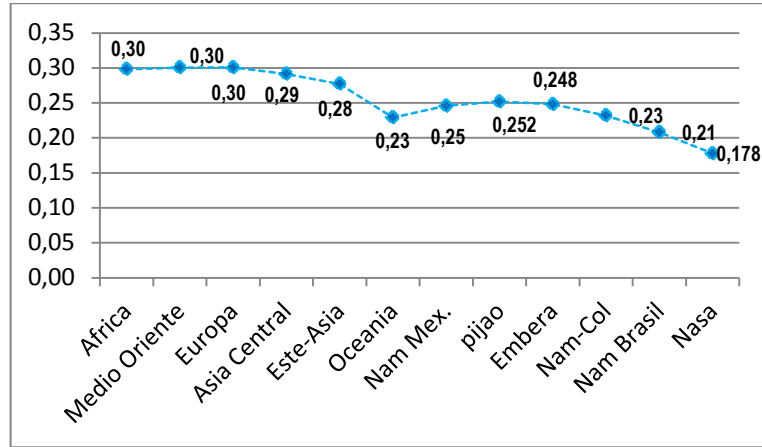
Los 100 AIMs, a diferencia de los marcadores uniparentales, ofrecen un panorama genómico global de la variación, pues están distribuidos por todo el genoma humano. El patrón de diversidad obtenido con los SNPs autosómicos define e integra con alta resolución, algunos patrones de diversidad detectados aisladamente por línea materna o paterna. Por ejemplo, todas las poblaciones indígenas –excepto los Zenues- muestran reducidos niveles de diversidad con relación a los mestizos; este patrón compartido de los indígenas los hace cercanos, formando un grupo separado y estrechamente vinculado entre sí, como lo demuestra el análisis MDS y las distancias *Fst*. Esta reducción de diversidad se puede haber originado como consecuencia de la deriva ocasionada por el aislamiento en regiones geográficas apartadas, como se ha observado en otros grupos indígenas de Suramérica (Catelli et al., 2011).

Por ejemplo en el casos de los nasas y emberas en las montañas del sur del Tolima -en el corregimiento de Gaitania-, y las cuchillas montañosas del sur de Córdoba respectivamente. Sin embargo otro factor importante pudo ser el aislamiento mediante la segregación cultural en los Pijaos, como medio de resistencia al sometimiento y el contacto con inmigrantes europeos, una vez los medios defensivos físicos de los indígenas no lograron evitar la invasión (V. Bedoya, 1950; Triana, 1992).

Estudios recientes indican que existe un patron global consistente con un decrecimiento gradual de la heterocigocidad cada vez que las poblaciones se alejan de África (Li et al., 2008). En este sentido, al comparar con los datos de algunos de estos estudios para poblaciones en África, el este de Asia, Oceanía (Papúa y Melanesia) y nativos americanos de Centro América (Mayas y Pimas) y Suramérica (Karitiana, Surui, y nativos de Colombia de origen no reportado); es interesante notar que las heterocigocidades de los Pijao, Nasa y Embera son cercanos con los patrones de otros amerindios, correspondiendo al patron a escala global (figura 26).

Éste explica que la tendencia puede ser consecuencia de un efecto fundador serial ocasionado por la historia de las expansiones poblacionales pasadas de la humanidad, producto de migraciones sucesivas de una pequeña fracción de individuos desde una locación previa, iniciadas en África hace 100 mil años, por parte de un grupo pequeño, el cual habría seguido un patrón geográfico en cadenas secuenciales de colonias cada vez más periféricas y jóvenes (Deshpande, Batzoglou, Feldman, & Luca Cavalli-Sforza, 2009; McVean, 2002; Ramachandran et al., 2005).

**Figura 26.** Tendencia global de la heterocigocidad.



Fuente: el autor.

\* Comparación de la heterocigocidad en los Pijao, Embera y Nasa, con relación a otros amerindios y poblaciones del mundo. Los datos de las demás poblaciones proceden de trabajos publicados por otros autores (Li et al., 2008).

Los mestizos de Ibagué, Ortega y Planadas, constituyen un grupo diferenciado, con diversidades intermedias, incluso es el de mayor cohesividad y cercanía genética, indicando que no existe algún proceso de estratificación entre ellos (análisis MDS y distancias Fst), mientras que los mestizos de Córdoba y Zenu, que tienen las diversidades más altas, constituyen un tercer grupo distante de los mestizos del Tolima y de los indígenas Nasa, Embera y Pijao.

Por ahora, no hay duda de que las poblaciones mestizas del Tolima y las indígenas tienen en común su propia historia fundacional ancestral, desde el punto de vista de la diversidad en los tres sistemas polimórficos evaluados. Así, la resultante de la interacción de las historias genéticas matrilineal y patrilineal, mediante el análisis de los AIMs, ha producido tres grupos poblacionales bien diferenciados por los grados de diversidad y corroborado por el análisis de las distancias pareadas con las que se elaboró el análisis multivariado, en orden de mayor a menor: indígenas (Embera, Nasa y Pijao) < mestizos Tolima < Mestizos Córdoba+zenu.

### 8.3 ORÍGENES ANCESTRALES, MEZCLA Y ESTRUCTURA GENÉTICA DE LAS POBLACIONES ANALIZADAS

**8.3.1** Linajes maternos y paternos en el Tolima. Solo tres poblaciones amerindias - Nasa, Pijao y Embera-, de las cuatro analizadas en este estudio, poseen una característica distintiva que consiste en una mayor frecuencia de Hgs nativos tanto del mDNA como del cromosoma Y. Estos linajes, al ser subconjuntos de la variación asiática (Schurr & Sherry, 2004b) corroboran, por un lado, el posible origen asiático ancestral –Siberia Central-, compartido por todos los nativos americanos (Dulik et al., 2012) y también en una medida, pueden explicar en parte la baja diversidad encontrada. Ya que, como se ha demostrado en muchos estudios, durante la migración a través de Bering se pudo generar un efecto de cuello de botella, ocasionando la baja diversidad observada en la mayoría de las poblaciones amerindias en Centro y Suramérica (Bisso-Machado et al., 2012). Sin embargo quizás no sea la única razón, pues también se conoce que la disminución del tamaño poblacional por el contacto europeo (debido a pestes, guerras, esclavitud, etc), igualmente pudo generar otro efecto de cuello de botella en los grupos nativos (O'Fallon & Fehren-Schmitz, 2011) así, el segundo evento pudo oscurecer el efecto del anterior, lo cual es difícil de establecer, dado que se requeriría de análisis adicionales, como por ejemplo con muestras de DNA antiguo de varias poblaciones y en un número apropiado de muestras para poder comparar con la estructura moderna (Raff et al., 2011), teniendo en cuenta modelos de mezcla en las últimas 20 generaciones. En algunos casos, se conoce que la mezcla incrementa la diversidad (Salas et al., 2008), como se ha observado con lo incidencia de linajes maternos africanos en los pijaos del presente estudio; o también el efecto de la ancestría paterna africana y europea en la diversidad de los mestizos de Córdoba y Zenues.

Los grupos indígenas que ancestralmente han habitado el Tolima, de los cuales persisten hoy los Nasa y Pijao son, en general, similares genéticamente y, en conjunto, cercanos a los emberas. La presencia en alta frecuencia de Hgs NRY europeos en los zenues y la ancestría mitocondrial africana materna, hacen de este grupo indígena el más diverso y también el de mayor mezcla con ambos componentes ancestrales, condiciones que ocasionan la alta diferenciación encontrada con relación a las demás poblaciones indígenas o mestizas del Tolima (Anexo 13).

Un factor común entre nasas, pijaos y emberas es la presencia de un componente paterno no-nativo alto (entre ~29 y 34%), en su mayoría del acervo genético típico europeo (R1b1b2 y R1 o R1a) y euroasiático (I y J), por ser frecuentes en la región del Cáucaso, en Medio Oriente (Karafet et al., 2008), y ausentes en los nativos americanos. Su presencia en las etnias analizadas probablemente se deba a la mezcla ocurrida con los españoles, en algún momento de la historia de cada población, posterior a 1492, ya que estos Hgs hacen parte de la diversidad ibérica actual (Adams et al., 2008). Por otro lado, el linaje Q (definido por la mutación derivada M242) fue el más frecuente -entre un ~66 y 71%- y es el característico de las poblaciones nativas americanas por predominar en altas frecuencias (Zegura, Karafet, Zhivotovsky, & Hammer, 2004), por ejemplo en las mexicanas

(~85%) (Sandoval et al., 2012), al igual que en algunas poblaciones siberianas (Malyarchuk et al., 2011; Mulligan et al., 2004) y relativamente ausente o en bajas frecuencias en otras poblaciones del mundo (Karafet et al., 2008), por lo anterior, se profundizó en los sublinajes de este clado.

La presencia de los subHgs Q1a3\* (mutación derivada en M242 y en M346 pero ancestral en M3) y Q1a3a\* (mutación derivada en M242, M346 y M3 pero ancestral en M19, M194, y M199) (Karafet et al., 2008) revela el origen paterno nativo principal en los Nasa, Pijao y Embera. De los dos sublinajes, Q1a3a\* fue el más frecuente, y está presente en las 4 etnias analizadas, como ya se ha reportado para muchos otros grupos amerindios (Bisso-Machado et al., 2012; Blanco-Verea et al., 2010).

La presencia exclusiva de Q1a3\* en pijaos y emberas (37 y 25% respectivamente), puede estar indicando una mayor estructuración y diversidad. Al compararse con otras poblaciones, se observa que también está presente en nativos mexicanos (46%) y quechuas en Perú y Bolivia (14%) (Sandoval et al., 2012), lo cual puede ser indicativo de un gradiente de distribución de mayor a menor frecuencia entre centro y Suramérica. Su presencia también en las poblaciones de Altai (25.8 %), Todjin (38%) y Tuvinia (38%) en el sur de Siberia, pero no en poblaciones del este de Asia como los grupos Mongoles (Malyarchuk et al., 2011), sugiere el posible origen de los nativos americanos en estas poblaciones siberianas, con evidencias también a nivel del mDNA (Dulik et al., 2012).

A pesar de la resolución para definir los Hgs fundadores nativos de las poblaciones amerindias del presente estudio, es necesario poder profundizar acerca de la heterogeneidad al interior de cada sublinaje y caracterizar los haplotipos acompañantes de cada uno, mediante microsatélites en la región NRY, como se ha hecho en otras poblaciones, para poder tener una mayor resolución de los patrones de diversidad y factores evolutivos como la deriva, que afectan la diferenciación poblacional (Palha et al., 2012; Sandoval et al., 2012; Toscanini et al., 2011).

Los principales linajes maternos en las poblaciones indígenas del Tolima y los Embera son de origen nativo. Esta ancestría materna común y uniforme para indígenas, se presenta como un entorno genético de poca mezcla europea o africana en el mDNA. La primera implicación importante es que a pesar del contacto europeo, se han conservado los linajes fundadores amerindios, sin embargo, como lo que se ha observado corresponde a las condiciones actuales de las tres poblaciones, de acuerdo con la metodología y objetivos del presente trabajo, no se puede conocer en qué medida este evento afectó las frecuencias, la diversidad o los sublinajes que estaban presentes en dichas comunidades en periodos prehispánicos. A pesar de esto, algunos estudios en mDNA antiguo sugieren que las observaciones realizadas hoy en día en muchas poblaciones nativas de América, no difieren de la estructura haplogrupal antigua (Raff et al.,

2011) y esto se ha corroborado comparando las frecuencias haplogrupales del mDNA entre poblaciones actuales y en restos humanos con aproximadamente 2000 años de antigüedad de la etnia Guane en el nor-oriental colombiano (Casas-Vargas et al., 2011). La disminución de los linajes nativos del mDNA y del Cromosoma Y en los Zenues y no-indígenas de Córdoba está relacionada con el mezcla africana, dada su cercanía con puertos que históricamente sirvieron como puerta de entrada para los africanos que llegaban desde las colonias españolas en África occidental (Salas et al., 2004).

La historia es diferente para los mestizos del Tolima, que no han experimentado un influjo africano reciente, lo cual es además evidente al revisar las cifras reportadas por los censos nacionales, donde tan solo el 1.2% de la población tolimense se autodenomina como afrocolombiana. A pesar de que la población indígena del Tolima ocupa una fracción regional <5%, aproximadamente ~93% de los mestizos son portadores de un Hg mitocondrial de origen nativo, estos datos sugieren que el aporte de las mujeres indígenas a la formación de las poblaciones tolimenses a lo largo de la historia puede tomarse como el principal.

Desde el punto de vista paterno, en los mestizos del Tolima, predominan los linajes típicos europeos o ibéricos (~50.1%) –principalmente R1b1b2- y euroasiáticos (~16.7%), en un orden decreciente de frecuencias así: R1b1b2 >J>I>K-R>G, considerando algunas variaciones de presencia o ausencia del G. En Ibagué y Ortega son similares los perfiles de frecuencias, los cuales revelan una historia común, caracterizada por una mayor influencia europea dado la facilidad de acceso geográfico a lo largo de la historia. Mientras que en zonas como las que habitan los mestizos de planadas, probablemente por estar aisladas de centros urbanos, se ha visto disminuido un poco el influjo de Hgs como el R1b1b2 a la vez que se incrementa componente nativo (Q1a3).

Dos aspectos interesantes que explican este relativo aislamiento en Planadas son, por un lado el aporte africano alto, y por el otro una fracción de >10% en Hgs del clado K-R, que no son Q o R dado que sus portadores presentaron las mutaciones ancestrales en M242 y M173 respectivamente, pero tampoco son O, L, o N porque para todos estos también presentaron la mutación ancestral; entonces, al tener la mutación derivada en M9, que determina el clado K-R, estos individuos pueden ser posiblemente de los linajes euroasiáticos K o T, tal y como se ha encontrado en otras poblaciones mestizas como en México, Nicaragua y Argentina (Corach et al., 2010; Martínez-Cortés et al., 2012; Núñez et al., 2012) donde también su frecuencia es similar. Este clado –K-R- a su vez parece que presenta una distribución clinal Sur-Norte, ya que se encuentra en una frecuencia intermedias en Ortega (4%) y menor del 2% en Ibagué.

La mayoría de los linajes paternos del Tolima de origen ibérico o euroasiático, posiblemente pudieron llegar a la población tolimense mediante los españoles, dado que la distribución de frecuencias haplogrupales actual en esta población presenta una relación similar a la población del Tolima, entre otras de Latino



América, que estuvieron bajo la influencia ibérica (Adams et al., 2008; Núñez et al., 2012), naturalmente tomando en cuenta la ausencia del componente nativo. Por otro lado los españoles no-R1b, al momento del arribo a las Américas, hace algo más de 500 años, ya venían de un proceso de convivencia con musulmanes moros, judíos y norte-africanos durante un periodo de aproximadamente 700 años en el Medioevo europeo, de manera que posiblemente por mezcla con estas poblaciones, obtuvieron Hgs típicos de la región de Medio Oriente, Cáucaso y de la costa nor-oriental del mediterráneo, como Turquía o Grecia (Adams et al., 2008; Bryc et al., 2010).

No obstante, al revisar algunas historias personales de participantes en el presente proyecto, y referenciadas en las entrevistas, la presencia de una fracción de linajes euroasiáticos puede tener otro origen igualmente válido; como una inmigración directa desde el medio Oriente, no obstante solo verificable en el momento de realizar análisis de microsatélites. Se conoce que a principios del siglo XX, cuando llegaba a su fin el imperio Otomano, una oleada pequeña de inmigrantes Sirio-libaneses llegaron a Colombia a través de los puertos de Barranquilla, en la Costa Caribe, muchos de ellos se establecieron allí, pero otros siguieron la ruta del Magdalena buscando los centros andinos de desarrollo, entre ellos Ibagué por ser un centro cultural musical reconocido (Fawcett & Posada, 1992; Giraldo, 2012) lo cual actualmente se puede evidenciar por la existencia de apellidos como Ayram, Hayek, Bechará, entre otros, de ancestría sirio-libanesa o turca autoreconocida, referenciada en las entrevistas de soporte para el presente proyecto o comunicaciones personales.

Las poblaciones no-indígenas crecieron independientemente de las nativas, segregadas por leyes que las limitaban a vivir en resguardos. Aunque en esta separación inicial pudo influir la diferenciación genética actual observada entre ambas poblaciones del Tolima, hoy en día se puede evidenciar un componente nativo patrilineal importante en los mestizos (~21.5% en promedio) con un comportamiento clinal sur-norte en su frecuencia así: Planadas>Ortega>Ibagué, lo cual refleja el efecto de la localización geográfica y el aislamiento gradual en el patrón de la distribución de las frecuencias paternas de origen nativo en estas poblaciones. Una característica notable es la presencia del subHg Q1a3\* en Planadas, el cual pudo provenir del aporte paterno de otros grupos indígenas distinto a los Nasa, pues está ausente en estos últimos, siendo los Pijaos el grupo fuente más probable, con el que pudo ocurrir una mezcla inicial temprana y luego sobrevino el aislamiento.

Muchas otras poblaciones no-indígenas latinoamericanas que han sido analizadas para los mismos marcadores también presentan influencia nativa patrilineal similar al Tolima, por ejemplo en Nicaragua (15%) (Núñez et al., 2012) y México (30.6%) (Martinez-Cortes et al., 2012), pero también hay poblaciones con un componente nativo relativamente bajo, por ejemplo en Antioquia-Colombia (~1%) (G. Bedoya et al., 2006; Carvajal-Carmona et al., 2003; Carvajal-Carmona et al., 2000), Costa

Rica (~8%) (Carvajal-Carmona et al., 2003) o Argentina (~5%) (Corach et al., 2010). Por consiguiente, esta evidencia sugiere patrones variables a lo largo de Latino América en cuanto al flujo de linajes paternos nativos en las poblaciones que no se autodefinen como indígenas, lo cual puede favorecerse por la alta densidad nativa inicial y la menor migración de hombres europeos, un grado de aislamiento o exclusión de indígenas con relación a las nuevas poblaciones fundadas (Catelli et al., 2011), o a la adaptabilidad, forzada o no de los nativos, a nuevas condiciones antropogénicas, que en el caso de México o en el Tolima pudo ser mayor.

Los datos genéticos indican que en la población no-indígena del Tolima, durante la historia demográfica de su poblamiento, ha ocurrido un flujo linajes maternos y paternos, por medio de la mezcla ocurrida entre mujeres nativas con una fracción importante de hombres españoles o de regiones euroasiáticas pero también con la participación de nativos, y en menor proporción de africanos. En Montelíbano la situación es más compleja, una influencia materna nativa/africana y del lado paterno europeo/africano/nativo, lo cual explica su alta diversidad, dados los patrones complejos de entrecruzamiento y mezcla que pueden emerger, como se discutirá más adelante, durante el análisis de la estructura genética de las poblaciones estudiadas.

**8.3.2 Estructura genética global de las poblaciones analizadas.** La información derivada del mDNA y del cromosoma Y ha puesto de manifiesto varios procesos como el patrón de mezcla genética en todas las poblaciones del presente estudio y la relación de la diversidad con los componentes ancestrales encontrados. Sin embargo hay diferencias entre la ancestría establecida a partir de los sistemas uniparentales y el autosómico.

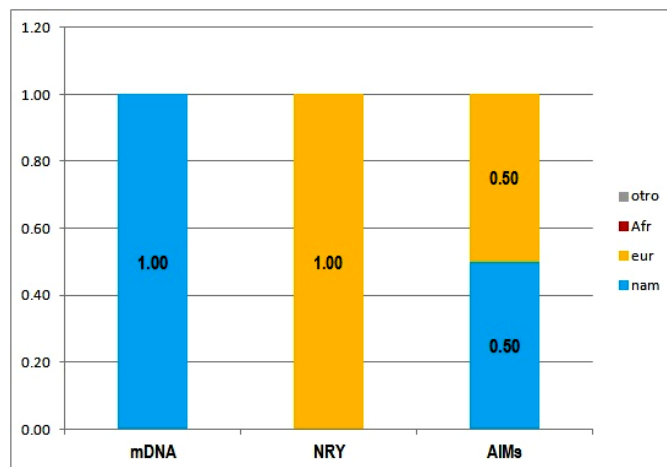
Supóngase que una población humana es originada a partir de la mezcla de otras dos, previamente separadas, se asume sesgo en la mezcla, así que la población parental 1, o europea, sólo aporta hombres de Hg R, con su par de cromosomas sexuales ( $XY^R$ ), el conjunto autosómico diploide (**AA**) y un genoma mitocondrial específico (**m**), de manera que en general, todo su genoma es representado así:  $XY^R AAm$ . De la población parental 2, o nativos americanos, se tienen solo mujeres nativas  $XXAAm$ .

Primero asúmase la situación más sencilla: el 100% de los hombres que aportan a la mezcla son europeos y todas las mujeres son nativas; bajo este panorama, asumiendo que los cruces se darán al azar, se tendría el siguiente cuadro de probabilidades:

	<b>NRY</b>	0	1
	<b>AIMs</b>	0.50	0.50
		<b>XA</b>	<b>Y<sup>RA</sup></b>
<b>1</b>	<b>XAm</b>	<b>XXAAm</b>	<b>XY<sup>RA</sup>AAm</b>

Fuente: el autor.

El aporte autosómico y uniparental europeo (eur) y nativo (nam) produce como descendencia un tipo de hombre y uno de mujer, que en conjunto aportan a los cálculos de las fracciones de ancestría para cada sistema genético: de cada 4 autosomas globales, dos de origen nativo y dos de origen europeo, en tanto que todas las mitocondrias de la población son indígenas y todos los cromosomas Y de los varones son europeos:

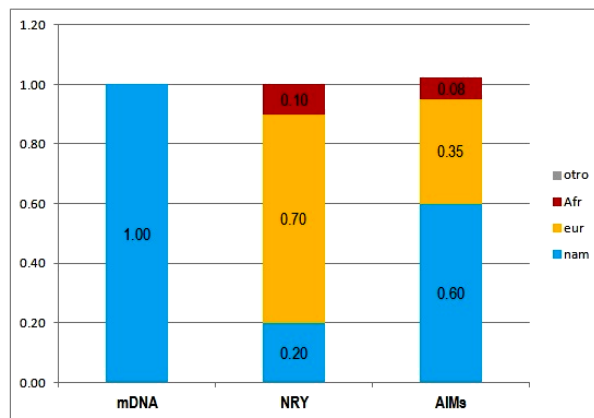


Fuente: el autor.

Nótese que la ancestría mitocondrial a primera vista está indicando que el origen materno de la población es 100% nativo americano, mientras el ancestro paterno, por su parte, indica un 100% de origen europeo (y solo se refiere a la mitad de la población), pero no se tiene una idea clara de la proporciones genómicas totales que cada población parental aporta, ya que en una población mezclada cada genoma uniparental es solo evidencia de una dimensión de la historia genética de la población. De manera que una forma más aproximada de estudiar la mezcla se obtiene al analizar los marcadores autosómicos, revelando el efecto real del patrón de mezcla entre hombres europeos y mujeres nativas, el cual genera individuos de genomas cuyo origen es de un 50% en cada población parental.

En un panorama más cercano a lo real, se pueden ejemplificar dos de los patrones distintivos obtenidos en el presente estudio. En este sentido, primero se simulará una aproximación a lo ocurrido con las poblaciones mestizas del Tolima, incorporando la participación de hombres nativos-Q, africanos-E y europeos-R:  $XY^QAAm$  (20%)  $XY^RAAm$  (80%),  $XY^EAAm$  (10%) y mujeres  $XXAAm$  (100%), de manera que se obtiene:

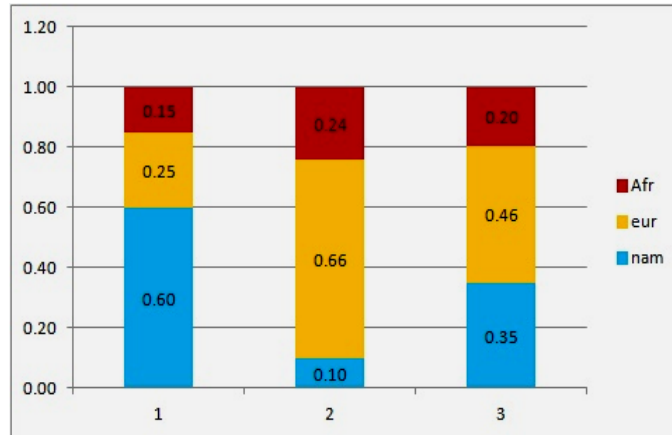
	<b>NRY</b>	0	0.2	0	0.7	0	0.1
	<b>AIMs</b>	0.10	0.10	0.35	0.35	0.05	0.05
		<b>XA</b>	<b>Y<sup>Q</sup>A</b>	<b>XA</b>	<b>Y<sup>R</sup>A</b>	<b>XA</b>	<b>Y<sup>E</sup>A</b>
<b>1</b>	<b>XAm</b>	<b>XXAAm</b>	<b>XY<sup>Q</sup>AAm</b>	<b>X XAAm</b>	<b>XY<sup>R</sup>AAm</b>	<b>XXAAm</b>	<b>XY<sup>E</sup>AAm</b>



Fuente: el autor.

Los valores calculados para las proporciones ancestrales a nivel autosómico, a partir de las frecuencias uniparentales preestablecidas, generan un patrón aproximado al observado en las poblaciones no indígenas del Tolima.

