

DIVERSIDAD GENÉTICA Y ANÁLISIS DE SUBESTRUCTURA POBLACIONAL
MEDIANTE 15 STR's AUTOSÓMICOS EN LA REGIÓN
DE TOLIMA-HUILA, COLOMBIA

INGRID GISELA PÉREZ PARADA

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGÍA
BUCARAMANGA

2013

DIVERSIDAD GENÉTICA Y ANÁLISIS DE SUBESTRUCTURA POBLACIONAL
MEDIANTE 15 STR's AUTOSÓMICOS EN LA REGIÓN
DE TOLIMA-HUILA, COLOMBIA

Estudiante

INGRID GISELA PEREZ PARADA

Trabajo de grado presentado para optar por el título de Bióloga

Director

WILLIAM USAQUÉN MARTÍNEZ Ph.D.

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGÍA
BUCARAMANGA

2013

AGRADECIMIENTOS

Me encuentro muy agradecida con el Grupo de Genética de Poblaciones e Identificación de la Universidad Nacional de Colombia, y en especial con el director y gran maestro William Usaquén Martínez por hacerme un miembro más de este grupo. A él todos mis agradecimientos por sus enseñanzas, consejos y regaños, y a cada uno de sus miembros por su dedicación y apoyo en cada una de las etapas para cumplir con este proyecto.

CONTENIDO

	Pág.
INTRODUCCIÓN	11
1. MATERIALES Y MÉTODOS	15
1.1 Población estudiada	15
1.2 Amplificación por PCR	15
1.3 Genotipificación de STR's	16
1.5 Análisis Descriptivo de Datos	16
1.6 Análisis de Estructura Poblacional	17
2. RESULTADOS	18
2.1 Análisis Descriptivo de Datos	18
2.2 Análisis de Estructura Poblacional	20
3. DISCUSIÓN	22
4. CONCLUSIONES	26
BIBLIOGRAFÍA	27

LISTA DE TABLAS

	Pág.
Tabla N°1. Frecuencias alélicas de los 15 sistemas STR's analizados en la población de los Departamentos de Tolima y Huila (Colombia).	32
Tabla N°2. Diversidad genética, Análisis descriptivo de cada polimorfismo y estimadores forenses y de paternidad de 15 loci STR's en la población de los Departamentos de Tolima y Huila (Colombia)	33
Tabla N°3. Análisis molecular de varianza (AMOVA) para la población de los Departamentos de Tolima y Huila (Colombia).	34
Tabla N°4. Índices de fijación locus por locus y total producto del análisis molecular de varianza.	34
Tabla N°5. Valores de la probabilidad (LnP) para cada número de agrupaciones (K).	35

LISTA DE FIGURAS

	Pág.
Figura N°1. Mapa de los departamentos de Tolima y Huila mostrando los 12 municipios muestreados y las seis zonas planteadas a priori.	31
Figura N°2. Análisis de estructura genética poblacional simulando de una a tres agrupaciones (K).	35

RESUMEN

TITULO: DIVERSIDAD GENÉTICA Y ANÁLISIS DE SUBESTRUCTURA POBLACIONAL MEDIANTE 15 STR's AUTOSÓMICOS EN LA REGIÓN DE TOLIMA-HUILA, COLOMBIA.

AUTOR: Ingrid Gisela Pérez Parada

PALABRAS CLAVES: STR's, Frecuencias alélicas, Diversidad, Alelos efectivos, Estructura genética.

CONTENIDO

Los departamentos de Tolima y Huila se han caracterizado por su importancia histórica y por las diferentes dinámicas poblacionales que a lo largo del tiempo la han influenciado en el contexto histórico colombiano. En el presente trabajo se evaluó a partir de 15 STR's autosómicos, la diversidad genética, el análisis descriptivo de los polimorfismos y la subestructura genética de la región. Se planteó una clasificación a priori de la población estudiada desde una perspectiva histórico-política, con la cual se propuso un modelo de estructura genética quedando la población dividida en seis zonas: Zona de ancestría antioqueña, Zona de corredor, Zona Pijao, Zona de población ancestral, Zona de comunidad indígena y Zona de ancestría caucana y nariñense. Se obtuvo la estimación de las frecuencias alélicas de la población; se observó una alta diversidad genética total, evidenciando que dichas poblaciones están sometidas a continuos procesos de mezcla. Se analizó cada sistema determinando el número de alelos efectivos y su relación con el número de alelos totales. Para esta población los marcadores FGA y D18S51 son los de mayor aporte en análisis forenses o de paternidad. Finalmente, no se encontró subestructura genética basada en un valor de F_{ST} de 0.001 (p -valor= 0.594) y el análisis molecular de varianza indicó que la mayor variación genética fue atribuible a los individuos dentro de la población total, demostrando que estos departamentos se están comportando como una sola población. Se concluye que Tolima y Huila es una población diversa debido a su origen multiétnico, aun así, no se encuentra diferenciación genética debido al proceso de homogenización de la región.

Proyecto de grado

Facultad de Ciencias, Escuela de Biología, Director, William Usaquén Martínez.

ABSTRACT

TITLE: GENETIC DIVERSITY AND SUBSTRUCTURE POPULATION ANALYSIS OF THE TOLIMA-HUILA REGION (COLOMBIA) USING 15 AUTOSOMAL STR's.

AUTHOR: Ingrid Gisela Pérez Parada

KEY WORDS: STR's, Allele frequencies, Diversity, effective alleles, genetic structure.

