

**PONTIFICIA UNIVERSIDAD JAVERIANA  
FACULTAD DE CIENCIAS  
CARRERA DE BIOLOGÍA**



**ESTUDIO GENÉTICO POBLACIONAL DE DOS POBLACIONES AFROCOLOMBIANAS  
DE CHOCÓ Y CAUCA Y UNA POBLACION “BLANCA” DE CÓRDOBA MEDIANTE EL  
USO DE RETROTRANSPONES DE LA FAMILIA HUMANA TA (LINE-1)**

**MYRIAM JOHANNA LIZARRALDE OLAVE**

**TRABAJO DE GRADO  
Presentado como requisito parcial  
para optar al título de  
BIOLOGA**

**DIRECTOR  
MANUEL RUIZ-GARCÍA Ph.D.  
Laboratorio Biología Evolutiva**

**Bogotá D.C. Colombia  
2002**

## **NOTA DE ADVERTENCIA**

Artículo 23 de la Resolución N° 13 de Julio de 1946

“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Solo velará por que no se publique nada contrario al dogma y a la moral católica y por que las tesis no contengan ataques personales contra persona alguna, antes bien se vea en ellas el anhelo de buscar la verdad y la justicia”.

**ESTUDIO GENÉTICO POBLACIONAL DE DOS POBLACIONES AFROCOLOMBIANAS  
DE CHOCÓ Y CAUCA Y UNA POBLACION “BLANCA” DE CÓRDOBA MEDIANTE EL  
USO DE RETROTRANSPONES DE LA FAMILIA HUMANA TA (LINE-1)**

**MIRYAM JOHANNA LIZARRALDE OLAVE**

**APROBADO**

---

**MANUEL RUÍZ-GARCÍA Ph. D.**  
Director

---

**JAIME BERNAL**  
Jurado

---

**MARIA IGNACIA CASTILLO**  
Jurado

**ESTUDIO GENÉTICO POBLACIONAL DE DOS POBLACIONES AFROCOLOMBIANAS  
DE CHOCÓ Y CAUCA Y UNA POBLACION “BLANCA” DE CÓRDOBA MEDIANTE EL  
USO DE RETROTRANSPOSONES DE LA FAMILIA HUMANA TA (LINE-1)**

**MIRYAM JOHANNA LIZARRALDE OLAVE**

---

**M. Phil ANGELA UMAÑA M.**  
Decana Académica

---

**LUZ MERCEDES SANTAMARÍA R.**  
Directora de Carrera

## FORMATO DESCRIPCIÓN TRABAJO DE GRADO

AUTOR O AUTORES

Apellidos	Nombres
LIZARRALDE OLAVE	MYRIAM JOHANNA

DIRECTOR (ES)

Apellidos	Nombres
RUIZ GARCIA	MANUEL

ASESOR (ES) O CODIRECTOR

Apellidos	Nombres

TRABAJO PARA OPTAR POR EL TÍTULO DE: BIOLOGA

TÍTULO COMPLETO DEL TRABAJO: ESTUDIO GENÉTICO POBLACIONAL DE DOS POBLACIONES AFROCOLOMBIANAS DE CHOCÓ Y CAUCA Y UNA POBLACION “BLANCA” DE CÓRDOBA MEDIANTE EL USO DE RETROTRANSPONES DE LA FAMILIA HUMANA TA (LINE-1)

SUBTÍTULO DEL TRABAJO: ESTUDIO GENÉTICO POBLACIONAL DE TRES POBLACIONES HUMANAS COLOMBIANAS MEDIANTE EL USO DE RETROTRANSPONES TA (LINE-1)

FACULTAD: CIENCIAS

PROGRAMA: Carrera  Especialización  Maestría  Doctorado

NOMBRE DEL PROGRAMA: BIOLOGÍA

CIUDAD: BOGOTA AÑO DE PRESENTACIÓN DEL TRABAJO: 2002

NÚMERO DE PÁGINAS: 82

TIPO DE ILUSTRACIONES:

- |  |                                      |
|--|--------------------------------------|
| <input type="checkbox"/> Ilustraciones                           | <input type="checkbox"/> Planos      |
| <input type="checkbox"/> Mapas                                   | <input type="checkbox"/> Láminas     |
| <input type="checkbox"/> Retratos                                | <input type="checkbox"/> Fotografías |
| <input checked="" type="checkbox"/> Tablas, gráficos y diagramas |                                      |

MATERIAL ANEXO (Vídeo, audio, multimedia o producción electrónica):

Duración del audiovisual: \_\_\_\_\_ Minutos.

Número de casetes de vídeo: \_\_\_\_\_ Formato: VHS \_\_\_\_ Beta Max \_\_\_\_  $\frac{3}{4}$  \_\_\_\_ Beta

Cam \_\_\_\_ Mini DV \_\_\_\_ DV Cam \_\_\_\_ DVC Pro \_\_\_\_ Vídeo 8 \_\_\_\_ Hi 8 \_\_\_\_

Otro. Cual? \_\_\_\_\_

Sistema: Americano NTSC \_\_\_\_\_ Europeo PAL \_\_\_\_\_ SECAM \_\_\_\_\_

Número de casetes de audio: \_\_\_\_\_

Número de archivos dentro del CD (En caso de incluirse un CD-ROM diferente al trabajo de grado: \_\_\_\_\_)

DESCRIPTORES O PALABRAS CLAVES.

Genética de poblaciones, afrocolombianos, LINE-1, historia, mestizaje, migración, retrotransposones Ta.

RESUMEN DEL CONTENIDO: Se realizo un estudio genético poblacional, mediante el uso de siete retrotransposones de la familia TA del tipo LINE-1 encontrados en humanos, en dos poblaciones afrocolombianas y una blanca mestiza, con el fin de medir sus diferencias y similitudes, así como los eventos genéticos que pudieran estar operando en ellas y dar una explicación de los resultados mediante datos históricos de Colombia. Se determinaron los estados de los loci bialélicos, ausencia o presencia de la inserción, en los 30 individuos que conformaban cada población. Contrario a lo que se esperaba la baja heterogeneidad encontrada mostró las estrechas relaciones entre las tres poblaciones pese a su fenotipo, a su distancia geográfica, y al aporte en su acervo genético proveniente de los diferentes grupos indígenas, africanos y españoles que las originaron; sugiriendo que predomina el aporte africano y que este provendría de pocas poblaciones de África como lo habían sugerido otros estudios que usaron marcadores moleculares diferentes. Entre las poblaciones se encontró además, un alto flujo génico, acorde con los procesos migratorios en Colombia y la posible intervención de un tipo de selección natural que esta unificando las frecuencias alélicas para dos marcadores. Finalmente se corrobora la utilidad de estos retrotransposones para estudios poblacionales, y al encontrar justificación histórica a los datos encontrados, se sugirió el uso de la genética de poblaciones como herramienta que puede suplir en parte los vacíos informativos de la historia de las poblaciones colombianas debidas a carencia de información y registros.

**Dedico este trabajo a todos**

**aquellos que creen que los sueños son realizables...**

**A mi familia, mis cómplices que siempre me han apoyado en todo...**

**A mi hijo y a Javier como un paso hacia todas las cosas grandes que vienen**

## **AGRADECIMIENTOS**

**A Manuel Ruíz García PhD. Laboratorio de Biología Evolutiva (genética de Poblaciones) del departamento de Biología de la Pontificia Universidad Javeriana por el seguimiento, asesoría y dirección del presente trabajo.**

**A Diana Álvarez por su asesoría en el manejo de las técnicas de laboratorio y estandarización de los procedimientos.**

**A Stéphane Boissinot por el envío de los marcadores utilizados.**

**A la Universidad de Antioquia por las muestras enviadas de Afrocolombianos.**

**Al grupo del laboratorio de Biología Evolutiva y Genética de Poblaciones del departamento de Biología de la Pontificia Universidad Javeriana por la facilitación de los equipos utilizados y por las muestras de la población de Córdoba.**

## TABLA DE CONTENIDOS

<b>1. INTRODUCCIÓN</b>	<b>15</b>
<b>2. MARCO TEÓRICO Y REVISIÓN DE LITERATURA</b>	<b>18</b>
<b>2.1 ASPECTOS GENÉTICOS:</b>	<b>18</b>
2.1.1 ELEMENTOS MÓVILES EN LOS GENOMAS:	18
2.1.2 RETROTRANSPOSONES DE LA FAMILIA HUMANA TA:	20
2.1.3 ESTUDIOS GENÉTICOS USANDO RETROTRANSPOSONES COMO MARCADORES:	21
2.1.4 ESTUDIOS GENÉTICOS CON POBLACIONES AMERICANAS UTILIZANDO MARCADORES DIFERENTES A LOS RETROTRANSPOSONES:	24
<b>2.2 ASPECTOS HISTÓRICOS:</b>	<b>27</b>
2.2.1 PROCEDENCIA DE LOS NEGROS ESCLAVOS:	27
2.2.2 CONFORMACIÓN HISTÓRICA DE LAS POBLACIONES DE ESTUDIO:	31
2.2.2.1 Población del Chocó:	31
2.2.2.2 Población de Cauca:	32
2.2.2.3 Población de Córdoba	32
2.2.3 MIGRACIÓN Y MESTIZAJE:	33
<b>3. FORMULACIÓN DEL PROBLEMA Y JUSTIFICACIÓN:</b>	<b>36</b>
<b>3.1 FORMULACIÓN DEL PROBLEMA:</b>	<b>36</b>
<b>3.2 PREGUNTAS DE INVESTIGACIÓN:</b>	<b>36</b>
<b>3.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN:</b>	<b>36</b>

<b>4. OBJETIVOS:</b>	<b>38</b>
4.1 OBJETIVO GENERAL:	38
4.2 OBJETIVOS ESPECÍFICOS:	38
<b><u>5. HIPOTESIS</u></b>	<b><u>39</u></b>
<b><u>6. MATERIALES Y METODOS</u></b>	<b><u>40</u></b>
6.1 DISEÑO DE LA INVESTIGACIÓN:	40
6.1.1 POBLACIÓN DE ESTUDIO Y MUESTRA:	40
6.1.2 VARIABLES DEL ESTUDIO:	41
6.2 MÉTODOS:	42
6.3 RECOLECCIÓN DE INFORMACIÓN:	42
6.4 ANÁLISIS DE LA INFORMACIÓN:	43
<b><u>7. RESULTADOS</u></b>	<b><u>51</u></b>
<b><u>8. DISCUSIÓN DE RESULTADOS</u></b>	<b><u>72</u></b>
<b><u>9. CONCLUSIONES</u></b>	<b><u>72</u></b>
<b><u>10. RECOMENDACIONES</u></b>	<b><u>74</u></b>
<b><u>11. REFERENCIAS</u></b>	<b><u>76</u></b>
<b><u>12. ANEXOS</u></b>	<b><u>81</u></b>

## INDICE DE TABLAS

[Tabla 1. Frecuencias alélicas para cada marcador en las tres poblaciones.](#)

[Tabla 2. Resultados test exacto equilibrio Hardy-Weinberg por población.](#)

[Tabla 3. Heterogeneidad Global.](#)

[Tabla 4. Diversidad Génica por locus y población.](#)

[Tabla 5. Estimaciones de heterocigocidad de Nei.](#)

[Tabla 6. Cálculo de los estadísticos F.](#)

[Tabla 7. Prueba de significancia de los estadísticos F.](#)

[Tabla 8. Distancias genéticas entre las tres poblaciones.](#)

[Tabla 9. Análisis de coordenadas principales.](#)

## INDICE DE FIGURAS

[Figura 1. Representación geométrica de las distancias Chord de Cavalli-Sforza & Edwards para dos alelos.](#)

[Figura 2. Árboles de distancias genéticas.](#)

[Figura 3. Análisis de coordenadas principales para las distancias de Nei \(1972\).](#)

[Figura 4. Análisis de coordenadas principales para las distancias de Cavalli-Sforza & Edwards \(1967\) en dos dimensiones.](#)

## RESUMEN

Se realizó un estudio genético poblacional en tres poblaciones Colombianas mediante el uso de retrotransposones de la familia TA del tipo LINE-1 encontrados en humanos, dos poblaciones eran afrocolombianas y una correspondía a “blancos mestizos”, con el fin de medir las diferencias genéticas entre ellas, así como los eventos genéticos que operan en esas poblaciones y dar una explicación de los resultados mediante datos históricos. Contrario a lo que se esperaba las tres poblaciones estuvieron estrechamente relacionadas pese a su distancia geográfica, y al aporte diferencial en su acervo genético proveniente de grupos indígenas, africanos y españoles que les dieron origen. Este resultado soportó lo encontrado en otros trabajos para los grupos afrocolombianos, con el uso de otros tipos de marcadores moleculares, con los que se sugirió que al aporte africano a dichas poblaciones no provendría de grupos muy diferenciados genéticamente entre sí. Se encontró además, que dentro de las poblaciones estudiadas se está presentando un alto flujo génico, acorde con los procesos migratorios ocurridos en Colombia y posiblemente, de forma concomitante, esté operando algún tipo de selección unificante, que esta llevando a las poblaciones a un estado frecuencial similar. Dos marcadores presentaron el estado ancestral (ausencia de la inserción retrotransposónica) en las tres poblaciones sugiriendo un mayoritario aporte africano en su seno debido a la antigüedad de este acervo genético. Finalmente se corroboró la utilidad de estos retrotransposones para estudios poblacionales, y al encontrar justificación histórica a los datos encontrados, se sugirió el uso de la genética de poblaciones como herramienta que puede suplir en parte los vacíos informativos de la historia de las poblaciones colombianas especialmente las afrocolombianas debidas a carencia de información y registros.

## 1. INTRODUCCIÓN

Varios trabajos previos apuntan a la utilidad de los retrotransposones como marcadores moleculares, para estudios de relaciones poblacionales. Por ser estos polimórficos, por que su probabilidad de generar homoplásias es mínima y porque su estado ancestral es conocido autores como Novick et al. (1998), Batzer et al. (1994), Furano y Usdin (1995) Boissinot et al. (2000) han sugerido su extraordinaria utilidad para la reconstrucción filogenética de las poblaciones humanas.

Los retrotransposones son elementos móviles insertos dentro de los genomas producto de virus contraídos en algún momento de la historia evolutiva de las especies, donde algunos aún conservan la maquinaria para replicarse e insertarse en nuevos loci, (Eline et al. 2000). El nombre de retrotransposón se debe, a que se unen al genoma siguiendo un proceso igual al de los retrovirus, por medio de una transcriptasa reversa. Es decir, que a partir de un RNA se traduce un DNA que se inserta al nuevo locus (Furano 2000).

En humanos los retrotransposones de tipo LINE-1 corresponden al 14% de su genoma y la mayoría de los que son activos pertenecen a una gran familia humana de retroelementos llamada TA. En ella se presentan las inserciones más polimórficas Boissinot (2000).

Se han realizado estudios genético poblacionales en humanos utilizando retrotransposones pero del tipo Alu (SINEs). (Batzer et al. 1994 y Novick et al. 1998). Los trabajos mas importantes usando retrotransposones de tipo LINE-1 se han realizado en ratones, (Novick et al. 1998, Furano y Usdin 1995, Usdin et al. 1995 y Verneau et al. 1997). En todos ellos se corroboró la utilidad de los retrotransposones para estudios de relaciones poblacionales o taxonómicas. El presente trabajo es el primero en usar retrotransposones LINE-1 de la familia humana TA en poblaciones latinoamericanas, específicamente colombianas.

Por otro lado, Jaramillo et al. (2001) utilizando otro tipo de marcadores, analizaron varias poblaciones colombianas amerindias y afrocolombianas y encontraron que entre las poblaciones amerindias la heterogeneidad genética era muy elevada, pero no encontraron heterogeneidad significativa entre las poblaciones afrocolombianas, atribuyéndolo a que el aporte africano al acervo genético de éstas provendría más bien de pocas poblaciones del occidente de África. Algunos historiadores y antropólogos colombianos han sugerido lo mismo, (Mina 1975 y Maya 1998).

Adicionalmente, las poblaciones afrocolombianas han sufrido un intenso proceso migratorio a lo largo de la costa pacífica y atlántica en respuesta al trabajo minero durante la colonia española. Actualmente se encuentran tres puntos donde converge aproximadamente la mitad de la población total de afrocolombianos. Estos son Quibdó, Tumaco y Buenaventura. (Pardo 1996).

Por todo lo anterior, se realizó un estudio genético poblacional usando siete marcadores pertenecientes a la familia humana TA inserciones LINE-1 non-LTR (Boissinot et al. 2000), en dos poblaciones afrocolombianas de Quibdó (Chocó) y Cauca y se compararon con una población “blanca” de Córdoba, con el fin de establecer sus relaciones y diferencias, así como también determinar los eventos genético poblacionales que pudieran estar ocurriendo en su interior, y comparar todo ello con los datos históricos de que se dispone. Se determinó como la procedencia y la migración afrocolombiana pudieran estar afectando a los marcadores estudiados.

Como hechos más trascendentes, se encontraron dos marcadores Tau-9 y Tau-10 que presentaron el estado ancestral de la especie humana, posiblemente por el aporte africano, que es el acervo genético más antiguo de nuestra especie, que tenían las tres poblaciones. Dos marcadores, Tau-1 y Tau1-22, no se encontraron en equilibrio Hardy-Weinberg por un exceso de homocigotos, posiblemente por algún tipo de selección unificante que opera en esas poblaciones. Adicionalmente, se encontró que las tres poblaciones están estrechamente relacionadas pese al fenotipo “blanco” de Córdoba debido al proceso de mestizaje y a un alto flujo génico entre ellas, acorde con los procesos migratorios que han tenido durante su historia.

Finalmente, se corroboró lo encontrado con otros trabajos (Jaramillo et al. 2001) que sugieren que la parte del acervo genético africano de las poblaciones afrocolombianas provendría de más bien pocas poblaciones de África.