En el caso de la población mestiza de Córdoba, con componente africano paterno (0.24) en mayor proporción que en el Tolima, por otro lado adicionando el materno africano (0.15), el paterno europeo (0.66) el nativo (0.1), y el materno nativo (0.60), se obtiene lo siguiente:



Fuente: el autor.

Se puede notar que los datos de las frecuencias amerindias mitocondriales (0.60) y del cromosoma Y (0.1), usados para el anterior modelamiento, se basaron en aquellos observados en Montelíbano y predicen el valor de mezcla autosómica nativa observada en esta población (0.35), lo mismo sucede cuando se observan los datos africanos. De manera que partiendo de este hecho conocido, y, en vista de que una fracción importante del mDNA en los mestizos de Montelíbano no pudo ser identificada; al simularse la ancestría europea autosómica generada asumiendo la proporción de mDNA no determinado como europeo, se predijo aproximadamente el valor de ancestría europea observada en los autosomas. Por lo tanto es altamente probable que el 25% de las mitocondrias a las que no se les pudo asignar uno de los linajes estudiados (A, B, C, D, L o J), puedan ser otros Hgs europeos, como el H, entre otros (para más detalle de los cálculos, ver anexo 14).

El modelo anterior calcula, de manera manual las proporciones ancestrales esperadas a nivel autosómico, bajo un patrón fundador donde se conocen las fracciones uniparentales, asumiendo que estas están directamente relacionadas con las proporciones iniciales al momento de la fundación de la población. En este sentido, el modelo calcula lo que se podría esperar bajo ciertas condiciones, por ejemplo, no más de un evento migratorio posterior al primero, entre otras condiciones como los el apareamiento aleatorio. Este podría constituir una primera aproximación novedosa e interesante para realizar estimaciones teóricas de las fracciones ancestrales esperadas en una población para la cual no se tengan datos autosómicos, o para establecer los sesgos con relación a lo esperado; sin embargo tendrían que realizarse algunos pruebas estadísticas con otros datos de poblaciones alrededor del mundo con el fin de establecer la confiabilidad del carácter predictivo.

Lo anterior indica que el uso de marcadores autosómicos como los AIMs, es clave para definir la estructura genética de las poblaciones, dado que los mismos se

seleccionan con base en las amplias diferencias en las frecuencias interpoblacionales y contemplando un alto grado de informatividad  $I_n$ , por ejemplo los 100 SNPs que se emplearon en el presente trabajo fueron tomados de un panel recientemente diseñado para establecer las proporciones ancestrales en poblaciones latinoamericanas, dado que fueron seleccionados con base en las tres principales poblaciones ancestrales de Latino América: nativos americanos, europeos y africanos, mostrando capacidad para detectar aportes de cada población parental en varias mestizas como en Chile, México y Colombia (Galanter et al., 2012).

Debido a que el factor determinante en la selección de los SNPs fue la alta diferencia intercontinental, mas no la intracontinental, no fue posible discernir la estructuración presentada entre las poblaciones indígenas Nasa, Pijao y Embera, pues todas presentaron el mismo componente nativo, compartido con las demás poblaciones amerindias que se tomaron como referencia. La subdivisión poblacional en nativos americanos se ha estudiado en varias ocasiones, en uno de ellos se reporta un leve patrón de agrupamiento de nativos con que comparten el grupo lingüístico chibcha-paez, al igual que una clara separación de los grupos amazónicos Karitiana y Surui (S. Wang et al., 2007). Un análisis reciente llevado a cabo con más de 300.000 SNPs fue direccionado para establecer el patrón de poblamiento de América por parte de los nativos, pero a pesar de la densidad de marcadores no se detectó subdivisión poblacional, no obstante se confirmó ampliamente el origen asiático de la variación nativa americana (Reich et al., 2012). La detección de subdivisión poblacional basada en patrones de frecuencias de SNPs en nativos americanos no es sencilla, un estudio en el que se emplearon 2.500 SNPs encontró un patrón ya observado con relación a grupos amazónicos pero poco definido o sin estructuración. Las poblaciones amerindias son parecidas y se requiere del diseño de paneles de marcadores que tengan diferencias intracontinentales marcadas (Kidd et al., 2011).

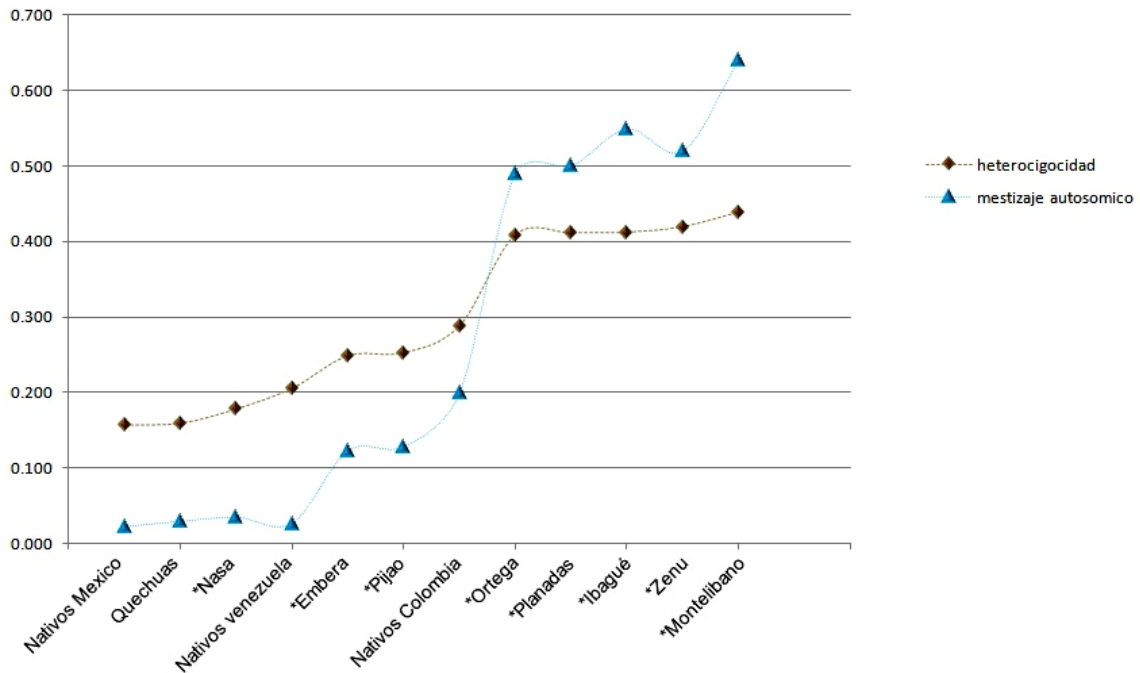
Pero por otro lado, se encontró diferenciación entre las poblaciones nativas y las mestizas, incluyendo en estas últimas a los Zenues dado su alto componente no-indígena, principalmente de origen europeo y africano.

En el entorno latino americano, los nasas constituyen un grupo étnico con un componente nativo altamente conservado y similar a indígenas centroamericanos como los zapotecas o suramericanos como los quechuas (referencias). Por su parte los pijaos y emberas al tener una ligera influencia africana y europea, son un poco más diversos que las demás etnias centro y suramericanas, por eso quizás en el análisis MDS son relativamente más cercanos entre sí que con los nasas, pero en términos generales las tres poblaciones obedece a un patrón de origen ancestral uniparental amerindio.

Como se discutió en apartados anteriores, al parecer las diferencias en los valores de diversidad genética entre grupos indígenas geográficamente aislados pueden ser explicados en parte por el grado de mezcla europea. Este comportamiento ha

sido también evidente en el presente estudio a nivel autosómico con los SNPs empleados, ya que los niveles de heterocigidad medidos con base en estos marcadores se incrementan también con el grado de componente no-nativo (figura 27).

**Figura 27.** Correlación entre la heterocigidad y la proporción de mezcla no nativa.



Fuente: el autor.

\*Correlación entre el componente europeo + africano y la heterocigidad observada en las poblaciones del presente estudio (\*) y en algunas poblaciones indígenas de Centroamérica y Suramérica. ( $r^2=0.97$ ).

El aislamiento geográfico puede estar jugando un rol importante en el alto grado de diferenciación poblacional observado en los nasas, se ha sugerido también que la baja heterocigidad en los zapotecas ( $He=0.229$ ) –que incluso es menor en los nasas- se debe a su aislamiento geográfico (Silva-Zolezzi et al., 2009), en general este aislamiento es identificado como un patrón recurrente en otras poblaciones nativas americanas (Li et al., 2008).

Los nasas del presente estudio habitan un resguardo localizado en una zona montañosa de difícil acceso, en el sur del Tolima –corregimiento de Gaitania-, tomando entre 12 y 16 horas poder llegar allí desde Ibagué. Por otro lado, de acuerdo con la historia del resguardo (comunicación personal), al parecer la población se ha constituido por algunas migraciones desde ciertos municipios

puntuales del Cauca, consecuentemente, la baja diversidad observada puede estar obedeciendo a la deriva que está experimentando la población.

Los Pijaos han ocupado una región geográfica ancestral, de fácil acceso por constituir en primer lugar, un valle interandino y segundo, una ruta transitada desde los tiempos de la colonia, comunicando el pacífico y centro del país. Por ello, desde la primera fundación de un centro urbano en el Tolima -Ibagué en 1550-, este grupo nativo ha estado expuesto directamente al contacto español (Triana, 1992) (anexo 15), pero es interesante notar que a pesar de ello, el componente europeo es solo del 12% en promedio, claro está que como hay variación individual, en la población se pueden encontrar algunos pocos individuos con ancestrías entre el 40 y 60%, por esta razón las desviaciones estándar de las fracciones europeas son más altas que en las demás poblaciones indígenas de referencia y que en los Nasa.

Lo anterior indica que un contacto prolongado no necesariamente implica una influencia mayor, ya que hay circunstancias socioculturales como las políticas tribales o la segregación social externa, que pueden generar la diferenciación observada en los pijaos con relación a las poblaciones urbanas no-nativas cercanas, como Ortega, donde la mezcla europea es alta (45%).

La etnia Embera habita principalmente las áreas ribereñas de los ríos en las selvas del pacífico, desde Panamá hasta Ecuador, así como también zonas del oeste de la cordillera occidental y del alto Sinú. Los Embera comparten la misma distribución geográfica que las etnias Wounana, Awa y Tule, también comparten sectores rurales y urbanos de población afrocolombiana y la población blanco-mestiza (Valencia & Hernandez, 2001). A pesar de su relativo aislamiento es posible que exista flujo genético entre los emberas y poblaciones no nativas, que en el presente estudio se evidencia por su componente africano (3%) y europeo (12%), similar al patrón de mezcla que presentaron los pijaos; aunque su componente africano es ligeramente mayor por la cercanía a centros de influencia afrocolombiana. Por esta razón ambas poblaciones son genéticamente más cercanas entre sí que con los nasas.

La alta desviación estándar del componente europeo en los Embera indica que ha ocurrido una mezcla o migración reciente de personas con alta ancestría europea, como se ha notado en algunas poblaciones de Costa Rica (Z. Wang et al., 2010), la razón no se conoce, pero posiblemente la situación social de desplazamiento esté relacionada.

Como se ha visto, los Zenues han mostrado ser el grupo étnico con mayor mezcla genética, similar a los mestizos de Córdoba, diferenciados de estos por el mayor aporte nativo (~35%) y el menor africano y europeo. Esta mezcla indica que si bien, existe una composición nativa ligeramente predominante, en todos los análisis de distancias genéticas la población ha estado vinculada con los mestizos debido a que básicamente obedecen a un patrón ancestral triparental. Hay que

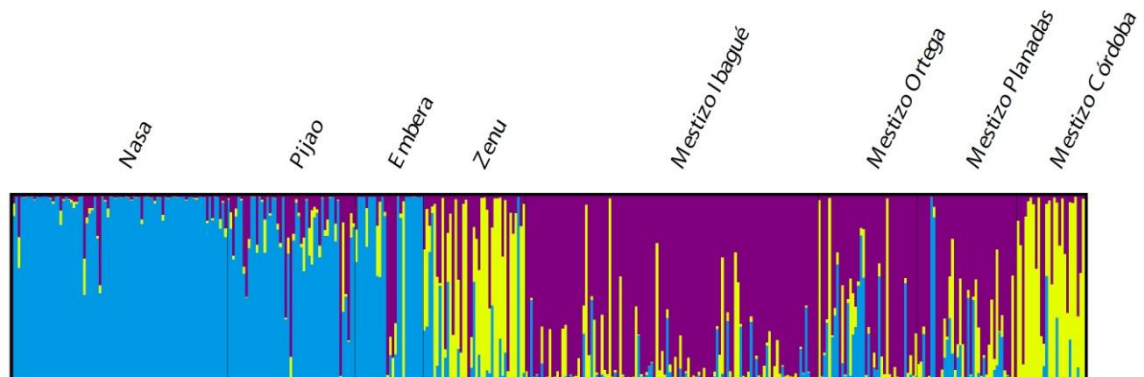


considerar que Montelíbano fue fundado por sirio-libaneses a principios del siglo XX en 1907, y según los historiadores estos se mezclaron rápidamente con la población local africana femenina que ya habitaba la región. También se conoce la migración temprana de europeos, de la cual hay evidencias actuales a través de la presencia de apellidos como Jaamman, descendientes de migrantes finlandeses (Mendoza, 2005; Vilorio, 2009).

En segundo lugar es un municipio minero (Castillo & Zapata, 1988), lo cual pudo haber estimulado nuevas inmigraciones, pues por ejemplo en algunas poblaciones de México se ha vinculado el desarrollo de la minería con la migración europea y el consecuente incremento en dicha ancestría detectada mediante las vías del cromosoma Y y autosómica (Martínez-Cortés et al., 2012). Por último, la migración interna debida a situaciones sociales de desplazamiento (Humanitaria, 2002). Todo este entorno histórico es consecuente con la alta diversidad y variación en las ancestrías africana, europea y nativa encontrada en los Zenues y mestizos de la región.

Si bien es cierto que en el análisis de estructuración poblacional global (Ensayo 2 de *Structure*) no existió una separación clara de las poblaciones mestizas tolimenses o cordobesas, cuando se realizó una simulación únicamente con las 8 poblaciones del estudio se observó una ligera segregación entre ambas poblaciones (figura 28). Esto indica que en el estudio se tienen tres tipos de poblaciones ligeramente diferenciadas: los nativos, mestizos del Tolima y la población de Córdoba. Entre Ibagué, Planadas y Ortega actualmente no hay ningún tipo de estructuración y su ancestría es básicamente biparental.

**Figura 28.** Estructuración genética poblacional observada en las poblaciones del estudio.



Fuente: el autor.

\*Se observan las poblaciones indígenas Nasa, Pijao, Embera y Zenú, y las mestizas del Tolima y Córdoba. La simulación se realizó empleando los genotipos de 100 AIMS y realizando las simulaciones en el programa STRUCTURE, bajo las mismas condiciones de la página 129, pero sin incluir las poblaciones de referencia, con el propósito de encontrar patrones de frecuencias propios.

La estratificación genética de toda la población indica que actualmente existen en los marcadores empleados, tres patrones de frecuencias génicas, distintivos y separados, posiblemente originados por las diferencias en la historia demográfica de las poblaciones analizadas. Los nativos de ancestría básicamente uniparental, los mestizos del Tolima de origen biparental y las dos poblaciones de Córdoba triparentales. Aunque para el presente estudio no hay forma de medir la correlación entre el grado de conservación de la lengua nativa y la mezcla genética, es interesante que los grupos indígenas, con un relativo aislamiento de centros urbanos grandes, como en los Nasa y Embera, aun conserven la lengua nativa como se pudo evidenciar durante la visita a dichas comunidades y por las referencias que existen (Mortensen, 1999; Yule-Yatacué, 1996). En tanto que para los pijaos, los historiadores aseguran que hacia 1943 vivió el último nativo que hablaba la lengua autóctona y la razón más clara ha sido la sustitución por la lengua española dada la fuerte y constante influencia de la colonización en su estructura cultural (V. Bedoya, 1950; Durbin & Seijas, 1973).

El análisis de la estructura genética en poblaciones latino americanas como las del presente trabajo ha recibido atención durante los últimos años (Bryc et al., 2010; González Burchard et al., 2005; Kidd et al., 2011; Wall et al., 2011), en primer lugar, por haber constituido, la última frontera en ser colonizada hace aproximadamente 15 mil años por algunas migraciones humanas provenientes del Siberia y establecidas en la región de Bering desde donde surgieron los primeros nativos que llegaron y poblaron el continente americano (Long & Bortolini, 2011; Reich et al., 2012). En segundo lugar, debido a los efectos de una migración más reciente, ocurrida hace un poco más de 500 años, procedente de Europa. Ambos eventos han dejado una huella genética en las poblaciones latinoamericanas actuales, unas que conservan el ancestro amerindio en alto grado y otras que reflejan la mezcla europea en varios grados, dependiendo de la región. Por ejemplo, las poblaciones como México y Argentina muestran patrones de mezcla diferentes, consistentes con sus respectivas historias y patrones demográficos; en México hay una composición nativa alta (~55%), europea menor (~41%) y africana muy baja (<5%); por su lado, en Argentina la ancestría amerindia es menor (31%) y la europea mayor (65%), prácticamente patrones opuestos debido al estímulo de la inmigración europea hacia Argentina a principios del siglo XX; por otra parte, el aporte africano es similar y refleja un punto común en su historia, correspondiente con la traída de africanos en tiempos de la colonia por los españoles (Avena et al., 2012; Corach et al., 2010; Martínez-Fierro et al., 2009; Silva-Zolezzi et al., 2009). Comparativamente, la población analizada del Tolima se acerca más a la mexicana y a la Costarricense (Z. Wang et al., 2010), teniendo en cuenta ambos

componentes (nam: 47%, eur: 49%). Los Zenues o los mestizos de Córdoba, debido a la influencia africana (~22%), tienen una similaridad a la población brasilera, donde se ha encontrado un componente africano del 24% (Giolo et al., 2012), lo cual está quizás reflejando en alguna medida, la proporción de africanos que llegaron a las costas atlántica colombiana y brasilera durante la colonia, pues en estudios recientes, donde se analizó la población de Bolívar y Magdalena, existe un componente africano alto (43% y 28% respectivamente) (Rojas et al., 2010), incluso mayor que el encontrado para los zenues y mestizos de Córdoba del presente estudio. Como se aprecia, a nivel latino americano el aporte africano sigue una misma pauta, aquellas poblaciones donde el flujo ha sido menor al 5%, en las cuales la interacción nativo-europea es la principal, siendo por lo tanto poblaciones fundamentalmente biparentales como en el Tolima. Las otras poblaciones con un flujo africano >20%, demuestran una estructura diferente obedeciendo a un modelo de mezcla triparental como en Brasil, y en la región de Córdoba.

Aunque el flujo genético entre poblaciones ha ocurrido a lo largo de la historia de la humanidad, solo recientemente se ha dado escala global entre poblaciones continentales. Y el ocurrido en América Latina es de especial interés debido a que el desequilibrio de ligamiento generado puede servir para detectar genes que aportan al riesgo para una enfermedad dada, mediante el mapeo de la mezcla (González Burchard et al., 2005; Winkler et al., 2010). Países de América Latina, como México han avanzado en un conocimiento profundo de la estructura genética de su población en términos de la diversidad, patrones de desequilibrio de ligamiento y de haplotipos compartidos entre la población nativa y mestiza, con el fin desarrollar la medicina genómica (Martínez-Cortés et al., 2012; Silva-Zolezzi et al., 2009). En Colombia hay un estudio en 24 poblaciones que también ha sido enfocado para el análisis de la diversidad (Rojas et al., 2010) y apenas se empiezan a desarrollar los estudios de asociación de la ancestría con enfermedades particulares, por ejemplo la africana con el riesgo al asma (D. Campbell et al., 2012; Vergara et al., 2009). Mientras que en el Tolima hay una ausencia de estudios genómicos comparativos entre sistemas genéticos, solo dos estudios localizados en poblaciones de Coyaima analizaron simultáneamente el mtDNA y algunos microsatélites autosómicos y de la región NRY (Rondon et al., 2008; Rondon et al., 1999). Por esta razón, se considera que los resultados del presente estudio puedan aportar al conocimiento, no solo de la diversidad y las relaciones genéticas poblacionales en el Tolima o con otras regiones de Colombia, sino también para futuras aplicaciones del mapeo de la mezcla en la búsqueda de nuevos genes relacionados por ejemplo con diferentes tipos de cáncer como los de colon, tiroides, gástrico o de glándula mamaria, ya que conociendo el patrón de mezcla ocurrido en la región es principalmente nativo-europeo, se pueden diseñar estrategias metodológicas para establecer la extensión del desequilibrio de ligamiento, con paneles más densos que también permitan conocer los bloques de haplotipos característicos de la población tolimense para los futuros estudios de asociación.

## 9. CONCLUSIONES

1. Las poblaciones nativas Nasa, Pijao y Embera son las de mayor conservación de la variación nativa ancestral, reflejada en la alta frecuencia de los Hgs mitocondriales A, B, C y D (98, 94 y 93% respectivamente), del cromosoma Y (71, 70 y 67%) y de los AIMs (96, 85 y 84%). En los Pijaos y Embera existe una mezcla, principalmente de origen europea en el cromosoma Y y en los autosomas, que ha influenciado el incremento de la diversidad con relación a otras poblaciones indígenas de Centro y Suramérica.
2. La amplia variabilidad observada en las proporciones nativa y europea en Pijao y Nasa, indica una influencia reciente. Estas consideraciones y las de la conclusión 1, han generado una diferenciación genética con respecto a la población mestiza analizada en general, dado que éstos poseen un componente europeo mayor a nivel del cromosoma Y y en los autosomas (en promedio, NRY: 69%, AIMs: 47%), o africano en el caso de los Zenu (mDNA: 6%, NRY: 17%) y Córdoba (mDNA: 13%, NRY: 24% y AIMs: 22%).
3. En la población mestiza del Tolima, se ha observado un proceso de mezcla genética principalmente entre mujeres nativas y hombres europeos, reflejado en los patrones haplogrupales uniparentales en dicho grupo poblacional (mDNA nativo: 93% y NRY europeo: 70%). Este sesgo pudo tener su origen en la historia de la población durante la época colonial, que es el mismo que se ha presentado en muchas otras poblaciones de Colombia y América Latina. Sin embargo, también se observó una influencia paterna de origen amerindio (21%) comparable con la presentada por poblaciones como la mexicana o en Costarica.
4. En la población mestiza de Córdoba, ocurre una variación nativa, europea y africana a nivel del mDNA (nam: 60%, afr: 13%), cromosoma Y (nam: 10%, eur: 67%, afr: 24%) y autosómico (nam: 35%, eur: 42%, afr: 22%), posiblemente debido a su historia fundacional de origen nativo, euroasiático y africano; adicionalmente a las condiciones socioculturales de desplazamiento y económicas actuales, incidiendo en su alta diversidad (mDNA:  $0.80 \pm 0.0348$ , NRY:  $0.77 \pm 0.0725$ , AIMs:  $0.44 \pm 0.007$ ), con relación a la media (mDNA: 0.73, NRY: 0.69, AIMs: 0.35). Por lo tanto, de acuerdo con los argumentos anteriores y de la conclusión 2, Córdoba y el Tolima tienen una estructura genética diferente, lo cual indica acerca de la influencia de la historia demográfica diferencial entre ambas poblaciones en la estratificación poblacional.
5. La diversidad hallada en la población nasa, contemplando el mDNA ( $0.6786 \pm 0.0267$ ), la región NRY ( $0.4611 \pm 0.0819$ ) y los AIMs ( $0.178 \pm 0.014$ ), fueron bajas, en comparación con la media (mDNA: 0.73, NRY: 0.69, AIMs: 0.34). Dicha observación puede estar relacionada con el aislamiento geográfico y la deriva. Sin embargo se requerirían estudios adicionales con otras poblaciones del mismo grupo étnico.

6. Los zenues constituyen el grupo indígena más diverso en el estudio (mDNA:  $0.81 \pm 0.024$ , NRY:  $0.70 \pm 0.075$ , AIMs:  $0.41 \pm 0.009$ ) si se le compara con la media presentada por todos los nativos (mDNA: 0.70, NRY: 0.66, AIMs: 0.27). Esto puede explicarse en parte a la alta variación en la ancestría europea, africana y nativa en los marcadores empleados, posiblemente por la influencia de varias condiciones socioculturales como la cercanía a centros urbanos y el estímulo a la inmigración durante los últimos 30 años debido a la economía minera de la región, entre otros factores como el desplazamiento humano.

## 10. RECOMENDACIONES

Ampliar los marcadores utilizados en el análisis de los haplogrupos mitocondriales, del cromosoma Y y autosómicos, con el objetivo de incrementar el nivel de resolución de los grados de diferenciación poblacional, tanto en nativos como en mestizos.

Aumentar el número de poblaciones nativas y mestizas, con el fin de incrementar la diversidad estudiada y enriquecer el mapa genético de las poblaciones colombianas. Especialmente contactar nuevas comunidades indígenas de la etnia Embera, dado que muy posiblemente el tamaño de la muestra que se empleó para el presente trabajo puede estar influenciando los datos observados para esta comunidad.

Establecer convenios de colaboración con otros proyectos de diversidad genética humana que se estén desarrollando a nivel local o internacional con el fin de ampliar el alcance de las observaciones que se reportaron en el presente trabajo.

Continuar investigando en temas que no pudieron ser abordados en el presente trabajo, con el fin de avanzar en el conocimiento de las poblaciones mezcladas o indígenas de la región y vincular nuevos estudiantes tanto de pregrado como de maestría.