CONTENT

The departments of Tolima and Huila have been characterized by their historical importance and by the different population dynamics that have influenced the Colombian historical context over time. In the current work, using 15 autosomal STRs, we analyzed the region's genetic diversity, the descriptive analysis of the polymorphisms and its genetic substructure. An *a priori* classification of the assessed population was proposed, using a historical-political perspective, with which a genetic structure model was created, rendering a population divided in six zones: the Antioquia ancestry zone, the "corridor" zone, the Pijao zone, the ancestry population zone, the indian community zone and the Cauca and Nariño ancestry zone. The estimation of the allelic frequency in the population was obtained; a high total genetic diversity was observed, showing that these populations are subject of continuous mixture processes. Each system was analyzed, obtaining the number of effective alleles and their relation with the number of total alleles. For this population, the FGA and the D18S51 markers were the ones that contributed the most to the forensic and paternity analyzes. Finally, no genetic substructure was found, based on an F_{ST} value of 0.001 (p -value= 0.594) and a molecular analysis of variance indicated that most of the genetic variation was attributable to the individuals in the total population, demonstrating that this departments are behaving like a single population. We conclude that Huila and Tolima are a diverse population owing its multiethnic origin. In spite of that, no genetic difference was found, owing to the homogenization process in the region.

Work of Degree

Faculty of Sciences, Biology School, Director, William Usaquén Martínez.

INTRODUCCIÓN

En Colombia dos de las regiones geográficas en donde se observa a nivel fenotípico, una aparente variación de la estructura genética de sus componentes poblacionales son el sur-occidente y el centro del país (Rondón *et al.*, 2008). Los estudios enfocados en la zona sur-occidental del país describen que ésta población está caracterizada por procesos de mezcla y migración (Rondón *et al.*, 2008; Porras *et al.*, 2009; Olarte *et al.*, 2010) además de presentar mayores consecuencias socioculturales generados durante la época de colonización española (Zapata, 1997). Sin embargo, se ha planteado que el impacto de ésta colonización en las migraciones de las comunidades indígenas y campesinas puede que no haya tenido la misma magnitud que se dió en los asentamientos principales (Triana, 1992). Los departamentos de Tolima y Huila se encuentran en la parte sur de la región andina en el valle del río Magdalena. Históricamente, ha sido una región que ha tenido la integración política, administrativa, cultural y socioeconómico (Tovar, 1998).

Las pocas investigaciones adelantadas en el sur del Tolima y en general, en la región del alto Magdalena, no han permitido establecer los procesos de migración y poblamiento por los que pasaron las comunidades de esta región; debido a que los estudios genéticos poblacionales de la región andina que existen en la actualidad, han sido enfocados en estudiar la diversidad genética y estructura de algunas poblaciones incluyendo pocas muestras del Departamento del Tolima (Rondón *et al.*, 2008; Porras *et al.*, 2009), mientras que para Huila se conoce un reporte forense de la región en el que fue incluido dicho departamento (Acosta *et al.*, 2002).

La Genética de Poblaciones humana evalúa y describe la variabilidad genética que existe en los diferentes grupos humanos desde una perspectiva evolutiva. Además, busca demostrar las causas determinantes de dicha variabilidad, a través

del análisis de los procesos microevolutivos que operan sobre la estructura genética de una población a lo largo de las generaciones (Hartl y Clarck, 1997). En las poblaciones humanas no solamente se debe tener en cuenta los factores biológicos y ecológicos, sino también el aspecto cultural y social (Usaquén, 2011). La diversidad genética puede aparecer especialmente estructurada a diferentes escalas: población, subpoblación ó en un grupo de individuos próximos, y su distribución se ve influenciada, no sólo por la acción de fuerzas evolutivas tales como selección natural, migración y deriva (Keyeux, 2001), sino también, por procesos demográficos y sucesos históricos.

En el contexto planteado, los estudios genéticos de las poblaciones establecidas en las cabeceras municipales y en las regiones campesinas características de los departamentos de Tolima y Huila, permitirían ampliar el conocimiento de su subestructura genética con relación a su identidad dentro del territorio Colombiano. Desde el punto de vista genético, el uso de los microsatélites como marcadores moleculares, en este proyecto, constituye una herramienta válida para caracterizar tanto el tipo como la frecuencia de alelos presentes, así como estimar la diversidad y el grado de estructura genética (Ruiz, 1999; Rondón *et al.*, 2008; Hincapié *et al.*, 2009).

Los microsatélites o repeticiones cortas en Tandem (STR's) son segmentos de ADN de uno a siete pares de bases (Edwards *et al.*, 1991) utilizados como marcadores moleculares, que se repiten generando un elevado grado de polimorfismo, de forma aleatoria encontrándose en mayor abundancia en organismos eucariontes. Estos marcadores se encuentran generalmente en regiones no codificantes del genoma y distribuidos uniformemente (Goldstein y Schlotterer, 1999). El tipo de herencia es codominante, es decir, los dos alelos se expresan mostrando el fenotipo del heterocigoto. Fáciles de medir, analizar, presentan una alta confiabilidad y los procesos de laboratorio para su detección se encuentran automatizados. Los STR's son secuencias altamente polimórficas suficientemente específicos para diferenciar individuos dentro de una población

(Goldstein y Schlotterer, 1999). Fueron introducidos en el campo forense a principios de la década de los noventa por Edwards et al (1991), además de ser utilizados en pruebas de paternidad y diagnóstico de enfermedades moleculares (Luque, 2001). Estos marcadores en medicina forense permiten reconocer el valor del análisis de ADN como herramienta de investigación debido al gran número de genotipos que existen en la población, el cual conduce a una alta probabilidad de encontrar patrones genéticos exclusivos de cada individuo (Morales *et al.*, 2004).