## 2. MARCO TEÓRICO Y REVISIÓN DE LITERATURA

### 2.1 Aspectos Genéticos:

#### 2.1.1 Elementos móviles en los genomas:

Dentro de los genomas eucariotes, existen secuencias repetitivas de ADN capaces de translocarse de un lugar a otro (elementos transponibles). Algunos de ellos lo hacen por si solos, ya que poseen la maquinaria completa para hacerlo o, por el contrario, hay otros que han perdido su capacidad translocativa y requieren ser transferidos con la mediación de genes cercanos. (Klug & Cummings 1993).

De acuerdo con, Eline et al. (2000), los elementos transponibles actualmente son clasificados en dos grandes grupos: los transposones DNA y los retrotransposones. Los primeros se caracterizan porque ellos mismos se cortan y se pegan en otro locus dentro del genoma mientras que, los segundos se mueven homológamente como lo hacen los retrovirus. Es decir, que el retrotransposón original se mantiene *"in situ"* haciendo una copia de si mismo en un RNA intermedio, el cual es transcrito a un nuevo DNA con la mediación de una transcriptasa reversa y finalmente la nueva secuencia de ADN se adhiere a otro locus diferente.

El grupo de retrotransposones, a su vez, es subdividido en otros dos grandes subgrupos, los LTR y los non-LTR, determinados por la presencia o ausencia de largas terminaciones repetitivas, respectivamente.

Los elementos transponibles pueden ser autónomos o no. Esto quiere decir que algunos tienen toda la maquinaria para moverse como los del tipo LINE-1 o son definitivamente dependientes de otros elementos transponibles para su movilidad, como los del tipo SINEs.

En el genoma humano los elementos móviles han sido útiles en tanto que han desarrollado defensas al hospedero. Están involucrados en la remodelación estructural de los cromosomas, (telomerasas e inactivación del cromosoma x), incrementan la plasticidad del genoma y facilitan la expresión de algunos genes. No obstante, algunas inserciones son responsables de enfermedades genéticas como carcinomas, hemofilia tipo A, distrofia muscular entre otras, pero la mayoría de las veces estos elementos se insertan en intrones por lo que en general su impacto como causales de enfermedad es mínima.

Las secuencias de tipo LINEs, pertenecen al grupo de retrotransposones autónomos y carecen de LTR (Long Terminal Repeated) por lo que se ubican en el subgrupo non-LTR. Ellas se encuentran en la mayoría de genomas eucariotes; en el caso de los mamíferos, se han estado replicando desde su radiación hace más de 100 millones de años llegando a conformar actualmente el 20% o más de la masa genómica de estos. (Furano 2000).

En el caso de los humanos, las secuencias LINE-1 o L1 son consideradas las más activas. Se han encontrado aproximadamente 500.000 con la peculiaridad de haber perdido su secuencia final 5' durante el proceso de inserción; tres mil de ellas mantienen su longitud completa y de un 40-60 aún están activas perteneciendo, estas últimas, en su mayoría, a la familia de inserciones humanas TA. (Eline et al. 2000).

Un retrotransposón de longitud completa, consta de cinco partes o regiones. La primera es la 5'- UTR que no es traducible y cuya función es de promotor; consecutivamente se encuentra la región ORF I (fragmento de apertura de lectura) y codifica la proteína ORF I<sub>p</sub> cuya función es enlazar ácidos nucleicos; la tercera es la ORF II (segundo fragmento de apertura) que codifica para la proteína ORFII<sub>p</sub> que presenta actividad endonucleasa y transcriptasa reversa; la cuarta es la región 3'- UTR que contiene una gran cadena de Guaninas seguidas de polipurinas y finalmente se encuentra una larga cola de poli adeninas. (Furano 2000).

Eline et al. (2000) reportaron, de acuerdo con otros estudios, que la endonucleasa, del retrotransposón, manifiesta una preferencia por las regiones del ADN ricas en uniones de A+T (Adenina-Timina). Tanto así, que en dichas regiones la secuencia de la endonucleasa está sobre representada así como también lo está dentro de otras inserciones L1, que a su vez son ricas en uniones A+T. Por esta preferencia se atribuye que al momento de codificar las L1 presentan abundantes uniones G+C. Las secuencias L1 non-LTR además, se han convertido durante su evolución en exitosos parásitos genómicos debido a que, en su mayoría, las inserciones se hacen dentro de regiones no codificantes del genoma (intrones) lo que disminuye las mutaciones deletéreas y favorece nuevas inserciones. No obstante, algunas pocas veces la inserción se realiza dentro de genes codificantes, causando enfermedades genéticas graves.

### **2.1.2 Retrotransposones de la familia humana Ta:**

Según Boissinot et al. (2000), hace aproximadamente cuatro millones de años aparece la familia Ta durante la divergencia entre humanos y chimpancés y se ha separado desde entonces en dos subfamilias, Ta0 y Ta1, siendo esta última la más reciente, la cual además está generando un nuevo subgrupo desde hace aproximadamente 1.4 millones de años. Este es el subgrupo Ta1-d.

Para las inserciones Ta, ocurre que los grupos más antiguos reducen su actividad replicativa mientras que los grupos más recientes son dominantes en cuanto a secuencias activas. Por lo tanto, Ta1-d contiene las inserciones más activas y polimórficas de la gran familia Ta.

Las tasas de acumulación por generación encontradas para la familia Ta son del orden de 0.004 por generación. Esta cifra revela tanto la transposición como la eliminación por selección, o deriva genética, y sus polimorfismos son reconocidos mediante marcadores bialélicos.

Todo lo anterior hace que estas secuencias por tener altas tasas de mutación y por su capacidad de hacer nuevas inserciones en los genomas sin poder salir de su hospedero, sean útiles para estudios filogenéticos, evolutivos, para mapeos y

librerías genéticas. Además, la posibilidad de encontrar homoplasias es mínima y su estado ancestral es conocido (Ausencia) lo que las hace más ventajosas que otros marcadores como microsatélites para la construcción de árboles de relaciones (Furano & Usdin 1995, Boissinot et al. 2000, Furano 2000, Eline et al. 2000).

### **2.1.3 Estudios genéticos usando Retrotransposones como marcadores:**

En poblaciones humanas han sido realizadas investigaciones utilizando retrotransposones como marcadores, pertenecientes al grupo SINEs (Short interspersed repetitive elements). Estos son los elementos llamados Alu que son cerca de 500.000 en primates por genoma haploide (Novick et al. 1998) y 1.000.000 en el genoma humano. Estos retrotransposones se caracterizan porque carecen de la región ORF (apertura de lectura) por lo cual no son autónomos y requieren de otros elementos móviles en el genoma para su translocación. (Eline et al. 2000).

Un estudio realizado por Batzer et al. (1994) sustentó el origen africano de los polimorfismos Alu pertenecientes a humanos. En él se analizaron 664 individuos provenientes de 16 poblaciones alrededor del mundo; las relaciones encontradas establecieron cuatro grandes grupos: uno con poblaciones africanas, otro con poblaciones europeas, otro que unía las poblaciones asiáticas con las americanas y por último un grupo de poblaciones australianas con poblaciones de Nueva Guinea.

Fue de destacar que para las inserciones Alu se conocía el estado ancestral el cual correspondía a la ausencia del retrotransposón. Este carácter se incluyó en el árbol de relaciones apareciendo las poblaciones Africanas como las más antiguas, por lo tanto, se corroboró el origen africano de estas inserciones. Adicionalmente, se demostró que los polimorfismos encontrados en estas inserciones son muy útiles en estudios evolutivos, debido a que se conoce su estado ancestral y porque la presencia de uno de estos elementos en un cromosoma particular refleja un único y particular evento en la evolución humana.

Novick et al. (1998) estudiaron una superfamilia de Alu específica para humanos cuyas inserciones son de hace 200.000 a 6'000.000 de años; unos pocos elementos

dentro de este grupo son transcripcionalmente activos y se adhieren a nuevas partes del genoma por retrotransposición mediada. La adquisición de nuevas mutaciones, a través del tiempo, los hacen bastante polimórficos convirtiéndolos en buenas herramientas para estudios genéticos. Adicionalmente, son secuencias bastante estables, sus sitios de inserción son distinguibles y su tasa de producción de nuevas inserciones es baja.

Este último trabajo, comparó 5 inserciones Alu en 815 individuos de 30 poblaciones de las cuales 24 eran del continente Americano (norte, centro y sur) y 6 de otras partes del mundo. Previo al estudio se apuntaba a que las poblaciones Americanas tenían origen asiático desde donde se desprendían tres familias lingüísticas, Esquimales, Na-dene y Amerindios; posterior al estudio se corroboró el origen asiático pero no se encontró sustento a las hipótesis de las tres familias lingüísticas, que estimaban la aparición de los amerindios hace 15.000 años, de los Na-dene entre 15.000 y 10.000 años y finalmente de los Esquimales hace 10.000 años.

Se estableció que las relaciones entre las poblaciones tenían una correlación con la distribución geográfica dentro del continente americano, estableciéndose tres grupos: uno en Norteamérica donde solo se incluye una población suramericana (Wayuu) posiblemente por el comercio que se mantenía con poblaciones antillanas, las cuales estaban bastante mezcladas; otro en Centroamérica y el tercero en Suramérica. Así como también se hizo evidente por la poca heterogeneidad entre las poblaciones el cuello de botella experimentado por los amerindios debido a que provenían de un grupo pequeño que se distribuyó desde Alaska hasta Tierra de Fuego. Este trabajo entre muchas otras cosas corrobora la sensibilidad con que los marcadores Alu permiten comparar y asociar poblaciones.

Por otro lado, los trabajos más importantes de relaciones filogenéticas y evolutivas usando retrotransposones de tipo LINE-1 non-LTR se han realizado principalmente en ratones.

De acuerdo con Furano y Usdin (1995), los elementos LINE-1 son óptimos para el estudio de relaciones entre grupos debido a su dinámica evolutiva, la replicación de éstas secuencias genera copias activas y copias defectuosas, donde las últimas,

pese a que no se expresan, son mantenidas en el genoma y acumulan nuevas mutaciones durante el tiempo (son más polimórficas). Por lo tanto, esas secuencias quedan como testigos de la historia del grupo y se pueden mantener incluso cuando en éste halla especiación. Además, si cada nuevo grupo sigue adquiriendo mutaciones y nuevas secuencias, éstas le serán exclusivas y no se compartirán con otros grupos, salvo por las inserciones que pudieran tener los ancestros comunes.

En resumen las L-1 pueden ser tomadas como un carácter filogenético, descartando en gran medida homoplasias y por su polimorfismo, pueden servir también para asociar poblaciones al igual que las secuencias Alu (Usdin et al. 1995, Verneau et al. 1997-1998, Boissinot et al. 2000).

Usdin et al. (1995) demostraron que las homoplasias pueden ser descartadas mediante el uso de secuencias LINE-1 como carácter taxonómico, por varias razones: Se puede estar seguro de que la presencia puede ser un carácter derivado compartido. Es decir, que lo pueden tener varios grupos porque provino de una transposición realizada durante la divergencia de dichos grupos, pero nunca será caso de paralelismo, convergencia, o reversión del estado ancestral, porque ¿qué posibilidad hay de que la transposición ocurra en el mismo locus en dos taxa diferentes? Mas aún la forma en que se generan nuevas inserciones demuestran que la copia original no puede salir del locus por lo tanto no hay reversión; por otra parte, la ausencia de la inserción en un taxón particular correspondería claramente al estado ancestral y finalmente, los ratrotransposones L1 non-LTR no son sensibles al numero de taxas estudiados, y no dependen de ninguna asunción evolutiva, ni del tipo de algoritmo usado al momento de construir los árboles.

Todo lo anterior fue soportado con un estudio filogenético en ratones murinos, donde se usaron secuencias L1 del tipo Lx. Estas secuencias se estima que coinciden con la radiación de la familia murinae. 24 de 27 especies que se incluían en este grupo tenían la presencia de la inserción por lo cual fueron excluidas las tres que manifestaron el estado ancestral. Estas tres especies eran mantenidas en el grupo sustentadas por supuestos patrones únicos en su primer molar, lo cual al final resulto ser una homoplasia. Siempre se encontraron problemas en la clasificación,

de la familia pues ésta se basaba en caracteres morfológicos del cráneo y dientes; con el uso de los marcadores Lx se pudieron organizar las 24 especies, que antes no tenían relaciones definidas otros datos moleculares y algunos morfológicos, corroboraron lo anterior.

Verneau et al. (1997) determinaron las relaciones evolutivas en *Rattus sensu lato* (Rodentia: Muridae) usando retrotransposones L1. Lo más interesante es que con estos marcadores, obtuvieron un solo árbol de relaciones donde en ninguna base de las ramas aparecía más de un carácter demostrando, también, la utilidad de estos elementos.

Boissinot et al. (2000) realizaron un trabajo de la evolución y amplificación de la familia humana de retrotransposones Ta (L1-non-LTR) en el cual, identificaron 91 nuevos Ta de los cuales el 34% de retrotransposones tenían todas sus regiones completas. Determinaron la tasa de sustitución nucleotídica por linaje (0.25% por millón de años) lo cual resultó ser más alto de un 15%-60% que las estimas de pseudogenes (genes que han perdido su capacidad de transcripción) y la tasa de acumulación por generación en 0.004 elementos. Encontraron, además, dos grandes subfamilias de inserciones donde la más reciente Ta1 emergió hace aproximadamente 1.6 millones de años. Finalmente, probaron la capacidad de las nuevas inserciones para estudios de poblaciones humanas, muestreándolas en 8 individuos de diferentes partes del mundo, encontrando los diferentes polimorfismos y estableciendo sus relaciones a grandes rasgos. Las inserciones Ta resultaron tan útiles como las Alu o las Lx para el establecimiento de relaciones.

Todos los anteriores trabajos, corroboran que los retrotransposones son adecuados para estudios de relaciones sea entre taxa o entre poblaciones.

#### **2.1.4 Estudios genéticos con poblaciones americanas utilizando marcadores diferentes a los retrotransposones:**

Otros marcadores diferentes de los LINEs o SINEs han sido utilizados en estudios de poblaciones resultando también útiles para esclarecer relaciones. No obstante, a la hora de escoger cual de ellos debe usarse hay que tener en cuenta algunas cosas como la cantidad de polimorfismo que se requiere para responder la pregunta de investigación. Esto es, si se tienen poblaciones muy relacionadas entre ellas es necesario que el marcador que se utilice sea altamente polimórfico para encontrar las sutiles diferencias entre una y otra; si el interés, por el contrario, es comparar tasas, el grado de polimorfismo de los marcadores no requiere ser tan alto. Otro aspecto que influye en la escogencia del marcador es la disponibilidad de acceso a los análisis de laboratorio y estadísticos y, finalmente, el más usual es la disponibilidad de recursos para costear los materiales (Parker et al. 1998).

Lo ideal es que los resultados obtenidos por cualquier método conlleven a una misma solución lo más cercana posible a la realidad entendiendo que uno u otro método están aportando solo una parte y un particular punto de vista de una misma realidad. Por lo tanto los estudios citados a continuación prueban algunas hipótesis planteadas para poblaciones humanas cuyos resultados podrían ser soportados mediante el uso de otros marcadores o, por el contrario, podrían ser debatidas.

Por ejemplo, utilizando secuencias de ADN mitocondrial (Bonatto & Salzano 1997) soportaron la hipótesis planteada por Szathmary (1978) que dió un vuelco total a lo que se había encontrado hasta entonces del proceso de poblamiento de América.

Muchos estudios antropológicos, lingüísticos y genéticos están de acuerdo en que el origen de las poblaciones americanas se encontraba en Asia. Sin embargo, se planteaban hipótesis diferentes acerca de las poblaciones que les dieron origen; según el análisis lingüístico, los indígenas americanos provenían de tres grupos diferentes que dieron origen a amerindios, Na-Dene y esquimales, respectivamente. Otros proponían que estos tres grupos lingüísticos se habían originado por tres migraciones diferentes provenientes de Siberia pero las dos últimas migraciones no correspondían con las edades de las poblaciones.

Fue así como analizando todas las secuencias disponibles del primer segmento hipervariable del DNA mitocondrial humano en la región de control de 544 americanos nativos, se encontró que esquimales, Na-Dene y amerindios estaban más estrechamente relacionados entre ellos que con cualquier población Asiática y que provendrían de una única migración a Beringia desde el este de Asia central soportando la hipótesis de Szathmary (quien sugirió la diferenciación de los grupos en Beringia); muchos de los que llegaron a Beringia atravesaron posteriormente el corredor de hielo de Alberta con rumbo a América, tiempo después, ocurrió un colapso de este corredor de hielo que los dejó aislados y desde allí migraron poblando todo el continente americano.

Chen et al. (2000) analizaron individuos del norte de Sudáfrica, de la familia lingüística Khoisan (Kung y Khwe) con análisis de alta resolución RFLPs del mtADN y los compararon con datos previos del sub-sahara encontrando que las muestras analizadas tenían presentes la mayoría de los linajes mtDNA de los que se encontraron en el sub-sahara sugiriendo que las dos poblaciones son muy antiguas. Los haplotipos encontrados en la población de Kung se posicionaron en las raíces del árbol filogenético del mtADN mostrando que esta es una de las más antiguas poblaciones encontradas en África. Adicionalmente, las estimas de las edades de los haplotipos de todas las poblaciones africanas analizadas fueron congruentes con el tiempo de origen de los humanos actuales soportando que provienen de África.

Otros tipos de estudios, usan diferentes marcadores para comparar su efectividad al momento de resolver relaciones, o para dar más peso a sus resultados. En un trabajo realizado por Jorde et al. (2000), se estudió la diversidad genética humana comparando datos mitocondriales, autosómicos y del cromosoma Y.