## REFERENCIAS

- Adams, S. M., Bosch, E., Balaesque, P. L., Ballereau, S. J., Lee, A. C., Arroyo, E., . . . Jobling, M. A. (2008). The Genetic Legacy of Religious Diversity and Intolerance: Paternal Lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *American journal of human genetics*, 83(6), 725-736.
- Alves-Silva, J., da Silva Santos, M., Guimarães, P. E. M., Ferreira, A. C. S., Bandelt, H.-J., Pena, S. D. J., & Prado, V. F. (2000). The Ancestry of Brazilian mtDNA Lineages. *American journal of human genetics*, 67(2), 444-461.
- Avena, S., Via, M., Ziv, E., Pérez-Stable, E. J., Gignoux, C. R., Dejean, C., . . . Fejerman, L. (2012). Heterogeneity in Genetic Admixture across Different Regions of Argentina. *PLoS ONE*, 7(4), e34695. doi: 10.1371/journal.pone.0034695
- Bailliet, G., Ramallo, V., Muzzio, M., García, A., Santos, M. R., Alfaro, E. L., . . . Demarchi, D. A. (2009). Brief communication: Restricted geographic distribution for Y-Q\* paragroup in South America. *American Journal of Physical Anthropology*, 140(3), 578-582. doi: 10.1002/ajpa.21133
- Bailliet, G., Rothhammer, F., Carnese, F., Bravi, C., & Bianchi, N. (1994). Founder mitochondrial haplotypes in amerindian populations. *The American Journal of Human Genetics*, 54, 27-33.
- Banco de la República, B. L. Á. A. (2012). "Tolima, un territorio con 13.000 años de historia" *Publicación digital en la página web de la Biblioteca Luis Ángel Arango del Banco de la República*. Retrieved 30 de julio, 2012
- Bedoya, G., Montoya, P., Garcia, J., Soto, I., Bourgeois, S., Carvajal-Carmona, L., . . . Ruiz, A. (2006). Admixture dynamics in Hispanics: A shift in the nuclear genetic ancestry of a South American population isolate. *Proceedings of the national Academy of sciences*, 103(19), 7234–7239.
- Bedoya, V. (1950). *Etnología y conquista del Tolima y la hoya del Quindío*. Ibagué.
- Bisso-Machado, R., Bortolini, M., & Salzano, F. M. (2012). Uniparental genetic markers in South Amerindians. *Genetics and Molecular Biology*, 35, 365-387.
- Blanco-Verea, A., Jaime, J. C., Brión, M., & Carracedo, A. (2010). Y-chromosome lineages in native South American population. *Forensic Science International: Genetics*, 4(3), 187-193. doi: 10.1016/j.fsigen.2009.08.008

- Bortolini, M., Salzano, F., Thomas, M., Stuart, S., Nasanen, S., Bau, C., . . . Ruiz, A. (2003). Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas. *The American Journal of Human Genetics*, 73, 524-539.
- Brace, C. L., Nelson, A. R., Seguchi, N., Oe, H., Sering, L., Qifeng, P., . . . Tumen, D. (2001). Old World sources of the first New World human inhabitants: A comparative craniofacial view. *Proceedings of the National Academy of Sciences*, 98(17), 10017-10022. doi: 10.1073/pnas.171305898
- Bryc, K., Velez, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., . . . Ostrer, H. (2010). Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proceedings of the National Academy of Sciences*, 107(Supplement 2), 8954-8961. doi: 10.1073/pnas.0914618107
- Campbell, D., Parra, M. V., Duque, C., Gallego, N., Franco, L., Tandon, A., . . . Ruiz-Linares, A. (2012). Amerind Ancestry, Socioeconomic Status and the Genetics of Type 2 Diabetes in a Colombian Population. *PLoS ONE*, 7(4), e33570. doi: 10.1371/journal.pone.0033570
- Campbell, M., & Tishkoff, S. (2008). African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. *The annual Review of Genomics and Human Genetics*, 9, 403-433.
- Cardona, L. (2001). La lengua perdida de los Zenú, *EL Tiempo*.
- Carvajal-Carmona, L., Ophoff, R., Service, S., Hartiala, J., Molina, J., Leon, P., . . . Ruiz, A. (2003). Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Human genetics*, 112 534-541.
- Carvajal-Carmona, L., Soto, I., Pineda, N., Ortiz, D., Duque, C., Ospina, J., . . . Ruiz, A. (2000). Strong Amerind/White Sex Bias and a Possible Sephardic Contribution among the Founders of a Population in Northwest Colombia. *The American Journal of Human Genetics*, 67, 1287-1295.
- Casas-Vargas, A., Gómez, A., Briceño, I., Díaz-Matallana, M., Bernal, J. E., & Rodríguez, J. V. (2011). High genetic diversity on a sample of pre-Columbian bone remains from Guane territories in northwestern Colombia. *American Journal of Physical Anthropology*, 146(4), 637-649. doi: 10.1002/ajpa.21626
- Castillo, H., & Zapata, B. (1988). *Impacto socio-económico de Cerromatoso S.A. en la comunidad de montelibano*. Montería: Universidad de Córdoba. Centro de Investigaciones.



- Catelli, M., Alvarez-Iglesias, V., Gomez-Carballa, A., Mosquera-Miguel, A., Romanini, C., Borosky, A., . . . Salas, A. (2011). The impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genetics*, *12*(1), 77.
- Cavalli-Sforza, L. (1990). Opinion: how can one study individual variation for 3 billion nucleotides of the human genome? (0002-9297 (Print)).
- Cavalli-Sforza, L. (2005). The Human Genome Diversity Project: past, present and future. [10.1038/nrg1596]. *Nat Rev Genet*, *6*(4), 333-340.
- Clark, G., Nielsen, R., Signorovitch, N., Matise, T., Glanowski, S., Heil, J., . . . Lai, E. (2003). Linkage Disequilibrium and Inference of Ancestral Recombination in 538 Single-Nucleotide Polymorphism Clusters across the Human Genome. *American Journal of Human Genetics*, *73*, 285-300.
- Cogent Study. (2008). Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nature Genetics*, *40* (12), 1426-1435.
- Collins, H., Chima, B., Morii, T., Wah, K., Figueroa, Y., Criswell, L., . . . Seldin, M. (2004). Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Human genetics*, *114*, 263-271.
- Corach, D., Lao, O., Bobillo, C., Van Der Gaag, K., Zuniga, S., Vermeulen, M., . . . Kayser, M. (2010). Inferring Continental Ancestry of Argentineans from Autosomal, Y-Chromosomal and Mitochondrial DNA. *Annals of Human Genetics*, *74*(1), 65-76. doi: 10.1111/j.1469-1809.2009.00556.x
- Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Dambueva, I., Perkova, M., . . . Zakharov, I. (2007). Phylogeographic Analysis of Mitochondrial DNA in Northern Asian Populations. *American journal of human genetics*, *81*(5), 1025-1041.
- Deshpande, O., Batzoglou, S., Feldman, M. W., & Luca Cavalli-Sforza, L. (2009). A serial founder effect model for human settlement out of Africa. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1655), 291-300. doi: 10.1098/rspb.2008.0750
- Dulik, Matthew C., Zhadanov, Sergey I., Osipova, Ludmila P., Askapuli, A., Gau, L., Gokcumen, O., . . . Schurr, Theodore G. (2012). Mitochondrial DNA and Y Chromosome Variation Provides Evidence for a Recent Common Ancestry between Native Americans and Indigenous Altaians. *American journal of human genetics*, *90*(2), 229-246.

- Durbin, M., & Seijas, H. (1973). A Note on Panche, Pijao, Pantagora (Palenque), Colima and Muzo. *International Journal of American Linguistics*, 39(1), 47-51.
- Ecue, M. (1989). *Aprendamos a leer en páez (cartilla básica) = leiya' piyaaga nasa yuwete*. Lomalinda- Meta: Editorial Townsend.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14(8), 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3(1), 87-112. doi: 10.1016/0040-5809(72)90035-4
- Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47-50.
- Fagundes, N., Kanitz, R., & Bonatto, S. (2008). A Reevaluation of the Native American MtDNA Genome Diversity and Its Bearing on the Models of Early Colonization of Beringia. *Plos One*, 3(9), 1-5.
- Falush, D., Stephens, M., & Pritchard, J. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, 164, 1567–1587.
- Fawcett, L., & Posada, E. (1992). En la tierra de las oportunidades: Los libaneses en Colombia *Boletín Cultural y Bibliográfico-Banco de la Republica*, 29(29).
- Galanter, J. M., Fernandez-Lopez, J. C., Gignoux, C. R., Barnholtz-Sloan, J., Fernandez-Rozadilla, C., Via, M., . . . for the, L. C. (2012). Development of a Panel of Genome-Wide Ancestry Informative Markers to Study Admixture Throughout the Americas. *PLoS Genet*, 8(3), e1002554. doi: 10.1371/journal.pgen.1002554
- Gardner, T. (1978). *Aprendamos a escribir en páez = Fi'ya' piyaaga nasa yuwete I*. Lomalinda (Meta): Editorial Townsend.
- Gayden, T., Regueiro, M., Martinez, L., Cadenas, A., & Herrera, R. (2008). Human Y-chromosome haplotyping by allele-specific polymerase chain reaction. *Electrophoresis*, 29, 2419-2423.
- Gerdel, F. (1989). *Nasa yuwete ew fi'jyá pu'chsa, la ortografía páez*. Lomalinda, Meta: Editorial Townsend.

- Giolo, S. R., Soler, J. M. P., Greenway, S. C., Almeida, M. A. A., de Andrade, M., Seidman, J. G., . . . Pereira, A. C. (2012). Brazilian urban population genetic structure reveals a high degree of admixture. *Eur J Hum Genet*, *20*(1), 111-116.
- Giraldo, J. (2012). Hakim Salomon. *Biografías* Retrieved agosto 2, 2012
- Goebel, T. (1999). Pleistocene human colonization of Siberia and peopling of the Americas: An ecological approach. *Evolutionary Anthropology: Issues, News, and Reviews*, *8*(6), 208-227. doi: 10.1002/(sici)1520-6505(1999)8:6<208::aid-evan2>3.0.co;2-m
- Goebel, T., Waters, M., & O'Rourke, D. (2008). The Late Pleistocene Dispersal of Modern Humans in the Americas. *Science* *319*, 1497-1502.
- Gómez-Carballa, A., Ignacio-Veiga, A., Álvarez-Iglesias, V., Pastoriza-Mourelle, A., Ruíz, Y., Pineda, L., . . . Salas, A. (2012). A melting pot of multicontinental mtDNA lineages in admixed Venezuelans. *American Journal of Physical Anthropology*, *147*(1), 78-87. doi: 10.1002/ajpa.21629
- González-Andrade, F., Sánchez, D., González-Solórzano, J., Gascón, S., & Martínez-Jarreta, B. (2007). Sex-Specific Genetic Admixture of Mestizos, Amerindian Kichwas, and Afro-Ecuadorans from Ecuador. *Human Biology*, *79*(1), 51-77.
- González-José, R., Bortolini, M., Santos, F. R., & Bonatto, S. L. (2008). The peopling of America: Craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *American Journal of Physical Anthropology*, *137*(2), 175-187. doi: 10.1002/ajpa.20854
- González Burchard, E., Borrell, L. N., Choudhry, S., Naqvi, M., Tsai, H.-J., Rodriguez-Santana, J. R., . . . Risch, N. (2005). Latino Populations: A Unique Opportunity for the Study of Race, Genetics, and Social Environment in Epidemiological Research. *American Journal of Public Health*, *95*(12), 2161-2168. doi: 10.2105/ajph.2005.068668
- Greenberg, J., Turner, C., & Zegura, S. (1986). The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. *Current Anthropology*, *27*, 477-497.
- Griffiths, A. J. F., & Murillo Araujo, F. J. (2000). *Genética moderna*. Madrid: McGraw-Hill : Interamericana.
- Guardado-Estrada, M., Juarez-Torres, E., Medina-Martinez, I., Wegier, A., Macias, A., Gomez, G., . . . Berumen, J. (2009). A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population. *J Hum Genet*, *54*(12), 695-705.

- Harris, H., & Hopkinson, D. A. (1972). Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphisms. *Annals of Human Genetics*, 36(1), 9-20. doi: 10.1111/j.1469-1809.1972.tb00578.x
- Hartl, D., & Clark, A. (2007). Inbreeding, Population Subdivision, and Migration *Principles of Population Genetics* (4 ed., pp. 257-311). Sunderland, Massachusetts: Sinauer Associates.
- Hewlett, B., van de Koppel, J., & Cavalli-Sforza, L. (1982). Exploration Ranges of Aka Pygmies of the Central African Republic. *The Journal of the Royal Anthropological Institute*, 17(3), 418-430.
- Hinds, D., Stuve, L., Nilsen, G., Halperin, E., Eskin, E., Ballinger, D., . . . Cox, D. (2005). Whole-Genome Patterns of Common DNA Variation in Three Human Populations. *Science*, 307, 1072-1079.
- Humanitaria, F. I. d. A. (2002). *Poblaciones desplazadas y cooperación : experiencias de desarrollo comunitario, memorias del encuentro de Montelíbano*. Montería Acción Contra el Hambre. Colombia Fundación Internacional de Ayuda Humanitaria.
- Hunley, K., & Healy, M. (2011). The impact of founder effects, gene flow, and European admixture on native American genetic diversity. *American Journal of Physical Anthropology*, 146(4), 530-538. doi: 10.1002/ajpa.21506
- Hunley, K., & Long, J. C. (2005). Gene flow across linguistic boundaries in Native North American populations. *Proceedings of the National Academy of Sciences of the United States of America*, 102(5), 1312-1317. doi: 10.1073/pnas.0409301102
- Hunley, K., Spence, J. E., & Merriwether, D. A. (2008). The impact of group fissions on genetic structure in Native South America and implications for human evolution. *American Journal of Physical Anthropology*, 135(2), 195-205. doi: 10.1002/ajpa.20720
- Jobling, M., Hurles, M., & Tyler, C. (2004). *Human evolutionary genetics, origins, peoples and disease*. Cambridge-UK: Garland Sciences.
- Jorde, L., Watkins, W., Bamshad, M., Dixon, M., & Ricker, C. (2000). The distribution of human genetic diversity: A comparison of mitochondrial, autosomal, and Y-chromosome data *The American Journal of Human Genetics*, 66, 979-988.
- Jota, M. S., Lacerda, D. R., Sandoval, J. R., Vieira, P. P. R., Santos-Lopes, S. S., Bisso-Machado, R., . . . and The Genographic, C. (2011). A new subhaplogroup of native American Y-Chromosomes from the Andes. *American Journal of Physical Anthropology*, 146(4), 553-559. doi: 10.1002/ajpa.21519

- Karafet, T., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., & Hammer, M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research*. doi: 10.1101/gr.7172008
- Karafet, T., Osipova, L., Gubina, M., Posukh, O., Zegura, S., & Hammer, M. (2002 ). High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Human Biology*, 74(6), 761-789.
- Keyeux, G., Rodas, C., Gelvez, N., & Carter, D. (2002). Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Human Biology*, 74(2), 221-233.
- Kidd, J. R., Friedlaender, F., Pakstis, A. J., Furtado, M., Fang, R., Wang, X., . . . Kidd, K. K. (2011). Single nucleotide polymorphisms and haplotypes in Native American populations. *American Journal of Physical Anthropology*, 146(4), 495-502. doi: 10.1002/ajpa.21560
- Kimura, M. (1969). THE RATE OF MOLECULAR EVOLUTION CONSIDERED FROM THE STANDPOINT OF POPULATION GENETICS. *Proceedings of the National Academy of Sciences*, 63(4), 1181-1188.
- Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.
- Kitchen, A., Miyamoto, M., & Mulligan, C. (2008). A Three-Stage Colonization Model for the Peopling of the Americas. *Plos One*, 3(2), 1-7.
- Kosoy, R., Nassir, R., Tian, C., White, P. A., Butler, L. M., Silva, G., . . . Seldin, M. F. (2009). Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Human Mutation*, 30(1), 69-78. doi: 10.1002/humu.20822
- Lewontin, R. (1972). The apportionment of human diversity. *Evolutionary Biology*, 6, 381-398.
- Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., . . . Myers, R. M. (2008). Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. *Science*, 319(5866), 1100-1104. doi: 10.1126/science.1153717
- Liu, H., Prugnolle, F., Manica, A., & Balloux, F. (2006). A Geographically Explicit Genetic Model of Worldwide Human-Settlement History. *American journal of human genetics*, 79(2), 230-237.

- Long, J. C., & Bortolini, M. (2011). New developments in the origins and evolution of Native American populations. *American Journal of Physical Anthropology*, 146(4), 491-494. doi: 10.1002/ajpa.21620
- Malyarchuk, B., Derenko, M., Denisova, G., Maksimov, A., Wozniak, M., Grzybowski, T., . . . Zakharov, I. (2011). Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. *J Hum Genet*, 56(8), 583-588. doi: <http://www.nature.com/jhg/journal/v56/n8/supinfo/jhg201164s1.html>
- Mao, X., Bigham, A., Mei, R., Gutierrez, G., Weiss, K., Brutsaert, T., . . . Parra, E. (2007). A Genomewide Admixture Mapping Panel for Hispanic/Latino Populations. *The American Journal of Human Genetics*, 80, 1171–1178.
- Martinez-Cortes, G., Salazar-Flores, J., Gabriela Fernandez-Rodriguez, L., Rubi-Castellanos, R., Rodriguez-Loya, C., Velarde-Felix, J. S., . . . Rangel-Villalobos, H. (2012). Admixture and population structure in Mexican-Mestizos based on paternal lineages. *J Hum Genet*.
- Martinez-Fierro, M. L., Beuten, J., Leach, R. J., Parra, E. J., Cruz-Lopez, M., Rangel-Villalobos, H., . . . Rojas-Martinez, A. (2009). Ancestry informative markers and admixture proportions in northeastern Mexico. *J Hum Genet*, 54(9), 504-509. doi: <http://www.nature.com/jhg/journal/v54/n9/supinfo/jhg200965s1.html>
- McVean, G. A. T. (2002). A Genealogical Interpretation of Linkage Disequilibrium. *Genetics*, 162(2), 987-991.
- Melton, P., Briceño, I., Gomez, A., Devor, E., Bernal, J., & Crawford, M. (2007). Biological Relationship Between Central and South American Chibchan Speaking Populations: Evidence From mtDNA. *American Journal of Physical Anthropology*, 133, 753-770.
- Mendizabal, I., Sandoval, K., Berniell-Lee, G., Calafell, F., Salas, A., Martinez-Fuentes, A., & Comas, D. (2008). Genetic origin, admixture, and asymmetry in maternal and paternal human lineages in Cuba. *BMC Evolutionary Biology*, 8(1), 213.
- Mendoza, J. (2005). *Montelíbano, fundación y desarrollo*. Montelíbano-Córdoba: Ediciones Tana.
- Mesa, N., Mondragon, M., Soto, I., Parra, M., Duque, C., Ortiz, D., . . . Ruiz, A. (2000). Autosomal, mtDNA, and Y-Chromosome Diversity in Amerinds: Pre- and Post-Columbian Patterns of Gene Flow in South America. *The American Journal of Human Genetics*, 67, 1277–1286.
- Moraga, M., Rocco, P., Miquel, F., Nervi, F., Llop, E., Chakraborty, R., . . . Carvalho, P. (2000). Mitochondrial DNA Polymorphisms in Chilean

Aboriginal Populations: Implications for the Peopling of the Southern. *American Journal of Physical Anthropology*, 113, 19-29.

Moraga, M., Santoro, C., Standen, V., Carvallo, P., & Rothhammer, F. (2005). Microevolution in Prehistoric Andean Populations: Chronologic mtDNA Variation in the Desert Valleys of Northern Chile. *American Journal of Physical Anthropology*, 127, 170-181.

Mortensen, C. (1999). In R. H. v. editor (Ed.), *A reference grammar of Northern embera languages : studies in the languages of Colombia*. Dallas International Academic Bookstore.

Mulligan, C., Hunley, K., Cole, S., & Long, J. (2004). Population genetics, history, and health patterns in native Americans. *Annual Review of Genomics and human Genetics*, 5, 295-315.

Núñez, C., Geppert, M., Baeta, M., Roewer, L., & Martínez-Jarreta, B. (2012). Y chromosome haplogroup diversity in a Mestizo population of Nicaragua. *Forensic Science International: Genetics*(0). doi: 10.1016/j.fsigen.2012.06.011

O'Fallon, B. D., & Fehren-Schmitz, L. (2011). Native Americans experienced a strong population bottleneck coincident with European contact. *Proceedings of the National Academy of Sciences*, 108(51), 20444-20448. doi: 10.1073/pnas.1112563108

O'Rourke, D. H., Hayes, M. G., & Carlyle, S. W. (2000). ANCIENT DNA STUDIES IN PHYSICAL ANTHROPOLOGY. *Annual Review of Anthropology*, 29(1), 217-242. doi: doi:10.1146/annurev.anthro.29.1.217

O'Rourke, D., & Raff, J. (2010). The Human Genetic History of the Americas: The Final Frontier. *Current Biology*, 20(4), R202-207.

Palha, T., Gusmão, L., Ribeiro-Rodrigues, E., Guerreiro, J. F., Ribeiro-dos-Santos, Â., & Santos, S. (2012). Disclosing the Genetic Structure of Brazil through Analysis of Male Lineages with Highly Discriminating Haplotypes. *PLoS ONE*, 7(7), e40007. doi: 10.1371/journal.pone.0040007

Pardo, M. (1984). *Fonología del idioma Embera del noroccidente antioqueño*. Medellín: Secretaría de Educación y Cultura de Antioquia.

Peakall, R., & Smouse, P. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288-295. doi: 10.1111/j.1471-8286.2005.01155.x

Peakall, R., & Smouse, P. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*. doi: 10.1093/bioinformatics/bts460

- Perego, U., Angerhofer, N., Pala, M., Olivieri, A., Lancioni, H., Kashani, B., . . . Torroni, A. (2010). The initial peopling of the Americas: A growing number of founding mitochondrial genomes from Beringia. *Genome Research*, 20, 1174-1179.
- Perego, U., Lancioni, H., Tribaldos, M., Angerhofer, N., Ekins, J., Olivieri, A., . . . Achilli, A. (2012). Decrypting the Mitochondrial Gene Pool of Modern Panamanians. *PLoS ONE*, 7(6), e38337. doi: 10.1371/journal.pone.0038337
- Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155, 945–959.
- Prugnolle, F., Manica, A., & Balloux, F. (2005). Geography predicts neutral genetic diversity of human populations. *Current biology : CB*, 15(5), R159-R160.
- Raff, J. A., Bolnick, D. A., Tackney, J., & O'Rourke, D. H. (2011). Ancient DNA perspectives on American colonization and population history. *American Journal of Physical Anthropology*, 146(4), 503-514. doi: 10.1002/ajpa.21594
- Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44), 15942-15947. doi: 10.1073/pnas.0507611102
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., . . . Ruiz-Linares, A. (2012). Reconstructing Native American population history. [10.1038/nature11258]. *Nature*, advance online publication. doi: <http://www.nature.com/nature/journal/vaop/ncurrent/abs/nature11258.html#supplementary-information>
- Relethford, J. H., & Harding, R. M. (2001). *Population Genetics of Modern Human Evolution* eLS: John Wiley & Sons, Ltd.
- Rodas, C., Gelvez, N., & Keyeux, G. (2003). Mitochondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and mestizo populations. *Human Biology*, 75(1), 13-30.
- Rodríguez, E., Hernández, A., Salamanca, L., Ruiz, F., Poveda, M., Machado, E., . . . Guerrero, M. (2007). Colombia una nación multicultural, su diversidad étnica. 1-49. Retrieved from DANE website:
- 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. *American Journal of Physical Anthropology*, 143(1), 13-20. doi: 10.1002/ajpa.21270

- Rondon, F., Osorio, J., & Peña, V. (2008). Diversidad genética en poblaciones humanas de dos regiones Colombianas. *Colombia Médica*, 39(2), 52-60.
- Rondon, F., Vallejo, G., Osorio, E., & Barreto, G. (1999). Frecuencias haplotípicas mitocondriales en la comunidad Indígena Coyaima del departamento del Tolima, Colombia. *Revista de la Asociación Colombiana de Ciencias Biológicas*, 11(1), 45-63.
- Rosenberg, N., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J. K., & Feldman, M. W. (2005). Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure. *PLoS Genet*, 1(6), e70. doi: 10.1371/journal.pgen.0010070
- Rosenberg, N., Pritchard, J., Weber, J., Cann, H., Kidd, K., Zhivotovsky, L., & M, F. (2002). Genetic Structure of Human Populations. *Science*, 298, 2381-2385.
- Ruiz-linares, A. o., A. Figueroa, M. Mesa, N. Munera, J. Bedoya, G. Velez, I. Garcia, L. Perez, L. Bertranpetit, J. Feldmani, M. y Goldstein, D. (1999). Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proceedings of the national Academy of sciences*, 96, 6312–6317.
- Rueda, L., Henrik, C., & Orlando, J. (1996). *Historia de Colombia: el establecimiento de la dominación española*. Bogotá: Presidencia de la República ; Imprenta Nacional de Colombia.
- Salas, A., Acosta, A., Alvarez-Iglesias, V., Cerezo, M., Phillips, C., Lareu, M., & Carracedo, A. (2008). The mtDNA Ancestry of Admixed Colombian Populations. *American Journal of Human Biology*, 20, 584-591.
- Salas, A., Lovo-Gómez, J., Álvarez-Iglesias, V., Cerezo, M., Lareu, M., Macaulay, V., . . . Carracedo, A. (2009). Mitochondrial Echoes of First Settlement and Genetic Continuity in El Salvador. *PLoS ONE*, 4(9), e6882. doi: 10.1371/journal.pone.0006882
- Salas, A., Richards, M., Lareu, M., Scozzari, R., Coppa, A., Torroni, A., . . . Carracedo, A. (2004). The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade. *The American Journal of Human Genetics*, 74(3), 454-465. doi: 10.1086/382194
- Salzano, F., & Bortolini, M. (2002). *The Evolution and Genetics of Latin American Populations*: Cambridge University Press

- Sandoval, K., Moreno-Estrada, A., Mendizabal, I., Underhill, P. A., Lopez-Valenzuela, M., Peñalosa-Espinosa, R., . . . Comas, D. (2012). Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. *American Journal of Physical Anthropology*, 148(3), 395-405. doi: 10.1002/ajpa.22062
- Sans, M. (2000). Admixture studies in Latin America: from the 20th to the 21st century. *Human Biology*, 72(1), 155-177.
- Santos, F., Bonatto, S., & Bortolini, M. (2007). Molecular evidence from contemporary indigenous populations to the peopling of America. In e. Lima M (Ed.), *Recent advances in molecular biology and evolution: applications to biological anthropology*. Lima: Kerala:Research Signpost.
- Schurr, T. (2004a). The peopling of the new world: Perspectives from Molecular Anthropology. *Annual Review Anthropology*, 33, 551–583.
- Schurr, T., & Sherry, S. (2004b). Mitochondrial DNA and Y Chromosome Diversity and the Peopling of the Americas: Evolutionary and Demographic Evidence. *American Journal of Human Biology*, 16, 420–439.
- Seielstad, M., Minch, E., & Cavalli, L. (1998). Genetic evidence for a higher female migration rate in humans. *Nature Genetics*, 20, 278-280.
- Seldin, M., Tian, C., Shigeta, R., Scherbarth, H., Silva, G., Belmont, J., . . . Alarcon, M. (2007). Argentine Population Genetic Structure: Large Variance in Amerindian Contribution. *American Journal of Physical Anthropology*, 132, 455–462.
- Sengupta, S., Zhivotovsky, L. A., King, R., Mehdi, S. Q., Edmonds, C. A., Chow, C.-E. T., . . . Underhill, P. A. (2006). Polarity and Temporality of High-Resolution Y-Chromosome Distributions in India Identify Both Indigenous and Exogenous Expansions and Reveal Minor Genetic Influence of Central Asian Pastoralists. *American journal of human genetics*, 78(2), 202-221.
- Serre, D., & Pääbo, S. (2004). Evidence for Gradients of Human Genetic Diversity Within and Among Continents. *Genome Research*, 14(9), 1679-1685. doi: 10.1101/gr.2529604
- Shastri, B. (2007a). Role of SNPs and Haplotypes in Human Disease and Drug Development. In S. US (Ed.), *Micro/Nano Technologies for Genomics and Proteomics* (Vol. 2).
- Shriver, M., Parra, E., Dios, D., Bonilla, C., Norton, H., Jovel, C., . . . Kittles, R. (2003). Skin pigmentation, biogeographical ancestry and admixture mapping. *Human genetics*, 112, 387–399.