La evaluación de la estructura genética poblacional para genotipos polimórficos se realiza con el índice de fijación, el cual guarda relación con la reducción de heterocigosidad esperada, respecto a una población con cruzamiento aleatorio, en un nivel de jerarquía de la población relativo a otro nivel menor de la estructura; con el fin de cuantificar el efecto de la endogamia en las poblaciones, estableciendo tres niveles jerárquicos: diferencia del individuo con respecto al total de la población (F_{IT}), diferencia entre subpoblaciones (F_{ST} y R_{ST}) y diferencia entre individuos de la misma subpoblación (F_{IS}), denominándolos estadísticos F. Siendo el F_{ST} de mayor importancia, ya que se utiliza para estimar la diferenciación en las poblaciones desde la frecuencia de los alelos; este índice puede ser entendido como una medida del grado de endogamia en una subpoblación en relación a toda la población, y refleja la probabilidad de que dos alelos tomados al azar desde la subpoblación sean idénticos por descendencia; mientras que el R_{ST} asume el modelo de mutación gradual de los STR's definido como una variación de la varianza total del tamaño de los alelos entre las poblaciones. Es decir, el R_{ST} indica el exceso de similitud entre los alelos elegidos al azar dentro de la misma subpoblación en relación con toda la población. Dichos estadísticos fueron modificados por Weir y Cockerham en 1984, para ser usados en análisis de STR's, ya que tienen en cuenta el modelo mutacional paso a paso de los microsatélites, quedando definidos como $Capf$, $Theta$, $Smallf$ respectivamente.

Por lo tanto, en el presente proyecto se llevó a cabo la realización del estudio genético poblacional de los departamentos del Tolima y Huila teniendo en cuenta las poblaciones campesinas e indígenas, donde se analizó la diversidad y estructura genética, para contribuir como parte fundamental a la construcción del conocimiento biológico y cultural del país.

1. MATERIALES Y MÉTODOS

1.1 Población estudiada.

Se realizó un muestreo en seis municipios de Tolima: Honda (HO), Líbano (LI), Ibagué (IB), Espinal (ES), Coyaima (CO) y Natagaima (NA) y en seis municipios de Huila: Villavieja (VV), Neiva (NE), Palermo (PA), Garzón (GA), La plata (LP) y San Agustín (SA) (Figura 1). En el cual se analizaron 191 personas, a las cuales se les realizó una reconstrucción genealógica a través de una encuesta.

De cada persona se tomaron 5 ml de sangre periférica con previo consentimiento informado. Las muestras fueron preservadas mediante tarjetas FTA (Whatman BioScience, USA). A partir de un fragmento de la matriz de la tarjeta fue extraído ADN utilizando los instructivos del sistema de calidad del Laboratorio del Grupo de Genética de Poblaciones e Identificación.

1.2 Amplificación por PCR.

Para cada muestra se amplificaron 15 sistemas genéticos microsatélites (*Short Tandem Repeats* – STR's) utilizados en el Combined DNA Index System CODIS (HUMCSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D3S1358, D5S818, D7S820, D8S1179, HUMFGA, HUMTH01, HUMTPOX y HUMVWA31), por PCR utilizando el kit comercial AmpFISTR® Identifier™ (Applied Biosystems, Warrington, U.K.), en un termociclador programable 2720 (Applied Biosystems, Foster City, California). Las amplificaciones fueron realizadas utilizando los siguientes parámetros por ciclo: denaturación inicial (95°C por 11 min) seguido por 26 ciclos de denaturación de las hebras (94°C por 1 min), anillamiento de los cebadores (59°C por 1 min), extensión del DNA (72°C por 1 min), y extensión final (60°C por 60 min).

1.3 Genotipificación de STR's.

Los amplificadores de ADN previamente fluoromarcados son mezclados con formamida (Applied Biosystems, Foster City, California) y LIZ 500 (Applied Biosystems, Warrington, U.K.) y genotipificadas en un analizador genético ABI Prism 310 (Applied Biosystems, Stafford, Texas) utilizando el programa Data Collection versión 3.1. Las asignaciones alélicas fueron hechas mediante la comparación de los fragmentos amplificados con la escalera alélica Ladder (Applied Biosystems, Warrington, U.K.) que contiene el kit y utilizando el programa Genemapper™ ID versión 3.1. (Applied Biosystems, Foster City, California).

1.4 Análisis Descriptivo de Datos.

Para el análisis descriptivo de datos se **tuvieron** en cuenta los siguientes parámetros: estimación de la diversidad genética, análisis descriptivo de los polimorfismos y estadísticos para práctica forense y de paternidad. Las pruebas descriptivas estimadoras básicas son las que tienen por objeto describir la población y sus principales características genéticas (Hartl y Clarck, 1997).

En el programa Genepop 4.09 (Raymond y Rousset, 1995; Rousset, 2008) se calcularon frecuencias alélicas locus por locus, heterocigosidad (Nei, 1987) observada (H_o) y el promedio de heterocigosidad esperada (H_e) bajo el modelo de Equilibrio Hardy Weinberg (EHW). El EHW fué calculado con los siguientes parámetros: 10000 períodos o ciclos de demorización, 1000 batches y 400 iteraciones.

El análisis descriptivo de los polimorfismos se realizó en base al planteamiento de Usaquén (2012), en el cual plantea cinco propuestas de unificación entre genética de poblaciones y genética forense para la próxima década.

Los estimadores forenses: Probabilidad de "Matching" (PM) y Poder de Discriminación (PD), el contenido de información polimórfica (CIP), y los índices de paternidad: Índice Típico de Paternidad (ITP) y poder de exclusión (PE) fueron calculados con el software PowerStats V.12 (Tereba, 2000).

1.5 Análisis de Estructura Poblacional.

Para el estudio genético poblacional, se realizó una clasificación a priori de la población estudiada desde una perspectiva histórico-política (Chala, 2012), teniendo en cuenta la división que diversos autores sociales han establecido a partir de procesos político-administrativos y las relaciones sociales que rodearon la construcción del territorio (Montañez y Delgado, 1998), quedando dividida la población en 6 zonas: Zona de ancestría antioqueña (Norte del Tolima), Zona de corredor (Nororiente de Tolima ribereña al Magdalena), Zona Pijao (municipios ancestría indígena), Zona de población ancestral (Tolima y Huila), Zona de comunidad indígena (Resguardos y Cabildos huilenses), y Zona de ancestría caucana y nariñense (Sur Huila).