Los datos autosómicos incluían, 30 RSPs (Restriction-site polymorphisms) 13 Inserciones Alu, 60 STRPs (Short tandem repeat polymorphisms) y un elemento de tipo LINE-1; los mitocondriales correspondían al análisis de 611 pb de la región de control y los del cromosoma Y correspondían a 10 polimorfismos. Se muestrearon seis poblaciones Africanas, cinco Asiáticas y cuatro Europeas.

Se encontró que los datos eran congruentes para todos los sistemas analizados. Las poblaciones africanas tenían la mayor diversidad en todos los marcadores excepto para RSPs, los cuales, aparentemente, son más polimórficos en las poblaciones Europeas, atribuyéndose ésto a procesos adaptativos a los diferentes climas. Los valores de  $G_{ST}$  fueron más altos al usar sistemas autosómicos que al usar datos mitocondriales o del cromosoma Y, pues estos últimos se ven afectados por los valores bajos en el tamaño efectivo de la población. Finalmente, se encontró que para todos los sistemas, excepto los del cromosoma Y, hay menos variación entre las poblaciones de un continente que entre continentes. Aquí también se soportó el origen africano de las poblaciones humanas actuales.

Para las poblaciones Colombianas se destaca un estudio reciente realizado por Jaramillo et al. (2000) en el que se analizaron los genes APOE, APOB y ACE en comunidades afrocolombianas y amerindias. Estos marcadores corresponden a loci polimórficos encontrados dentro de los genes antes mencionados, los cuales están involucrados con el riesgo de desarrollar enfermedades arteriocoronarias.

Se muestrearon diez poblaciones amerindias, seis afrocolombianas y la población de Bogotá (mestiza). Los marcadores fueron útiles para establecer las relaciones entre esas poblaciones, donde las afrocolombianas no mostraban relaciones claras entre ellas debido entre otras cosas a un proceso de homogenización entre las mismas acorde con altas tasas de flujo génico y las amerindias no solo fueron organizadas en tres ramas importantes, sino que, además, mostraron una correlación entre las distancias geográficas y genéticas. Se encontró además, una heterogeneidad más alta en las poblaciones amerindias que en las afrocolombianas.

## ***2.2 Aspectos históricos:***

### **2.2.1 Procedencia de los negros esclavos:**

Las poblaciones colombianas están demarcadas por un altísimo mestizaje debido a que su acervo genético se compone de la mezcla de tres razas humanas, negros,

caucásicos e indígenas. Por lo tanto, es necesario comprender la historia de la conformación de las poblaciones del Chocó, Cauca y Córdoba para explicar su composición genética.

Las zonas de Cauca y Chocó antes de la llegada de los españoles estaban pobladas por grupos indígenas. Una vez colonizadas, éstos eran utilizados como esclavos para trabajos en las minas (en el Chocó) y en los cultivos (en el Cauca), pero dadas las difíciles condiciones geográficas del Chocó, los indígenas enfermaban fácilmente y morían; por ello los españoles decidieron adquirir mano de obra africana por medio de comercio con negros esclavos, lo cual ya era común para ellos desde antes del descubrimiento de América (Gutiérrez 1980).

Según Gutiérrez (1980), a los africanos capturados para su comercio se les asignaba un sobrenombre que indicaba el lugar de procedencia. Sin embargo, no siempre fue así ya que a muchos se les asignó el nombre del puerto africano desde donde eran embarcados o el nombre de la hacienda de sus propietarios españoles. Es por tal razón, que solo hay indicios generalizados de los lugares de procedencia. Con el tiempo estos sobrenombres adquirieron el título de castas y cada una de ellas describía una región costera del África Occidental y Central; para el caso de Colombia a mediados del siglo XVII, la mayoría de los esclavos correspondían a la casta Mina cuya procedencia correspondía a la fortaleza portuguesa Elmina de la costa de Oro. Posteriormente llegó la casta Arará procedente del golfo de Benín, luego los Carabalí del golfo de Biafra.

Mina (1975), por su parte, argumentó que la mayoría de los esclavos venían de: Senegal, Guinea y Angola (costa occidental de África); aunque mayoritariamente provenían de Guinea. Los grupos, según este autor, más comunes eran los Minas, Arará, Carabalí, Mandinga, Biáfara, Lucumí, Chala, Bran, Popo, Cetre, Angola, Cuagui, Bibi, Satinga, Cambara, Bane, Yolofo y las naciones de procedencia eran: Ashanti, Fanti, Yoruba, Ibo, Congo, Iwu, entre otras.

Sin embargo, Mina (1975) argumenta que muchos negros ya se habían entre mezclado en la costa africana o en los barcos que los trasladaban.

Maya (1998) hace una recopilación de datos de los Afrocolombianos llegados a Cartagena entre 1533 a 1810.

Desde 1533 a 1580, los españoles que exploraban en busca de oro introdujeron aproximadamente 3000 negros procedentes de la región comprendida entre Senegal y Sierra Leona (costa occidental de África). Los denominados Negros de Ley, que pertenecían a 16 grupos diferentes. Los Yolofofos eran mayoría a finales del siglo XVI, su lengua era ampliamente conocida por los demás esclavos, pues en el reino Yolofo en África habían personas que hablaban otras lenguas y eran llamados Berbesies, Tuculoros y Mandingas.

En 1580, Felipe II anexa a Portugal a la corona de Castilla logrando así un tratado para abastecer América de esclavos ofreciendo el monopolio de este negocio. Fue así como los portugueses llevaron a Cartagena hasta 1640, africanos del antiguo reino del Kongo (Congos, Monicongos, Anzicos y Angolas). Durante la primera mitad del siglo XVII Angola, Loanda y la isla de Santo Tomé (todos en África central) se convirtieron en los principales puertos exportadores. Entre 1580 a 1640 llegaron a Cartagena 192.397 negros que fueron distribuidos a toda la América española. La casta africana predominante fue la de los Bantú.

En el periodo comprendido entre 1640 a 1810 ocurrió en principio una parálisis en la importaciones de africanos que provocó una crisis en las minas colombianas llegando al punto de cerrar varias de ellas. Esto ocurrió por una guerra entre España y Portugal; entonces Holanda toma el control del comercio y nutrió de Esclavos a América mediante el contrabando utilizando diferentes puertos.

La cerrada de las minas originó expediciones pacificadoras desde Popayán hacia el Chocó donde se encontraron nuevas minas y esta zona se convirtió en centro del segundo ciclo minero. Hasta la segunda mitad del siglo XVII, se reactivó el comercio lícito y Cartagena logró abastecerse de africanos traídos de Jamaica y principalmente de Curazao.

Entre 1676 a 1679, llegaron a Cartagena aproximadamente 600 esclavos legales. Los datos del contrabando Holandés no se conocen pero se sabe que estos se abastecían en la costa centro occidental de África de los grupos Ararás y Popos. Cuando el comercio esclavista se reactivó, llegaron a Cartagena 9.853 negros entre 1698-1702 donde el 53% provenían directamente de África, el 25% de Jamaica, Curazao y Barbados. Para finales del siglo XVII, los negros predominantes eran Araras y Minas.

Los últimos países en tomar el dominio de la trata de esclavos fueron Francia e Inglaterra de 1704 a 1713. Los Franceses llevaron a Cartagena 3.913 esclavos del puerto de Ovilah y Mina (en el golfo de Benin); de 1714 a 1740 los ingleses introdujeron 10.475 esclavos procedentes de Jamaica con un promedio de 1.000 esclavos de contrabando al año. 30.000 esclavos provenían de la Costa de Oro y del Golfo de Benin con mayor porcentaje de de negros Minas y Araras.

Finalmente, de 1740 a 1810 a Cartagena llegaron 15.176 esclavos en libre comercio, traídos por los ingleses principalmente del Golfo de Biafra estos pertenecían a los grupos Carabali, Mina y Arará.

De acuerdo con Gutiérrez (1980), un millón y medio de esclavos fueron llevados a Hispanoamérica de los cuales 150.000 fueron introducidos por Cartagena, pues esta ciudad se había convertido en el puerto más importante además de ser legal (en cuanto a permisos otorgados por la corona española) desde donde no solo se distribuían los esclavos hacia el interior del país, sino que de allí se abastecían los mercados de la Capitanía de Venezuela, la audiencia de Quito y el Virreinato del Perú. Se calcula que en Colombia quedaron 80.000 esclavos ubicados en las regiones mineras.

Una vez llegaban a Cartagena se distribuían a terminales ubicadas en Popayán, Santa Fe de Antioquia, Honda, Anzerma, Zaragoza y Cali. Esto desarrolló un mercado de oferta y demanda, con fluctuaciones en los precios e, incluso, se presentó bastante contrabando; Los contrabandistas entraban evadiendo impuestos por el Darién, Tolú, Santa Marta y Riohacha (Gutiérrez 1980).

## **2.2.2 Conformación histórica de las poblaciones de estudio:**

### **2.2.2.1 Población del Chocó:**

El establecimiento de la población chocona se produjo en diferentes etapas iniciándose con las expediciones españolas por la costa atlántica y el río Atrato a comienzos del siglo XVI. Luego en el siglo XVII con el asentamiento de las colonias y a finales del siglo XVIII con el establecimiento de la economía minera (Vargas 1999).

A la llegada de los españoles, el Chocó estaba ocupado por los grupos indígenas Cuna, Embera y Waunana. Dentro del grupo de los Embera estaban los Eubida (gente de montaña) que se ubicaban en los cursos altos de los ríos Atrato y San Juan y los Dobidá ubicados en los afluentes orientales del río Baudó. Por su parte, los Waunana tenían su asentamiento por la costa pacífica en la desembocadura del río Baudó y los Cuna (gente de llanura) habitaban las zonas planas de las cuencas del Atrato y el río Baudó. (Zambrano 1998, Vargas 1999).

Los esclavos que llegaron a esta región desde 1680 provenían de África central y la costa de oro. Eran en su mayoría de las castas Congo y Mina. Por otra parte, los que llegaron a esta zona por contrabando, muy seguramente eran Araras y Popos de la costa centro occidental de África, de acuerdo con los sitios de aprovisionamiento de los Holandeses, quienes eran los que manejaban el contrabando en esa época.

De acuerdo con Zambrano (1998) la primera incursión española al Pacífico se dió en 1509, por el golfo de Urabá fundando las poblaciones de San Sebastián de Urabá y Santa María la Antigua del Darién.

En 1510, Pascual de Andagoya estableció un puerto en Buenaventura que se conectaba con Cali por el cañón del Dagua. Los siguientes ríos explorados fueron

San Juan y Atrato. Este último fue poblado por mineros antioqueños y del alto Cauca que vivían en las urbes de Anserma, Cartago, Cali y Popayán. Los españoles llamaron Chocó a los Embera y Waunana y a comienzos del siglo XVII con las poblaciones indígenas fragmentadas y con la economía minera se inició una introducción masiva de Africanos desde Anserma y Popayán.

#### **2.2.2.2 Población de Cauca:**

El área que hoy comprende el departamento del Cauca a la llegada de los españoles estaba ocupada principalmente por los indígenas Páez y Guambianos, que pertenecen a la etnia de los caribes. Estos se ubicaban en la cordillera central donde los últimos, ocupaban la vertiente occidental al noreste del departamento, allí también se encontraban los grupos: Aviramas, Totoroes, Polindaras y paniquitae (Todos de la familia chibcha); Los Kokonukos estaban al oriente de Popayán y los Patias, Bojoles, Chapanchicas y sindaguas estaban en el sur del departamento y en el valle del río Patía; mientras que por el lado del Pacífico estaban los grupos indígenas que se consideran de la familia del Choco y se les llaman cholos, Estos son los Emberas.

En cuanto a los esclavos, éstos provenían de Senegal, Angola y Guinea y los grupos más comunes que ocuparon la zona fueron los Carabalí, Mina, Arará, Mandruga, Biáfara y Lucumí (Beltrán y Hernández 1993).

El principal centro de abastecimiento de la costa pacífica de Cauca fue Zaragoza (Gutiérrez 1980).

#### **2.2.2.3 Población de Córdoba**

El área que actualmente ocupa Córdoba estaba habitada por los indígenas Zenues antes de la llegada de los españoles. Estos eran una tribu de la familia Caribe. Estos indígenas se dividían en tres provincias: los Finzenu que habitaban el valle del río Sinú hasta la serranía de San Jacinto, los Panzenu que se ubicaban en el valle del río San Jorge hasta la desembocadura del río Magdalena y los Senufana que

ocupaban los valles del bajo Cauca y el Nechí. Estos indígenas trabajaban el oro por lo cual, a la llegada de los españoles se vieron diezmados, algunas tribus fueron exterminadas y otras se dispersaron. Actualmente los pobladores indígenas corresponden a Chocó Emberá y Katios, donde algunos pocos aún se conservan sin mezclas, pero ninguna de las tribus actuales parecen ser herederos de los Zinues.

Se cree que los Embera llegaron del Chocó avanzando por las selvas del norte del departamento cuando ya estaba ocupado por los españoles. La historia data invasiones de estos indígenas hacia 1628 a la zona de Antioquia. (Archivos de EL TIEMPO y EL ESPECTADOR).

Parte de la economía de la zona se ha basado en la explotación minera. Los españoles trajeron consigo negros esclavos para trabajar las minas. Su central principal de abastecimiento era Cartagena.

La población actual, se compone de afrocolombianos y amerindios Emberas y Katios principalmente en las áreas rurales, y de “blancos” mestizos en las cabeceras del departamento. En los últimos años se ha presentado una incursión del elemento árabe especialmente de sirios y libaneses.

### **2.2.3 Migración y mestizaje:**

Durante la colonia, los negros africanos vieron amenazada su identidad, debido entre otras cosas a la ideología del blanqueamiento, asociando el color de la piel con la posibilidad de encontrar mayor libertad (Wade 1996).

Al respecto, Gutiérrez (1980) cita un texto de D. Jorge Juan y Antonio de Ulloa, capitanes de la fragata, los cuales, hacen una descripción del mestizaje en Cartagena comentando la existencia de blancos europeos y criollos (hijos de aquel país). La mezcla entre un blanco y un negro era llamada mulato, mulato y blanco formaba el tercerón; tercerón y blanco formaba el cuarterón; cuarterón con blanco

forma el Quiterón, y por ultimo, la mezcla de blanco y Quiterón era donde se creía producir ya un español considerado libre de todo antecedente negro.

Las mezclas que podían ocurrir en sentido contrario, es decir que desde el grado de Mulato se mezclaran con negros, eran denominadas “salto atrás” porque en lugar de adelantarse a ser blancos retrocedían a ser negros. Finalmente, todas las mezclas hasta Quinterón entre negros e indígenas se denominaban Zambos.

Por otra parte, según Mina (1975), durante las etapas de mayor crisis en el comercio de la trata de negros esclavos los costos de éstos se incrementaron y una de las estrategias españolas para conseguir más esclavos y evitar las rebeliones, fue promover las relaciones familiares entre ellos. Es decir, que éstos se casaran y mantuvieran una familia estable ya que esto los hacia más sumisos. Dichas familias podían ya estar establecidas desde África o se podían establecer entre los esclavos pertenecientes a un mismo dueño. Esto promovió el mestizaje al interior de los diferentes grupos africanos.

Después de la colonia, las relaciones entre indígenas y negros en la costa pacifica fue cordial. Los indígenas se asentaban en las partes altas de los ríos y los segundos en las partes bajas y en las llanuras debido entre otras cosas a los diferentes nichos ocupados por ambos grupos, pues los afrocolombianos se dedicaban a la extracción de minerales y a la tala y los indígenas al cultivo; sus relaciones se basaban en el intercambio y el compadrazgo (parentesco que contrae el padrino con los padres de una criatura). Esto ultimo evitaba los vínculos sexuales entre las dos nuevas familias (Wade 1996 y Pardo 1996).

Según Pardo, los blancos mestizos (criollos) ocuparon el nicho dejado por los españoles en las funciones de prestancia pero no se arraigaron mucho en estas zonas.

A mediados del siglo XIX el Chocó recibió importantes migraciones de afrocolombianos del Atlántico a través del Darién en respuesta a la presión demográfica.

A partir de los 50, aumentó la demanda maderera y el Pacífico se convirtió en la fuente principal de este recurso; los aserrios se consolidaron como la actividad capitalista más importante y las poblaciones afrocolombianas, se conglomeraron en torno a estas fuentes de trabajo perdiendo su distribución dispersa lineal, que les caracterizaba hasta entonces.

Los afrocolombianos desde el fin de la colonia hasta ahora han tenido constantes migraciones hacia centros urbanos en busca de trabajo en minería, aserrios y centros agrícolas, por lo cual, la mitad de la población total de afrocolombianos se concentra en Quibdó, Buenaventura y Tumaco.

En Quibdó por la promoción de proyectos de nuevas vías de acceso, la inmigración de colonos blancos mestizos a aumentado pues éstos buscan monopolizar el comercio. Es por ello que Quibdó es una de las poblaciones más mestizas ( Wade 1996 y Vargas 1999).

La costa pacífica mantiene unos 800.000 habitantes de los cuales el 85% son afrocolombianos, el 5% indígenas Embera y Waunana, que viven exclusivamente en áreas rurales, y 10% mestizos blancos, que ocupan las cabeceras municipales (Pardo 1996).

### **3. FORMULACIÓN DEL PROBLEMA Y JUSTIFICACIÓN:**

#### ***3.1 Formulación del problema:***

Se partió del hecho de que existían diferencias para marcadores retrotransposones entre las poblaciones, debidas al aporte genético particular de los diferentes grupos indígenas, africanos y españoles que conformaron a cada una.

El trabajo buscó aportar información genética que midiera las diferencias entre Cauca, Choco y Córdoba además, de obtener información de los eventos genéticos poblacionales que estuvieran ocurriendo al interior de ellas. Con retrotransposones que nunca antes se habían aplicado a poblaciones humanas Colombianas.

#### ***3.2 Preguntas de investigación:***

¿Cuales son las medidas de las diferencias y relaciones genéticas entre las tres poblaciones en cuanto a los marcadores estudiados?