- Silva-Zolezzi, I., Hidalgo-Miranda, A., Estrada-Gil, J., Fernandez-Lopez, J. C., Uribe-Figueroa, L., Contreras, A., . . . Jimenez-Sanchez, G. (2009). Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0903045106
- Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D., Mulligan, C., . . . Malhi, R. (2007). Beringian Standstill and Spread of Native American Founders. *Plos one*(9), 1-6.
- Tang, H., Choudhry, S., Mei, R., Morgan, M., Rodriguez-Cintron, W., Burchard, E. G., & Risch, N. J. (2007). Recent Genetic Selection in the Ancestral Admixture of Puerto Ricans. *American journal of human genetics*, 81(3), 626-633.
- The International HapMap Consortium. (2003). The International HapMap Project. *Nature*, 426, 789-796.
- The International HapMap consortium. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature*, 449 (18), 851-862.
- The Y Chromosome Consortium. (2002). A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups. *Genome Research*, 12, 339-348.
- Torroni, A., Schurr, T., Cabell, M., Brown, M., Neel, J., Larsen, M., . . . Wallace, D. (1993). Asian Affinities and Continental Radiation of the Four Founding Native American mtDNAs. *The American Journal of Human Genetics*, 53, 563-590
- Torroni, A., Schurr, T., Yang, C., Szathmary, E., Williams, R., Schanfield, M., . . . Wallace, D. (1992). Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics*, 130, 153-162.
- Toscanini, U., Gusmão, L., Berardi, G., Gomes, V., Amorim, A., Salas, A., & Raimondi, E. (2011). Male lineages in South American native groups: Evidence of M19 traveling south. *American Journal of Physical Anthropology*, 146(2), 188-196. doi: 10.1002/ajpa.21562
- Triana, A. (1992). *La colonización española en el Tolima, siglos 16 y 17* (Vol. 1). Santa Fe de Bogotá, D.C.: Funcol-Cuadernos del Jaguar.
- Underhill, P., Jin L, L., A., Mehdi, S., Jenkins, T., Vollrath, D., Davis, R., . . . Oefner, P. (1997). Detection of numerous Y-chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography *Genome Research*, 7, 996-1005.

- Underhill, P., Jin, L., Zeman, R., Oefner, P., & Cavalli, L. (1996). A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proceedings of the national Academy of sciences*, 93, 196-200.
- Valencia, M., & Hernandez, C. (2001). *Emberás territorio y biodiversidad : estrategias de control en escenarios de conflicto* (Hernandez, C. ed.). Bogotá: Arfo Editores e Impresores-Fundación Swissaid Colombia.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., . . . Zhu, X. (2001). The Sequence of the Human Genome. *Science*, 291(5507), 1304-1351. doi: 10.1126/science.1058040
- Vergara, C., Caraballo, L., Mercado, D., Jimenez, S., Rojas, W., Rafaels, N., . . . Barnes, K. (2009). African ancestry is associated with risk of asthma and high total serum IgE in a population from the Caribbean Coast of Colombia. *Human Genetics*, 125(5), 565-579. doi: 10.1007/s00439-009-0649-2
- Viloria, J. (2009). *El ferroníquel de Cerro Matoso: aspectos económicos de Montelíbano y el Alto San Jorge*. Bogotá: Banco de la República, Centro de estudios económicos regionales-Cartagena.
- Wall, J. D., Jiang, R., Gignoux, C., Chen, G. K., Eng, C., Huntsman, S., & Marjoram, P. (2011). Genetic variation in Native Americans, inferred from Latino SNP and resequencing data. *Molecular Biology and Evolution*. doi: 10.1093/molbev/msr049
- Wallace, D. C., Garrison, K., & Knowler, W. C. (1985). Dramatic founder effects in Amerindian mitochondrial DNAs. *American Journal of Physical Anthropology*, 68(2), 149-155. doi: 10.1002/ajpa.1330680202
- Wang, S., Lewis, C. M., Jr., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., . . . Ruiz-Linares, A. (2007). Genetic Variation and Population Structure in Native Americans. *PLoS Genet*, 3(11), e185. doi: 10.1371/journal.pgen.0030185
- Wang, S., Ray, N., Rojas, W., Parra, M., Bedoya, G., Gallo, C., . . . Ruiz, A. (2008). Geographic Patterns of Genome Admixture in Latin American Mestizos. *Plos Genetics*, 4(3), 1-9.
- Wang, Z., Hildesheim, A., Wang, S., Herrero, R., Gonzalez, P., Burdette, L., . . . Yu, K. (2010). Genetic Admixture and Population Substructure in Guanacaste Costa Rica. *PLoS ONE*, 5(10), e13336. doi: 10.1371/journal.pone.0013336
- Watkins, W. S., Xing, J., Huff, C., Witherspoon, D., Zhang, Y., Perego, U., . . . Jorde, L. (2012). Genetic analysis of ancestry, admixture and selection in

- Bolivian and Totonac populations of the New World. *BMC Genetics*, 13(1), 39.
- Weir, B., & Cockerham, C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
- Wilder, J. A., Kingan, S. B., Mobasher, Z., Pilkington, M. M., & Hammer, M. F. (2004). Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. [10.1038/ng1428]. *Nat Genet*, 36(10), 1122-1125. doi: [http://www.nature.com/ng/journal/v36/n10/supinfo/ng1428\\_S1.html](http://www.nature.com/ng/journal/v36/n10/supinfo/ng1428_S1.html)
- Winkler, C., Nelson, G., & Smith, M. (2010). Admixture mapping comes of age. *Annual Review of Genomics Human Genetics*, 11, 65-89.
- Xing, J., Watkins, W. S., Shlien, A., Walker, E., Huff, C. D., Witherspoon, D. J., . . . Jorde, L. B. (2010). Toward a more uniform sampling of human genetic diversity: A survey of worldwide populations by high-density genotyping. *Genomics*, 96(4), 199-210. doi: 10.1016/j.ygeno.2010.07.004
- Yang, N., Li, H., Criswell, L., Gregersen, P., Alarcon, M., Kittles, R., . . . Seldin, M. (2005). Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Human genetics*, 118 (382-392).
- Yule-Yatacué, M. (1996). *Nasa üus yaht'n u'hun'i "por los senderos de la memoria y el sentimiento Paez" : Nasa Yuwe Twehn'i "tradición oral nasa Paez*. Toribio-Cauca: Programa de Educación Bilingüe Intercultural Pebin: Proyecto Nasa Toribío.
- Zegura, S. L., Karafet, T., Zhivotovsky, L. A., & Hammer, M. F. (2004). High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas. *Molecular Biology and Evolution*, 21(1), 164-175. doi: 10.1093/molbev/msh009

## **ANEXOS**

**Anexo A1.** Consentimiento Comunitario. Fuente: Basado en el formato de consentimiento informado #3 del proyecto: *1000 Genomes Project, Developing a Research Resource for Studies of Human Genetic Variation*.

## CONSENTIMIENTO COMUNITARIO PARA PARTICIPAR

---

Señor: \_\_\_\_\_ Rango en la comunidad:

Nombre de la comunidad y/o etnia:

---

**RESPETADO SEÑOR:** *Por favor lea detenidamente este documento, durante el tiempo que considere necesario. Al final, después haber aclarado sus dudas y logrado comprenderlo, decida si está de acuerdo con que los miembros de su comunidad participen en la investigación.*

### 1. Introducción:

El grupo de investigación de Citogenética, filogenia y evolución de poblaciones (LCFEP), de la Universidad del Tolima está desarrollando un proyecto de investigación denominado “*Estudio de la estructura genética de las poblaciones Amerindias Latino Americanas*”, el cual pretende estudiar las diversas variantes genéticas que existen entre los pobladores nativos del continente, con el objetivo de entender los procesos de poblamiento Americano, las relaciones ancestrales entre las diversas comunidades indígenas de la región, su relación con otros grupos étnicos y el impacto de la mezcla reciente con poblaciones foráneas. La participación de su comunidad en el proyecto depende de su autorización. Si usted está de acuerdo en que su comunidad forme parte del proyecto, la participación de los miembros de la misma implica:

- Autorizar a nuestro grupo de investigación a tomarle la muestra de sangre mediante la firma de una forma de consentimiento informado.
- Ser mayor de edad o, en caso de ser menor de edad, tener la autorización de uno de sus padres.
- Donar una muestra de sangre de la cual se extraerá material genético (ADN y ARN) y se establecerá una línea celular.

**El proyecto no producirá patentes, vacunas o algún conocimiento que genere ganancias económicas:** El proyecto sólo generará conocimientos básicos acerca de la variación genética en las comunidades indígenas participantes. Los datos derivados no se emplearán para patentar secuencias genómicas. No obstante es posible que la muestra se use en futuros estudios que intenten aclarar las bases genéticas de enfermedades en la región, aun así su aplicabilidad no será inmediata. La única ganancia es el conocimiento científico que estará disponible y de libre acceso en Internet y en revistas científicas.

Su nombre solo se mencionará en la forma de consentimiento informado. A esta información solo tendrán acceso los coordinadores del LCFEP o una persona autorizada por ellos. Una vez recolectada la muestra de sangre, la directora del LCFEP o la persona autorizada, le asignará un código único que no permitirá reconstruir información personal del donante como: nombre, apellidos, fecha de nacimiento o número de cédula. En la investigación no se accederá tampoco a los datos médicos particulares. Los tubos con las muestras de sangre solo tendrán el código de la muestra. A los investigadores que están estudiando las mismas se les suministrará el código único, la edad, la información del género (hombre o mujer) y la etnia o grupo geográfico del cual proviene la muestra. Si el donante piensa que puede ser parte de este proyecto, por favor lea el resto de este documento y tome el tiempo que sea necesario para hacer las preguntas que desee. La decisión de participar o no es completamente suya.

### 2. ¿Cuál es el propósito de la presente investigación?

El propósito del proyecto *Estudio de la estructura genética de las poblaciones Amerindias Latino Americanas* es analizar la variación genética que existe en las poblaciones Nativas de Latino América, mediante el estudio del ADN, extraído de una muestra de sangre donada por los participantes. Los genes contienen las instrucciones básicas para las células que constituyen el cuerpo. Nunca dos personas poseen exactamente el mismo ADN, exceptuando los gemelos idénticos. Los cambios en el ADN se denominan mutaciones o variaciones genéticas. Estas explican parcialmente algunas de las diferencias físicas entre las personas o, el por qué algunos seres humanos adquieren enfermedades como el cáncer o la diabetes, aunque estas enfermedades también pueden ser afectadas por factores como la dieta o la polución ambiental. Las mutaciones genéticas se presentan en las personas de todas partes del mundo, pero existen diferencias entre los grupos étnicos en cuanto al tipo y la forma en que ocurren. Por ejemplo, todos los grupos sanguíneos son encontrados en las poblaciones humanas, pero se presentan diferencias en la frecuencia de un tipo sanguíneo en una determinada población. Al estudiar el ADN de personas de orígenes geográficos y filiaciones étnicas diferentes, los investigadores encontrarán la mayoría de las variaciones genéticas, inter-grupos, lo cual permitirá en futuros estudios encontrar genes relacionados con enfermedades como el cáncer. Los resultados obtenidos serán colocados en bases de datos científicas de acceso libre en Internet, de igual manera, se generarán artículos científicos a los que cada comunidad participante podrá tener acceso. La información que estará en las bases de datos permanecerá por largo tiempo y muchos otros investigadores interesados la podrán consultar y emplear en sus investigaciones sobre la historia genética de las poblaciones o para encontrar variantes genéticas relacionadas con la salud y la enfermedad.

**3. ¿Qué tipo de preguntas me harán si decido participar?** Para determinar si el donante es elegible se le preguntará su edad (en el proyecto solo se incluirán mayores de edad o menores que tengan autorización de uno de sus padres) y el grupo étnico al que pertenece. De ser elegible, se le pedirá consentimiento para tomarle una muestra de sangre. A dicha muestra se le asignará un código único específico que no permitirá en ningún momento identificar sus datos personales.

**4. ¿Qué sucederá con la muestra que he decidido donar?** Las muestras de sangre y las formas de consentimiento informado serán enviadas al LCFEP en la Universidad de Tolima para ser procesadas y almacenadas. En el LCFEP, se extraerá su material genético y se establecerá una línea celular a partir de su muestra. Con dichas muestras, en el futuro inmediato se hará lo siguiente:

- Se leerá parcial o totalmente la información genética incluida en ella. Los experimentos para leer dicha información serán llevados a cabo en la Universidad del Tolima o en Instituciones de investigación colombianas o extranjeras que no tengan ánimo de lucro. El uso de estas muestras en otras Instituciones deberá contar con la autorización del Director de LCFEP.
- Se estudiará y analizará la variación genética de todas las muestras.
- Se publicarán los datos genéticos obtenidos y los resultados de las investigaciones en bases científicas en el Internet, en eventos y revistas científicas.

En futuras investigaciones se podrán usar las muestras y las bases de datos para estudiar otros aspectos, tales como que genes y/o variantes genéticas que afectan la forma en que los genes trabajan y los productos que estos generan (esto se denomina expresión génica o estudios proteómicos). Las investigaciones podrán girar en torno a los siguientes tipos de estudios:

- Biología del ADN
- Cómo surgen las nuevas variaciones
- Cómo trabaja el proceso evolutivo
- La composición y el tamaño de los grupos humanos; y
- Cómo las diferentes personas de Latino América se relacionan unas con otras y con otras poblaciones del mundo.

Las personas que realizan pruebas de identidad, como las de paternidad podrán usar los resultados depositados en las bases de datos científicas para hacer investigaciones generales acerca de los patrones



de variación genética humana. Algunos de los resultados de estas investigaciones también podrán ser colocados en las bases de datos en internet. Sin embargo, será prácticamente imposible para cualquiera encontrar alguna información personal suya en alguna de estas investigaciones debido a que las muestras y las bases de datos científicas sólo incluirán el código de la muestra y no tendrán algún tipo de información asociada como el nombre u otra adicional (dirección, lugar de nacimiento, número de identificación personal, etc.).

Cualquier investigador, que en un futuro quiera estudiar su muestra, deberá pertenecer a una entidad sin ánimo de lucro como una universidad o instituto gubernamental de investigación y realizar una solicitud formal al director del LCFEP, por medio de una “*declaración escrita de intención para la investigación*”. El Director del LCFEP revisará cada declaración para estar seguros de que el propósito para el cual la muestra será usada es consistente con esta forma de consentimiento. Adicionalmente, el Director de LCFEP creará un grupo de asesoramiento a la comunidad, el cual estará constituido por personas de la comunidad que recibirán reportes anuales, durante los primeros tres años de la investigación, del LCFEP en los cuales se describirá la forma en que están siendo usadas las muestras y los resultados que se estén obteniendo.

**5. ¿Se comercializarán las muestras?** No. Las muestras obtenidas no tendrán ningún valor comercial y no se comercializarán.

**6. ¿Existe algún beneficio directo para los participantes por donar la muestra?** Los participantes del estudio probablemente no se beneficiarán personalmente por donar las muestras al presente proyecto, debido a que este tipo de investigaciones conlleva un largo período de tiempo en producir resultados. No obstante la participación ayudará a los investigadores de este y otros proyectos futuros a entender más acerca de la variación genética y la forma en que se relaciona dicha variación con la salud y la enfermedad. Los donantes y su comunidad podrán entender más acerca de las relaciones genéticas con otras culturas o étnias a lo largo de América Latina y del resto del mundo y conocer el grado de mezcla que pueda existir con otros grupos étnicos. Una vez los investigadores analicen los datos de su comunidad, y si sus líderes lo consideran pertinente, los investigadores del LCFEP preparan seminarios informativos donde se presenten los resultados publicados.

**7. ¿Cómo se protegerá mi privacidad?** La privacidad de los participantes en la investigación será protegida de varias formas: Protegeremos su privacidad de muchas formas:

1. Se almacenará el documento de consentimiento firmado por usted en un archivo guardado bajo seguridad y sólo los miembros del grupo de Citogenética de la Universidad del Tolima tendrán acceso a este archivo.

2. No se registrarán su nombre ni cualquier otra información personal (dirección, teléfonos, lugar o fecha de nacimiento, número de identificación), tampoco se le asignará a su muestra un número o código que pueda identificarlo.

3. Nadie –ni siquiera usted o nosotros- sabrá con seguridad si su muestra fue empleada o si alguna información en la base de datos proviene de ella.

**Debido a las anteriores medidas de seguridad, será muy difícil para cualquier persona que busque en las bases de datos científicas, conocer cual información proviene del donante, o incluso si toda la información proviene de él.**

**8. ¿Existe algún tipo de riesgo por donar la muestra?** Protegeremos su privacidad de muchas formas:

1. La toma de la muestra no genera un riesgo mayor que el de cualquier otro examen de sangre. Probablemente usted sólo experimentará un pequeño pinchazo durante la extracción. Este proceso será realizado por personal médico certificado por una entidad gubernamental reguladora de salud, el cual tendrá en cuenta todas las medidas de bioseguridad.

2. Aunque pueden existir riesgos por tener información depositada en bases de datos científicas, es muy improbable que alguien relacione la información que allí encuentre directamente con los participantes de

la investigación. Sin embargo hay dos formas de riesgo, la primera es que alguien sospeche de que exista información de algún participante en una bases de datos, que tome, con engaño, una muestra del donante, desarrollar muchas pruebas sobre dicha muestra nueva y comparar los datos nuevos con los que ya existían en la base de datos. La otra vía ocurre si alguien compara la información en la base de datos con información genética previa relacionada con sus datos personales, disponible en otras bases de datos, descubriéndolo. La posibilidad de que estos riesgos ocurran es demasiado pequeña en nuestro país, pero puede crecer en el futuro. Nosotros no podemos predecir los resultados de investigaciones futuras (a largo plazo), de manera que pueden emerger nuevos riesgos en el futuro que no podemos predecir ahora. Por ello tratamos de hacer el procedimiento de la manera más confidencial posible, lo cual también depende del donante, en la medida en que restrinja el número de personas que conoce que ha donado una muestra de sangre. El papel del grupo asesor será el velar porque los dos riesgos mencionados anteriormente se minimicen.

**9. ¿Hay algún tipo de riesgo para mi comunidad?** Los nombres de los grupos geográficos o étnicos de los cuales provienen las muestras se incluirán tanto en las muestras como en las bases de datos. En futuros estudios, los investigadores podrán encontrar que ciertas variantes genéticas aparecen con mayor frecuencia en su grupo en comparación con otros, y que estas variaciones son más comunes en personas con ciertas enfermedades o estados saludables. Esto podría generar problemas de discriminación genética. Esta práctica, sin embargo, es ilegal en la mayoría de los países del mundo.

En América Latina existen políticas de protección de comunidades indígenas que pueden contrarrestar el uso inapropiado de la información derivada de este proyecto por parte de terceros. Estos mecanismos de protección se basan en **la declaración de Helsinki sobre Principios éticos para las investigaciones médicas en seres humanos; la Declaración de las Naciones Unidas sobre los derechos de los pueblos Indígenas, y también a la ley 21 de 1991 de la constitución Nacional de Colombia, sobre la ratificación del convenio 169 acerca de los pueblos indígenas y tribales en países independientes.**

Algunas personas pueden usar esta información de las bases de datos científicas, o de estudios futuros para exagerar las diferencias entre los grupos por prejuicios, entre otras razones negativas. Otros pueden usar esta información para minimizar las diferencias entre los grupos humanos para decir que todos los genes de las personas son los mismos, de manera que no es necesario preocuparse por las diferencias entre los grupos. La biología no proporciona una razón para prejuicios, pero la discriminación existe. Por lo anterior, se trabajará meticulosamente para asegurar que la identidad geográfica o étnica de su comunidad sea descrita de la manera más cuidadosa posible durante la recolección de la muestra, en las bases de datos científica y en los artículos que los investigadores escriban basados en esta investigación, pero no podemos controlar como será interpretada la información por otros autores en sus artículos.

**10. ¿Puedo cambiar mi forma de pensar después de donar la muestra?** Debido a que no asociaremos nombres a las muestras, no se conocerá la identidad de la persona que ha donado una muestra particular. Por esta razón una vez el participante dona la muestra de sangre no será posible retirarla del proyecto. También, una vez haya sido analizada, el donante no podrá retirar la información acerca de la muestra que esté disponible en las bases de datos, debido a que no sabrá si le pertenece o no. En el caso en que se decida crear grupos de asesoría de la comunidad, ellos pueden solicitar retirar todas las muestras del LCFEP, en el caso de que las muestras usadas en estudios futuros estén siendo empleadas de manera inconsistente con los parámetros descritos en este consentimiento informado y por lo tanto la comunidad no esté de acuerdo con estos nuevos estudios.

**11. ¿Cómo puedo saber acerca de los resultados del proyecto de investigación?** Es probable que pase mucho tiempo antes de que los resultados de este proyecto y estudios futuros, pueda generar conocimientos relacionados con la salud. Por esta razón y debido a que no se conoce quien ha sido el donante, no será posible dar resultados individuales sobre esta investigación. Sin embargo hay varias formas en que el donante podrá recibir noticias sobre los resultados que se estén obteniendo, una de ellas es por medio de la socialización que los investigadores programen en su comunidad a través del grupo de asesoría que se ha creado, dándose a conocer también la forma en que se están usando las muestras, la información en las bases de datos y como están ayudando en el conocimiento de las enfermedades y la salud. Otra forma adicional es la revisión de las páginas web donde estarán disponibles los datos generados con las muestras de su comunidad, lo cual se informará en el momento oportuno.