Se evaluó subestructura poblacional a partir del cálculo de los estadísticos F (F_{ST} , Θ , S_{smallf}) (Weir y Cockerham, 1984) locus por locus, con intervalos de confianza estimados basándose en técnicas de remuestreo Jackknifing y bootstrapping al 95% sobre los loci, en el software Fstat 2.9.3.2 (Goudet, 1995). La estimación de los F_{ST} locus por locus y el análisis molecular de varianza (AMOVA) con F_{ST} usando 10000 repeticiones y un nivel de significancia de 0.05 (Excoffier *et al.*, 1992) se realizó con el software Arlequin 3.5.1.2 (Schneider *et al.*, 2000).

Finalmente fue realizado un análisis de asignación de individuos a poblaciones en el software Structure 2.3.3 (Pritchard *et al.*, 2000) asumiendo un modelo de mezcla. Con un periodo de quemado de 10.000 y 10.000 MCMC repeticiones sobre el período de quemado.

2. RESULTADOS

2.1 Análisis Descriptivo de Datos.

El análisis descriptivo a partir de la muestra para las poblaciones de Tolima y Huila, arrojó los valores de las frecuencias alélicas para los 15 sistemas STR's autosómicos (Tabla 1). La heterocigosidad esperada promedio para todos los loci incluidos fué 0.7848 y la heterocigosidad observada promedio fue 0.7829, lo que indica que no hay una diferencia significativa entre los heterocigotos observados y esperados en la población total, esto es corroborado con el análisis de equilibrio de Hardy-Weinberg, que para el total de la población no muestra diferencias genotípicas significativamente diferentes a las esperadas (p-valor 0.6793), por lo tanto la población se encuentra en equilibrio, y para el análisis de cada sistema sólo uno de los 15 loci evaluados presentó diferencias significativas (D2S1338 p-valor 0.0257) encontrándose en desequilibrio. El sistema más diverso fue D2S1338 (0.895) y el menos diverso fué D3S1358 (0.681) (Tabla 2). Por lo tanto, se observa que Tolima y Huila tienen alta diversidad genética total, lo que puede evidenciar que dichas poblaciones están sometidas a continuos procesos de mezcla.

El sistema que presentó un mayor porcentaje de alelos efectivos con respecto al número de alelos fué D13S317, ya que de 7, 5.6 alelos están representados en la mayoría de los individuos estudiados, y sólo el 20% corresponden a alelos en baja frecuencia. Mientras, en el sistema con el mayor número de alelos, FGA, sólo el 59.2% de los alelos son efectivos, es decir, de 14 alelos presentes en el loci para la muestra total, solo 8.3 son los alelos de mayor frecuencia. El sistema con una mayor diferencia entre el número de alelos efectivos y el número de alelos totales fué el TPOX (Tabla 2).

Respecto al sistema total, es decir, teniendo en cuenta los 15 loci analizados, cinco de ellos (CSF1PO, D19S433, D3S1358, D8S1179 y TPOX) presentan mayor número de alelos en baja frecuencia en comparación con los alelos efectivos.

El sistema D19S433 mostró el mayor porcentaje (52%) para la relación del número de variantes alélicas encontradas en el estudio con las reportadas a nivel mundial, donde se observaron de 25 variantes reportadas, 13 para la población de Tolima y Huila; y de los 325 posibles genotipos teóricos, sólo se observaron 48 para este sistema en esta población, es decir, el 14.77%. En general, es bajo el porcentaje de genotipos observados para todos los loci STR's evaluados respecto al teórico posible de genotipos (Tabla 2).

Al realizar la distribución genotípica, se observan las diferencias entre los genotipos observados y esperados para cada locus evaluado, dicha información no es relevante al momento de comparar las heterocigosidades y aclarar porque la mayor diferencia entre los heterocigotos observados y esperados la presenta el sistema D21S11, que no es el que se encuentra en desequilibrio de HW. Una posible explicación a lo anterior podría ser las diferencias relativas que hay en cada sistema, ya que debido al tamaño del polimorfismo, la variabilidad o la efectividad, las diferencias entre las heterocigosidades puede variar de un sistema a otro.

En cuanto a los Índices forenses para las muestras estudiadas la probabilidad de Matching arrojó valores bajos de concordancia genética en cada uno de los locus, entendiéndose que hay poca similitud entre los individuos incluidos en el estudio. FGA fué el sistema con menor probabilidad de coincidencia (0.0292) mientras el locus TPOX presentó la más alta probabilidad de coincidencia (0.1608), indicando que la probabilidad de encontrar dos personas con el mismo genotipo para este sistema es de 16.08%. La probabilidad de discriminación mostró un alto valor de diferencia entre los perfiles genéticos de los individuos de

Tolima y Huila, siendo los sistemas FGA y D18S51 los de mayor probabilidad con valores de 0.9708 y 0.9661 respectivamente, por lo cual, se entiende que de 191 individuos aproximadamente 185 tendrán diferentes genotipos para el locus FGA y 184 para el D18S51; el contenido de información polimórfica muestra que en la población hay una alta heterocigosidad esperada, debido al polimorfismo que hay en los loci estudiados, predominando el sistema FGA (0.8684), mientras que el sistema TPOX presenta menor valor, determinado no por el número de alelos presentes sino por el número de alelos efectivos ya que coincide con que este sistema presenta el menor número de alelos efectivos en la población, y es el segundo con menor número de alelos (Tabla 2).

Los índices de paternidad fueron evaluados en la población, los mejores sistemas para tener en cuenta en casos de paternidad o de identificación según este estudio son D2S1338 y D21S11, debido a los altos valores de poder de exclusión e índice típico de paternidad.