¿Se pueden explicar las diferencias y/o relaciones, encontradas, con datos históricos de esas poblaciones humanas?

¿Serán útiles los datos genéticos obtenidos para aportar nuevos conocimientos a la reconstrucción de la historia de las poblaciones afrocolombianas, de las cuales falta información debido a perdida de documentos, contrabando de los negros esclavos y desconocimiento de la procedencia exacta de estos?

¿Qué tan diferente puede ser la población blanca mestiza de Córdoba con respecto a las otras dos afrocolombianas?

#### ***3.3 Justificación de la investigación:***

Colombia es un país con un alto mestizaje y con poco conocimiento de la historia del establecimiento de sus poblaciones, no tanto por desinterés de los historiadores sino porque muchas veces la información es incompleta, esta perdida o no fue registrada nunca.

Esta carencia es aún más fuerte en cuanto a las poblaciones afrocolombianas, que pese a que comparten una relativa cercanía geográfica entre ellas, su mayor concentración está en las costas sobretodo en la Pacífica. Estas poblaciones deberían presentar diferencias genéticas remarcables, respecto a los grupos indígenas y españoles que también convergieron en el territorio que hoy en día llamamos Colombia. Adicionalmente, muchos africanos llegaron a estas regiones por vía ilegal, pues el contrabando se presentó en aquella época (siglo XVII).

En este punto, la genética de poblaciones se muestra como una herramienta muy útil para medir las relaciones y diferencias entre las tres poblaciones; de lo cual en futuras comparaciones con poblaciones africanas y españolas o indígenas, se podría esclarecer el aporte que cada grupo étnico les dió y quizá las rutas históricas por las que llegaron esas secuencias si se rastrearan.

Es por tanto, que el presente estudio cobra importancia, como un acercamiento al conocimiento de la dinámica genética interna que puede estar ocurriendo al interior de las poblaciones que se estudiarán, así como a la medición genética de sus diferencias y relaciones con marcadores moleculares de nuevo uso como son los retrotransposones.

Si los resultados obtenidos pueden de alguna manera justificarse con los datos históricos de que se dispone se abrirá una nueva manera de corroborar o de reconstruir la historia Colombiana por medio de la genética de poblaciones y en particular por medio del uso de marcadores como los retrotransposones Ta que son supremamente útiles en el establecimiento de relaciones filogenéticas por su mínima posibilidad de producir homoplasias.

## **4. OBJETIVOS:**

### ***4.1 Objetivo general:***

Obtener las frecuencias alélicas para tres poblaciones, dos afrocolombianas de Cauca y Quibdo (Choco) y una fenotípicamente blanca mestiza de Córdoba utilizando la presencia o ausencia de inserciones en retrotransposones de la familia humana TA y someterlas a análisis estadísticos, para tratar de establecer las diferencias y similitudes poblacionales, así como los eventos genéticos que estén actuando en ellas.

### ***4.2 Objetivos específicos:***

- Calcular el equilibrio Hardy-Weinberg para los marcadores Ta.
- Calcular distancias genéticas, el desequilibrio gamético y el grado de heterogeneidad entre las tres poblaciones estudiadas.
- Construir dendrogramas de relaciones entre esas tres poblaciones.
- Aplicar un métodos multivariantes a los datos para comparar sus resultados respecto a los otros análisis genético poblacionales desarrollados.
- Explicar los resultados encontrados mediante datos históricos.

## **5. HIPOTESIS**

Existen diferencias muy significativas en las frecuencias alélicas de las inserciones de los retrotransposones entre las tres poblaciones que permiten separarlas unas de otras, consiguiendo una estructura muy marcada de la población total.

## **6. MATERIALES Y METODOS**

### ***6.1 Diseño de la investigación:***

La investigación que se realizó fue de tipo analítica y descriptiva, cuyo factor de diseño fueron siete marcadores retrotransposones de tipo Line-1 pertenecientes a la familia humana Ta; los niveles de factor de diseño fueron las frecuencias alélicas de cada población para dichos marcadores. La unidad de respuesta fue cada población muestreada: Chocó, Cauca y Córdoba. La variable de respuesta correspondió con cada uno de los estados monomórficos o polimórficos de cada marcador para los individuos muestreados dentro de las poblaciones. Por último, la unidad de muestreo fue un grupo de 90 individuos de los cuales 30 pertenecían a la población de Choco, 30 a la población del Cauca y 30 a la población de Córdoba. Las muestras de las dos primeras poblaciones correspondían a personas fenotípicamente afrocolombianas mientras que las de Córdoba eran personas de fenotipo “blanco”.

#### **6.1.1 Población de estudio y muestra:**

Se obtuvieron muestras de ADN de 90 individuos, provenientes de 3 poblaciones diferentes de Colombia, correspondiendo a: 30 individuos afro colombianos de Quibdó-Chocó, de los cuales se conocía que 15 eran mujeres, 30 individuos afro colombianos de la costa pacífica de Cauca y 30 individuos blancos de cabeceras municipales de Córdoba. Las muestras de Chocó y Cauca fueron enviadas por la Universidad de Antioquia y las de Córdoba son parte de una investigación poblacional más extensa que se realiza en el seno del grupo de Biología Evolutiva (Genética de Poblaciones) del departamento de Biología de la Pontificia universidad Javeriana. El ADN se había extraído previamente de muestras de sangre mediante el método del fenol cloroformo, las muestras purificadas eran mantenidas a una temperatura de -20°C durante el tiempo de realización del presente trabajo.

### 6.1.2 Variables del estudio:

Se utilizaron siete variables dentro del estudio, correspondiendo a siete retrotransposones de tipo LINE-1, específicamente de la familia TA que es exclusiva de humanos (Skowronski et al 1988). Se amplificaron arreglos o inserciones dentro de cada uno de ellos utilizando los marcadores desarrollados por Boissinot et al. (2000). Estos fueron: Tau-1, ubicado dentro de un retrotransposón en el cromosoma cinco humano; la presencia de la inserción tiene aproximadamente 970 pb (pares de bases) y su ausencia 430 bp; Tau-3, ubicado en un retrotransposón del cromosoma siete, con una presencia de la inserción de 660pb y una ausencia de 500 pb; Tau-9, ubicado en un retrotransposón del cromosoma veintiuno, con una presencia de la inserción de 1080pb y una ausencia de 800pb; Tau-10, ubicado en un retrotransposón del cromosoma seis, cuya inserción tiene un tamaño de 1560 pb y su ausencia es de 800 pb; Tau-11, localizado en un retrotransposón del cromosoma cinco cuya presencia tiene 2400pb por su tamaño no amplifica en una primera PCR en esta solo se puede determinar su ausencia que mide 500pb, se requiere de una segunda PCR que revela la presencia de la inserción con una banda de 350pb aproximadamente. Los marcadores hasta aquí mencionados no han sido clasificados aún dentro de ninguna de las dos subfamilias de la familia TA Boissinot et al (2000).

TaO-11, pertenece a la subfamilia antigua TA0, la inserción se encuentra en el cromosoma X, por lo cual, solo pudieron ser muestreadas las mujeres para evitar la distorsión de los datos, su presencia equivale a 3400 pb por lo que requiere al igual que Tau11 de doble PCR, en la primera solo se revela la ausencia del arreglo con una banda de 480 pb y en la segunda la presencia con una banda de 780 pb; y finalmente se utilizó Ta1-22, perteneciente a la subfamilia mas actual Ta1, se ubica en el cromosoma 11 y requiere de doble PCR pues su presencia es de 2500pb por lo que, en la primera PCR solo se distingue su ausencia con 475 pb y en la segunda su presencia con 900 pb.

A cada variable se le determinó su estado en cada individuo como Homocigotos para la ausencia o para la presencia de la inserción en el retrotransposón y heterocigotos.

## **6.2 Métodos:**

Las muestras de ADN de cada individuo se amplificaron en un termociclador Perkin-Elmer usando condiciones de reacción estándar con variación de la temperatura de anillamiento de acuerdo con las descritas por el fabricante Boissinot et al. (2000) para los retrotransposones del estudio: Tau-1/52°C, Tau-3/50°C, Tau-9/48°C, Tau-10/48°C, Tau-11/48°C en ambas PCR, Ta0-11/48°C en ambas PCR y Ta1-22/48°C en ambas PCR.

Tau-11, Ta0-11 y Ta1-22 requerían de dos PCR para la obtención de resultados debido al tamaño del retrotransposón entre 2400-3400 bp.

Los productos de PCR se separaron mediante electroforesis en gel de agarosa al 0.8% a 100v durante 40 minutos, con tinción de Bromuro de Etidio y se visualizaron las presencias o ausencias de los retrotransposones en cámara UV teniendo como patrón de tamaños moleculares *hind* III. Algunos geles fueron fotografiados con cámara Polaroid.

## **6.3 Recolección de información:**

Se llevó un registro por población, de cada individuo y su estado (polimórfico o monomórfico) para cada uno de los marcadores utilizados, tomando (a) como la ausencia del retrotransposon y (A) como la presencia, de tal manera que un genotipo (aa) correspondía a un homocigoto para la ausencia, (AA) a un homocigoto para la presencia y (Aa) a un heterocigoto. Los datos se organizaron dentro de un formato que contenía el nombre de la población, el código del individuo, el marcador utilizado y su estado por marcador.

Como Tau-11, Ta0-11 y Ta1-22 requerían doble PCR los datos se obtenían teniendo en cuenta lo siguiente: en la primera PCR las muestras que presentaban una banda del tamaño esperado era porque, en al menos uno de los alelos estaba ausente la inserción, mientras que las muestras sin bandas sugerían la presencia de la inserción en ambos alelos; con la segunda PCR se visualizaban las muestras que presentaban la presencia de la inserción . Es decir, que una muestra que presentaba bandas en las dos PCR correspondía a un Heterocigoto, uno que amplificaba en la primera pero no en la segunda correspondía a un homocigoto para la ausencia y uno que solo amplificaba para la segunda PCR correspondía a un homocigoto para la presencia de la inserción.

#### **6.4 Análisis de la información:**

Una vez visualizados los estados en cada individuo para cada marcador se procedió a determinar las frecuencias alélicas por conteo directo. Posteriormente se elaboraron matrices con los datos obtenidos asignándoles valores de 1 al alelo de ausencia y 2 al de presencia, las cuales fueron corridas en el programa GENPOP versión 3.1.

Primero, se calculó el equilibrio Hardy-Weinberg, que es una regla que relaciona las frecuencias alélicas con las frecuencias genotípicas donde dichas frecuencias serán las mismas generación tras generación si se cumplen ciertas condiciones: apareamiento aleatorio, tamaño infinito de la población, no mutación, no migración, no selección natural, reproducción sexual y no solapamiento de generaciones. Si esto se cumple se mantendrá el equilibrio.

El programa GENPOP calcula las frecuencias esperadas para cada marcador y las compara con las observadas basado en la hipótesis de la existencia de equilibrio, si encuentra que estas no son iguales rechaza el equilibrio evaluando la significancia estadística de los resultados mediante un test exacto. Se tomaron dos hipótesis alternativas diferentes para aumentar el poder estadístico del análisis (Goudet 1996) Exceso de homocigotos y exceso de heterocigotos.

El test de equilibrio se practicó por marcador para cada población y por marcador para todas las poblaciones juntas, según los métodos Weir & Cockerham (1984) y Robertson & Hill (1984) presentando valores entre cero y uno donde valores menores de 0.05 rechazan la hipótesis nula y por encima de esta cifra, se mantiene la hipótesis nula.

Posteriormente se dió significación a los valores de la chi-cuadrado por el método de Fisher.

En segundo lugar, se practicaron dos test multilocus que se basan en el método de las cadenas de Markov y determinan si el no equilibrio se presenta por déficit o exceso de heterocigotos. Esto es un método de análisis alternativo al anterior.

En tercer lugar se determinó si existía desequilibrio gamético (D) entre pares de marcadores. Este se presenta cuando dos genes que no están ligados físicamente muestran una mayor frecuencia de individuos que portan o no las dos características diferente a lo que se esperaría para dos genes totalmente desligados.

Para esto, se compararon al azar pares de marcadores dentro de cada población con el método de las cadenas de Markov con mil permutaciones. Este arroja las probabilidades de que se presente desequilibrio gamético entre dichos pares; posteriormente se compararon pares de marcador a lo largo de la población total con el método de Fisher el cual evalúa la significancia de las chi-cuadrado resultantes. Esto también permite saber si la deriva genética puede ser un evento evolutivo trascendentalmente significativo dentro de la población.

Se calculó también el flujo génico entre las poblaciones estudiadas a partir de los estadísticos de heterogeneidad genética ( $F_{ST}$ ). Wright (1951) mostró que con un modelo isla donde se haya alcanzado el equilibrio entre deriva genética y flujo genético, entonces  $Nm = (1 / F_{ST} - 1) / 4$ .

El siguiente paso fue estimar la heterogeneidad genética entre pares de poblaciones y en las poblaciones tomadas como un todo. Este análisis determina si hay diferencias significativas entre las poblaciones basándose en las frecuencias alélicas y genotípicas.

La heterogeneidad o diferenciación genética pone de manifiesto si existen o no diferencias entre una población que permita dividirla en sub-poblaciones. Se determino primero usando el programa GENEPOP por el método de las cadenas de Markov.

Posterior a esto la primera matriz construida con los resultados obtenidos en el laboratorio, fue transformada para el calculo estadístico en el programa FSTAT 2.91 Goudet (2001). Con este se determinó el grado de heterogeneidad con el método de Nei y la estructura poblacional con los estadísticos F.

El análisis de la diversidad genética de Nei (1973) permite calcular el grado de diferencias genéticas (medidas por las frecuencias de heterocigotos) al jerarquizar una población total en sub-poblaciones. Este se calcula conociendo las frecuencias alélicas entre cada jerarquía. El análisis tiene la ventaja de ser independiente de cualquier evento genético poblacional (deriva, endogamia, cuellos de botella etc.).

En él se calcula la heterogeneidad esperada  $H_s$ , la observada  $H_o$  y la total  $H_t$  para luego calcular el estadístico  $D_{st} = H_t - H_s$ , el cual determina las diferencias genéticas entre las sub-poblaciones y la población total, pero que incluye la identidad genética propia, por tanto es corregido por la  $G_{ST} = D_{ST} / H_T$ . Este último estadístico tiene un valor relativo pues depende del valor grande o pequeño de  $D_{ST}$ . Valores de  $G_{ST}$  significativamente diferentes de cero corresponderán a una alta diferenciación genética entre las sub-poblaciones. La  $G_{ST}$  se mide por locus y por el total de loci.

Para determinar la estructura de las poblaciones se estimaron los estadísticos F según Weir & Cockerham (1984) siguiendo las estimas de relaciones de Queller &

Goodnight (1989) y las correcciones de relaciones endogámicas de Pamilo (1984, 1985)

Se calcularon las tres F ( $F_{IT}$ ,  $F_{IS}$  y  $F_{ST}$ ) donde  $F_{IS}$  es la medida de desviación de las frecuencias genotípicas esperadas para una reproducción panmíctica en términos de deficiencia o exceso de heterocigotos. Esto es lo que se conoce como el coeficiente de endogamia, el cual es convencionalmente definido como la probabilidad de que dos alelos en un individuo sean idénticos por descendencia. Es decir, la correlación de unir gametos parentales a los gametos dados al azar dentro de una sub población (individuos/sub población) promediado sobre sub poblaciones.

$F_{IS} = H_s - H_o / H_s$  donde  $H_o$  es la heterogeneidad observada y  $H_s$  la heterogeneidad esperada, asumiendo entrecruzamiento aleatorio.  $F_{IS}$  puede tomar valores entre -1 y +1 valores negativos indican un exceso de heterocigotos (no endogamia) y valores positivos un déficit (endogamia).

$F_{IT}$  funciona igual que  $F_{IS}$ , solo que a nivel de individuos con respecto a la población total es decir, todas las sub-poblaciones juntas como una sola gran población. Sus valores también oscilan entre -1 y +1 y se interpretan igual que en el caso anterior.

$F_{ST}$  mide el efecto de la subdivisión poblacional, la cual es la reducción en la heterogeneidad de una sub-población debido a la deriva genética. Este estadístico también es llamado coeficiente de coancestralidad y se define como la correlación de gametos dentro de las sub-poblaciones con respecto a los gametos sacados al azar de la población total (sub población/población total).

$$F_{ST} = H_T - H_s / H_T$$

Los valores de  $F_{ST}$  oscilan entre 0 y 1, donde 0 quiere decir que no hay estructura (subdivisión) dentro de la población, y 1 quiere decir que se presenta un aislamiento total entre las sub-poblaciones (subdivisión extrema).

Posteriormente, fueron estimadas las distancias genéticas, para lo cual se agruparon los individuos en poblaciones con sus respectivas frecuencias alélicas y esta matriz se corrió en el programa NTSYS; dos distancias genéticas de

propiedades matemáticas diferentes fueron aplicadas: La distancia de Nei (1972) y la distancia de la cuerda de Cavalli-Sforza & Edwards (1967).

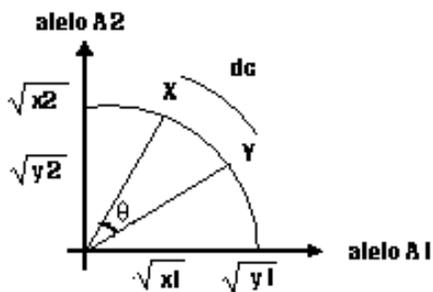
Estas dos se diferencian en que la segunda es una distancia geométrica muy usada para clasificación de poblaciones por lo que sus valores no se ven afectados por ningún evento poblacional ya que no hace ninguna asunción biológica, mientras que las distancias de Nei (1972) son usadas para estudios de evolución y sus valores se ven afectados por algunos eventos poblacionales clásicos.