## CONSENTIMIENTO COMUNITARIO DE PARTICIPACION

### ¿A quién me puedo dirigir si tengo alguna pregunta o inconveniente?

Si tiene alguna duda acerca de este proceso de recolección de muestras, dirigirse a:

Investigador1: \_\_\_\_\_ Teléfono: \_\_\_\_\_

Investigador2: \_\_\_\_\_ Teléfono: \_\_\_\_\_

Si tiene alguna duda acerca de sus derechos como parte de este proyecto investigativo, contactar a:

Nombre: \_\_\_\_\_ teléfono: \_\_\_\_\_

[Opcional; se usará sólo cuando exista un grupo asesor de la comunidad]: si el donante desea consultar el grupo asesor de la comunidad para aclarar alguna duda acerca de las muestras colectadas en su comunidad, contactar a:

Nombre: \_\_\_\_\_ teléfono: \_\_\_\_\_

Título o Rango en la comunidad: Gobernador \_\_\_\_; otro \_\_\_\_, ¿Cuál? \_\_\_\_\_

[Opcional; se usará sólo cuando sea necesario un traductor que haya leído este documento al responsable de la comunidad]

Nombre del traductor: \_\_\_\_\_ Teléfono: \_\_\_\_\_

### Consentimiento y firma

**Por favor lea la siguiente Información, piense acerca de su elección y firme si está de acuerdo. Yo estoy de acuerdo con:**

- Informar a mi comunidad sobre los objetivos de la investigación
- Invitar a participar a mi comunidad en la investigación.
- Hacer claro a los miembros de la comunidad que la participación es voluntaria.

Yo he leído o escuchado la información de este documento, yo he hecho preguntas que tenía, y todas mis dudas fueron aclaradas.

**Nombre:** \_\_\_\_\_ **CC:** \_\_\_\_\_

**Firma** \_\_\_\_\_ **Fecha** \_\_\_\_\_

**Copia dada al participante:** \_\_\_\_\_ **si**

**Anexo A2:** Consentimiento Individual. Fuente: Basado en el formato de consentimiento informado #3 del proyecto: *1000 Genomes Project, Developing a Research Resource for Studies of Human Genetic Variation*.

## CONSENTIMIENTO INDIVIDUAL PARA PARTICIPAR

---

RESPETADO SEÑOR: *Por favor lea detenidamente este documento, durante el tiempo que considere necesario. Al final, después haber aclarado sus dudas y logrado comprenderlo, decida si está de acuerdo en participar o no.*

### 1) Introducción:

El grupo de investigación de Citogenética, filogenia y evolución de poblaciones (LCFEP), de la Universidad del Tolima está desarrollando un proyecto de investigación denominado “*Estudio de la estructura genética de las poblaciones Amerindias Latino Americanas*”, el cual pretende estudiar las diversas variantes genéticas que existen entre los pobladores nativos del continente, con el objetivo de entender los procesos de poblamiento Americano, las relaciones ancestrales entre las diversas comunidades indígenas de la región, su relación con otros grupos étnicos y el impacto de la mezcla reciente con poblaciones foráneas. La participación de su comunidad en el proyecto es de gran importancia.

Si usted está de acuerdo formar parte del proyecto, su participación implica que:

1. No se producirán patentes, vacunas o algún conocimiento que genere ganancias económicas.
2. Su muestra no se empleará para realizar un clon de usted o patentar secuencias genómicas provenientes de la misma.
3. No se anotarán el nombre ni los datos médicos particulares (excepto la consideración de ser sano y no tener antecedentes de enfermedades familiares).
4. La decisión de participar o no es completamente suya. Si decide no firmar este documento, esto no afectará cualquier beneficio al cual usted tenga derecho y no implicará sanciones de ninguna clase.
5. Usted no incurrirá en gastos por participar en este proyecto. Lo suyo es una colaboración y, por tanto, nosotros cubriremos los gastos de la toma de muestra y de todos los análisis.
6. No existe algún beneficio directo para los participantes por donar la muestra. Usted probablemente no se beneficiará personalmente por donar la muestra, debido a que este tipo de investigaciones requiere un largo período de tiempo en producir resultados, que sean aplicables en la medicina. No obstante su participación ayudará a los investigadores de este y otros proyectos futuros a entender más, acerca de la variación genética y la forma como se relaciona con la salud y la enfermedad.
7. Usted y su comunidad podrán entender más acerca de las relaciones genéticas con otras culturas o étnias a lo largo de América Latina y conocer el grado de mezcla que pueda existir.

**¿Cómo se protegerá mi privacidad?** Protegeremos su privacidad de muchas formas:

1. Se almacenará el documento de consentimiento firmado por usted en un archivo guardado bajo seguridad y sólo los miembros del grupo de Citogenética de la Universidad del Tolima tendrán acceso a este archivo.

2. No se registrarán su nombre ni cualquier otra información personal (dirección, teléfonos, lugar o fecha de nacimiento, número de identificación), tampoco se le asignará a su muestra un número o código que pueda identificarlo.
3. Nadie –ni siquiera usted o nosotros- sabrá con seguridad si su muestra fue empleada o si alguna información en la base de datos proviene de ella.
4. La toma de la muestra no genera un riesgo mayor que el de cualquier otro examen de sangre. Probablemente usted sólo experimentará un pequeño pinchazo durante la extracción. Este proceso será realizado por personal médico certificado por una entidad gubernamental reguladora de salud, el cual tendrá en cuenta todas las medidas de bioseguridad.
5. Debido a que no se asociarán nombres a las muestras, no se conocerá la identidad de la persona que ha donado alguna en particular. Por esta razón, una vez proporcionada la muestra de sangre, no podremos retirarla del proyecto. De la misma manera, una vez haya sido analizada, usted no podrá retirar la información que esté disponible en las bases de datos, debido a que no sabrá si le pertenece o no a usted.

**¿Que se hará con las muestras de sangre?**

Las muestras se llevarán al laboratorio de Citogenética, Filogenia y Evolución de Poblaciones de la Universidad del Tolima, en Ibagué. Allí se procesarán para extraer DNA y llevar a cabo los análisis en el mismo laboratorio o en otros laboratorios fuera del país.

**¿Cómo puedo saber acerca de los resultados del proyecto de investigación?** Los resultados serán socializados en su comunidad una vez concluya el proyecto.

## CONSENTIMIENTO INDIVIDUAL DE PARTICIPACION

Edad: \_\_\_\_, Género: M\_\_ F\_\_, etnia:

---

### ¿A quién me puedo dirigir si tengo alguna pregunta o inconveniente?

Si tiene alguna duda acerca de este proceso de recolección de muestra, dirigirse a:

Investigador1: *María Magdalena Echeverry* Teléfono: 316 3033776

Investigador2: *Rodrigo Prieto Sánchez* Teléfono: 3152430924

Si tiene alguna duda acerca de sus derechos como parte de este proyecto investigativo, contactar a:

**Nombre:** *Ulises Tique Esquivel* teléfono: 2625085 (CRIT-Ibagué).

[Opcional; se usará sólo cuando exista un grupo asesor de la comunidad]: si usted desea consultar el grupo asesor de la comunidad para aclarar alguna duda acerca de las muestras colectadas en su comunidad, contactar a:

**Nombre:** *Álvaro Yara* teléfono: 3114895828 **cargo en la Institución** y/o rango en la comunidad:  
*Gobernador*

[Opcional: cuando se requiera que alguien traduzca el presente documento al participante, colocar su nombre, identificación persona y teléfono en el siguiente espacio:

---

\_\_\_\_\_

### Consentimiento y firma

**Por favor lea la siguiente información, piense acerca de su elección y firme si está de acuerdo.**

**Yo estoy de acuerdo con:**

- Donar una muestra de sangre.
- Que se obtenga una línea celular a partir de la muestra, lo cual permitirá a los investigadores obtener ADN por un largo periodo de tiempo.
- La línea celular y el ADN se usarán tanto en el Programa de investigación: “Análisis genético de enfermedades humanas” como en otros proyectos investigativos. El DNA obtenido de las muestras de sangre se podrá analizar en laboratorios de Colombia o en el extranjero.
- Obtener el código genético completo a partir de la muestra, el cual será depositado en bases de datos científicas de libre acceso en internet.
- Que una vez la muestra se haya estudiado, no puedo tomar y retirar la información de la muestra que está en internet.

Yo he leído o escuchado la información de este documento, yo he hecho preguntas que tenía, y todas mis dudas fueron aclaradas. Yo se que donar la muestra es mi elección.

**Firma** \_\_\_\_\_ **Fecha** \_\_\_\_\_

**Copia dada al participante:** \_\_\_\_si

**Anexo B.** Detalles para cada uno de los 100 AIMs seleccionados, en orden aparece: IDrs (código del SNP), el cromosoma y la posición, el alelo 1 (A1) y el alelo 2 (A2), las frecuencias alélicas reportadas para amerindios (FA-Ami), europeos (FA-Eur), y africanos (FA-Afr), la población para la cual es informativo, los valores LSBL(Fst) y LSBL(ln).

IDrs SNP	Cromosoma	posición	A1	A2	FA-Ami	FA-Eur	FA-Afr	población	LSBL(Fst)	LSBL(ln)
rs10779334	1	217173027	G	A	0,008	0,097	0,837	AFR	0,606	0,36
rs12021830	1	119363905	T	C	0,107	0,838	0,995	NAM	0,626	0,376
rs2025236	1	203049828	A	G	0,081	0,887	0,277	EUR	0,484	0,275
rs6684063	1	30471927	T	G	0,183	0,841	0,218	EUR	0,411	0,223
rs10175357	2	35015176	C	T	0,038	0,059	0,82	AFR	0,605	0,352
rs16851773	2	214897054	G	A	0,126	0,647	0,024	EUR	0,342	0,195
rs305163	2	36163166	G	A	0,172	0,712	0,063	EUR	0,356	0,196
rs1809119	2	174234175	C	A	0,046	0,207	0,917	AFR	0,608	0,357
rs6710083	2	192480837	T	G	0,004	0,078	0,84	AFR	0,633	0,38
rs6711746	2	29292053	T	C	0,102	0,809	0,874	NAM	0,546	0,308
rs1196705	2	149927468	G	A	0,015	0,638	0,117	EUR	0,344	0,198
rs1567803	2	100709450	T	C	0,019	0,642	0,029	EUR	0,429	0,253
rs1606237	2	205955695	T	C	0,019	0,334	0,971	AFR	0,591	0,371
rs2140701	2	9725153	G	A	0,034	0,586	0,937	NAM	0,501	0,307
rs298259	2	157033071	C	T	0,16	0,761	0,131	EUR	0,381	0,206
rs3927978	2	178152100	G	A	0,027	0,288	0,99	AFR	0,667	0,429
rs7594727	2	96853597	A	C	0,008	0,136	0,869	AFR	0,615	0,366
rs7598069	2	98127823	G	A	0,011	0,61	0,971	NAM	0,571	0,368
rs13097560	3	85212019	A	G	0,046	0,658	0,068	EUR	0,392	0,221
rs862500	3	64247689	A	T	0,134	0,821	0,117	EUR	0,485	0,268
rs2937673	3	118715077	G	A	0,019	0,553	0,044	EUR	0,327	0,19
rs6768750	3	7215675	G	A	0,073	0,695	0,976	NAM	0,542	0,325
rs6803290	3	59667208	C	T	0,023	0,636	0,01	EUR	0,435	0,262
rs7630522	3	108635778	C	T	0,058	0,787	0,262	EUR	0,372	0,208
rs7630111	3	115648591	A	C	0,186	0,756	0,117	EUR	0,366	0,198
rs924780	4	148905447	A	C	0,115	0,19	0,942	AFR	0,625	0,364
rs7665516	4	11243690	G	A	0,023	0,637	0,888	NAM	0,546	0,333
rs10026397	4	41260644	A	G	0,046	0,194	0,932	AFR	0,643	0,381
rs7664076	4	172625303	A	G	0,344	0,803	0,034	EUR	0,334	0,191
rs920559	4	102437738	C	G	0,275	0,856	0,211	EUR	0,378	0,205
rs1107154	5	165319891	T	C	0,061	0,39	0,985	AFR	0,557	0,352
rs11960137	5	155270659	C	G	0,122	0,875	0,942	NAM	0,614	0,353
rs17157450	5	104384818	C	T	0,145	0,864	0,903	NAM	0,544	0,305
rs11743823	5	168097322	C	T	0,103	0,814	0,947	NAM	0,59	0,342
rs6866970	5	100978603	G	A	0,015	0,06	0,858	AFR	0,674	0,403
rs1875017	5	10504154	A	G	0,286	0,894	0,301	EUR	0,374	0,204

rs2042314	5	170110158	G	A	0,061	0,86	0,854	NAM	0,638	0,37
rs349323	5	100804840	T	A	0,031	0,629	0,015	EUR	0,417	0,247
rs353373	5	40261722	G	A	0,191	0,863	1,000	NAM	0,53	0,315
rs10946113	6	158504679	C	A	0,151	0,896	0,971	NAM	0,607	0,354
rs12701745	7	4558764	G	A	0,012	0,526	0,024	EUR	0,325	0,194
rs2707575	7	147269656	G	A	0,031	0,574	0,034	EUR	0,348	0,201
rs7808899	7	14779699	T	G	0,298	0,83	0,146	EUR	0,362	0,197
rs17170516	7	32785284	G	A	0,164	0,841	0,981	NAM	0,54	0,315
rs2384982	7	103493636	C	T	0,187	0,938	0,398	EUR	0,424	0,242
rs17150066	8	9347921	G	A	0,088	0,858	0,985	NAM	0,674	0,405
rs16880715	8	111524781	T	G	0,019	0,091	0,859	AFR	0,641	0,379
rs2450219	8	119372638	T	C	0,057	0,641	0,058	EUR	0,374	0,21
rs2736340	8	11381382	C	T	0,089	0,792	0,879	NAM	0,556	0,315
rs956969	8	62227592	T	C	0,034	0,623	0,049	EUR	0,381	0,218
rs9642819	8	114231199	G	A	0,031	0,731	0,806	NAM	0,565	0,333
rs13280988	8	112439692	A	G	0,031	0,581	0,058	EUR	0,333	0,19
rs6994396	8	114881166	G	T	0,115	0,865	0,85	NAM	0,552	0,31
rs12380601	9	73838451	A	G	0,031	0,583	0,029	EUR	0,36	0,21
rs2809393	9	119931458	C	T	0,103	0,867	0,35	EUR	0,389	0,218
rs10751867	10	2339064	T	C	0,351	0,916	0,175	EUR	0,429	0,241
rs2795918	10	55847021	T	C	0,046	0,692	0,85	NAM	0,533	0,311
rs9422913	10	127335597	C	T	0,771	0,118	0,718	EUR	0,399	0,218
rs11034734	11	38427247	C	T	0,019	0,512	0,01	EUR	0,318	0,192
rs1790740	11	66886698	T	C	0,225	0,929	0,17	EUR	0,542	0,308
rs878874	11	94228113	A	G	0,107	0,843	0,985	NAM	0,628	0,374
rs485645	11	125938189	A	G	0,046	0,641	0,966	NAM	0,536	0,328
rs7124676	11	64069867	A	G	0,008	0,521	0,029	EUR	0,318	0,191
rs7927234	11	27392932	C	G	0,012	0,11	0,852	AFR	0,614	0,363
rs874189	11	15838275	A	G	0,023	0,689	0,107	EUR	0,405	0,232
rs1396883	11	34714023	A	G	0,023	0,559	0,029	EUR	0,343	0,201
rs7103088	11	129514516	A	G	0,019	0,111	0,917	AFR	0,713	0,429
rs7961436	12	10134132	G	A	0,057	0,64	0,084	EUR	0,353	0,196
rs7958163	12	112759625	C	A	0,034	0,702	0,172	EUR	0,357	0,201
rs314580	13	88709705	A	G	0,366	0,953	0,403	EUR	0,365	0,207
rs9576996	13	39813272	G	C	0,223	0,953	0,971	NAM	0,564	0,329
rs9533623	13	43278753	A	G	0,05	0,552	0,01	EUR	0,325	0,192
rs17359176	13	22565334	G	A	0,204	0,941	0,995	NAM	0,592	0,354
rs2585897	13	20296979	G	A	0,076	0,857	0,884	NAM	0,631	0,364
rs10145908	14	62823082	C	T	0,011	0,086	0,83	AFR	0,608	0,36
rs1244379	14	33750102	C	T	0,088	0,635	0,063	EUR	0,341	0,189
rs6573746	14	66576406	A	C	0,512	0,056	0,743	EUR	0,344	0,196
rs11625446	14	47314308	C	A	0,043	0,769	0,922	NAM	0,638	0,38



rs1952790	14	37412427	T	C	0,02	0,545	0,034	EUR	0,327	0,192
rs1958053	14	33253458	T	A	0,023	0,692	0,789	NAM	0,542	0,323
rs2251244	14	61093017	C	A	0,717	0,181	0,985	EUR	0,407	0,237
rs4402506	15	91931362	C	T	0,05	0,588	0,015	EUR	0,357	0,209
rs7164838	15	32754866	G	A	0,08	0,702	0,976	NAM	0,536	0,32
rs1453858	15	46091402	T	A	0,206	0,911	0,228	EUR	0,489	0,273
rs2304580	15	38816959	T	C	0,136	0,842	0,927	NAM	0,555	0,315
rs7173885	15	93608135	G	C	0,008	0,249	0,917	AFR	0,581	0,351
rs746655	15	71894730	A	G	0,004	0,122	0,854	AFR	0,608	0,363
rs8041147	15	35085650	T	C	0,092	0,902	0,223	EUR	0,546	0,311
rs4787040	16	7500981	T	A	0,08	0,714	0,039	EUR	0,448	0,255
rs4347621	16	50241402	G	A	0,065	0,152	0,903	AFR	0,625	0,361
rs9923547	16	12685472	A	C	0,011	0,258	0,942	AFR	0,612	0,374
rs2052074	17	56516428	C	T	0,118	0,177	0,966	AFR	0,677	0,403
rs7211426	17	51009547	A	G	0,092	0,212	0,951	AFR	0,637	0,376
rs4812831	20	42451674	G	A	0,168	0,904	0,961	NAM	0,586	0,338
rs6101991	20	38515722	G	A	0,038	0,104	0,853	AFR	0,608	0,352
rs4812381	20	37420974	A	G	0,046	0,694	0,845	NAM	0,532	0,31
rs2187239	21	42293624	A	C	0,023	0,684	0,819	NAM	0,55	0,329
rs2300395	21	35167068	T	C	0,011	0,294	0,951	AFR	0,594	0,366
rs2836181	21	38477680	T	C	0,016	0,049	0,801	AFR	0,604	0,358
rs132663	22	34892970	A	T	0,027	0,684	0,893	NAM	0,581	0,351

Fuente: CEGEN.

**Anexo C.** Secuencias de los cebadores para la reacción de amplificación, para las regiones de los 100 SNPs.

Ensayo	rs-ID SNP	Forward	Reverse
E1	rs10145908	ACGTTGGATGTGCTCGACCTTCTGAATTC	ACGTTGGATGATTGCAATTTAGTGGCCAAC
E1	rs10751867	ACGTTGGATGCCTCTCAGAATAAAATTAC	ACGTTGGATGGCAGTTCTATTTCTTTCAAC
E1	rs11034734	ACGTTGGATGTGTAGTTCTCCTTGTGGAGC	ACGTTGGATGATCACCACACAAAAGCCCTG
E1	rs1107154	ACGTTGGATGACACTTCTGCCACCTAACTC	ACGTTGGATGAAGACAGCACAGATGGTTAG
E1	rs11960137	ACGTTGGATGTGGAGCCATGGCCTTTCTAA	ACGTTGGATGATGTTAGCGGAATAGCAATG
E1	rs12380601	ACGTTGGATGAGACGGACAATTCTAGGAGG	ACGTTGGATGGATTGCTCCTACCTCCTATC
E1	rs1244379	ACGTTGGATGAAAGGATATAGACCGATTG	ACGTTGGATGGCTACCTATGATTTCTTTGC
E1	rs12701745	ACGTTGGATGGAACCACCACATGCACACTC	ACGTTGGATGAACTGGTGCAGCTTTGCTGG
E1	rs13097560	ACGTTGGATGCTCAGGTATATCAATTAAC	ACGTTGGATGTAAAGATTGGTCTGACCAGC
E1	rs16851773	ACGTTGGATGCTAGAGGCAAAGTCCTAAGC	ACGTTGGATGGCATGGAACACATTTACATC
E1	rs17150066	ACGTTGGATGTCTCTGTATCATAACCTCG	ACGTTGGATGTGCTCACTCAGCACACATCT
E1	rs17157450	ACGTTGGATGGACTGTTTTCTTATCTGTTT	ACGTTGGATGTGCAAAGACCACATGTCCAG
E1	rs2450219	ACGTTGGATGCTTTATACATACACCTTTTC	ACGTTGGATGCAGGCGAATGAAAGAAGTTG
E1	rs2937673	ACGTTGGATGGACAACCTGAAGACTTAGCGG	ACGTTGGATGGACATCACAATATGATAGCAG
E1	rs4402506	ACGTTGGATGAGCCTGTCTCATTGCTGAAC	ACGTTGGATGAGGCCAATTAACAAGAGGG
E1	rs4812831	ACGTTGGATGCACGCAGAAGGTGCTGAATA	ACGTTGGATGGCTGGATGAATAACTGAGTG
E1	rs6101991	ACGTTGGATGCCTGTGCTATTTCTTTGTT	ACGTTGGATGGCAGAATCTATAGTCCCCTG
E1	rs6710083	ACGTTGGATGAATGGACCTGTACTGCCTTC	ACGTTGGATGGCAGCAGGAAGTCAGAAGTA
E1	rs6768750	ACGTTGGATGGAACCTTGAGATACAGTCC	ACGTTGGATGATGGGATTCATAACAAGG
E1	rs6866970	ACGTTGGATGGAACCAATTAGAGGAGTAG	ACGTTGGATGTTAGAAGGCATTCCCAGGAC
E1	rs7665516	ACGTTGGATGAGCAATAGGGTCTGCATAGG	ACGTTGGATGGCTTCATGCTTGTTTTTAGG
E1	rs7808899	ACGTTGGATGGAAGAGACCACAAGAGTGAG	ACGTTGGATGCAGTTTAGTTGACTGTCCC
E2	rs10175357	ACGTTGGATGAGCTTCTTGTCTCACCATC	ACGTTGGATGTGGTGGTAGTCAGTGATCAG
E2	rs10946113	ACGTTGGATGGATAAATAGATCGATGGAGTG	ACGTTGGATGCAAACAAAAACCCTAATGG
E2	rs11743823	ACGTTGGATGGTTCTGGAAACAACTCTGC	ACGTTGGATGTGTTACCAGCACTGGATTC

E2	rs16880715	ACGTTGGATGAGCTGGTGATACCTGGTTAG	ACGTTGGATGGATCCCAGAATTTCTGTCAAC
E2	rs1790740	ACGTTGGATGCTCCCATGCTAGAACTCCA	ACGTTGGATGATGAATGAATAGCTCCCAGG
E2	rs1809119	ACGTTGGATGTTGTTTAGCCTGGGTAGGAG	ACGTTGGATGAGCTGTTTCTGGAAATCGCC
E2	rs2707575	ACGTTGGATGCTGTAACCAAAATAGAAGC	ACGTTGGATGGTGGGTACACAAACATGTGC
E2	rs2736340	ACGTTGGATGCATGGCCACCATTAGTGATG	ACGTTGGATGTCTTGACCCCCATCGCATTG
E2	rs2795918	ACGTTGGATGGCAAATCAGTATACAAATTCC	ACGTTGGATGGCTGCTATTAAGAATTAC
E2	rs305163	ACGTTGGATGGTATGGTAGTGTATTTTCAGTC	ACGTTGGATGGTTATTTCAAGTCCTAAACC
E2	rs314580	ACGTTGGATGTTAAAATTTGATAGCATGAG	ACGTTGGATGTATTTACTTGTGCTTTTCC
E2	rs4787040	ACGTTGGATGCAAACCTTCTGAGCACACC	ACGTTGGATGCACAGCAGGAAATAAACTGG
E2	rs485645	ACGTTGGATGGTTCCAGGCATAAGTTTAAGC	ACGTTGGATGGGAAACATCCTTCACAGGGT
E2	rs6573746	ACGTTGGATGAGGCCTTAACATCAAGTCC	ACGTTGGATGCAACCGTTCTACTTTGTGAG
E2	rs6711746	ACGTTGGATGCATTGGGCACTGTAGTCAC	ACGTTGGATGCCTGGGATCCAAGGAGTTAA
E2	rs6803290	ACGTTGGATGCATGGAATCCTCGATCTGAC	ACGTTGGATGGGAAGGTATTAGTTCATGC
E2	rs7124676	ACGTTGGATGGATGTGTGTCACCATGCAGG	ACGTTGGATGCTTGCTCCATTGCCTGATG
E2	rs7164838	ACGTTGGATGGTGTCTAAGAAGAAGGCCAG	ACGTTGGATGGAGTCCCTGAGAGTCTTTTG
E2	rs7630522	ACGTTGGATGTTTTCTGACCCTCTGCTAC	ACGTTGGATGAAGGAACAGAAGAGGTCTGC
E2	rs7927234	ACGTTGGATGACACTGCCACATAGCCATTG	ACGTTGGATGAAGCCACCCAAGCAATCAGC
E2	rs7961436	ACGTTGGATGACTCCAGCCTGGAAAATCTC	ACGTTGGATGCTTTTCCTTCTAGCAAGGTC
E2	rs862500	ACGTTGGATGTTACTCTCTCATATCAC	ACGTTGGATGATCTTTACCTGTCAATTGCTG
E2	rs874189	ACGTTGGATGATCAGACAGCCCTCAGACAC	ACGTTGGATGACTTGATGAGAGCCTGGATG
E2	rs878874	ACGTTGGATGTATCCACCAAGCCAAGACAG	ACGTTGGATGTGGAGAAGGAGGTGTTAGTG
E2	rs924780	ACGTTGGATGCACTGCACCCAGCCCAGTA	ACGTTGGATGAACTTACAGCTCCTCAGCAC
E2	rs9422913	ACGTTGGATGGTTGTGAACAAACAGAATG	ACGTTGGATGCTCATTCACTCTCAGCACCT
E2	rs9533623	ACGTTGGATGTTAGATGCCACAGGTGGTC	ACGTTGGATGTGGTTTGGATGCTACATTAC
E2	rs956969	ACGTTGGATGTGGTGATAAGGCGCCATCTT	ACGTTGGATGAAGTCCAAGATCAAGGCTCC
E2	rs9576996	ACGTTGGATGACATTTTGCCTGTCTAAAGC	ACGTTGGATGCAGCCAGTTTCTTTGTGGAC
E2	rs9642819	ACGTTGGATGTGGAACTTGATTCAGCTCAC	ACGTTGGATGTGCTATCCCTCAAAGGTTCC
E3	rs10026397	ACGTTGGATGGGAGACTATAACATTTTGCC	ACGTTGGATGTTGACAGGAGATCCCTGATG