2.2 Análisis de Estructura Poblacional.

Para el análisis de subestructura poblacional se evaluó la hipótesis inicial de agrupamiento de los municipios por zonas teniendo en cuenta una perspectiva histórico-política. Las estimaciones de los Estadísticos-F (Weir y Cockerham, 1984) tanto para la población total (F_{IT} o $Capf$ 0.003 \pm 0.010) como para los individuos dentro de los municipios (F_{IS} o $Smallf$ 0.002 \pm 0.01) muestra un bajo valor de endogamia, asumiendo que debido a que no hay probabilidad significativa, los individuos no comparten alelos idénticos por descendencia. El valor de diferenciación subpoblacional o F_{ST} (Theta) fue 0.001 con un intervalo de confianza de 0.002 y un p-valor de 0.59444, el valor de R_{ST} fue 0.0105. Adicionalmente, con el análisis molecular de varianza (Tabla 3) se obtuvo un alto porcentaje de variación (99.99%) entre individuos dentro de los municipios, mostrando el componente de varianza evaluado; donde se refleja el carácter

polimórfico de los sistemas STR's incluidos en el estudio, concordante con la alta diversidad genética promedio, y no hubo diferencias estadísticas significativas entre zonas, ni entre los municipios dentro de las zonas. No se observó subestructura genética significativa teniendo en cuenta los bajos valores de F_{ST} y R_{ST} , por lo cual se infiere que las zonas estudiadas del Tolima y Huila se comportan como una sola población desde el punto de vista genético. Además, se obtuvo el valor de los índices de fijación para cada uno de los loci (Tabla 4), en cada nivel jerárquico, se encuentra que sólo un locus está aportando significativamente a la poca subestructura de la población del Tolima y Huila. Dentro de los municipios respecto a la población total el locus TPOX muestra un F_{ST} de 0.0145 (p 0.0318), entre municipios dentro de las zonas no se encuentran valores significativos y entre las zonas un F_{CT} de 0.0221 (p 0.0063) para el D21S11 (Tabla 4).

Finalmente, se realizó un análisis de asignación de individuos a poblaciones mediante un análisis bayesiano, sobre la información genotípica. Tomando la clasificación como una agrupación a priori se calculó la distribución de frecuencias alélicas a posteriori asumiendo que la población está en equilibrio de Hardy Weinberg. Evaluando los valores de probabilidad ($\ln P$) (Tabla 5), se escogió el mayor valor que corresponde al mejor ajuste del número de agrupaciones (K). En la población de Tolima y Huila teniendo en cuenta la agrupación de regiones se encontró un mejor ajuste en $K=1$ (Figura 2); por lo tanto, se confirma que no hay subestructura poblacional y la agrupación en zonas planteada desde el punto de vista genético no es válida.

3. DISCUSIÓN

En este estudio fueron tipificados 15 sistemas STR's autosómicos empleados en estudios de genética de poblaciones y forense; fueron estimadas las frecuencias alélicas de cada sistema, además del cálculo de varios parámetros estadísticos que indican la utilidad a priori de dichos marcadores (Cal, 2001), como la diversidad genética, EHW, análisis descriptivo de los polimorfismos (Usaquén, 2011), los índices forenses y de paternidad.

Para la población de Tolima y Huila se observó una diversidad genética de 0.7848, la cual evidencia que la población se encuentre en equilibrio de HW para los marcadores usados, ya que no se observan diferencias significativas entre las heterocigosidades observadas y esperadas. Entendiéndose, no como la no evolución de la población; sino que dicha población se encuentra en un equilibrio dinámico, posiblemente porque se están presentando apareamientos de forma aleatoria, manteniendo una variabilidad genotípica amplia; esta diversidad puede estar influenciada por los fuertes procesos históricos tanto en sus linajes ancestrales, como en las migraciones y procesos de mestizaje que se han venido dando en el país (Olarte *et al.*, 2010), lo que ha llevado a las poblaciones a un proceso de mezcla (Rondón *et al.*, 2008).

La distribución de las frecuencias alélicas para cada sistema, muestra el número de alelos presentes en la población estudiada; si bien, los sistemas STR's se encuentran entre los más polimórficos hasta el momento; es importante determinar la eficacia de la variabilidad del polimorfismo de cada sistema en la población (Usaquén, 2011). El cálculo de alelos efectivos sobre el número de alelos del sistema en la población, determina dicho polimorfismo. En este caso, aunque el FGA fué el sistema con mayor número de variantes alélicas, no presentó el mayor porcentaje de alelos efectivos respecto a los totales, corroborándose, que a pesar de que las poblaciones muestren un gran número de variantes genéticas, muchas

de estas variantes tienen una baja frecuencia para la población. Siendo el sistema D19S433 el que aporta mayor variabilidad o posee un mayor polimorfismo.

Además, el presente trabajo aporta la estimación de los índices forenses y de paternidad no evaluados anteriormente para la población de Tolima y Huila, incluidos en estudios anteriores para otras regiones de Colombia (Paredes *et al.*, 2003; Rey *et al.*, 2003; Vargas *et al.*, 2003; Bravo *et al.*, 2004; Martínez *et al.*, 2005; Olarte *et al.*, 2010), los cuales son usados en pruebas de identificación humana. Gracias a los buenos valores de éstos índices, se puede inferir que la caracterización de los 15 loci utilizados en este estudio; son válidos, en cuanto a la informatividad y utilidad en casos de paternidad y en estudios de relaciones familiares de individuos afrodescendientes de la región suroccidente de Colombia (Olarte *et al.*, 2010); ya que las pruebas de paternidad y de identificación de individuos deben basar sus análisis estadísticos relacionados con las frecuencias alélicas de los diferentes marcadores moleculares en las poblaciones de las cuales provienen (Moroni *et al.*, 2008).

La probabilidad de Matching, muestra similitud genética para cada locus comparando los perfiles genéticos entre dos individuos de la misma población tomados al azar (Gaismaier y Schooler, 2008) mientras que la probabilidad de discriminación se define como la probabilidad de que dos individuos tomados al azar, en la misma población, posean diferentes genotipos, dependiendo de la heterocigosidad del sistema (Jones, 1972), entendiéndose que estas se complementan, ya que entre mayor sea el poder de discriminación menor es la probabilidad de concordancia para cada locus.