La distancia de Cavalli-Sforza & Edwards (1967), consiste en la ubicación dentro del hiperespacio m-dimensional puntos que representan la raíz cuadrada de las frecuencias alélicas, los cuales son proyectados desde un origen (de radio máximo 1) desde el cual, se forman ángulos entre un punto y otro; a partir del valor de se puede medir la distancia entre un punto y otro por el largo de la proyección o cuerda entre ellos. La distancia se calcula matemáticamente así:

$$dc = \sqrt{2 - 2\cos(\theta)}$$

para el caso de dos alelos es posible graficar el espacio de forma circular como muestra la figura 1.

Figura 1. representación geométrica de las distancias Chord de Cavalli-Sforza & Edwards para dos alelos



X es la población 1  
 Y es la población 2  
 $x_1$  y  $x_2$  son las raíces cuadradas de los alelos de la población X  
 $Y_1$  y  $Y_2$  son las raíces cuadradas de los alelos de la población Y  
 dc es la distancia de la cuerda  
 $\theta$  es el ángulo entre X y Y

La distancia de Nei (1972), por su parte, se basa en la medición de las diferencias codónicas mínimas en varios loci para muchos individuos y su posterior computo con el promedio del mínimo número de diferencias codónicas entre y dentro de las poblaciones. Donde una diferencia codónica consiste en que se saquen dos alelos al azar, ya sea de una misma población o de dos poblaciones distintas y se encuentre que ambos son diferentes. Esto representaría una diferencia mínima de un codon entre ellas.

El mínimo de diferencias codónicas entre dos alelos sacados de la población X y de la población Y están dados por:

$$dx = 1 - \sum Xi^2 \quad y \quad dy = \sum Yi^2 \text{ respectivamente.}$$

Por lo tanto, el número mínimo de diferencias codónicas entre las poblaciones estará dado por:

$$d = dx - (dx + dy) / 2 = \sum (xi - yi)^2$$

Finalmente, la mínima distancia genética se obtiene promediando el total de las d sobre todos los loci:

$$D(m) = Dxy(m) - [Dx(m) - Dy(m)] / 2$$

La desventaja de este tipo de distancia es que si las dos poblaciones presentan alelos muy diferentes (como en el caso de especies o géneros diferentes) se subestima el valor de Dm siendo este más pequeño que uno cuando debería ser de uno que representa la máxima distancia, por lo tanto es utilizada solo en casos de comparaciones intraespecíficas. No obstante, Nei (1975) demostró que sus propiedades matemáticas las hacen muy confiables para estudios de polimorfismos genéticos intraespecíficos.

Posteriormente, se realizó un análisis multivariante. Para cada distancia se construyeron dendrográmas de relaciones usando sus matrices y dos algoritmos diferentes UPGMA Sokal y Michener (1985) y Neighbor-joining Saitou y Nei (1987), con el fin de visualizar las relaciones entre las poblaciones.

UPGMA (Unweight pair group médium arithmetic) busca en la matriz de distancias, la mas pequeña entre dos poblaciones constituyendo esta como la primera rama del árbol, luego toma cada unidad taxonómica y le calcula el promedio de similitud que tiene con dicha rama, a la nueva matriz de distancia obtenida por este procedimiento le extrae nuevamente la mínima distancia formando así la segunda rama y nuevamente repite el proceso hasta que se obtengan los árboles.

Neighbor – Joining es un método de descomposición, contrario a lo que hace un análisis de ramas, este algoritmo rastrea nodos en un árbol en cambio de rastrear taxa o ramas de taxa. En el procedimiento, se generan dos nuevos valores a la matriz original de distancias  $r$  y  $r/2$  donde  $r$  es la suma de las distancias entre dos OTUS (taxa, poblaciones, grupos etc.) con respecto a las demás. Luego se genera otra matriz en la que se encuentran las distancias originales con unas nuevas distancias transformadas que se consiguen restando a la distancia original entre dos OTUs el promedio de los valores  $r$  de ambos. Luego se crea el primer nodo (valor que justifica una rama) escogiendo el valor mas pequeño de la matriz transformada. Se genera entonces una nueva matriz reducida a la que se le aplican todos los pasos anteriores hasta obtener el árbol o árboles (en el caso de que en la matriz transformada hayan dos valores menores iguales) finales. Este es un algoritmo que genera árboles aditivos.

Finalmente se realizó un análisis de coordenadas principales (Pcord) para cada distancia, esta técnica desarrollada por Gower (1966) relaciona matrices de distancias genéticas con matrices reconstruidas de coeficientes de similitud en el espacio multidimensional, luego genera representaciones gráficas del parecido entre poblaciones en 2 y 3 dimensiones.

Para obtener las coordenadas principales se construye primero la matriz S de coeficiente de similitud para obtener posteriormente la matriz de agrupamiento T:

$$T = [t_{ij}], \text{ donde } t_{ij} = \overline{S_{ij}} - \overline{S_i} - \overline{S_j} + S \text{ Donde,}$$

$\overline{S_{ij}}$ : Es cada uno de los elementos de la matriz S

$\overline{S_i}$ : Es el promedio de las i-ésimas filas que se encuentran en la matriz S

$\overline{S_j}$ : Es el promedio de las j-ésimas columnas de la matriz S

S: Es el promedio de todos los elementos que conforman la mencionada matriz

Luego se diagonaliza la matriz T:

$$T = M D u B' \text{ Donde,}$$

M es la matriz de vectores propios

B' es la matriz transpuesta de vectores propios.

Du es la matriz diagonal de vectores propios.

Así de esta ultima matriz se escogen las coordenadas principales de interés.

## 7. RESULTADOS

Las frecuencias alélicas que se obtuvieron para cada marcador en cada población fueron las que se muestran en la Tabla 1. Nótese que para cinco loci estuvo presente la inserción del retrotransposón dentro de las poblaciones mientras que, para Tau-9 y Tau-10 estuvo ausente,  $a = 1$ . Es decir, que los individuos analizados resultaron ser homocigotos para la ausencia (aa) en las tres poblaciones. Los datos para Cauca y Córdoba de Ta0-11 están ausentes porque se desconocía el sexo de los individuos de la muestra. Este es un retrotransposón ligado al sexo y por lo tanto, para evitar distorsión de los resultados solo se pueden muestrear las mujeres.

**Tabla 1. Frecuencias alélicas para cada marcador en las tres poblaciones**

	Tau-1		Tau-3		Tau-9		Tau-10		Tau-11		Ta0-11		Ta1-22	
	A	a	A	a	A	a	A	a	A	a	A	a	A	a
Cauca	0,672	0,328	0,800	0,200	<u>0,000</u>	<u>1,000</u>	<u>0,000</u>	<u>1,000</u>	0,683	0,317	N.D.		0,519	0,481
Choco	0,685	0,315	0,750	0,250	<u>0,000</u>	<u>1,000</u>	<u>0,000</u>	<u>1,000</u>	0,667	0,333	0,467	0,533	0,367	0,366
Córdoba	0,800	0,200	0,648	0,352	<u>0,000</u>	<u>1,000</u>	<u>0,000</u>	<u>1,000</u>	0,800	0,200	N.D.		N.D.	

A= Presencia del retrotransposon; a= Ausencia del retrotransposon ND= No hay datos

El test exacto para el equilibrio Hardy-Weinberg por población mostró valores inferiores a 0.05 en Ta1-22 para Cauca y en Tau-1 para las tres poblaciones estudiadas lo cual es significativo para rechazar el equilibrio en estos dos marcadores (ver tabla 2)

Al calcular el equilibrio global para cada marcador, es decir en las tres poblaciones simultáneamente, se encontraron valores de chi-cuadrado de 44.1 para Tau-1 con 6 g.l. y una probabilidad de 0. Adicionalmente, Ta1-22 presento valores de chi-cuadrado de 12 para 4 g.l. con probabilidad de 0.017. Estas dos probabilidades son significativas para rechazar la hipótesis por lo tanto, estos dos marcadores no estuvieron en equilibrio Hardy-Weinberg para el conjunto de poblaciones. En el caso de Ta1-22 el hecho de que solo la población de Cauca no se encontrara en

equilibrio, fue suficiente para que al evaluar las poblaciones en conjunto se rechazara también el equilibrio.

Tau-3 tuvo valores de de chi cuadrado de 5.5 con 6 g.l. y probabilidad de 0.4778; Tau-11 tuvo valores de 0.0 con 6 g.l. y probabilidad de 1. Estas probabilidades no son significativas para rechazar el equilibrio.

**Tabla 2. Resultados test exacto equilibrio Hardy-Weinberg por población**

	<b>Marcador</b>	<b>P-val</b>	<b>W&amp;C<sup>a</sup></b>	<b>R&amp;H<sup>o</sup></b>
<b>Cauca</b>	Tau-1	<u>0,0014</u>	0,620	0,637
	Tau-3	0,3071	0,183	0,187
	Tau-11	1,0000	0,016	0,016
	Ta1-22	<u>0,0056</u>	0,568	0,585
<b>Choco</b>	Tau-1	<u>0,0002</u>	0,751	0,776
	Tau-3	1,0000	0,039	0,040
	Tau-11	1,0000	-0,033	-0,034
	Ta0-11	0,1145	0,491	0,517
	Ta1-22	0,4485	0,155	0,158
<b>Córdoba</b>	Tau-1	<u>0,0010</u>	0,759	0,786
	Tau-3	0,2050	0,287	0,294
	Tau-11	1,0000	-0,025	-0,025

<sup>a</sup> Weir & Cockerham 1984; <sup>o</sup> Robertson & Hill 1984

El test multilocus cuya hipótesis nula es déficit de heterocigotos arrojó probabilidades de 1 con desviación estándar de 0 tanto por población como por locus. Lo cual, no permite rechazar dicha hipótesis. Lo anterior fue constatado por el test multilocus cuya hipótesis nula es un exceso de heterocigotos, los valores de probabilidad fueron de 0 con desviación estándar de 0. Este valor obliga a rechazar la hipótesis nula. Estos resultados fueron contundentes para determinar que la falta de equilibrio Hardy-Weinberg en los marcadores se debió a un déficit de heterocigotos en las poblaciones.

No se encontró desequilibrio gamético por pares de marcadores dentro de cada población. Tampoco por pares de marcadores para todas las poblaciones cogidas como una sola. Es decir que la segregación gamética de cada Ta estudiado es independiente de los demás analizados. Este resultado permite descartar la deriva genética como causante del déficit de heterocigotos.

La heterogeneidad por pares de poblaciones no fue significativa (probabilidades superiores a 0.05 con desviaciones estándar entre 0 y 0.006), como tampoco lo fue la heterogeneidad global, (ver tabla 3). Es decir que para cada locus analizado entre pares de poblaciones no habían diferencias significativas así como tampoco las había para las tres poblaciones en conjunto, ni por locus, ni entre los loci estudiados.

**Tabla 3. Heterogeneidad Global**

<b>Locus</b>	<b>Valor probabilidad</b>	<b>Desviación Estándar</b>
Tau-1	0,31806	0,01061
Tau-3	0,18736	0,00751
Tau-11	0,21898	0,00864
Ta1-22	0,12906	0,00644
<b>Test combinado (Método de Fisher)</b>		
<b>chi-cuadrado</b>	<b>g.l.</b>	<b>Probabilidad</b>
12,77298	8	0,1118

En los cálculos obtenidos con el programa FSTAT V2.9.1 y GENPOP 3.1 no se encontraron diferencias ostensibles en los valores de diversidades genéticas para las inserciones al evaluarlas por locus y población (Tabla 4). Nótese que entre las tres poblaciones el valor de diversidad genética entre las inserciones Ta es casi el mismo, lo que revela que las poblaciones son muy cercanas entre sí. Esto lo confirmaron los valores de  $G_{ST}$  obtenidos a partir de dichas diversidades con la estimación del método de Nei (1987). En la tabla 5 no se observan valores

significativos para una alta diferenciación genética entre las poblaciones al evaluarlas por loci ni para el conjunto de éstos. Valores de  $G_{ST}$  significativamente diferentes de cero corresponderían a una alta diferenciación genética entre las poblaciones, sin embargo, esto no ocurrió. Por lo tanto, las tres poblaciones no fueron significativamente diferentes entre ellas para los marcadores estudiados. Es decir, que las tres poblaciones estudiadas se comportan como si fuera una sola.

**Tabla 4. Diversidad Génica por locus y población**

Inserción	Cauca	Choco	Córdoba
Tau-1	0.453	0.446	0.332
Tau-3	0.326	0.382	0.467
Tau-9	0.000	0.000	0.000
Tau-10	0.000	0.000	0.000
Tau-11	0.440	0.452	0.325
Tau-22	0.514	0.474	ND

ND No hay datos para Tau-22 en esta población

**Tabla 5. Estimaciones de heterocigocidad de Nei**

	$H_o$	$H_s$	$H_t$	$D_{st}$	$D_{st}^*$	$H_t^*$	$G_{ST}$	$G_{ST}^*$	$G_{is}$
Tau-1	0,121	0,41	0,408	-0,002	-0,003	0,407	-0,005	-0,008	0,705
Tau-3	0,322	0,392	0,394	0,003	0,004	0,396	0,007	0,010	0,177
Tau-9	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Tau-10	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Tau-11	0,411	0,406	0,408	0,003	0,004	0,410	0,006	0,009	-0,013
Ta1-22	0,311	0,494	0,499	0,006	0,011	0,505	0,011	0,022	0,370
TOTAL	0,194	0,284	0,285	0,001	0,002	0,286	0,005	0,008	0,315

A partir de los estadísticos F (tabla 6), se encontró que no había estructuración poblacional significativa, pues los valores de  $F_{ST}$  estuvieron muy cercanos a 0, lo cual indica que las tres poblaciones son prácticamente iguales para los marcadores Ta estudiados.

El valor  $F_{ST}$  para Tau-1 presentó valores negativos lo cual, conceptualmente no es posible, pues este estadístico solo oscila entre valores de cero a uno. Sin embargo, al ser la corrección por tamaño muestral y por el número de poblaciones con el método de Weir & Cockerham (1984) mayor que el propio valor de la  $F_{ST}$  obtenida, eso induce a que el valor de este estadístico sea negativo. Sin embargo, podemos interpretar ese valor como igual a cero.

La  $F_{IS}$  tuvo valores positivos, revelando un déficit general de heterocigotos a nivel de las subpoblaciones especialmente en los marcadores Tau-1 y Ta1-22.  $F_{IS}$  presentó valores negativos en Tau-11 demostrando así un ligero exceso de heterocigotos, sin embargo ese valor no alcanza a ser significativo.

El estadístico  $F_{IT}$  mostró una tendencia similar a lo observado para  $F_{IS}$  solo que a nivel de individuos con respecto a la población total.

**Tabla 6. Cálculo de los estadísticos F**

Estadísticos F				95% intervalo de confianza con el método "bootstrap".		
Locus	$F_{IT}$	$F_{ST}$	$F_{IS}$	$F_{IT}$	$F_{ST}$	$F_{IS}$
Tau-1	0.699	-0.009	0.701			
Tau-3	0.181	0.010	0.173			
Tau-11	-0.004	0.009	-0.013	0.056	-0.004	0.047
Ta1-22	0.373	0.022	0.359	0.575	0.020	0.575
<b>Total loci</b>	0.317	0.009	0.311			
<b>corte sobre el total de loci con el método Jackknife</b>				<b>99% intervalo de confianza.</b>		
	$F_{IT}$	$F_{ST}$	$F_{IS}$	$F_{IT}$	$F_{ST}$	$F_{IS}$
<b>Total</b>	+ 0.319	+ 0.009	+ 0.313	-0.004	-0.009	-0.013
	- 0.149	- 0.007	- 0.153	0.699	0.022	0.701

En las pruebas de significancia para cada una de las F (Tabla 7) los valores significativos es decir, por debajo de 0.05, para la  $F_{IS}$  revelaron una desviación del equilibrio Hardy-Weinberg en los marcadores Tau-1 y Ta1-22 y por ende en el total de los loci de los individuos con respecto a las sub poblaciones.

En el caso de los valores de  $F_{IT}$ , se encontró significancia y por ende una desviación del equilibrio en los marcadores Tau-1, Tau-3 y Ta1-22 así como también en el total de los loci, estos resultados son para los individuos con respecto a la población total.

Finalmente los valores de significancia para las  $F_{ST}$  asumiendo o no un equilibrio Hardy-Weinberg, no fueron en ningún caso significativos y su proximidad a cero constata que las tres poblaciones se encuentran estrechamente relacionadas y que esta condición no cambiaría así todos los marcadores se encontraran en equilibrio.

**Tabla 7. Prueba de significancia de los estadísticos F**  
valores de significancia

Marcador	$F_{IS}$	$F_{IT}$	$F_{ST}$ Asumiendo equilibrio H-W	$F_{ST}$ Sin asumir equilibrio H-W
Tau-1	<u>0,00010</u>	<u>0,00010</u>	0,31560 - 0,30730	0,48400 - 0,48290
Tau-3	0,09530 - 0,03270	<u>0,03910 - 0,03870</u>	0,18980 - 0,18030	0,24180 - 0,24130
Tau-11	0,46514 - 0,43990	0,46540 - 0,45250	0,22350 - 0,19570	0,22530 - 0,19610
Ta1-22	<u>0,00820 - 0,00170</u>	<u>0,00170 - 0,00150</u>	0,12690 - 0,08940	0,39700 - 0,39700
Total loci	<u>0,00010</u>	<u>0,00010</u>	0,11480	0,34510

El calculo del flujo génico arrojó valores infinitos es decir, que el flujo génico entre las tres poblaciones fue muy alto.

Las distancias genéticas (ver tabla 8) tanto para el método de Nei como para el de Cavalli-Sforza & Edwards revelaron que Chocó y Cauca estaban muy relacionadas entre ellas y que de las dos la más cercana a Córdoba fue Chocó. Sin embargo, en general esos valores de distancias genéticas tan bajos mostraron que, globalmente, las tres poblaciones estaban estrechamente relacionadas desde la perspectiva genética analizada. Sus diferencias fueron mínimas.