E3	rs10779334	ACGTTGGATGTCCACTCTTTTATTGTGGG	ACGTTGGATGCCTTTCTGATTATGAACTGG
E3	rs1196705	ACGTTGGATGAATGAGTTCTAGCCCCCTTC	ACGTTGGATGATCCCATGGCAGAAGGTAGA
E3	rs12021830	ACGTTGGATGAAAGAGACTGGATACTGGCG	ACGTTGGATGCACTGTACACGTATATTTGC
E3	rs132663	ACGTTGGATGCTAAACGGATTGAGATCTGTC	ACGTTGGATGTGGCTTTCAACCAGCAACTC
E3	rs13280988	ACGTTGGATGAAGAGGGAGTTCTTTCTGGC	ACGTTGGATGGAAGGGACTGAGGCATTCTG
E3	rs1396883	ACGTTGGATGCAGTAAAGATGCAAGCTGCC	ACGTTGGATGTTTCTCTGCAGCATGCTACC
E3	rs1453858	ACGTTGGATGTGGAGTTTGAATTTGAGGG	ACGTTGGATGGGGGAAAAAAGCATTTCCTT
E3	rs1567803	ACGTTGGATGTCCAGGTGACTGCTGGTTTG	ACGTTGGATGTTTCATCATCTCAGTGAGTGG
E3	rs2187239	ACGTTGGATGGGGAATCCAAACTAAGACAC	ACGTTGGATGGGAAGGGAGTTTTAGTGTG
E3	rs2251244	ACGTTGGATGATTCCAGACTGACACTAACC	ACGTTGGATGATTTCCCTTATGGCCAAGCTG
E3	rs2304580	ACGTTGGATGCAGTGGGTCTGAAGAAACCT	ACGTTGGATGGTGTTCCTGTTCTATGGG
E3	rs2384982	ACGTTGGATGTTGGAAACAGGAGAACAGGC	ACGTTGGATGAGTAGGTCCTTGTTGGGT
E3	rs2585897	ACGTTGGATGGGAAGGAGAAGCAAGTAAGG	ACGTTGGATGAATTTGGATGGCCAACAGAC
E3	rs2836181	ACGTTGGATGTATAACAGATGCCAGGAGAC	ACGTTGGATGCTTGGCAGAATCTAAGCCAC
E3	rs298259	ACGTTGGATGTTACAGACCCGTTTGCTAGG	ACGTTGGATGAAAAGTGCATGAGTCTTAC
E3	rs349323	ACGTTGGATGCAAAAATAACAAAGTGCC	ACGTTGGATGCTGATTGGTCCAAAACAGAA
E3	rs353373	ACGTTGGATGGGACCATGAAAAGATGCACG	ACGTTGGATGAGGTGATACTTTCTCAGTG
E3	rs4812381	ACGTTGGATGGAAGGAAGCCAATACATGTG	ACGTTGGATGGTAATTCTGGTATGATTAGC
E3	rs6684063	ACGTTGGATGGGTTTTGGATCATTAGGTGC	ACGTTGGATGGAAGTCAGAATAGAACGTCTC
E3	rs6994396	ACGTTGGATGGGGTCTCTCAATCATATGGG	ACGTTGGATGGGGTCCATGTTTCAGTTTA
E3	rs7103088	ACGTTGGATGTATAACTGGATCAACCCGAC	ACGTTGGATGTGCATGTCTCATGTTGAAG
E3	rs7173885	ACGTTGGATGCTGACCTCTACCCACCAGAT	ACGTTGGATGTGGCACTGTCTGCAAATAG
E3	rs7211426	ACGTTGGATGAGTCATCATCACTTCCAGCC	ACGTTGGATGTGAAAGTTCATTCCACGGGC
E3	rs7594727	ACGTTGGATGACTATGTCTCAGATAAGTGC	ACGTTGGATGCTTAGGTGGCTTGCATTCTG
E3	rs7664076	ACGTTGGATGTTGCAGGTTGTACCTCTTG	ACGTTGGATGGCATCACCACCAAGCATATC
E3	rs7958163	ACGTTGGATGACAAATAGGATTGTTGAGGG	ACGTTGGATGGTGGGAGTATTTTTGGCAAC
E3	rs8041147	ACGTTGGATGGCCTGTGCATAATACACATTT	ACGTTGGATGCACTATCAAAGTCAGCATTC
E3	rs9923547	ACGTTGGATGACTAGCATGGCAGGCTGAAA	ACGTTGGATGCCTGAACCTAGGTCATGCAG

E4	rs11625446	ACGTTGGATGTATCTCCATGATGATGGCCC	ACGTTGGATGGACCCATGAGGTTATTGTGC
E4	rs1606237	ACGTTGGATGGAACATACTTTCTCCCACC	ACGTTGGATGCCATGAATCTCATAATGAC
E4	rs17170516	ACGTTGGATGCAGGGAAGCATGGATTAGAC	ACGTTGGATGTGAATGACTGGGCCTTTATG
E4	rs17359176	ACGTTGGATGCCTTTGCTCTCCAGAGAACT	ACGTTGGATGTCAGGACCAAGGGTTGTTAG
E4	rs1875017	ACGTTGGATGTGCCTCCTTTCTCATAAGC	ACGTTGGATGTCTTGCTGTGAGACATGGTG
E4	rs1952790	ACGTTGGATGAAGGAAGGATGGTATTCAGG	ACGTTGGATGGAGGCATCTTTCAGATTATTC
E4	rs1958053	ACGTTGGATGTATCCATCTTCCTGTGCCCA	ACGTTGGATGGCTTACACCAATTACGATCC
E4	rs2025236	ACGTTGGATGACTGGACACTCACACCTGTT	ACGTTGGATGACTAGAAGGACCTGGAGAAC
E4	rs2042314	ACGTTGGATGGGGAAGAGATTCATGGCTTT	ACGTTGGATGCAGAGCCCAGATGGTTTCTA
E4	rs2052074	ACGTTGGATGAAGTTAGAAGGGTCCAGAG	ACGTTGGATGGTCCCTAATTGCTGCACAAG
E4	rs2140701	ACGTTGGATGTCTGCCTCACTAGGGTTTTG	ACGTTGGATGTGGTGCAGAGAGGATATCCC
E4	rs2300395	ACGTTGGATGTGGGAGTCAAATCACCCCAG	ACGTTGGATGCCACTGGGCAGTTTATTATC
E4	rs2809393	ACGTTGGATGAGCAGCAGGTTTCACAGTAG	ACGTTGGATGCTCACAGGGTATTCCTTGA
E4	rs3927978	ACGTTGGATGTATGCTCTAGCTTGGGACTG	ACGTTGGATGTGGCCTCAAGGCCAATATTT
E4	rs4347621	ACGTTGGATGGCGACAAGAAGAAGGGAGAT	ACGTTGGATGGTGACCTGAAAGCAGTGAAT
E4	rs746655	ACGTTGGATGGTCTGATTTTAAACTCTGCC	ACGTTGGATGTCTGCCTGTAGCCTTATAG
E4	rs7598069	ACGTTGGATGTTACACACTGAAAGGACAGC	ACGTTGGATGACAGCTCCTGAGCCACTCAA
E4	rs7630111	ACGTTGGATGGAAATGAATAATTGCCACAG	ACGTTGGATGGGAAGAGACATGATCAAAC
E4	rs920559	ACGTTGGATGTAGTTTGACCATGCAGAGGG	ACGTTGGATGGATGGTCATCTAGTGGGTTTC

Fuente: CEGEN.

**Anexo D.** Secuencias de los cebadores para la reacción de extensión MassEXTEND, durante la tipificación de los 100 SNPs mediante la metodología de Sequenom MassARRAY iPLEX.

ID-SNP	EXT1	EXT2	Hebra de alineamiento	Ensayo	UEP_DIR	UEP_SEQ
rs10145908	T	C	+	E1	R	tCTAGAAATGTGTTTCCAAACT
rs10751867	C	T	+	E1	F	AACTAAAATTTTGCACCTG
rs11034734	C	T	+	E1	F	aAGCTCTTTCATCCATTGG
rs1107154	T	C	-	E1	R	ccCCTAACTCAGAGACTTAATAGC
rs11960137	G	C	+	E1	R	catccAAATGGTAGTATCATGTTTGT
rs12380601	A	G	+	E1	F	GGACAATTCTAGGAGGCATTTT
rs1244379	T	C	+	E1	R	gaacGGAATGGAAAAAATACACCAGTT
rs12701745	A	G	+	E1	F	GCTTGAGAACCAGCTTA
rs13097560	A	G	+	E1	F	cccgcTGTACTTCTTTATACCTTGGC
rs16851773	G	A	+	E1	R	TTTCATTTTTCTATCAAGCCCA
rs17150066	G	A	+	E1	R	G TTCAGAATCTGGTGTACTTAG
rs17157450	C	T	+	E1	F	CTAAGTATAGTTAATTTAAGTGGAAATG
rs2450219	C	T	+	E1	F	tACCTTTTCTTTTTCTTTAGATTAA
rs2937673	C	T	-	E1	F	CGGTAAGAATTATCTGGAGA
rs4402506	C	T	+	E1	F	aAATCACCAACATCCCTTA
rs4812831	A	G	+	E1	F	cctGTGCTGAATAAATGTTGACTGTTTC
rs6101991	G	A	+	E1	R	cTCATTGTTAAATGGAAAGACATA
rs6710083	G	T	+	E1	F	CTGACATCTTTTCTCTTTGTTTC
rs6768750	G	A	+	E1	R	gACAGTCCACTTTAATAACAGAAA
rs6866970	A	G	+	E1	F	GAGGAGTAGAATGTGGAG
rs7665516	A	G	+	E1	F	ATTTTAATAGAGCTATCATAATGGAG
rs7808899	G	T	+	E1	F	tgGAAAAGCAAAAAATAGCATTG
rs10175357	C	T	+	E2	F	CCTCTGTCTCTCTTCAAT

rs10946113	C	A	+	E2	R	tATCGATGGAGTGAATAG
rs11743823	T	C	+	E2	R	agCTCTGCAAAGCCTACATCTTCATAAC
rs16880715	G	T	+	E2	R	tgggGATACCTGGTTAGTTGAAAGA
rs1790740	T	C	-	E2	R	ccCAAACAAGCTCTCAAG
rs1809119	C	A	+	E2	R	ccGGGTAGGAGGATTCCTTT
rs2707575	G	A	+	E2	R	AAGCAAACCTGGCAA
rs2736340	T	C	+	E2	R	tATTAGTGATGGGTCTGATTG
rs2795918	T	C	+	E2	R	ccTATACAAATTCCTAAAGTAGATCGAC
rs305163	G	A	+	E2	R	AGTCATTTATTACTCCATGACAAA
rs314580	G	A	+	E2	R	agTAAAATTTGATAGCATGAGTCCTACA
rs4787040	T	A	+	E2	R	TCCCTTTCTGCATCTGGAATT
rs485645	G	A	-	E2	R	ggGGCATAAGTTTAAGCCAGCAA
rs6573746	C	A	+	E2	R	ACATCAAGGTCCAAGATGTTA
rs6711746	T	C	+	E2	R	acatACATACGCTCCAACCCCTTCAGGC
rs6803290	T	C	+	E2	R	ggtaCTCGATCTGACCAAGTTC
rs7124676	G	A	+	E2	R	accacGGCCAAGTCCATTCTAT
rs7164838	G	A	+	E2	R	GAAGGCCAGGACACA
rs7630522	C	T	+	E2	F	cacccCCCACAGAGCAATAACCTCG
rs7927234	C	G	+	E2	F	gACATAGCCATTGTACGATA
rs7961436	G	A	+	E2	R	AATCTCTGTACCTAGAATGTATTA
rs862500	A	T	-	E2	F	TCTCTCATATCACTGACTCT
rs874189	C	T	-	E2	F	CAGACACCACCTCTTC
rs878874	C	T	-	E2	F	aggctGAGGTGGGTTTCTTACTTTTC
rs924780	C	A	-	E2	R	aaCCAGCCCAGTAAACATT
rs9422913	T	C	+	E2	R	agtgcGTACACAGTAATTATCAGGTTTC
rs9533623	A	G	+	E2	F	attCCCACAGGTGGTCTATAAAATCT
rs956969	T	C	+	E2	R	aaatCAGCTTTCTCCATACAAAC
rs9576996	G	C	+	E2	R	tcgAAAGCAGGCGTGTTTC

rs9642819	G	A	+	E2	R	TGGCTTCACAGCTGTTA
rs10026397	A	G	+	E3	F	ttttCCTACTTCTTCATTCTTGGACCA
rs10779334	G	A	+	E3	R	CATAGTTTGGCTCCAGA
rs1196705	G	A	+	E3	R	CCCCTTCACTTCTCAC
rs12021830	C	T	+	E3	F	GCGGATAGTTGGCAGT
rs132663	T	A	+	E3	R	ctggcGATCTGTCTCAGACGCA
rs13280988	G	A	+	E3	R	GGCTCCATGCCAATC
rs1396883	G	A	+	E3	R	ACCCACACCATCTAAAA
rs1453858	T	A	+	E3	R	ATGAAGAGATCTTGACACT
rs1567803	T	C	+	E3	R	ATGCTTTATTTTGACCTAATCATTC
rs2187239	C	A	+	E3	F	ccccAACAGAACTCGAAGCACAGGTAA
rs2251244	C	A	+	E3	R	TCCCCACATTTTCAGTACCAT
rs2304580	C	T	+	E3	F	TGGGTCTGAAGAAACCTATATTT
rs2384982	C	T	+	E3	F	gGGCCCCTGGCTTTAGATCAA
rs2585897	A	G	+	E3	F	ggggcAGCAAGTAAGGAAGATGAC
rs2836181	T	C	+	E3	R	gggaAGATGCCAGGAGACTAAAAACC
rs298259	C	T	+	E3	F	CGTTTGCTAGGACCTTC
rs349323	T	A	+	E3	R	caatcAACAAAGTGCCATCTTCT
rs353373	A	G	+	E3	F	gcaagCAACTCAGCTTGAACAGTTA
rs4812381	G	A	+	E3	R	gagGTGTTTTTAAGCCAGCTACCA
rs6684063	G	T	+	E3	R	attcGGTGCTTCTTGACTAACTAGTCTA
rs6994396	G	T	+	E3	R	agGGGAAGATTTTCTGTATTAACCA
rs7103088	G	A	+	E3	R	AACCCGACAGCGTAT
rs7173885	C	G	+	E3	F	cccgACCCACCAGATGCCAGTA
rs7211426	A	G	+	E3	F	ccccCCAGCCTCGCCACCAT
rs7594727	C	A	+	E3	R	cGCTTTTCATCCACACTAC
rs7664076	A	G	+	E3	F	ccccCCTCTTGCCCATGTGATTTCCATTG
rs7958163	C	A	+	E3	R	aagGTTGAGGGGCCTCAT



rs8041147	C	T	+	E3	F	gacagTTTTTTTGCCTATTGGAC
rs9923547	C	A	+	E3	R	cctaAGGCTGAAACTCC
rs11625446	C	A	+	E4	F	cccctCCCAAATGAACTGAACTCCCTAG
rs1606237	C	T	+	E4	F	CACTAATGTCTCTATTTCGC
rs17170516	A	G	+	E4	F	gaataATGGATTAGACAATGAAACA
rs17359176	G	A	+	E4	R	gGAACTGAAGGCTGCTAC
rs1875017	G	A	+	E4	R	GCAGATGCCTGGTTA
rs1952790	T	C	-	E4	R	ccccTGTATGGGATCCAATGCAAAACA
rs1958053	T	A	+	E4	R	ctagCCATCTTCTGTGCCAGCTACG
rs2025236	A	G	+	E4	F	ccGTTCTCTCTGGCCC
rs2042314	A	G	-	E4	F	cGCAAGACCTTAAATCCTATGG
rs2052074	T	C	-	E4	R	GGGGTCCAGAGTGTAGAGAAT
rs2140701	C	T	-	E4	F	GCTGCCGTTTCCTATCG
rs2300395	C	T	+	E4	F	cttccAATCACCCCAGCTGGGAAACA
rs2809393	C	T	+	E4	F	GGTCCAAAGGAAAGTG
rs3927978	T	C	-	E4	R	ccccAGACTCTTGAGCTCCA
rs4347621	G	A	+	E4	R	gggatAAGGGAGATTTATCTGCC
rs746655	C	T	-	E4	F	tACTCTGCCCCGATTC
rs7598069	A	G	+	E4	F	ttCAGCCTGGCTGCTAC
rs7630111	C	A	+	E4	R	aatgGCCACAGAATTTTTGTCA
rs920559	G	C	-	E4	R	gagtGCATCATGGCCAATTGTACCTG

Fuente: CEGEN.

Anexo E. Cebadores empleados para la tipificación de los Hgs de la región NRY.

rs150173	Secuencia
<b>AleloA</b>	GAAGGTGACCAAGTTCATGCTAATTATCTGCCTGAAACCTGCCTGA
<b>AleloC</b>	GAAGGTCTGGAGTCAACGGATTATCTGCCTGAAACCTGCCTGC
<b>Común</b>	CACATCACATACATGATGGACCGAGAT
rs9786184	
<b>AleloA</b>	GAAGGTGACCAAGTTCATGCTCCCCACATATCTCCAGGTGTT
<b>AleloC</b>	GAAGGTCTGGAGTCAACGGATTCCCCACATATCTCCAGGTGTG
<b>Común</b>	GCCACCCTAGCCTTTTAAATATGCAAAAT
rs9786153	
<b>AleloC</b>	GAAGGTGACCAAGTTCATGCTGGGAATGATCAGGGTTTGGTTAATC
<b>AleloT</b>	GAAGGTCTGGAGTCAACGGATTGGGGAATGATCAGGGTTTGGTTAATT
<b>Común</b>	GCCTTCTGAGGCACATATGATAAAAAAAAA
rs2032631	
<b>AleloA</b>	GAAGGTGACCAAGTTCATGCTATCAAAAATTGGCAGTGAAAATTATAGATAA
<b>AleloG</b>	GAAGGTCTGGAGTCAACGGATTATCAAAAATTGGCAGTGAAAATTATAGATAG
<b>Común</b>	CTATCTCCTGGCCTGGACCTCA
rs3900	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTAGAAACGGCCTAAGATGGTTGAATC
<b>aleloG</b>	GAAGGTCTGGAGTCAACGGATTAGAAACGGCCTAAGATGGTTGAATG
<b>Común</b>	GAACGTTTGAACATGTCTAAATTAAGAAA
rs3913	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTGTGCCATTCTGGTGGCTCTA
<b>aleloG</b>	GAAGGTCTGGAGTCAACGGATTGTGCCATTCTGGTGGCTCTG
<b>Común</b>	TAAGGGCCCATATGGGCACTGG
rs9306841	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTGGTTTTTAATATTATACCTGAGTGTTTTATCC
<b>aleloG</b>	GAAGGTCTGGAGTCAACGGATTGGTTTTTAATATTATACCTGAGTGTTTTATCG
<b>Común</b>	ACAGAGCACTTTAAAGTGAGCTGTGAT
rs2032677	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTAATACTTTAAACATCCAGGGAGTGTTTTTC
<b>aleloT</b>	GAAGGTCTGGAGTCAACGGATTGAATACTTTAAACATACAGGGAGTGTTTTTT
<b>Común</b>	CAGGGTTGCAAACGGTAGTATTTCTAAAA
rs3894	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTGGGTCACCTCTGGGACTGAC
<b>aleloT</b>	GAAGGTCTGGAGTCAACGGATTATGGGTCACCTCTGGGACTGAT
<b>Común</b>	GTGAATCTGAAATTTAAGGGCATCTTTTCAT
rs3910	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTGATGTAGAGACATCTGAAACCACT
<b>aleloT</b>	GAAGGTCTGGAGTCAACGGATTGATGTAGAGACATCTGAAACCCACA
<b>Común</b>	TTTTGGGGTTTGTTAAGTGGCCAAACTAT

rs8179021	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTGCAAAAAGGTGACCAAGGTGCTC
<b>aleloT</b>	GAAGGTCTGGAGTCAACGGATTGCAAAAAGGTGACCAAGGTGCTT
<b>Común</b>	CCATGTACGTAAAAATAGATTTTTTTTCAA
rs2032624	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTTTTTCTTACAATTCAAGGGCATTTAGAACA
<b>aleloC</b>	GAAGGTCTGGAGTCAACGGATTCTTACAATTCAAGGGCATTTAGAACC
<b>Común</b>	CTGGCTTATCATTCTGAATATTAACAGAT
rs2032597	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTTACTATTTTTATTTACTTAAAAATCATTGTTCA
<b>aleloC</b>	GAAGGTCTGGAGTCAACGGATTACTATTTTTATTTACTTAAAAATCATTGTTCC
<b>Común</b>	AGGTCCTCATTTTACAGTGAGACACAA
rs13447352	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTGATGTGTTCAATTTGAAAGTAACTTGTGAA
<b>aleloC</b>	GAAGGTCTGGAGTCAACGGATTATGTGTTCAATTTGAAAGTAACTTGTGAC
<b>Común</b>	CAAATAACTTTCAAACGTCTTATACCAA
rs2032636	
<b>aleloG</b>	GAAGGTGACCAAGTTCATGCTATCTAATAATCCAGTATCAACTGAGGG
<b>AleloT</b>	GAAGGTCTGGAGTCAACGGATTGATCTAATAATCCAGTATCAACTGAGGT
<b>Común</b>	CCAGCATCCTATCAGCTTCATCAA
rs2032604	
<b>aleloG</b>	GAAGGTGACCAAGTTCATGCTCCCCAAACCCATTTTGATGCTTG
<b>AleloT</b>	GAAGGTCTGGAGTCAACGGATTCCCCAAACCCATTTTGATGCTTT
<b>Común</b>	TTAAATATCAGCCAGGTACAGAGAAAGTTT
rs9341278	
<b>aleloA</b>	GAAGGTGACCAAGTTCATGCTAAACAACATTTACTGTTTCTACTGCTTTCA
<b>aleloG</b>	GAAGGTCTGGAGTCAACGGATTCAACATTTACTGTTTCTACTGCTTTG
<b>Común</b>	GTGTTTGACACCACAGAAATTACAGGTAT
M346	
<b>aleloC</b>	GAAGGTGACCAAGTTCATGCTGCAGCCAAGAGGACAGTAAGAC
<b>aleloG</b>	GAAGGTCTGGAGTCAACGGATTGCAGCCAAGAGGACAGTAAGAG
<b>Común</b>	CTGATAGTAGGACAGCAATTTACAAAGAC

Fuente: el autor.

**Anexo F.** Valores P del Fst calculado a partir de las frecuencias de los Hgs mitocondriales en las poblaciones del presente estudio.