Otros índices importantes para los análisis de identificación, son los índices de paternidad (Carracedo y Barros, 1996): índice típico de paternidad (ITP) y probabilidad de exclusión (PE) reportados en este estudio para los 15 loci. El poder de exclusión es una probabilidad muy acertada en la actualidad (Moroni *et al.*, 2008), ya que es definida como la fracción de individuos que tienen un perfil de

ADN que es diferente del de un individuo seleccionado aleatoriamente en una muestra. Por otra parte, en casos de paternidad se tienen en cuenta el tipo de exclusión, dependiendo si en mínimo tres de los sistemas evaluados de un individuo no se coincide con el perfil estudiado o si esta exclusión se debe a errores técnicos en el laboratorio, para confirmar la exclusión.

Un sistema con alta probabilidad de discriminación permite individualizar a cada miembro de la población, mientras que un alto poder de exclusión tenido en cuenta para cálculos en pruebas de paternidad permite excluir ó no, a un individuo de un caso, por ejemplo, para este estudio los marcadores FGA y D18S51 son los de mayor aporte en análisis forenses ó de paternidad en los departamentos de Tolima y Huila, además de ser los de mayor contenido polimórfico.

A partir del análisis genético de la estructura poblacional para Tolima y Huila, se demuestra que las zonas planteadas a priori; no se están comportando como poblaciones separadas y que no hay subpoblaciones dentro de estas que las diferencien de las otras, por el contrario, debido a que no presentan diferenciación genética, se comportan como una sólo población. Lo anterior es corroborado con el análisis molecular de varianza, el cuál indicó que la mayor variación genética fue atribuible a los individuos, tanto en el análisis realizado para las zonas, como para los municipios, evidenciado en los altos valores de heterocigosidad observada por cada sistema en la población.

Asimismo, los bajos índices de endogamia (Hartl y Clarck, 1997) evaluados para la población total y para los individuos dentro de los municipios, demuestran que la población no está presentando cruzamientos entre individuos emparentados, por ello, se observan pocos genotipos homocigotos y una alta proporción de genotipos heterocigotos. Además, estos índices sirven para entender la estructura de los cruzamientos de las poblaciones, y el patrón de selección asociado con los alelos polimórficos (Nei, 1987).

Los análisis de poblaciones humanas requieren de clasificaciones sólidas y cuidadosamente realizadas que permitan hacer inferencias más adecuadas correspondiendo a criterios que den cuenta de la dinámica de poblamiento y a los procesos históricos específicos de la población (Usaquén, 2011). Es por ello, que gracias al extenso trabajo de campo realizado por el grupo de Genética de Poblaciones e Identificación, este estudio corrobora que los departamentos de Tolima y Huila pueden diferenciarse entre zonas debido a características históricas y socioculturales, pero al evaluar subestructura en dicha población no se encuentran diferencias genéticas entre los municipios, ni en las zonas planteadas a priori evidenciado por la falta de significancia de los valores de F_{ST} y en la asignación individual para formar agrupaciones, indicando que la población es mezclada, y soportando la idea de que se han presentado procesos de migración desde la época de la conquista española (Rondón *et al.*, 2008), ya que el intercambio de genes entre poblaciones, es una fuente de diversidad genética. A pesar, de que se presenta una alta variabilidad genotípica, las frecuencias alélicas para dichas poblaciones no varían, describiendo una población homogénea.

4. CONCLUSIONES

Los resultados obtenidos concluyen que Tolima y Huila están estrechamente relacionados tanto en su componente histórico como genético, hecho que se ve reflejado en el análisis descriptivo de los datos.

Los marcadores STR's que dan un mayor aporte para análisis forenses y pruebas de paternidad y de identificación, por su poder de discriminación, poder de exclusión, número de alelos efectivos y contenido de información polimórfica son: D2S1338, D21S11 y FGA.

La población de Tolima y Huila se encuentra en equilibrio de Hardy-Weinberg y presenta una alta diversidad genética, lo cuál demuestra que dichas poblaciones están sometidas a continuos procesos de mezcla, por eventos migratorios históricos y actuales. La evaluación de la estructura genética de la población arrojó bajos valores de F_{ST} para los marcadores usados, indicando que los departamentos de Tolima y Huila se están comportando como una sólo población, posiblemente por sus ubicaciones geográficas y además porque se está presentando un constante intercambio de individuos entre las poblaciones.

BIBLIOGRAFIA

ACOSTA M., BRION M., LAREU M., CARRACEDO A. 2002. Genetic data on eight STR's (D5S818, D7S820, F13B, LPL, TH01, TPOX, VWA31, CSF1PO) from a Colombian population. *Forensic Science International* 129:216-218 p.

BRAVO M., BUILES J., PANCORBO M., MORENO M. 2004. Analysis of 12 STR loci in Antioquia (Colombia) population simple. *International Congress Series* 1261: 151-153 p.

CAL M. 2001. Analisis de polimorfismos de AND microsatélite de cromosoma Y. Estudio de la población de Galicia y aplicaciones forenses. Tesis de Grado. Universidad Santiago de Compostela-España.

CARRACEDO A., BARROS F. 1996. Problemas bioestadísticos en genética forense. Universidad Santiago de Compostela-España.

CHALA D. 2012. Lo Histórico, lo Genético y lo Sociocultural, 3 perspectivas para diferenciar poblaciones: El Caso de los departamentos del Tolima y el Huila (Colombia). Informe joven investigador. Colciencias.

EDWARDS A., CIVITELLO A., HAMMOND H., CASKEY C. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* 49:746-756 p.

EXCOFFIER L., SMOUSE P., QUATTRO J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479–491 p.