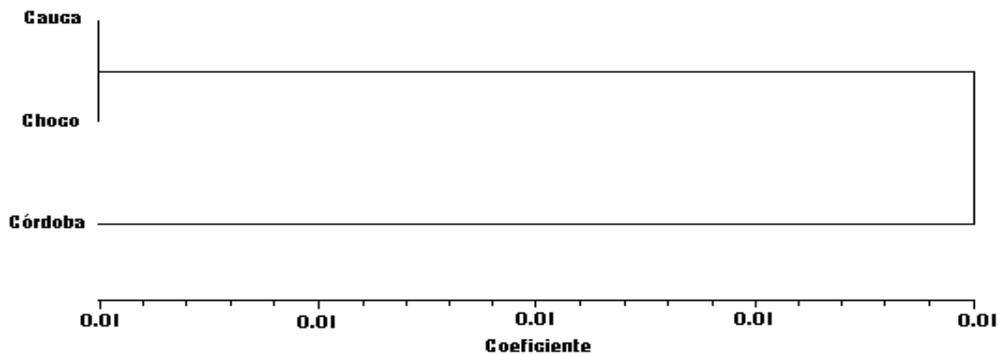
Todos los árboles obtenidos tanto con el algoritmo UPGMA como con el algoritmo NJ para ambas distancias (Fig.1) mostraron clara y contundentemente dichas relaciones para las tres poblaciones. Por lo tanto, pese al fenotipo “blanco” de la población muestreada en Córdoba el aporte Afro colombiano de ésta parece sustancialmente alto.

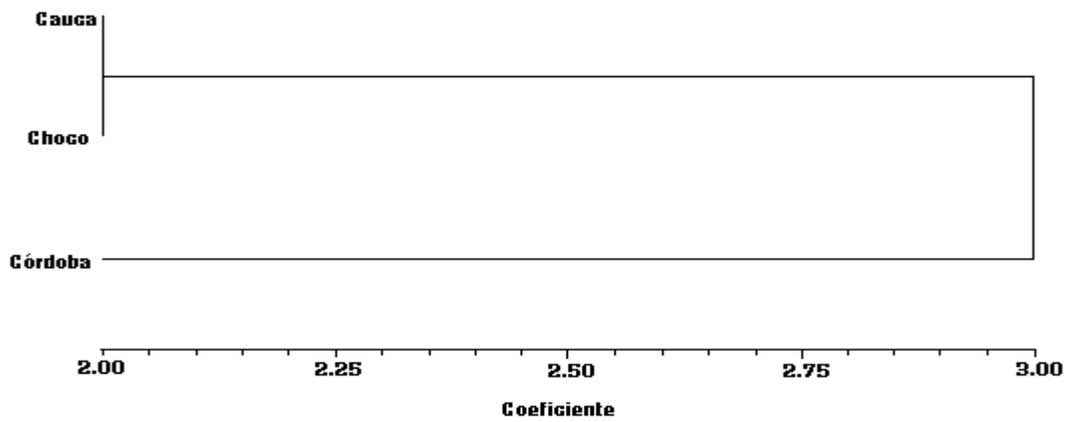
**Tabla 8. Distancias genéticas entre las tres poblaciones estudiadas**

	Distancias Genéticas Nei (1972)	Distancias genéticas Cavalli-Sforza & Edwards
Cauca-Choco	6.07433766457479E-003	5.52656877643258E-002
Cauca-Córdoba	1.38101690277067E-002	4.13740250135193E+000
Choco-Córdoba	1.06524939159530E-002	4.10559110633792E+000

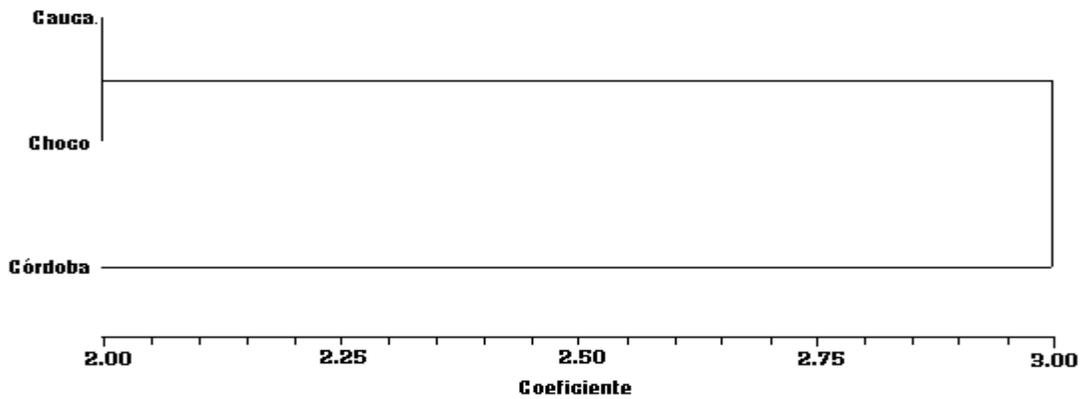
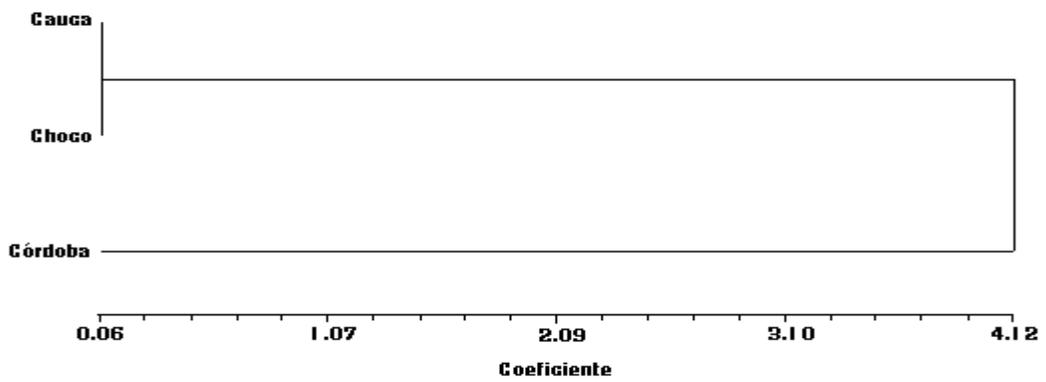
**Figura 2. Árboles de distancias genéticas**

Árboles obtenidos con las distancias de Nei (1972) usando los métodos UPGMA y Neighbor -joining respectivamente.





Árboles obtenidos de las distancias de Cavalli-Sforza & Edwards con los métodos UPGMA y Neighbor-Joining respectivamente



Como todos los árboles resultaron idénticos, no se consideró necesario obtener árboles consensos de los mismos.

Los análisis de coordenadas principales, para los dos tipos de distancias utilizadas, pueden verse en la tabla 9. La tabla 9a muestra que el 89% de las relaciones entre las poblaciones eran explicadas principalmente por la población de Cauca y el 10.9% por Chocó, mientras que Córdoba no participo significativamente en esos porcentajes.

Para el caso de las distancias de Cavalli-Sforza & Edwards (Tabla 9b) prácticamente el 100% de las relaciones las explicaba la población de Cauca, apenas un 0.009% correspondía a Chocó. Esto significa que Chocó y Córdoba están contenidas dentro de la población de Cauca.

**Tabla 9. Análisis de coordenadas principales**

9A) Análisis con Nei 1972

	valor propio	Porcentaje	% acumulativo
Cauca	0,00010125	89,0553	89,0553
Choco	0,00001244	10,9447	100,000
Córdoba	0,00000000	0,00000000	100,000

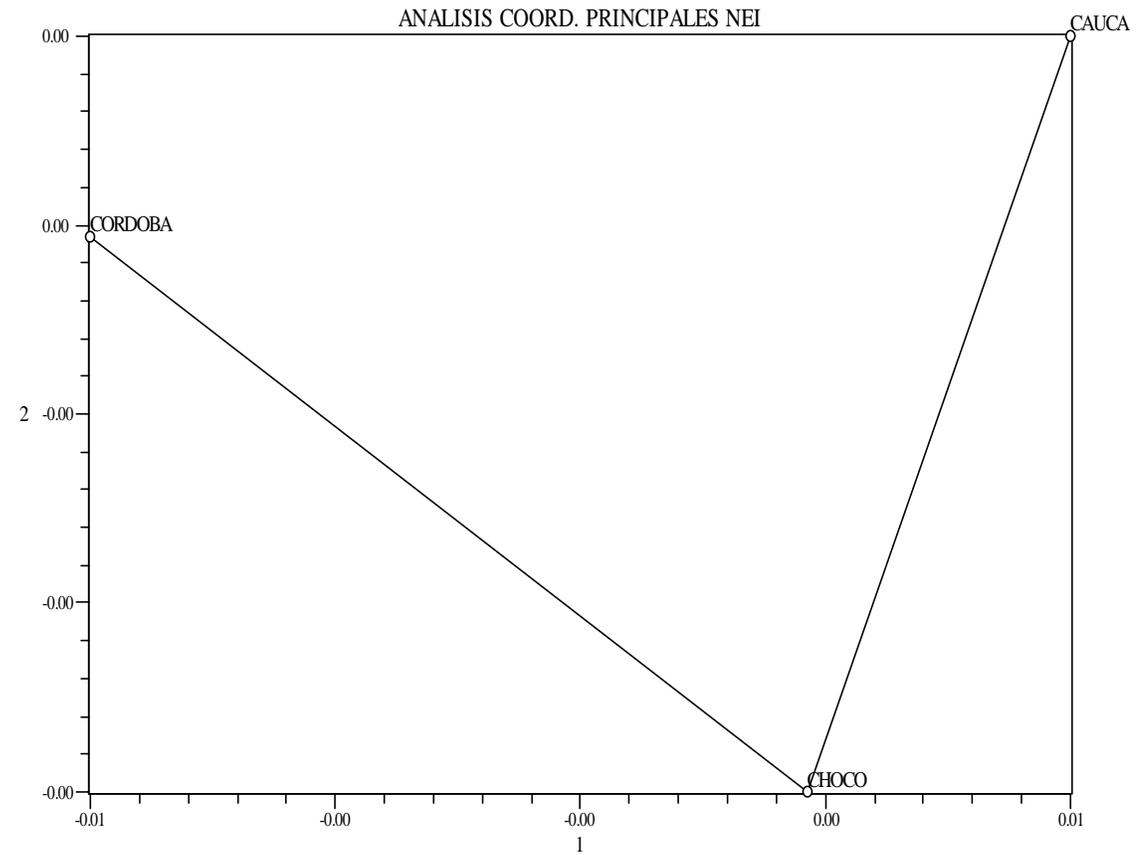
9B) Análisis con Cavalli-Sforza & Edwards

	valor propio	porcentaje	% acumulativo
Cauca	11,3246563	99,991	99,991
Choco	0,0010211	0,009	100,000
Córdoba	0	0	100,000

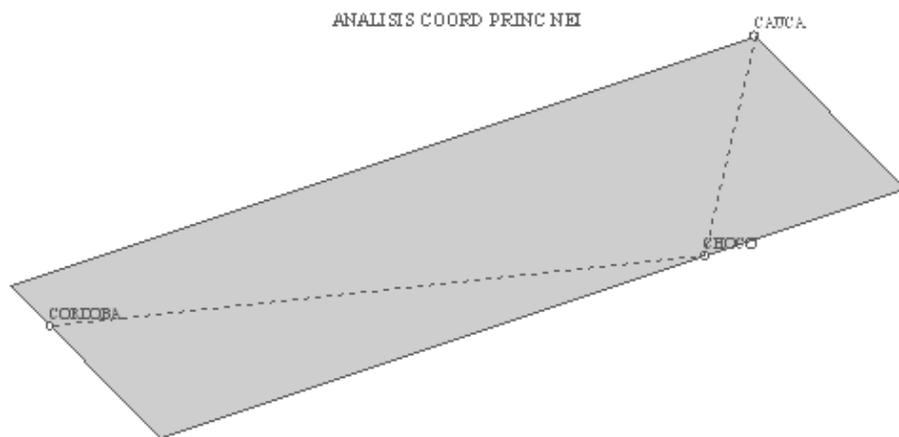
Al sobreponer los árboles minimum spanning al espacio multidimensional se obtuvieron los gráficos que se muestran en las figuras 3 A y B y en la figura 4. Todos estos gráficos muestran una relación conspicua entre Cauca y Chocó, las dos poblaciones afrocolombianas, y como la población mas cercana a Córdoba en este caso se encontró Chocó. No obstante, la estrecha relación fue más marcada para la distancia de Cavalli-Sforza & Edwards (1967) que para la de Nei (1972).

**Figura 3. Análisis de coordenadas principales para las distancias de Nei (1972)**

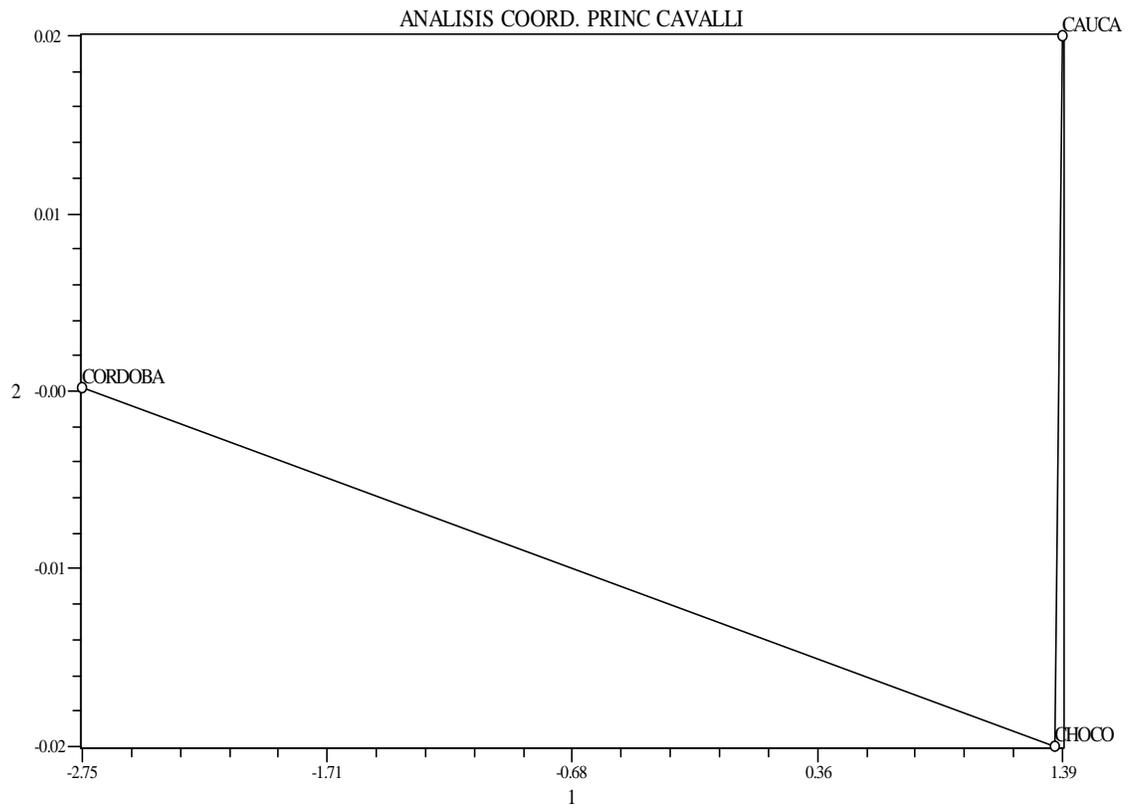
3 A) En dos dimensiones



3B) En tres dimensiones



**Figura 4. Análisis de coordenadas principales para las distancias de Cavalli-Sforza & Edwards (1967) en dos dimensiones**



## 8. DISCUSIÓN DE RESULTADOS

### Frecuencias alélicas

Uno de los resultados más llamativos fue que las inserciones retrotransposónicas estuvieron ausentes para los marcadores Tau-9 y Tau-10 en las tres poblaciones colombianas analizadas, entendiéndose que la presencia de la inserción es simplemente un arreglo o alteración dentro del Line-1 muestreado. Boissinot et al. (2000). En este momento es imposible saber si otras poblaciones latinoamericanas son portadoras o no de esta inserción ya que las poblaciones que aquí se reportan son las primeras analizadas en Latinoamérica para este tipo de marcadores moleculares.

Los mismos autores encontraron una correlación entre la edad del retrotransposón y su polimorfismo. Cuanto más jóvenes eran, como los pertenecientes a la subfamilia Ta1 y específicamente los del subgrupo Ta1d, había más probabilidad de encontrar polimorfismos para la presencia o ausencia de la inserción comparado con los loci que contienen inserciones viejas como la mayoría de los Ta-0. Sin embargo, dentro de ambas subfamilias existen inserciones sin polimorfismos en los subgrupos más antiguos de cada una e incluso, los mismos autores reportan que el estado ancestral de las inserciones Ta es la ausencia del arreglo.

Entonces es probable que Tau-9 y Tau-10 pertenezcan a cualquiera de los subgrupos antiguos contenidos ya sea en Ta-0 o en Ta-1 y por esta razón no presenten el arreglo siendo así esta la condición ancestral.