	Nasa	Pijao	Embera	Zenu	Ibagué	Ortega	Planadas	Córdoba
Nasa	*							
Pijao	0.12613+-0.0364	*						
Embera	0.00000+-0.0000	0.00000+-0.0000	*					
Zenú	0.04505+-0.0203	0.12613+-0.0364	0.00000+-0.0000	*				
Ibagué	0.01802+-0.0121	0.13514+-0.0244	0.00901+-0.0091	0.08108+-0.0212	*			
Ortega	0.01802+-0.0121	0.51351+-0.0508	0.01802+-0.0121	0.00901+-0.0091	0.27027+-0.0470	*		
Planadas	0.00000+-0.0000	0.09910+-0.0252	0.06306+-0.0273	0.03604+-0.0148	0.50450+-0.0562	0.29730+-0.0273	*	
Córdoba	0.00000+-0.0000	0.00000+-0.0000	0.02703+-0.0139	0.12613+-0.0278	0.04505+-0.0203	0.00901+-0.0091	0.02703+-0.0139	*

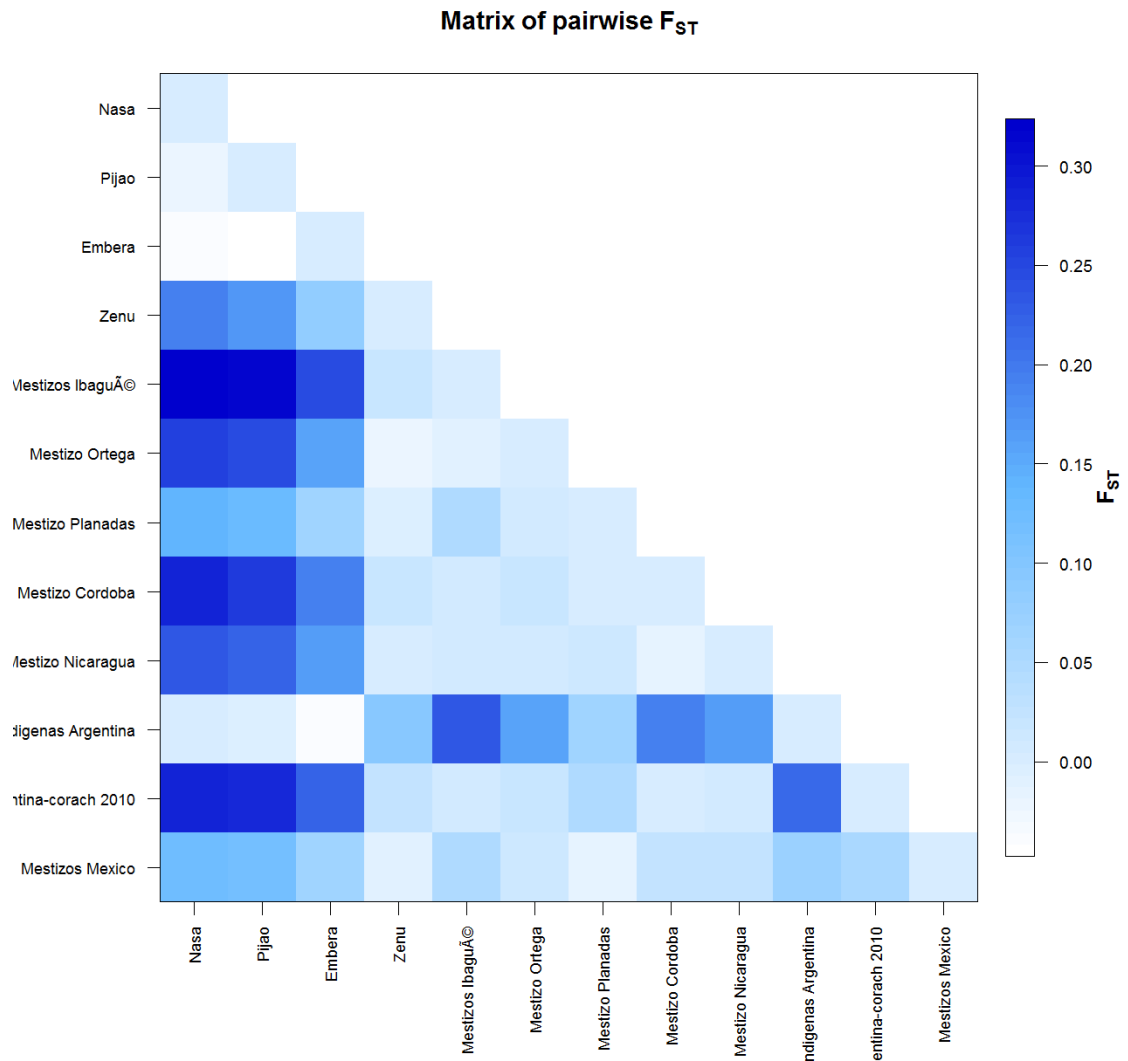
**Fuente:** El autor.

**Anexo G.** Matriz de datos para los valores de  $F_{st}$  pareados entre poblaciones, calculado a partir de las frecuencias de los Hgs mitocondriales en las poblaciones del estudio y de otras regiones de Colombia y Latinoamérica. Para la fuente de los datos en las poblaciones de referencia, ver pág.124. Fuente, el autor.

	Nasa	Pijao	Embera	Zenu	Mayas	Kogi	Nukak	Waunana	Ibagué	Ortega	Planadas	Cordoba	Quibdo	Bogota	Cali	Antioquia	Mexico	Argentina	Caldas	Santander	Nariño	Cauca	
Nasa	0																						
Pijao	0.00934	0																					
Embera	0.20913	0.14504	0																				
Zenu	0.03893	0.01369	0.1503	0																			
Mayas	0.01182	0.04284	0.1697	0.06473	0																		
Kogi	0.04537	0.12721	0.34988	0.13131	0.03298	0																	
Nukak	0.20591	0.17195	0.48443	0.2498	0.28244	0.42123	0																
Waunana	0.24246	0.14014	0.14634	0.20499	0.26159	0.45016	0.31092	0															
Ibagué	0.02976	0.00942	0.07746	0.01272	0.02595	0.10813	0.25217	0.15561	0														
Ortega	0.0367	-0.00642	0.08523	0.04008	0.05128	0.16986	0.20298	0.08519	0.00499	0													
Planadas	0.05971	0.02581	0.04625	0.0359	0.04221	0.15251	0.3291	0.16129	-0.0035	0.00967	0												
Cordoba	0.14536	0.07531	0.0807	0.02151	0.15527	0.27561	0.35475	0.16779	0.04406	0.06851	0.04564	0											
Quibdo	0.22519	0.14196	0.15102	0.07523	0.2461	0.37309	0.38749	0.19612	0.11699	0.13433	0.13133	-0.01127	0										
Bogota	0.0848	0.0546	0.06858	0.01466	0.07297	0.16021	0.33319	0.19709	0.0102	0.05184	0.01129	0.01174	0.07986	0									
Cali	-0.01054	0.00151	0.18205	0.01849	0.00813	0.05183	0.21451	0.22596	0.01383	0.0258	0.04113	0.11142	0.18955	0.05798	0								
Antioquia	0.12197	0.10786	0.03688	0.12103	0.0668	0.18332	0.41914	0.22736	0.03905	0.06731	0.01256	0.12114	0.21934	0.05205	0.10451	0							
Mexico	0.02653	0.01095	0.0886	0.04826	0.02112	0.10617	0.23647	0.1465	0.00275	-0.00285	0.00514	0.09361	0.1739	0.0455	0.01745	0.03864	0						
Argentina	0.04789	0.00968	0.17318	0.03421	0.08459	0.14767	0.20927	0.1716	0.04595	0.03943	0.04942	0.09787	0.169	0.07787	0.0391	0.13508	0.05465	0					
Caldas	0.11847	0.11464	0.05752	0.12231	0.05926	0.16021	0.41911	0.25182	0.04221	0.07872	0.01772	0.12998	0.22941	0.05531	0.10201	-0.00478	0.0433	0.14023	0				
Santander	0.12881	0.05414	0.09012	0.07164	0.14243	0.25422	0.32211	0.12972	0.05552	0.04463	0.02393	0.05428	0.12303	0.06703	0.10879	0.10497	0.07025	0.04328	0.11885	0			
Nariño	0.05114	0.00779	0.0694	0.03615	0.05723	0.15309	0.24027	0.11072	0.00615	-0.00276	-0.00239	0.05661	0.13103	0.0331	0.03673	0.05148	0.00905	0.0297	0.06157	0.02295	0		
Cauca	-0.00622	0.00436	0.19838	0.04178	0.02002	0.07168	0.16948	0.20957	0.02824	0.01982	0.0627	0.13485	0.20437	0.08918	-0.00587	0.12425	0.02122	0.05918	0.12299	0.13026	0.04935	0	
Huila	0.06876	0.01929	0.01437	0.02552	0.05611	0.19403	0.33196	0.11947	-0.01166	-0.00536	-0.02486	0.01581	0.09103	-0.00458	0.04525	0.01151	0.00092	0.05152	0.02277	0.01913	-0.01206	0.06372	

**Anexo H.** (a) Matriz de valores pareados  $F_{ST}$  con base en las frecuencias de los marcadores del cromosoma Y. Para la fuente de los datos en las poblaciones de referencia, ver pág.132. (b) Valores p de significancia de los  $F_{ST}$  calculados.

**a.**



**Fuente:** El autor.

B

	Nasa	Pijao	Embera	Zenu	Ibagué	Ortega	Planadas	Córdoba	Nicaragua	Nativos. Argentina	Argentina	Mexico
Nasa	*											
Pijao	0.84685+- 0.0389	*										
Embera	0.76577+- 0.0413	0.83784+- 0.0274	*									
Zenu	0.00901+- 0.0091	0.00000+- 0.0000	0.05405+- 0.0201	*								
Ibagué	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.13514+- 0.0244	*							
Ortega	0.00000+- 0.0000	0.00000+- 0.0000	0.00901+- 0.0091	0.65766+- 0.0334	0.62162+- 0.0317	*						
Planadas	0.00000+- 0.0000	0.00901+- 0.0091	0.12613+- 0.0278	0.43243+- 0.0446	0.04505+- 0.0152	0.36036+- 0.0359	*					
Córdoba	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.11712+- 0.0237	0.32432+- 0.0433	0.17117+- 0.0316	0.38739+- 0.0385	*				
Nicaragua	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.39640+- 0.0338	0.15315+- 0.0273	0.27027+- 0.0525	0.23423+- 0.0243	0.79279+- 0.0425	*			
Nam. Argentina	0.27928+- 0.0566	0.41441+- 0.0454	0.93694+- 0.0203	0.00901+- 0.0091	0.00000+- 0.0000	0.00000+- 0.0000	0.05405+- 0.0278	0.00000+- 0.0000	0.00000+- 0.0000	*		
Argentina	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	0.09910+- 0.0212	0.24324+- 0.0451	0.10811+- 0.0264	0.04505+- 0.0152	0.46847+- 0.0466	0.01802+- 0.0121	0.00000+- 0.0000	*	
Mexico	0.00000+- 0.0000	0.00000+- 0.0000	0.12613+- 0.0411	0.67568+- 0.0434	0.00000+- 0.0000	0.14414+- 0.0278	0.81081+- 0.0429	0.09009+- 0.0271	0.00000+- 0.0000	0.00000+- 0.0000	0.00000+- 0.0000	*

Fuente: El autor.

**Anexo I.** Frecuencias génicas de los 100 AIMs. MAF: frecuencia alélica del alelo menor.

<b>Cromosoma</b>	<b>SNP</b>	<b>Alelo 1</b>	<b>Alelo 2</b>	<b>MAF</b>
1	rs6684063	G	T	0.4897
1	rs12021830	C	T	0.3704
1	rs2025236	G	A	0.497
1	rs10779334	G	A	0.1899
2	rs2140701	G	A	0.4827
2	rs6711746	C	T	0.4177
2	rs10175357	C	T	0.2135
2	rs305163	G	A	0.3998
2	rs7594727	A	C	0.221
2	rs7598069	A	G	0.4927
2	rs1567803	T	C	0.3257
2	rs1196705	G	A	0.3392
2	rs298259	C	T	0.4719
2	rs1809119	C	A	0.3295
2	rs3927978	G	A	0.3581
2	rs67183	T	G	0.1734
2	rs1606237	T	C	0.3359
2	rs16851773	G	A	0.3774
3	rs6768750	A	G	0.4711
3	rs6803290	C	T	0.3182
3	rs8625	A	T	0.3772
3	rs13097560	A	G	0.3632
3	rs7630522	C	T	0.4525
3	rs7630111	A	C	0.493
3	rs2937673	G	A	0.2981
4	rs7665516	G	A	0.4873
4	rs126397	A	G	0.2736
4	rs920559	G	C	0.4479
4	rs924780	T	G	0.3041
4	rs7664076	G	A	0.4918
5	rs1875017	G	A	0.424
5	rs353373	A	G	0.3494
5	rs349323	T	A	0.3218
5	rs6866970	G	A	0.1753
5	rs17157450	T	C	0.3778
5	rs11960137	G	C	0.3211
5	rs1107154	A	G	0.3916



5	rs11743823	T	C	0.3905
5	rs2042314	T	C	0.4211
6	rs10946113	A	C	0.3207
7	rs12701745	G	A	0.3048
7	rs7808899	G	T	0.4545
7	rs17170516	A	G	0.3247
7	rs2384982	T	C	0.3822
7	rs2707575	G	A	0.3029
8	rs171566	A	G	0.3866
8	rs2736340	T	C	0.4382
8	rs956969	T	C	0.3118
8	rs16880715	T	G	0.1892
8	rs13280988	A	G	0.2987
8	rs9642819	A	G	0.4857
8	rs6994396	T	G	0.371
8	rs2450219	T	C	0.3391
9	rs12380601	A	G	0.2972
9	rs2809393	T	C	0.4472
10	rs10751867	C	T	0.4195
10	rs2795918	T	C	0.4756
10	rs9422913	T	C	0.4085
11	rs874189	A	G	0.3699
11	rs7927234	C	G	0.1865
11	rs1396883	A	G	0.29
11	rs11034734	C	T	0.2497
11	rs7124676	A	G	0.2539
11	rs1790740	G	A	0.3616
11	rs878874	G	A	0.3351
11	rs485645	T	C	0.4757
11	rs7103088	A	G	0.2445
12	rs7961436	G	A	0.376
12	rs7958163	C	A	0.4223
13	rs2585897	A	G	0.4246
13	rs17359176	A	G	0.2495
13	rs9576996	C	G	0.2812
13	rs9533623	A	G	0.2948
13	rs314580	G	A	0.3677
14	rs1958053	A	T	0.4648
14	rs1244379	C	T	0.3816
14	rs1952790	A	G	0.3096
14	rs11625446	A	C	0.3951

14	rs2251244	C	A	0.4989
14	rs10145908	C	T	0.2129
14	rs6573746	C	A	0.3549
15	rs7164838	A	G	0.4259
15	rs8041147	C	T	0.4117
15	rs2304580	C	T	0.374
15	rs1453858	A	T	0.4529
15	rs746655	A	G	0.2107
15	rs4402506	C	T	0.3063
15	rs7173885	G	C	0.2734
16	rs4787040	T	A	0.3757
16	rs9923547	A	C	0.3148
16	rs4347621	G	A	0.245
17	rs7211426	A	G	0.3055
17	rs2052074	G	A	0.2982
20	rs4812381	G	A	0.4451
20	rs6101991	G	A	0.2113
20	rs4812831	A	G	0.2893
21	rs23395	T	C	0.3036
21	rs2836181	T	C	0.1661
21	rs2187239	A	C	0.4631
22	rs132663	A	T	0.4878

Fuente: El autor.

**Anexo J.** Prueba de *Mann-Whitney* para analizar estadísticamente las similitudes o diferencias inter (a) o intrapoblacionales (b) en razón de las proporciones ancestrales nativa americana, europea y africana representadas por los grupos de poblacionales del presente estudio. El programa estadístico empleado fue la versión libre de Infostat, v.2008 (las diferencias significativas están resaltadas en rojo).

a.

Ancestría	Grupo 1	Grupo 2	n (1)	n(2)	Media(1)	Media(2)	DE(1)	DE(2)	W	p(2 colas)
Africana	Indígenas Tolima	MestizoTolima	131	188	0.01	0.03	0.02	0.05	20646.50	0.6962
Europea	Indígenas Tolima	MestizoTolima	131	188	0.07	0.51	0.10	0.13	9109.00	<0.0001
Nativa	Indígenas Tolima	MestizoTolima	131	188	0.92	0.46	0.11	0.13	32836.00	<0.0001
Africana	Córdoba	Zenu	26	39	0.23	0.17	0.15	0.13	973.00	0.1236
Europea	Córdoba	Zenu	26	39	0.42	0.35	0.15	0.15	981.50	0.0982
Nativa	Córdoba	Zenu	26	39	0.35	0.48	0.15	0.19	644.00	0.0042
Africana	Embera	Indígenas Tolima	27	131	0.03	0.01	0.07	0.02	2258.50	0.5399
Europea	Embera	Indígenas Tolima	27	131	0.12	0.07	0.19	0.10	2024.00	0.5661
Nativa	Embera	Indígenas Tolima	27	131	0.84	0.92	0.22	0.11	2282.50	0.5260
Africana	Nasa	Pijao	82	49	0.01	0.02	0.01	0.03	3935.00	0.0008
Europea	Nasa	Pijao	82	49	0.03	0.13	0.06	0.13	4347.50	<0.0001
Nativa	Nasa	Pijao	82	49	0.96	0.85	0.06	0.14	2070.50	<0.0001
Africana	Ibagué	Ortega	112	38	0.02	0.04	0.04	0.07	3189.50	0.1659
Europea	Ibagué	Ortega	112	38	0.55	0.45	0.10	0.14	2014.00	0.0002
Nativa	Ibagué	Ortega	112	38	0.43	0.51	0.11	0.15	3505.50	0.0060
Africana	Ibagué	Planadas	112	38	0.02	0.04	0.04	0.06	3255.00	0.0952
Europea	Ibagué	Planadas	112	38	0.55	0.47	0.10	0.15	2285.50	0.0117
Nativa	Ibagué	Planadas	112	38	0.43	0.49	0.11	0.15	3360.50	0.0337
Africana	Ortega	Planadas	38	38	0.04	0.04	0.07	0.06	1451.00	0.9008
Europea	Ortega	Planadas	38	38	0.45	0.47	0.14	0.15	1369.00	0.3288
Nativa	Ortega	Planadas	38	38	0.51	0.49	0.15	0.15	1525.00	0.5195

Africana	Embera	Nasa	27	82	0.03	0.01	0.07	0.01	1517.50	0.8146
europaea	Embera	Nasa	27	82	0.12	0.03	0.18	0.06	1469.00	0.9104
Nativa	Embera	Nasa	27	82	0.84	0.96	0.22	0.06	1502.00	0.9050
Africana	Embera	Pijao	27	49	0.03	0.02	0.07	0.03	889.50	0.1011
europaea	Embera	Pijao	27	49	0.12	0.13	0.18	0.13	837.00	0.0279
Nativa	Embera	Pijao	27	49	0.84	0.85	0.22	0.14	1211.00	0.0626
Africana	Embera	Zenu	27	39	0.03	0.17	0.07	0.13	514.50	<0.0001
europaea	Embera	Zenu	27	39	0.12	0.35	0.18	0.15	553.50	<0.0001
Nativa	Embera	Zenu	27	39	0.84	0.48	0.22	0.19	1301.50	<0.0001
Africana	Nasa	Zenu	82	39	0.01	0.17	0.01	0.13	3788.50	<0.0001
europaea	Nasa	Zenu	82	39	0.03	0.35	0.06	0.15	3847.50	<0.0001
Nativa	Nasa	Zenu	82	39	0.96	0.48	0.06	0.19	866.50	<0.0001
Africana	Pijao	Zenu	49	39	0.02	0.17	0.03	0.13	2505.00	<0.0001
europaea	Pijao	Zenu	49	39	0.13	0.35	0.13	0.15	2412.00	<0.0001
Nativa	Pijao	Zenu	49	39	0.85	0.48	0.14	0.19	921.00	<0.0001
Africana	Córdoba	Ibagué	26	112	0.23	0.02	0.15	0.04	3139.50	<0.0001
europaea	Córdoba	Ibagué	26	112	0.42	0.55	0.15	0.10	1096.00	0.0001
Nativa	Córdoba	Ibagué	26	112	0.35	0.43	0.15	0.11	1310.00	0.0068
Africana	Córdoba	Ortega	26	38	0.23	0.04	0.15	0.07	1256.00	<0.0001
europaea	Córdoba	Ortega	26	38	0.42	0.45	0.15	0.14	785.00	0.4121
Nativa	Córdoba	Ortega	26	38	0.35	0.51	0.15	0.15	571.00	0.0002
Africana	Córdoba	Planadas	26	38	0.23	0.04	0.15	0.06	1265.00	<0.0001
europaea	Córdoba	Planadas	26	38	0.42	0.47	0.15	0.15	732.00	0.1224
Nativa	Córdoba	Planadas	26	38	0.35	0.49	0.15	0.15	600.00	0.0008
Africana	Córdoba	Nasa	26	81	0.23	0.01	0.15	0.01	2441.00	<0.0001
europaea	Córdoba	Nasa	26	81	0.42	0.03	0.15	0.06	2443.00	<0.0001
Nativa	Córdoba	Nasa	26	81	0.35	0.96	0.15	0.06	353.00	<0.0001
Africana	Córdoba	Pijao	26	49	0.23	0.02	0.15	0.03	1592.00	<0.0001

Europea	Córdoba	Pijao	26	49	0.42	0.13	0.15	0.13	1530.00	<0.0001
Nativa	Córdoba	Pijao	26	49	0.35	0.85	0.15	0.14	367.00	<0.0001
Africana	Córdoba	Embera	26	27	0.23	0.03	0.15	0.07	1001.00	<0.0001
Europea	Córdoba	Embera	26	27	0.42	0.12	0.15	0.18	972.00	<0.0001
Nativa	Córdoba	Embera	26	27	0.35	0.84	0.15	0.22	394.00	<0.0001
Africana	Ibagué	Zenu	112	39	0.02	0.17	0.04	0.13	4593.50	<0.0001
Europea	Ibagué	Zenu	112	39	0.55	0.35	0.10	0.15	1313.50	<0.0001
Nativa	Ibagué	Zenu	112	39	0.43	0.48	0.11	0.19	3289.50	0.1664
Africana	Ortega	Zenu	38	39	0.04	0.17	0.07	0.13	990.50	<0.0001
Europea	Ortega	Zenu	38	39	0.45	0.35	0.14	0.15	1764.00	0.0041
Nativa	Ortega	Zenu	38	39	0.51	0.48	0.15	0.19	1578.00	0.3280
Africana	Planadas	Zenu	38	39	0.04	0.17	0.06	0.13	990.00	<0.0001
Europea	Planadas	Zenu	38	39	0.47	0.35	0.15	0.15	1847.00	0.0002
Nativa	Planadas	Zenu	38	39	0.49	0.48	0.15	0.19	1524.50	0.6650
Africana	Ibagué	Nasa	112	82	0.02	0.01	0.04	0.01	5875.50	<0.0001
Europea	Ibagué	Nasa	112	82	0.55	0.03	0.10	0.06	3403.00	<0.0001
Nativa	Ibagué	Nasa	112	82	0.43	0.96	0.11	0.06	12586.00	<0.0001
Africana	Nasa	Ortega	82	38	0.01	0.04	0.01	0.07	3169.50	<0.0001
Europea	Nasa	Ortega	82	38	0.03	0.45	0.06	0.14	3844.00	<0.0001
Nativa	Nasa	Ortega	82	38	0.96	0.51	0.06	0.15	755.00	<0.0001
Africana	Nasa	Planadas	81	38	0.01	0.04	0.01	0.06	3151.00	<0.0001
Europea	Nasa	Planadas	81	38	0.03	0.47	0.06	0.15	3722.00	<0.0001
Nativa	Nasa	Planadas	81	38	0.96	0.49	0.06	0.15	834.00	<0.0001
Africana	Ibagué	Pijao	112	49	0.02	0.02	0.04	0.03	3671.00	0.2732
Europea	Ibagué	Pijao	112	49	0.55	0.13	0.10	0.13	1322.00	<0.0001
Nativa	Ibagué	Pijao	112	49	0.43	0.85	0.11	0.14	6610.00	<0.0001
Africana	Ortega	Pijao	38	49	0.04	0.02	0.07	0.03	1894.00	0.0573
Europea	Ortega	Pijao	38	49	0.45	0.13	0.14	0.13	2496.00	<0.0001

Nativa	Ortega	Pijao	38	49	0.51	0.85	0.15	0.14	840.00	<0.0001
Africana	Pijao	Planadas	49	38	0.02	0.04	0.03	0.06	1917.50	0.0356
europa	Pijao	Planadas	49	38	0.13	0.47	0.13	0.15	2461.00	<0.0001
Nativa	Pijao	Planadas	49	38	0.85	0.49	0.14	0.15	864.00	<0.0001
Africana	Embera	Ibagué	26	112	0.03	0.02	0.07	0.04	1342.50	0.0113
europa	Embera	Ibagué	26	112	0.12	0.55	0.19	0.10	522.50	<0.0001
Nativa	Embera	Ibagué	26	112	0.84	0.43	0.23	0.11	2995.00	<0.0001
Africana	Embera	Ortega	27	38	0.03	0.04	0.07	0.07	695.00	0.0087
europa	Embera	Ortega	27	38	0.12	0.45	0.18	0.14	482.00	<0.0001
Nativa	Embera	Ortega	27	38	0.84	0.51	0.22	0.15	1269.00	<0.0001
Africana	Embera	Planadas	27	38	0.03	0.04	0.07	0.06	697.50	0.0097
europa	Embera	Planadas	27	38	0.12	0.47	0.18	0.15	488.50	<0.0001
Nativa	Embera	Planadas	27	38	0.84	0.49	0.22	0.15	1263.50	<0.0001

**b.**

Población	valor p (eur vs nam)	valor p (eur vs afr)	valor p (nam vs afr)
Nasa	<0.0001	<0.0001	<0.0001
Pijao	<0.0001	<0.0001	<0.0001
Embera	<0.0001	0.0664	<0.0001
Zenu	0.0023	<0.0001	<0.0001
Ibagué	<0.0001	<0.0001	<0.0001
Ortega	0.1502	<0.0001	<0.0001
Planadas	0.7871	<0.0001	<0.0001
Cordoba	0.0854	<0.0001	0.0022

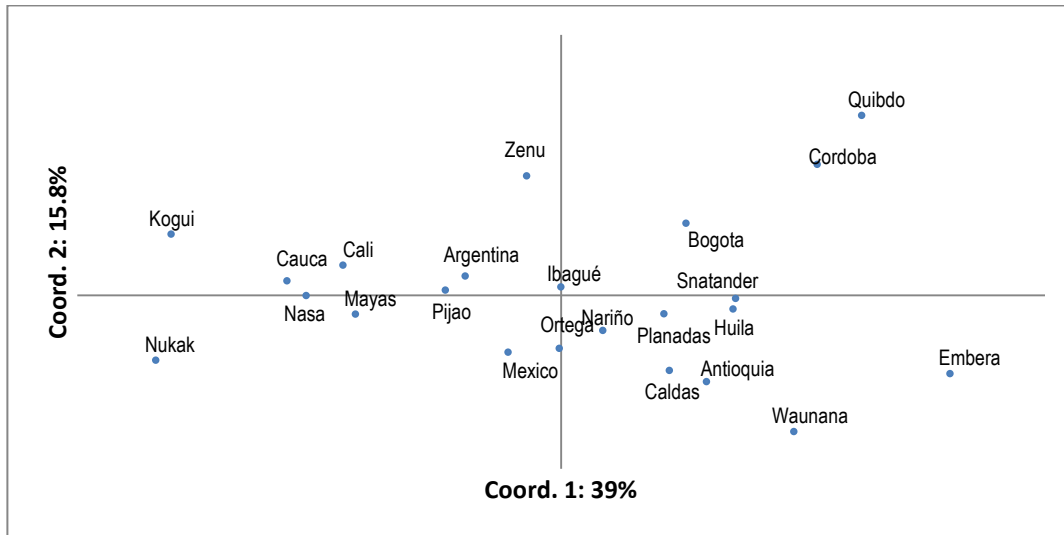
**Fuente:** el autor.