GAISSMAIER W., SCHOOLER L. 2008. The Smart Potential Behind Probability Matching". *Cognition* 109: 416-422 p.

GOLDSTEIN D., SCHLOTTERER C. 1999. Microsatellites: Evolution and Applications. Londres. Oxford University press. 352 p.

GOUDET J. 2002. Fstat 2.9.3.2. Retrieved from <http://www2.unil.ch/popgen/softwares/fstat.htm>.

HARTL D., CLARCK A. 1997. Principles of population genetics. Sunderland, Massachusetts. Sinauer Associates Inc. 542 p.

JONES D. 1972. Blood simple: probability of discriminations. Journal of Forensic Science Society. 12: 355-359 p.

KEYEUX G. 2001. Contribución del estudio del genoma humano al conocimiento de la diversidad y evolución de los grupos humanos. Innovación y Ciencia. Bogotá: Asociación Colombiana para el Avance de la Ciencia. 9: 66-71 p.

LUQUE J. 2001. Texto ilustrado de biología molecular e ingeniería genética. 1ª ed Madrid: Ediciones Hartcourt S.A.

MARTÍNEZ B., CARABALLO L., BARÓN F., GUSMAO L., AMORIM A., CARRACEDO A. 2005. Analysis of STR loci in Cartagena, a Caribbean city of Colombia. Forensic Science International-Genetics xxx:xxx-xxx p.

MONTAÑEZ G., DELGADO O. 1998. Espacio, territorio y región: conceptos básicos para un proyecto nacional. Cuadernos de Geografía: 120-134 p.

MORALES A., MORERA B., JIMENEZ G. 2004. La implementación forense de la tecnología del ADN en Costa Rica: Un análisis retrospectivo. Biología tropical 52 (3): 695-712 p.

MORONI R., GASBARRA D., ARJAS E., LUKKA M., ULMANEN I. 2008. Effects of reference population and number of STR markers on paternity testing. Forensic Science International-Genetics. Supplement Series 10:102 p.

NEI M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

OLARTE S., FUENTES A., CÁRDENAS H., BARRETO G. 2010. Diversidad Y Estructura Genética de Tres poblaciones Afrodescendientes del Suroccidente Colombiano a partir de 8 STR's. Acta biológica Colombiana. vol.15 no.2.

PAREDES M, GALINDO A, BERNAL M, AVILA S, ANDRADE D, VERGARA C et al. 2003. Analysis of the CODIS autosomal STR loci in four main Colombian regions. Forensic Science International-Genetics 137:67-73 p.

PRITCHARD J., STEPHENS M., DONNELLY P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959 p.

RAYMOND M., ROUSSET F. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. J Heredity 86:248-249 p.

REY M., GUTIÉRREZ A., SCHROEDERA B., USAQUÉN W., CARRACEDO A., BUSTOS I., GIRALDO A. 2003. Allele frequencies for 13 STR's from two Colombian populations: Bogota and Boyaca. Forensic Science International-Genetics 136:83-85 p.

RONDÓN F., OSORIO J., PEÑA A., GARCÉS H., BARRETO G. 2008. Diversidad genética en poblaciones humanas de dos regiones colombianas. Colombia Médica. 39(2) 1: 52-60 p.

SCHNEIDER S., ROESSLI D., EXCOFFIER L. 2000. Arlequin: A software for population genetics data analysis. User manual ver 2:2496–2497 p.

TEREBA A. 1999 Tools for analysis of population statistics, profiles in DNA. Promega Corporation 2:14–6 p.

TOVAR B. 1998. El Huila al final del siglo: historia, identidad y utopía. En observatorio del Huila. Secretaría de cultura departamental. Neiva. Trabajo docente UN.

USAQUEN W. 2011. Validación y consistencia de información en estudios de diversidad genética humana a partir de marcadores microsatélites. Tesis de Doctorado. Universidad Nacional de Colombia.

VARGAS C., CASTILLO A., GIL A., PICO A., GARCIA O. 2003. Population genetic data for 13 STR loci in a northeast Colombian (department of Santander) Population. International Congress Series1293:197-200 p.

WEIR B., COCKERHAM C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370 p.

ZAPATA M. 1997. La rebelión de los genes: El mestizaje americano en la sociedad futura. Colombia: Ediciones Altamir.

TABLAS Y FIGURAS

Figura N°1. Mapa de los departamentos de Tolima y Huila mostrando los 12 municipios muestreados y las seis zonas planteadas a priori.