Teniendo en cuenta que las tres muestras tienen un gran aporte africano en los genomas de los individuos estudiados, especialmente los de Chocó y Cauca y que en la historia humana las poblaciones africanas son consideradas como las raíces más antiguas (esto no solo ha sido soportado con datos morfológicos sino también con datos moleculares ver Batzer et al. (1994) quienes encontraron un origen africano a las inserciones polimorfitas Alu, pertenecientes al grupo de repeticiones SINES muy cercanas a las del tipo LINE. Igualmente Chen et al. (2000) encontraron

haplotipos mitocondriales en la población africana Kung que se posicionaban en la raíz de la filogenia del mtDNA humano sugiriendo también que esta población era una de las más antiguas. Corroborando estos hechos Jorde et al. (2000) soportaron el origen africano de las poblaciones humanas con datos mitocondriales, autosómicos y ligados al cromosoma Y. No sería de extrañar que el estado monomórfico para la ausencia de la inserción en Tau-9 y Tau-10 sea un reflejo de la antigüedad del componente africano en las poblaciones colombianas estudiadas y por ende carezcan de polimorfismos en esos marcadores ya que el estado ancestral para los retrotransposones Ta es la ausencia de la inserción. Boissinot et al. (2000).

Adicionalmente pudiera ser que las frecuencias de las inserciones dentro de las poblaciones sean demasiado bajas y por tanto la probabilidad de encontrarlas en un muestreo aleatorio de 90 individuos sea mínima. Sin embargo, los autores de los marcadores habían hecho pruebas para ambos marcadores y encontraron el estado polimórfico fácilmente entre pocos individuos en la población “blanca” norteamericana. (comunicación personal).

#### Equilibrio Hardy-Weinberg

Dos de los siete loci muestreados no se encontraron en equilibrio Hardy Weinberg; esto pudo ocurrir por dos eventos básicos: selección natural o efecto Wahlund, ya que estos dos eventos no afectan por igual todo el genoma, como lo hace la endogamia, sino únicamente a partes puntuales de él. Además, ambos pueden producir un aumento de homocigotos y una disminución de heterocigotos lo cual, está en concordancia con los resultados del test multilocus que mostró que el no equilibrio Hardy Weinberg se debió a un déficit de heterocigotos.

La selección natural puede estar actuando a su vez de dos formas diferentes: primero podría estar actuando directamente sobre el retrotransposón estudiado, o segundo podría estar actuando sobre un gen ligado al retrotransposón. Se descarta que esto último se deba a un desequilibrio gamético con el resto de retrotransposones muestreados debido a que el test para desequilibrio gamético no

mostró resultados significativos para aceptar un desequilibrio de tal naturaleza; es decir que los retrotransposones estudiados segregan independientemente entre sí.

Boissinot et al. (2000) encontraron tasas de acumulación por generación en humanos de estos elementos, de 0.004 por generación, cifra que revela tanto la transposición como su eliminación (por selección –en caso de inserciones deletéreas o letales- o por deriva genética).

Sin embargo, es probable que la razón por la que no existiera equilibrio Hardy-Weinberg en los dos loci sea porque se presentó un cierto efecto Wahlund (1928), pues cuando dos acervos genéticos, que se diferencian ligeramente en sus frecuencias alélicas, se mezclan se puede producir un aumento de homocigotos y una disminución de heterocigotos en algunos loci específicos, pero no en todo el genoma (solo en aquellos loci donde se presentan diferencias en sus frecuencias alélicas).

Sin embargo, existen algunos elementos que pueden poner en duda la importancia del efecto Wahlund como el causante de esta desviación. En primer lugar las diferencias genéticas entre las poblaciones analizadas fueron muy pequeñas tal como lo reflejaron los valores de  $G_{ST}$ , de las  $F$  de Wright (1951) y de las distancias genéticas.

En segundo lugar, el efecto wahlund podría ser aun menos parsimonioso dado el alto flujo génico que han tenido las poblaciones afrocolombianas a lo largo de su historia no solo las analizadas en el presente estudio sino en otros estudios moleculares donde las actuales poblaciones afrocolombianas presentaron una similitud genética extraordinariamente importante pese a las distancias geográficas que las separan. Jaramillo et al. (2001).

Solo entre 1988 y 1993 el departamento de Córdoba tuvo 39.302 inmigrantes y 76.215 emigrantes, Chocó 13.386 inmigrantes y 27.524 emigrantes y Cauca 42.899 inmigrantes y 74.407 emigrantes según el DANE. Desafortunadamente esta entidad no tiene datos discriminados que permitan conocer cuantos de todos éstos

corresponden a afrocolombianos. Sin embargo, muchas de estas personas que se movilizaron entre los departamentos se establecieron y formaron familias con personas residentes según el propio DANE. Por tanto, la no existencia de equilibrio Hardy-Weinberg por efecto Wahlund no parece ser la causa principal de este fenómeno. Sin embargo no se puede descartar totalmente ya que siempre existe la posibilidad de micro estructuras geográficas por componentes familiares que no son recogidos en las muestras. Igualmente, Durán & Ruiz García (2001) encontraron marcadores VNTRs en la población humana de Bogotá que tampoco estuvieron en equilibrio y Novick et al. (1998) tampoco encontraron equilibrio en algunas poblaciones amerindias colombianas para marcadores retrotransposones Alu.

Esto significa que los eventos selectivos pueden estar jugando un papel determinante en los sesgos observados respecto al equilibrio Hardy-Weinberg

Es interesante la correspondencia de los datos moleculares con los eventos sociales que se presentan en el país reafirmando que la genética de poblaciones puede ser una herramienta de mucha ayuda para esclarecer los vacíos informativos en la historia de la población colombiana.

#### Heterogeneidad y estructura poblacional

La baja heterogeneidad encontrada usando tanto cadenas de Markov como estimaciones del análisis genético de Nei (1973) fue ocasionada por el alto flujo génico que se presentó entre las poblaciones muestreadas. Sin embargo, para tres departamentos, cuya población fue originada por el mestizaje de diferentes grupos indígenas y españoles con esas poblaciones africanas, no se esperarían valores tan bajos de  $G_{ST}$ . Esto está corroborando que los afrocolombianos tienen un aporte africano muy grande, con relativa poca mezcla con otras poblaciones no africanas, el cual a su vez no provendría de poblaciones africanas muy diferentes entre sí, haciendo pensar que los actuales afrocolombianos del Cauca y del Chocó son descendientes de más bien pocas poblaciones africanas originales.

Esto a su vez, se ve complementado con el trabajo de Jaramillo et al. (2001) donde se presentaban diferencias claras al interior de poblaciones amerindias y de éstas con las poblaciones afrocolombianas, pero al interior de las poblaciones negras no se encontraron diferencias marcadas entre unas y otras, sugiriendo que el aporte africano provendría de unas pocas poblaciones de África.

El presente estudio corrobora efectivamente esto. Historiadores y antropólogos colombianos han sugerido este fenómeno desde hace tiempo Mina 1975 por ejemplo, pensaba que la mayor parte de los esclavos provenían de la costa occidental de África concretamente de Senegal, Angola y Guinea (la mayoría de esta última) y pertenecían a comunidades específicas, argumentando también que estos esclavos se mezclaban entre sí en los puertos y en los barcos antes de arribar a América.

En cuanto a la estructura poblacional encontrada con los estadísticos  $F$  se muestra que tal estructura prácticamente no existe entre las tres muestras analizadas, valores de  $F_{ST}$  tan cercanos a cero no muestran una subdivisión poblacional importante, no hay divergencia genética entre las poblaciones, es como si todos los individuos muestreados formaran una única población.

Lo anterior se podría deber también, a que los lugares de procedencia de los africanos fueran muy cercanos originalmente, a un alto flujo génico previo a la llegada a Colombia de estos individuos y a procesos migratorios que se han llevado a cabo en Colombia que han permitido una mayor homogeneización todavía.

$F_{IT}$  y  $F_{IS}$  por su parte corroboraron con sus valores positivos un déficit de heterocigotos. Esos resultados podrían ser explicados hipotéticamente por endogamia, pero este es un proceso que afecta, en principio, a todo el genoma, por lo que se debería esperar que todos los marcadores analizados presentaran el exceso de homocigotos. Eso volvería a potenciar la posibilidad de fenómenos selectivos afectando a esos marcadores o de estructuras microgeográficas, que a su vez sería incompatible con el elevado flujo génico encontrado.

Por otra parte, la similitud de la población “blanca” de Córdoba con las dos poblaciones afrocolombianas, mostró que esta última tiene un gran aporte africano pese a su fenotipo. Podría haberse dado un alto flujo génico entre “blancos y “negros” desde la colonia española. Mas aun, existen reportes de los puertos usados para el contrabando de esclavos como Santa Marta y Riohacha incorporando africanos hacia y desde la costa atlántica. Gutiérrez (1980). Esto se une a las importaciones legales de esclavos desde Cartagena. Muchos de estos africanos llegaron a trabajar en las minas, incluso, en las de Córdoba de las que se extraía cobalto, hierro y calizas.

Una idea de la importancia del proceso de mestizaje en el atlántico especialmente para Cartagena fue descrito por Gutiérrez (1980), donde se cuenta como la población negra era rechazada socialmente y como los mulatos preferían establecer sus familias con blancos para ser más aceptados socialmente y después de tres a cuatro generaciones los descendientes eran tan blancos como los españoles. Pero en sus genes siempre quedaría el registro de su historia.

No obstante, otra explicación alternativa es un importante aporte “blanco” en las poblaciones afrocolombianas estudiadas, sabemos por otros estudios moleculares (Jaramillo et al. 2001) que ciertos marcadores típicamente amerindios se han encontrado en la población negra del chocó en unas frecuencias relativamente elevadas. Eso significaría que ha existido flujo génico entre ambas comunidades. Es muy posible que las comunidades “blancas” y amerindias de las áreas implicadas, también se hayan mezclado intensamente. Por lo tanto, la similitud entre esas poblaciones africanas y mestizas podría darse por tener un origen amerindio compartido.

#### Distancias genéticas

Las distancias genéticas encontradas entre las tres poblaciones pueden ser justificadas por el alto flujo génico que presentan entre ellas, desde la época colonial hasta hoy. Otra posibilidad es que los procesos selectivos sean unificantes en todas las poblaciones y la similitud sea simplemente un fenómeno de convergencia

evolutiva. Sin embargo la homogeneización por flujo genético parece muy factible. Por ejemplo, Jaramillo et al. (2001) mostraron para los marcadores APOE, ACE, APOB (VNTRs) una fuerte asociación entre algunas poblaciones afrocolombianas y una muestra de la población de Bogotá. Esto pone en evidencia que seguramente ha existido una fuerte intromisión “blanca” en el acervo genético propiamente de origen africano.

Según el DANE la mayoría de emigrantes lo hacen desde las zonas rurales a las capitales de departamento o a las ciudades mas importantes del país convirtiéndose éstas en los centros mas importantes de intercambios étnicos. Los inmigrantes se mantienen constantes en esas poblaciones ya que acceden mas fácilmente al trabajo, educación y salud. Esto explica, por ejemplo, lo reportado por Vargas (1999) al presentar a Quibdó como la población con mayor mestizaje de todo el chocó.

Por otro lado, los dos tipos de distancias genéticas utilizadas y sus respectivos árboles obtenidos con los algoritmos UPGMA y Nj llegaron al mismo resultado a pesar de que sus bases matemáticas son muy diferentes. Esto da entonces, más credibilidad a las relaciones encontradas. Ambas fueron muy útiles para establecer las relaciones teniendo en cuenta que para los otros análisis las diferencias entre las tres poblaciones fueron casi nulas. Además de ello, corroboraron la utilidad sugerida por Boissinot et al. (2000) de los retrotransposones Ta a la hora de establecer relaciones entre poblaciones.

#### Análisis de coordenadas principales

Se esperaba que la población muestreada del Chocó, por ser el departamento que históricamente ha contenido desde finales del siglo XVII mas afrocolombianos Maya (1998) , explicara la mayor parte de las distancias genéticas encontradas en el análisis de coordenadas principales entre las tres poblaciones estudiadas. Sin embargo, esto no ocurrió.

Fue la población del Cauca quién respondió por más del 90% de las relaciones genéticas (tabla 9). Esto no quiere decir que históricamente Cauca haya sido el foco

mas importante de distribución de afrocolombianos. Puede ocurrir que la muestra obtenida del Cauca fue menos híbrida que la del Chocó. Al ser la población del Chocó (muestreada en Quibdó) mas híbrida que la del Cauca juega un papel intermedio mas destacado entre esta última y la población “blanca” de Córdoba. Si la muestra para choco se tomara en las zonas rurales o en cabeceras municipales de poca importancia comercial se esperaría que esta población explicara en un mayor o equivalente porcentaje las relaciones entre los grupos afrocolombianos de lo que explicó Cauca. El hecho de que esta última población fuera la que mayor varianza explicara es fácilmente interpretable partiendo de la base que esta población tiene un acervo africano más “puro” que la del Chocó.

Antes de justificar el resultado hay dos eventos que analizar: en primer lugar la ubicación histórica de los esclavos legales durante la colonia. Estos llegaban a Cartagena y de allí se distribuían a diferentes terminales; hacia el atlántico directamente por Cartagena; hacia el interior a Santa Fe de Antioquia, Honda (Tolima), Anzerma (Caldas) y hacia el Cauca llegaban a Popayán y Zaragoza y desde estos lugares la mayoría iban a parar a Venezuela, Ecuador y Perú y unos pocos eran distribuidos en Colombia a las minas principalmente del Pacifico. Gutiérrez (1980).

Zaragoza y Popayán se encontraban aislados geográficamente dentro del departamento de Cauca, pues a Popayán solo se ingresaba desde Buenaventura Guhl (1976). Por tanto, la presencia de afrocolombianos en la costa pacifica de cauca se debió especialmente al aporte realizado desde Zaragoza.

En segundo lugar, la consolidación más importante de la distribución de las comunidades negras se dió desde los años cincuenta. De acuerdo con Pardo (1996), durante este tiempo las poblaciones afrocolombianas del Pacifico vivieron un proceso intenso de desplazamientos y urbanización por el advenimiento de la demanda maderera para toda esta zona. Muchos migraron por la oferta de trabajo y se concentraron en tres ciudades principales Quibdó (Chocó), Buenaventura (Valle del Cauca) y Tumaco (Nariño), presentando estas tres cabeceras un mayor grado de mestizaje con “blancos” y amerindios.

Actualmente de 800.000 personas que habitan la costa Pacífica, aproximadamente la mitad viven en estos tres municipios; el 47% de los restantes viven en zonas rurales y el 53% lo hace en cabeceras municipales solo el 10% es mestizo con blanco y vive en poblados mayores y los indígenas que corresponden a un 4% solo viven en áreas rurales. En Cauca la población afrocolombiana con respecto a todo el departamento es más baja y se concentra en la costa principalmente en tres municipios Guapi, López y Timbiquí (Guhl 1976). Estos tres municipios son cercanos a Zaragoza.

Las dos anteriores características históricas de distribución llevarían a pensar que la población afrocolombiana muestreada en Cauca correspondería especialmente al porcentaje de la población negra del Pacífico que habita en zonas rurales con eventos migratorios no tan importantes, ya que estas zonas no han sido tan atractiva para los inmigrantes históricamente, como lo son las tres ciudades de mayor concentración de negros.

El grado de mestizaje de Cauca debe ser más bajo que el de la población muestreada de Quibdó (Chocó), que como ya se dijo más arriba es considerada una de las más mestizas. Además, su componente africano debe estar más relacionado con el que se forjó durante la distribución colonial de negros y por ello ocupa la posición más extrema en el análisis de coordenadas principales y es la población que en su eje explica la mayor parte de varianza.

En este punto se esperaría que el componente africano fuera el mismo al de la población de Cauca, pero en las poblaciones rurales de Chocó, debido a que éstas han recibido un menor aporte amerindio o blanco, como lo sugiere Wade (1996), quién describe como las relaciones interétnicas de afrocolombianos e indígenas se limitan al compadrazgo (padrinaje de los hijos) y rara vez se presentan matrimonios entre ellos. Contrariamente, en las capitales de departamento la mezcla entre afrocolombiano y mestizo blanco es más común.

Ese es el motivo principal por el que Cauca aparecería conteniendo a las otras dos poblaciones por presentar un mayor componente africano, y Quibdó por su mestizaje sería la más cercana a la población “blanca” de Córdoba. No obstante, las tres se encontrarían altamente relacionadas debido a que el componente africano es predominante en todas.

Al observar los gráficos de los análisis de coordenadas (figuras 3 y 4) se observa que al usar las distancias de Nei (1972) las diferencias entre las poblaciones se ven más claramente que con los análisis utilizando la distancia de Cavalli-Sforza & Edwards (1967). Sin embargo, esas diferencias son sutiles y no muestran el grado de diferencias entre ellas como lo encontrado en otros estudios con otros organismos. Ruiz García (1994), Ruiz García et al. (2002).

Sin embargo, la distancia de Nei (1972) ha demostrado ser muy útil para medir polimorfismos intraespecíficos. Además, está diseñada para estudios de evolución, ya que miden la sustitución codónica con lo que puede estar relacionada con el tiempo de divergencia evolutivo. Nei (1987).

## 9. CONCLUSIONES

No se encontraron diferencias significativas entre las tres poblaciones colombianas estudiadas, pese a que Cauca y Chocó correspondieron a individuos fenotípicamente afrocolombianos y Córdoba a individuos fenotípicamente “blancos” o ligeramente mestizos, lo cual permitió determinar que el aporte africano de para esta población es notablemente alta. Una explicación alternativa es que existan fuerzas selectivas constrictivas que hayan llevado a las poblaciones a un estado frecuencial similar.

Al formularse el problema de investigación se esperaba que por el hecho de que en cada una de las poblaciones, había un aporte indígena, caucásico y africano diferente, cada una iba a tener un componente genético intrínseco diferente. Al no ocurrir esto, se concluye que el principal componente de las tres es el africano y que el flujo génico ha sido considerado elevado entre ellas. Sería muy importante poder incluir poblaciones amerindias con este tipo de marcadores en futuros trabajos para establecer si existe algún grado de mezcla entre estas y las poblaciones afrocolombianas, pues en otros estudios como los Novick et al. (1998) un componente amerindio permitía establecer diferencias más claras al estar relacionado con orígenes asiáticos.