**Anexo K.** Medidas estadísticas de resúmen, calculadas en el programa *Infostat* v.8, a partir de las proporciones ancestrales que se obtuvieron mediante el análisis de los datos genotípicos de 100 Alms en STRUCTURE v.2. Las poblaciones analizadas fueron las indígenas (Nasa, Pijao, Embera, Zenu) y mestizas (Ibague, Ortega, Planadas y Córdoba).

población	Ancestría	n	Media	D.E.	Var(n)	E.E.	CV	Min	Máx
Nasa	Nam	82	0.96	0.06	3.8E-03	0.01	6.46	0.67	1.00
	Eur	82	0.03	0.06	3.3E-03	0.01	168.46	1.0E-03	0.31
	Afr	82	0.01	0.01	1.4E-04	1.3E-03	176.29	6.0E-04	0.07
Pijao	Nam	49	0.85	0.14	0.02	0.02	15.98	0.46	1.00
	Eur	49	0.13	0.13	0.02	0.02	103.12	1.8E-03	0.52
	Afr	49	0.02	0.03	6.6E-04	3.7E-03	147.53	0.00	0.10
Embera	Nam	27	0.84	0.22	0.05	0.04	26.25	0.33	1.00
	Eur	27	0.12	0.18	0.03	0.04	150.57	1.0E-03	0.58
	Afr	27	0.03	0.07	4.2E-03	0.01	199.02	0.00	0.28
Zenu	Nam	39	0.48	0.19	0.04	0.03	39.57	0.01	1.00
	Eur	39	0.35	0.15	0.02	0.02	43.88	4.0E-03	0.73
	Afr	39	0.17	0.13	0.02	0.02	76.22	1.0E-03	0.44
Ibague	Nam	112	0.43	<b>0.11</b>	0.01	0.01	25.25	0.19	0.67
	Eur	112	0.55	<b>0.10</b>	0.01	0.01	19.03	0.33	0.80
	Afr	112	0.02	<b>0.04</b>	2.0E-03	4.2E-03	189.64	1.0E-03	0.24
Ortega	Nam	38	0.51	0.15	0.02	0.02	28.74	0.27	0.85
	Eur	38	0.45	0.14	0.02	0.02	30.49	0.15	0.71
	Afr	38	0.04	0.07	4.7E-03	0.01	177.77	1.0E-03	0.27
Planadas	Nam	38	0.49	0.15	0.02	0.03	31.50	0.26	1.00
	Eur	38	0.47	0.15	0.02	0.02	32.45	2.0E-03	0.69
	Afr	38	0.04	0.06	3.4E-03	0.01	150.03	6.0E-04	0.20
Córdoba	Nam	26	0.35	0.15	0.02	0.03	41.95	0.10	0.71
	Eur	26	0.42	0.15	0.02	0.03	35.27	0.16	0.68
	Afr	26	0.23	0.15	0.02	0.03	65.58	0.02	0.66

**Fuente:** el autor.

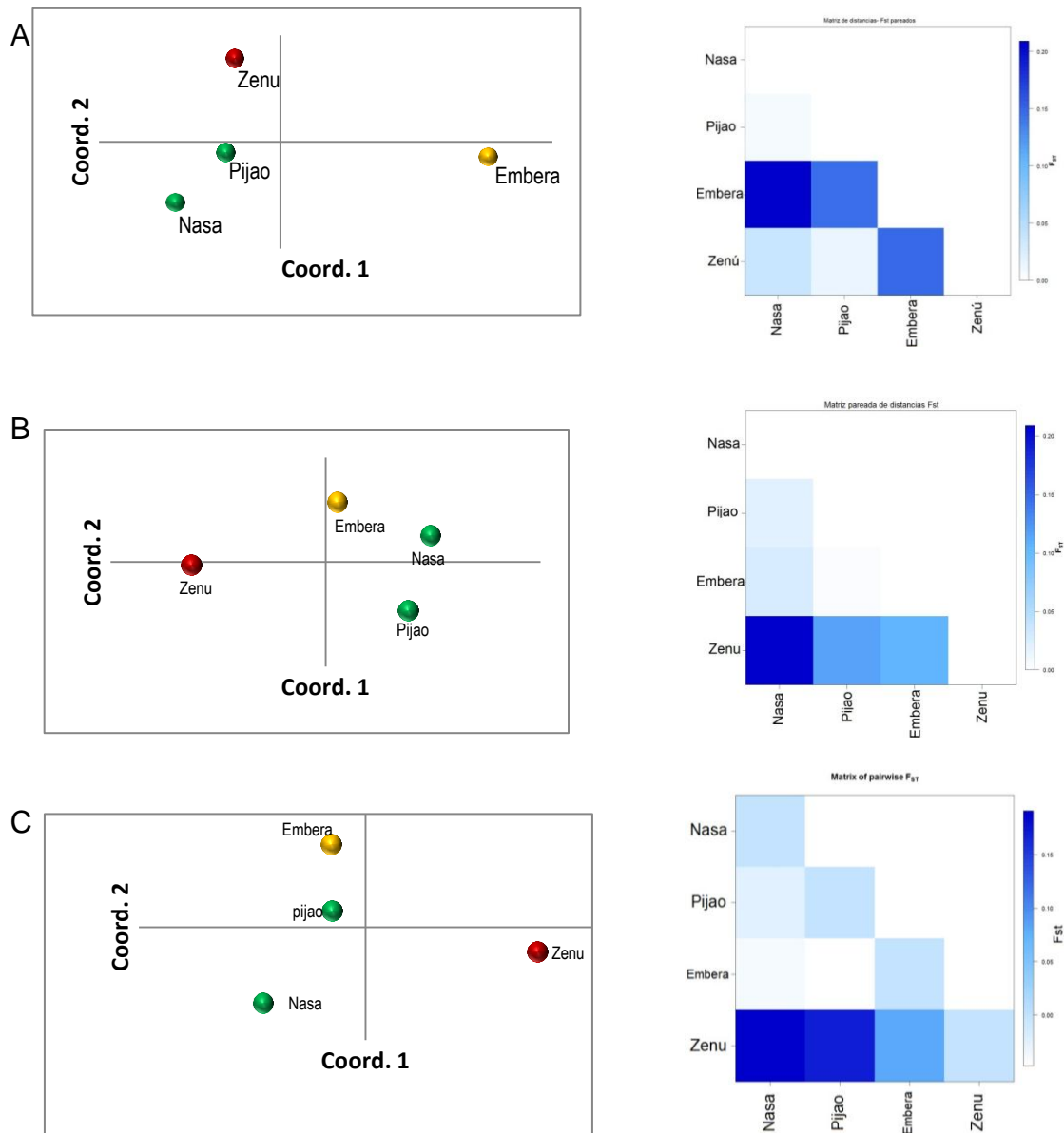
**Anexo L.** Análisis MDS con base en la matriz pareada de distancias Fst obtenido de las frecuencias haplogrupales del mDNA en las poblaciones objeto de estudio, entre otras poblaciones de Colombia y Latino América. Solo se representa la primera y tercera coordenada. Para la fuente de los datos en las poblaciones de referencia, ver pág.124.



**Fuente:** el autor.



**Anexo M.** Análisis multidimensional con base en la matriz de distancias  $F_{st}$  pareadas en el (a) mDNA, (b) cromosoma Y y (c) SNPs autosómicos solo en las poblaciones indígenas del estudio. Fuente: El autor.



Fuente: el autor.

**Anexo N.** Tabla basada en un diseño de cuadro *Punnet* para el cálculo de las frecuencias de ancestría autosómica esperada, con base en el conocimiento previo de las proporciones ancestrales en el mDNA y el cromosoma Y, en una población mezclada a partir de tres orígenes ancestrales y sesgo en la mezcla y sin aporte materno africano. (a) Las proporciones ancestrales uniparentales asumidas, (b) la matriz de *Punnet* con los aporte de cada sistema genético en cada cruce con su respectiva probabilidad, y (c) suma total de las probabilidades de los aportes de cada cruce a las proporciones ancestrales con los cuales de hace la grafica. Fuente: el autor.

**A**

madre3 (Afr)	XXAAm	0
madre2 (nam)	XXAAm	1
padre1 (Eur)	XY <sup>R</sup> AA	0.7
padre2 (Nam)	XY <sup>Q</sup> AA	0.2
padre 3 (Afr)	XY <sup>E</sup> AA	0.1

**B**

	NRY	0	0.2	0	0.7	0	0.1
	AIMs	0.10	0.10	0.35	0.35	0.05	0.05
		X A	Y <sup>Q</sup> A	X A	Y <sup>R</sup> A	XA	Y <sup>E</sup> A
1	XAm	XX-AA-m	XY <sup>Q</sup> AA-m	XX-AA-m	XY <sup>R</sup> AAm	XXAAm	XY <sup>E</sup> AAm
0	XAm	XX-AA-m	XY <sup>Q</sup> AA-m	XX-AA-m	XY <sup>R</sup> AAm	XX-AA-m	XY <sup>E</sup> AAm
aporte materno	mDNA	0.10 0.00	0.10 0.00	0.35 0.00	0.35 0.00	0.05 0.00	0.05 0.00
Aporte autosómico	AIMs	0.10 0.00	0.10 0.00	0.18 0.18 0.00	0.18 0.18 0.00	0.03 0.03 0.00	0.03 0.03 0.00
aporte paterno	NRY		0.20 0.00		0.70 0.00		0.10 0.00

**C**

	mDNA	NRY	AIMs
nam	1.00	0.20	0.60
eur	0.00	0.70	0.35
Afr	0.00	0.10	0.05

Fuente: el autor.

**Anexo O.** Tabla basada en un diseño de cuadro *Punnet* para el cálculo de las frecuencias de ancestría autosómica esperada, con base en el conocimiento previo de las proporciones ancestrales en el mDNA y el cromosoma Y, en una población de tres orígenes ancestrales con sesgo relativo en la mezcla y todos los aportes posibles. (a) Las proporciones ancestrales uniparentales asumidas, (b) la matriz de *Punnet* con los aporte de cada sistema genético en cada cruce con su respectiva probabilidad, y (c) suma total de las probabilidades de los aportes de cada cruce a las proporciones ancestrales con los cuales de hace la gráfica.

**A**

madre2 (Afr)	XX-AA-m	0.15
madre1 (nam)	XX-AA-m	0.6
padre3(Eur)	XY <sup>R</sup> AA	0.66
padre1 (Nam)	XY <sup>Q</sup> AA	0.1
padre2 (Afr)	XY <sup>E</sup> AA	0.24
madre3 (eur)	XXAAm	0.25

**B**

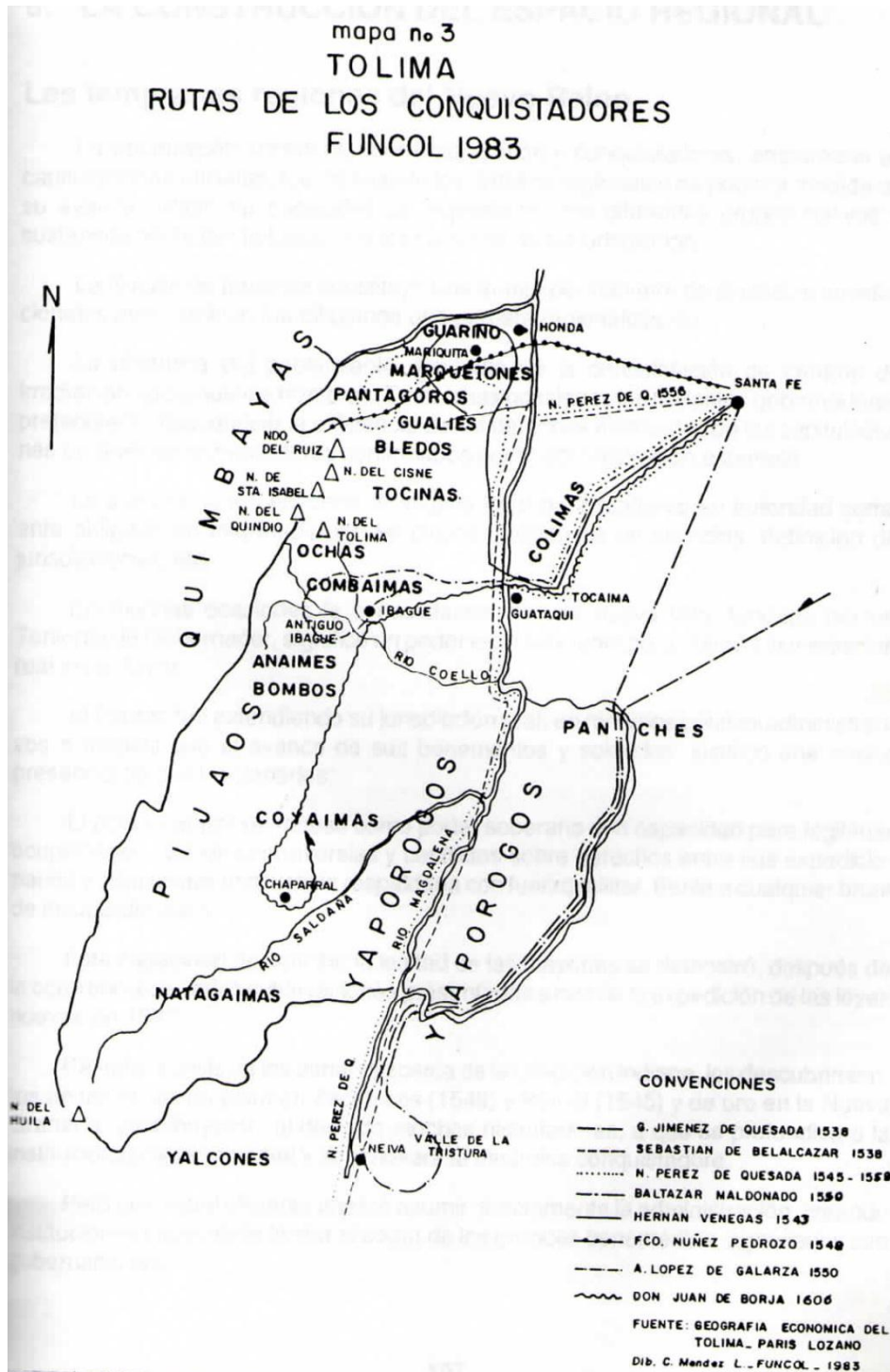
NRY	0	0.1	0	0.66	0	0.24	
AIMs	0.05	0.05	0.33	0.33	0.12	0.12	
	X A	Y <sup>Q</sup> A	X A	Y <sup>R</sup> A	XA	Y <sup>E</sup> A	
0.6	XAm	XXAAm	XY <sup>Q</sup> AAm	X XAAm	XY <sup>R</sup> AAm	XXAAm	XY <sup>E</sup> AAm
0.15	XAm	XXAAm	XY <sup>Q</sup> AAm	X XAAm	XY <sup>R</sup> AAm	XXAAm	XY <sup>E</sup> AAm
0.25	XAm	XXAAm	XY <sup>Q</sup> AAm	X XAAm	XY <sup>R</sup> AAm	XXAAm	XY <sup>E</sup> AAm
aporte materno	mDNA	0.0 0.0 0.0	0.0 0.0 0.0	0.2 0.0 0.0	0.2 0.0 0.0	0.0 0.0 0.0	0.0 0.0 0.0
	NA	3 1 1	3 1 1	0 5 8	0 5 8	7 2 3	7 2 3
aporte autosómico	AIMs	0.0 0.0 0.0	0.0 0.0 0.0	0.1 0.2 0.0	0.1 0.2 0.0	0.0 0.0 0.0	0.0 0.0 0.0
		4 0 1	4 0 1	0 1 2	0 1 2	4 7 2	4 7 2
aporte paterno	NRY		0.0 0.0 0.0		0.4 0.1 0.1		0.1 0.0 0.0
			6 2 3		0 0 7		4 4 6

**C**

	mDNA	NRY	AIMs
nam	0.60	0.10	0.35
eur	0.25	0.66	0.46
Afr	0.15	0.24	0.20

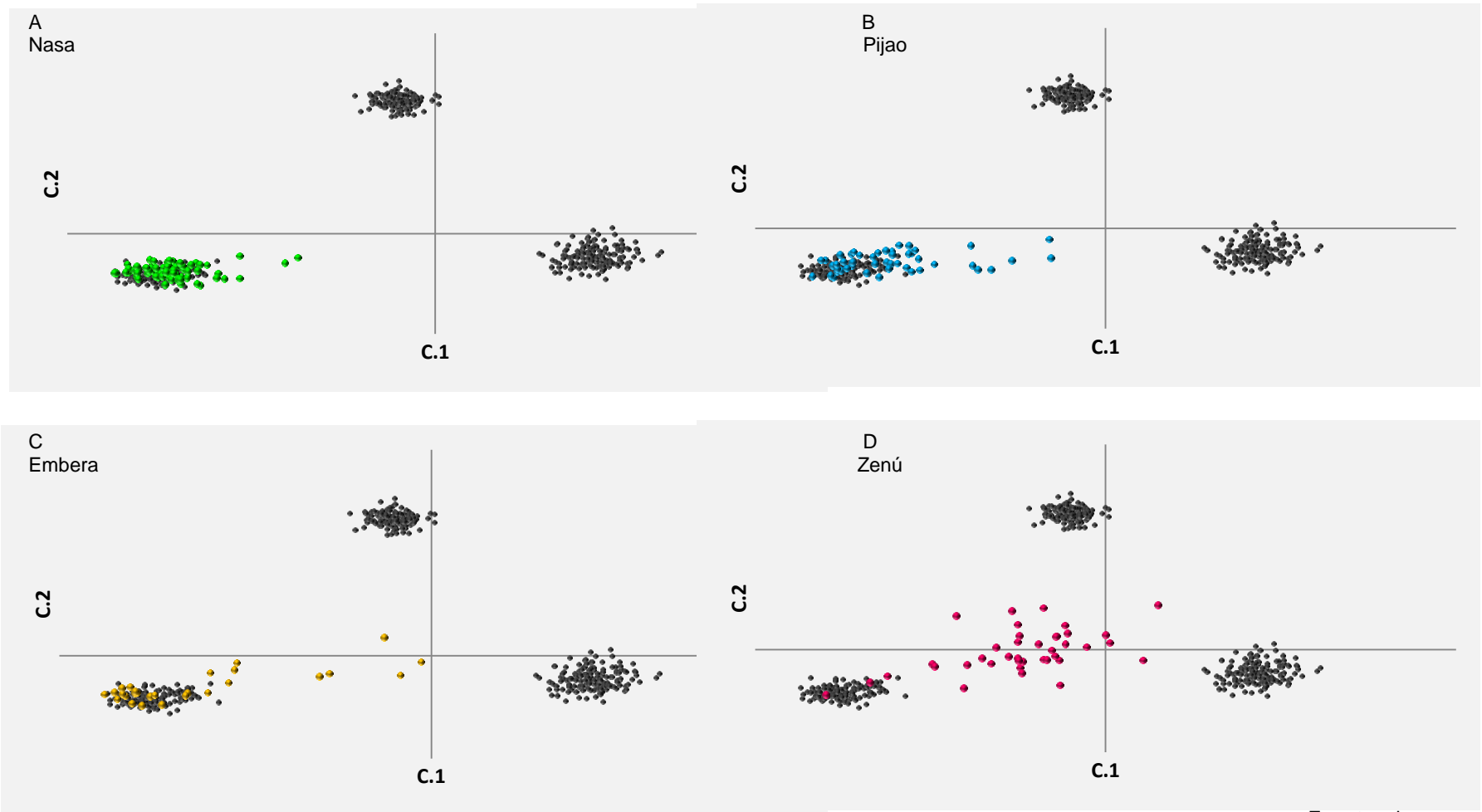
Fuente: el autor.

Anexo P. Ruta de los conquistadores en la región del Tolima hacia 1550.

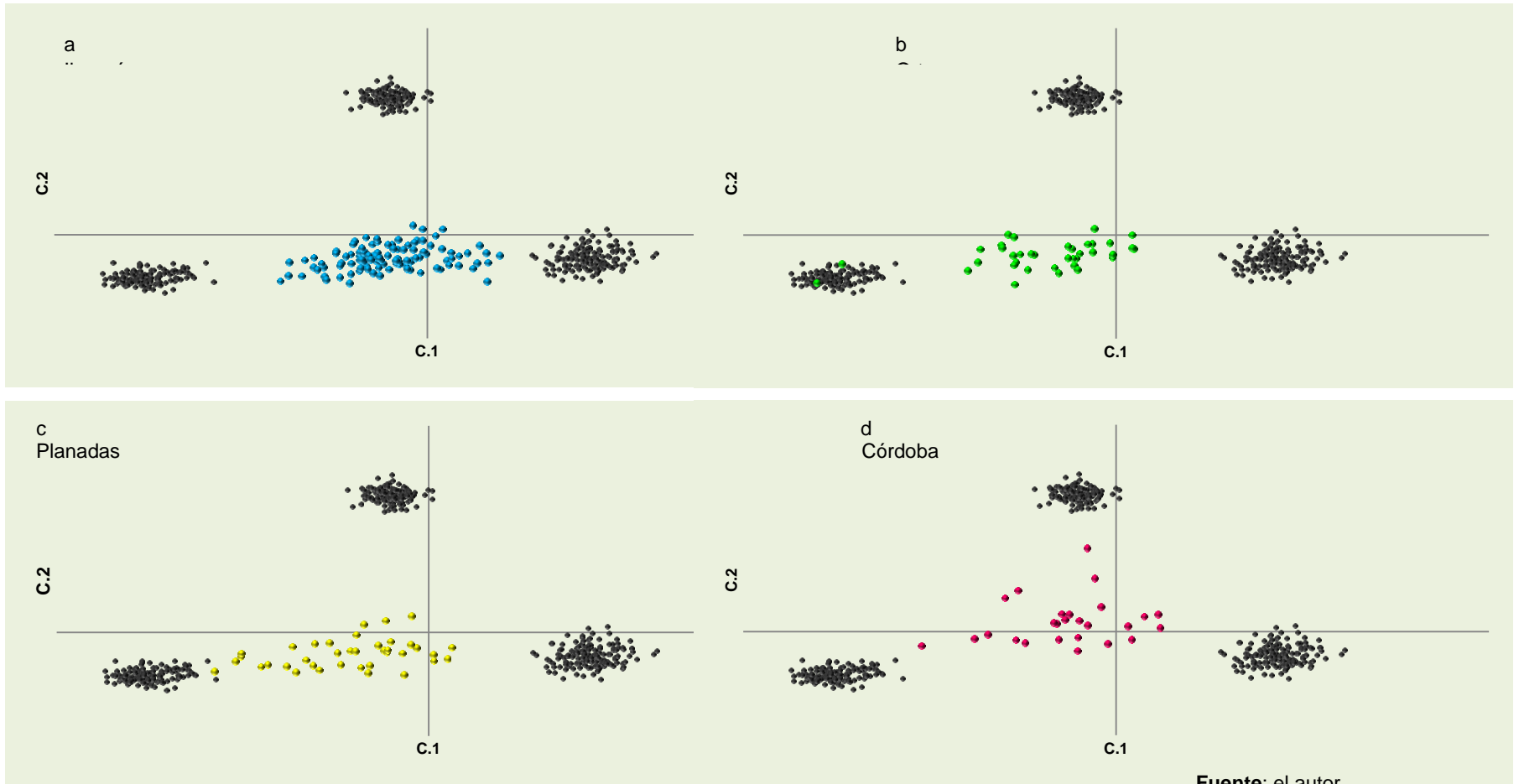


Fuente: Triana, 1992.

**Anexo Q1.** Patrón de agrupamiento de los indígenas Nasa (A: verde), Pijao (B: azul), Embera (C: amarillo) y Zenu (D: fucsia), con relación a las poblaciones de referencia (en gris): africanos (superior), europeos (izquierda) y nativos de LA (derecha). La grafica se obtuvo mediante un análisis MDS a partir de los genotipos de cada individuo.



**Anexo R.** Patrón de agrupamiento de los mestizos de Ibagué (A), Ortega (B), Planadas (C) y Córdoba (D), con relación a las poblaciones de referencia (en gris): africanos (superior), europeos (izquierda) y nativos de LA (derecha). La gráfica se obtuvo mediante un análisis MDS a partir de los genotipos de cada individuo.



Fuente: el autor.