En cuanto a los esclavos Africanos que llegaron a Colombia se encontró que estos provenían de pocas poblaciones africanas entre las cuales debió existir un alto flujo génico previo y que este se continuo en el territorio colombiano desde la colonia hasta hoy en día. Existen dos premisas que sostienen esta afirmación: a) Los valores de  $G_{ST}$ ,  $F_{ST}$  y las distancias genéticas fueron muy pequeñas b) la diversidad génica encontrada (heterocigocis esperada) es relativamente pequeña también, lo cual indica que el número de acervos genéticos diferentes que dieron lugar a esas actuales poblaciones afrocolombianas es limitado.

En cuanto a los eventos poblacionales internos, se encontró que predomina un altísimo flujo génico entre las tres poblaciones, por lo cual estas no están

presentando procesos de aislamiento o endogamia, o bien éstos están mitigados por fenómenos selectivos constrictivos y unificantes.

Las relaciones genéticas encontradas ratifican muchas circunstancias de la documentación histórica obtenida para dichas poblaciones, corroborándose la utilidad de los estudios genéticos en el esclarecimiento de los vacíos informativos de la historia de Colombia, que por falta de documentación los historiadores no tienen forma de suplir.

La hipótesis de la que partió el trabajo fue rechazada porque las poblaciones mostraron relaciones muy estrechas entre sus frecuencias alélicas para los retrotransposones y se esperaban diferencias mas considerables. Sin embargo, los retrotransposones Ta si son útiles para comparar poblaciones desde una perspectiva filogenética ya que son eventos únicos donde la posibilidad de homoplasias es prácticamente nula.

Finalmente, el objetivo general fue cumplido, pues se obtuvo la distribución de las frecuencias alélicas y éstas permitieron realizar las comparaciones y los análisis estadísticos correspondientes. Además, se encontraron justificaciones históricas como se pretendía desde un inicio.

## 10. RECOMENDACIONES

Novick et al. (1998) utilizando inserciones Alu establecieron que las poblaciones americanas (amerindias) manifestaban un origen asiático, pero que las africanas se ubicaban como las más antiguas de todas. Pese al mestizaje, ocurrido en Colombia, al muestrear las poblaciones amerindias y afrocolombianas por otros marcadores moleculares se diferencian claramente ambas etnias. El trabajo de Jaramillo et al. (2001), posiblemente por la misma razón, muestra que no existen relaciones claras y conspicuamente diferenciadas entre las diferentes poblaciones afrocolombianas ya que todas ellas provendrían de unos pocos acervos genéticos, africano mezclado entre ellos.

Así pues, para distinguir más claramente las relaciones poblacionales en Colombia y tratar de esclarecer el proceso de mestizaje. Se muestra que son útiles los retrotransposones Line-1 Ta como lo fueron previamente los Alu, pero se sugiere muestrear poblaciones rurales tanto de afrocolombianos como de amerindios, por tener menores tasas de inmigración que las poblaciones citadinas (DANE 2000).

Estos datos, comparados con poblaciones de Capitales de departamento cuyas tasas de inmigración son más grandes, podrían arrojar resultados que no solo muestren las relaciones filogenéticas entre las poblaciones sino que además, permitan inferir el aporte amerindio y afrocolombiano en los actuales focos de mestizaje, pues sabiendo de qué zonas del país provienen los principales aportes genéticos, se puede llenar el vacío informativo de los procesos migratorios en Colombia que presenta el DANE para antes de 1988.

Así mismo, se sugiere aplicar los marcadores Ta en las poblaciones indígenas colombianas, para determinar si, en éstas, las relaciones filogenéticas para los marcadores son más claras que lo encontrado para poblaciones afrocolombianas; y con ello soportar aun más los datos encontrados para otros marcadores, donde las relaciones entre poblaciones amerindias son más claras que las de afrocolombianos, o si por el contrario se podrían rebatir dichos datos.

Al momento de escoger las inserciones Ta, son mas útiles las pertenecientes a la subfamilia Ta-1, por ser mas polimorficas Boissinot et al. (2000) sin embargo usarlas simultáneamente con algunas inserciones Ta0 que son mas antiguas podrían establecer mejor las relaciones filogenéticas entre poblaciones.

No solo se deberían tener en cuenta los asentamientos de la población sino las áreas de desplazamiento de sus individuos. Para ello es útil referirse a datos interesantes encontrados por el DANE (2000) para las modalidades migratorias y las distancias recorridas. Mas del 50% de personas migran de las zonas rurales hacia capitales de departamento o cabeceras municipales, existiendo una correlación con las distancias geográficas. En todos los casos, los movimientos hacia cabeceras municipales no sobrepasan un radio de 110 Km de distancia y la mayoría solo lo hacen a distancias de 56 kilómetros.

Por otra parte, mas del 90% de los movimientos migratorios entre municipios no sobrepasa los 330 kilometros de distancia. El resto de los emigrantes guarda una relación inversa con la distancia entre los municipios.

La migración también es selectiva. Las mujeres migran distancias más cortas a las recorridas por los hombres. En el futuro se podría investigar si se detectan diferencias en las relaciones entre poblaciones al utilizar frecuencias de los retrotransposones en virtud del sexo de sus poseedores.

Adicionalmente, lo anterior permitiría establecer las subpoblaciones de la población colombiana teniendo en cuenta las migraciones para hacer un muestreo más preciso dependiendo de las hipótesis que se quieran probar.

## 11. REFERENCIAS

Arocha Jaime. 1996. Afrogenesis, eurogenesis y saberes expertos. 283-298 en: Escobar, A. y A. Pedrosa (ed). Pacifico ¿Desarrollo o Diversidad? Ecofondo-Cerec. Bogotá.

Arocha Jaime. 1998. El pacifico colombiano diverso y plural. 15-41 en: Zambrano Pantoja Fabio (ed). Colombia país de regiones volumen 4. Conciencias-Cinep. Bogotá.

Batzer, M. A., Stoneking, M., Hartman, M. A., Bazan, H., Kass, D. H., Shaikh, T. H.,

Novick, G. E., Ioannou, P. A., Scheer, W. D., Herrera, R. J., & Deininger, P. L. 1994. African origin of human-specific polymorphic Alu insertions. Proc. Natl. Acad. Sci. 91:12288-12292.

Batzer, M., Stoneking, M., Alegria, M., Bazan, H., Kass, D., Shaikh, T., Novick, G., Ioannou, P., Scheer, D., Herrera, R., and Deininger, P. 1994. African origin of human-specific polymorphic Alu insertions. Evolution. 91:12288-12292.

Beltran, G., Hernandez, T. 1993. Cauca: características geográficas. Instituto geográfico Agustín Codazzi. Colombia.

Boissinot, S., Chevret, P., Furano, A. 2000. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. Mol. Biol. Evol. 17(6):915-928.

Bonatto, S. L. & Salzano, F. M. 1997. A single and early migration for the peopling of the americas supported by mitochondrial DNA sequence data. Proc. Natl. Acad. Sci. 94:1866-1871.

- Casavant, N., Lee, R. N., Sherman, A. N., & Wichman, H. A. 1998. Molecular evolution of two lineages of L1(LINE1) retrotransposons in the California mouse, *peromyscus californicus*. *Genetics* 150:345-357.
- Cavalli-Sforza, L.L. & Edwards. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 21:550-570.
- Chen, Y., Olckers, A., Schurr, T. G., Kogelnik, A. M., Huoponen, K. & Wallace, D. 2000. MtDNA variation in the south african kung and khwe-and their genetic relationships to other african populations. *Am. J. Hum. Genet.* 66:1362-1383.
- Cockerham, C. C. 1993. Estimation of gene flow from F-statistics. *Evolution* 47:855-863.
- Colmenares Germán. 1997. Cali: terratenientes mineros y comerciantes siglo XVIII. TM editores, Universidad del Valle, Banco de la república y Colciencias. Capitulo 3 elementos de las haciendas. Colombia.
- DANE. 2000. Las migraciones internas en Colombia, 1998-1993. editorial División ediciones del DANE. Volumen 13. Bogotá-Colombia.
- Dombroski, B. A., Scott, A. F., & Kazazian, H. H. 1993. Two additional potential retrotransposons isolated from a human L1subfamily that contains an active retrotransposable element. *Proc. Natl. Acad. Sci.* 90:6513-6517.
- Eline, T., Prak, L. & Kazazian, H. 2000. Mobile elements and the human genome. *Macmillian Magazines Ltd* 1:134-144.
- Furano Anthony V. 2000. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Molecular Biology.* 64:255-293.
- Furano, A., & Usdin, K. 1995. DNA fossils and phylogenetic analysis. *The Journal of Biological Chemistry.* 270(43):25301-25304.

Gower, J. C., Ross, G. J. 1969. Minimum spanning trees and single cluster analysis. *Appl Stat* 18:54-64.

Gutierrez Idelfonso. 1980. *Historia del negro en Colombia: ¿sumisión o rebeldía?*. Editorial Nueva América. Bogotá – Colombia.

Jaramillo Correa, J. P., Keyeux, G., Ruiz García, M., Rodas, C., Bernal, J. 2000. Population genetic analysis of the genes APOE, APOB (3'VNTR) and ACE in some black and Amerindian communities from Colombia. *Hum Hered* 52:14-33.

Jorde, L. B., Watkins, W. S., Bamshad, M. J., Dixon, M. E., Ricker, C. E., Seielstad, M. T., & Batzer, M. A. 2000. The distribution of human genetic diversity: a comparison of mitochondrial, autosomal and y-chromosome data. *Am. J. Hum. Genet.* 66:979-988.

Jurka, Jerzy. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci.* 94:1872-1877.

Klug, W. S. & Cummings, M. R. 1996. *Essentials of Genetics*. Second edition. Prentice Hall. New Jersey.

Maya Luz A. 1998. Demografía histórica de la trata por Cartagena 1533-1810. p 11-52. en: Maya Restrepo, L. A (ed). *Geografía humana de Colombia*. Tomo IV. Instituto Colombiano de Cultura Hispánica. Bogotá-Colombia.

Mina Mateo. 1975. *Esclavitud y libertad en el valle del río Cauca*. Publicaciones La Rosa. Capítulo dos: la esclavitud de los africanos. Colombia.

Nei, M. 1972. Genetic distance between populations. *Am Naturalist* 106:283-292.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Nei, Masatoshi. 1987. *Molecular Evolutionary Genetics*. Columbia university press. New York.

Novick, G., Novick, C., Yunis, J., Yunis, E., Antunez, P., Sceer, D., Deininger, P., Stoneking, M., York, D., Batzer, M., & Herrera, R. 1998. Polymorphic Alu insertions and the asian origin of native american population. *Human Biology*. 70(1):23-39.

Pardo Mauricio. 1996. Movimientos sociales y relaciones interétnicas. 298-315 en: Escobar, A. y A. Pedrosa (ed). *Pacífico ¿Desarrollo o Diversidad?* Ecofondo-Cerec. Bogotá.

Parker, P., Snow, A., Senog, M., Booton, C. & Fuerst, P. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* 2:361-382.

Reynolds, J. B., Cockerham, C.C. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105:767-779.

Saitou, N. Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.

Usdin, K., Chevret, P., Catzefflis, F., Verona, R., & Furano, A. 1995. L1 (LINE-1) retrotransposable elements provide a fossil record of the phylogenetic history of Murid rodents. *Mol. Biol. Evol.* 12(1):73-82.

Vargas Patricia. 1999. *Construcción territorial en el choco. Volumen I. Programa de historia local del ICAN y PNR*. Bogotá-Colombia.

Verneau, O., Catzefflis, F., & Furano, A. 1997. Determination of the evolutionary relationship in *Rattus sensu lato* (Rodentia: Muridae) using L1 (LINE-1) amplification events. *Journal of Molecular Evolution* 45:424-436.

Verneau, O., Catzeflis, F., & Furano, A. 1998. Determining and dating recent rodent speciation events by using L1(LINE-1) retrotransposons. Proc. Natl. Acad. Sci. 95:11284-11289.

Wade Peter. 1996. Identidad y etnicidad. 283-298 en: Escobar, A. y A. Pedrosa (ed). Pacifico ¿Desarrollo o Diversidad? Ecofondo-Cerec. Bogotá.

Weir, B.S., Cockerham C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

## 12. ANEXOS

Registro por población de cada individuo y su estado polimórfico o monomórfico para cada uno de los marcadores

### MARCADORES

INDIVIDUOS	CAUCA	Tau-1	Tau-3	Tau-9	Tau-10	Tau-11	Ta1-22
	BPL 161	AA	AA	aa	aa	Aa	AA
BPL 163	AA	AA	aa	aa	Aa	Aa	
BPL 166	aa	Aa	aa	aa	AA		
BPL 170	Aa	Aa	aa	aa	AA	aa	
BPL 173	Aa	Aa	aa	aa	Aa		
BPL 176	AA	AA	aa		aa	AA	
BPL 182	aa	AA		aa	AA	aa	
BPL 192	AA	AA	aa	aa	Aa	aa	
BPL 193	AA	Aa	aa		aa	Aa	
BPL 194	AA	AA		aa	aa	AA	
BPL 195	Aa	AA	aa	aa	AA	Aa	
BPL 196	aa	AA	aa	aa	Aa	Aa	
BPL 197	AA	AA	aa	aa	Aa	AA	
BPL 198	AA	AA	aa	aa	Aa	Aa	
BPL 199	AA	AA	aa		Aa	AA	
BPL 200	AA	Aa	aa	aa	Aa	Aa	
BPL 201	AA	AA			AA		
BPL 202	AA	AA	aa	aa	AA	aa	
BPL 203		Aa			AA	AA	
BPL 204	aa	AA	aa	aa	AA	AA	
BPL 205	AA	Aa		aa	AA	aa	
BPL 206	aa	AA	aa	aa	Aa	AA	
BPL 207	aa	Aa	aa		Aa	AA	
BPL 208	aa	AA	aa	aa	AA	aa	
BPL 209	Aa	aa	aa	aa	Aa	aa	
BPL 210	AA	AA	aa	aa	AA	aa	
BPL 211	Aa	AA	aa	aa	AA	AA	
BPL 212	AA	Aa	aa	aa	AA	AA	
BPL 213	AA	AA	aa	aa	Aa	aa	
BPL 214	AA	AA	aa	aa	AA	aa	

A = presencia de la inserción  
a = ausencia de la inserción

### MARCADORES

	CHOCO	Tau-1	Tau-3	Tau-9	Tau-10	Tau-11	Ta0-11
	P 139	Aa	AA	aa	aa	Aa	-
	P 140	aa	AA	aa		Aa	-
	P 141	Aa	AA	aa	aa	Aa	-
	P 142	aa	Aa	aa	aa	Aa	-
	P 143	AA	aa	aa	aa	AA	-
	P 144	AA	Aa	aa	aa	Aa	-
	P 145	AA	AA	aa	aa	AA	-
	P 146	aa	AA	aa		Aa	-
	P 147	AA	Aa		aa	Aa	-
	P 148	AA	Aa	aa	aa	AA	-
	P 149	AA	AA	aa	aa	Aa	-
	P 150	aa	Aa	aa	aa	Aa	-
INDIVIDUOS	P 151	AA	AA	aa	aa	Aa	-
	P 152	aa	Aa	aa		Aa	-
	P 153	aa	AA	aa	aa	AA	-
	M 154		aa	aa		AA	aa
	M 155	AA	AA	aa	aa	AA	AA
	M 156	Aa	Aa	aa		AA	aa
	M 157	AA	AA	aa	aa	AA	Aa
	M 158	AA	Aa	aa	aa	AA	aa
	M 159	AA	Aa	aa	aa	aa	AA
	M 160	aa	AA	aa	aa	Aa	aa
	M 161	AA	AA	aa	aa	aa	Aa
	M 162	AA	AA	aa	aa	aa	Aa
	M 163		Aa	aa	aa	Aa	Aa
	M 164		AA	aa	aa	AA	AA
	M 165	AA	AA	aa	aa	AA	AA
	M 166	AA	Aa		aa	AA	AA
	M 167	AA	AA	aa	aa	AA	aa
	M 168	AA	AA	aa	aa	Aa	aa

A = presencia de la inserción  
a = ausencia de la inserción

## MARCADORES

<b>CORDOBA</b>	<b>Tau-1</b>	<b>Tau-3</b>	<b>Tau-9</b>	<b>Tau-10</b>	<b>Tau-11</b>
CO 1	AA	aa	aa	aa	AA
CO 2	AA	AA	aa	aa	AA
CO 3		AA		aa	AA
CO 4	AA		aa	aa	AA
CO 5	AA	AA	aa	aa	aa
CO 6	AA	aa	aa	aa	Aa
CO 7	Aa	AA	aa	aa	AA
CO 8	aa	AA	aa	aa	AA
CO 9	AA	AA			AA
CO 10	Aa	Aa	aa	aa	AA
CO 11	aa	Aa	aa	aa	AA
CO 12	AA	Aa	aa	aa	AA
CO 13	AA	Aa	aa	aa	AA
CO 14	AA	Aa	aa	aa	Aa
CO 15	aa	Aa	aa	aa	Aa
CO 16		AA	aa	aa	Aa
CO 17		AA			AA
CO 18	AA	aa	aa	aa	Aa
CO 19	AA	Aa	aa	aa	AA
CO 20	AA	aa	aa	aa	Aa
CO 21	AA	Aa	aa	aa	AA
CO 22	AA	AA	aa	aa	Aa
CO 23			aa	aa	AA
CO 24	AA	AA	aa		AA
CO 25	AA	aa	aa	aa	AA
CO 26	AA	AA	aa	aa	Aa
CO 27	AA	AA	aa	aa	Aa
CO 28	aa		aa	aa	AA
CO 29	AA	AA	aa	aa	AA
CO 30		Aa		aa	Aa

A = presencia de la inserción  
a = ausencia de la inserción