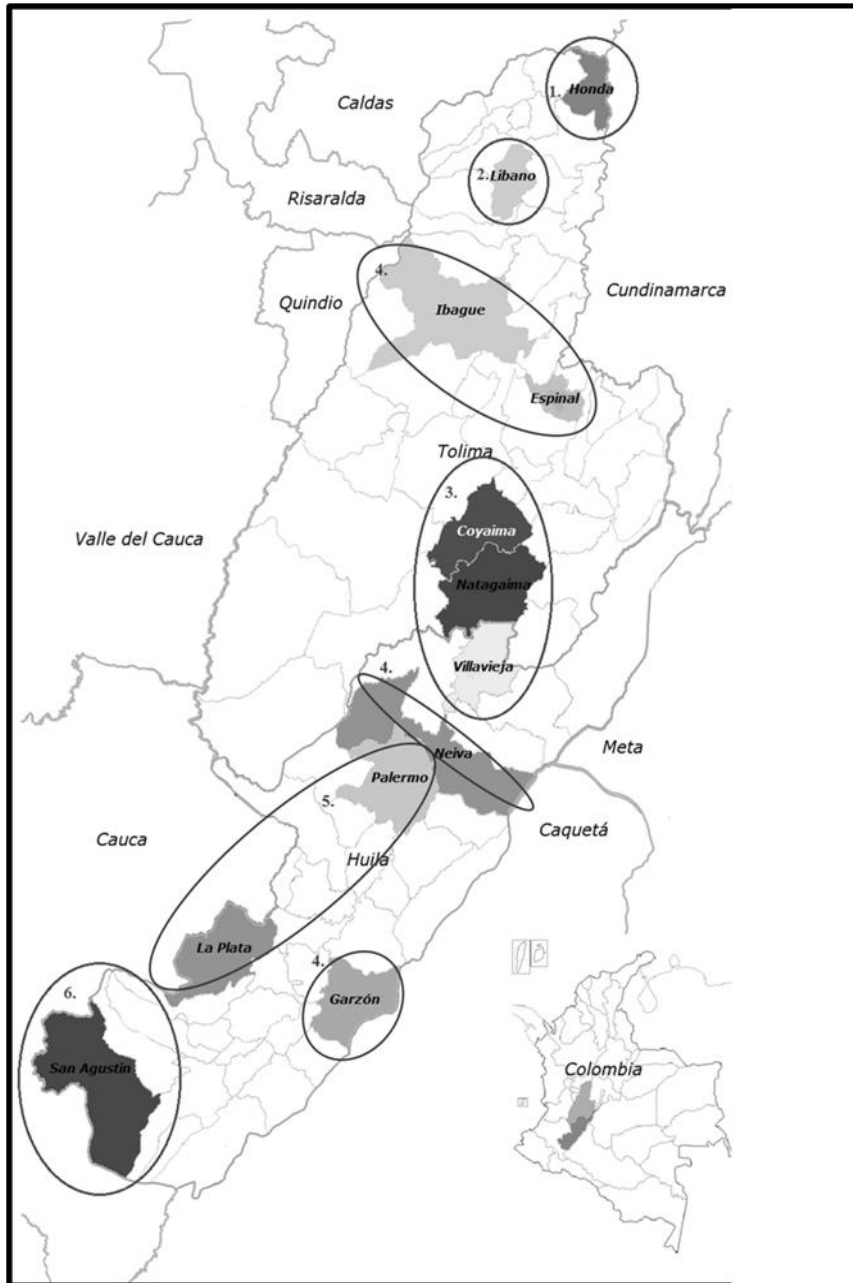


Tabla N°1. Frecuencias alélicas de los 15 sistemas STR's analizados en la población de los Departamentos de Tolima y Huila (Colombia).

ALELO	CSF1PO	D13S317	D16S539	D18S51	D19S433	D21S11	D2S1338	D3S1358	D5S818	D7S820	D8S1179	FGA	TH01	TPOX	vWA
6	-	-	-	-	-	-	-	-	-	-	-	-	0.4058	-	-
7	-	-	-	-	-	-	-	-	0.0288	0.0183	-	-	0.2225	0.0026	-
8	0.0209	0.0759	0.0105	-	-	-	-	-	0.0052	0.0759	0.0052	-	0.0759	0.4921	-
9	0.0131	0.1911	0.1571	-	-	-	-	-	0.0838	0.0654	0.0079	-	0.1073	0.1047	-
9.3	-	-	-	-	-	-	-	-	-	-	-	-	0.178	-	-
10	0.2382	0.0759	0.2042	0.0026	-	-	-	-	0.0628	0.2906	0.0785	0.0026	0.0105	0.0288	-
11	0.2251	0.199	0.2042	0.0079	0.0052	-	-	-	0.3874	0.3298	0.1047	-	-	0.233	-
11.2	-	-	-	-	0.0209	-	-	-	-	-	-	-	-	-	-
12	0.4319	0.2618	0.2696	0.0969	0.055	-	-	-	0.2984	0.1832	0.0759	-	-	0.1335	-
12.2	-	-	-	-	0.0209	-	-	-	-	-	-	-	-	-	-
13	0.0602	0.144	0.1335	0.1204	0.1859	-	-	0.0026	0.1283	0.0288	0.3743	-	-	0.0052	-
13.2	-	-	-	-	0.1309	-	-	-	-	-	-	-	-	-	-
14	0.0105	0.0524	0.0209	0.1623	0.2513	-	-	0.0864	0.0052	0.0079	0.2251	-	-	-	0.0628
14.2	-	-	-	-	0.0288	-	-	-	-	-	-	-	-	-	-
15	-	-	-	0.1571	0.1806	-	-	0.4372	-	-	0.0995	-	-	-	0.1099
15.2	-	-	-	-	0.0654	-	-	-	-	-	-	-	-	-	-
16	-	-	-	0.1466	0.034	-	0.0262	0.267	-	-	0.0288	-	-	-	0.3665
16.2	-	-	-	-	0.0131	-	-	-	-	-	-	-	-	-	-
17	-	-	-	0.1571	0.0079	-	0.2382	0.1047	-	-	-	0.0105	-	-	0.2984
18	-	-	-	0.055	-	-	0.0628	0.0969	-	-	-	0.0105	-	-	0.1126
19	-	-	-	0.0445	-	-	0.1518	0.0052	-	-	-	0.055	-	-	0.0419
20	-	-	-	0.0183	-	-	0.1178	-	-	-	-	0.1073	-	-	0.0052
21	-	-	-	0.0236	-	-	0.0157	-	-	-	-	0.1178	-	-	0.0026
22	-	-	-	0.0079	-	-	0.1126	-	-	-	-	0.1492	-	-	-
23	-	-	-	-	-	-	0.1387	-	-	-	-	0.1126	-	-	-
24	-	-	-	-	-	-	0.0707	-	-	-	-	0.178	-	-	-
25	-	-	-	-	-	-	0.0524	-	-	-	-	0.1361	-	-	-
26	-	-	-	-	-	-	0.0131	-	-	-	-	0.0707	-	-	-
27	-	-	-	-	-	0.0157	-	-	-	-	-	0.0445	-	-	-
28	-	-	-	-	-	0.1047	-	-	-	-	-	0.0026	-	-	-
29	-	-	-	-	-	0.199	-	-	-	-	-	0.0026	-	-	-
30	-	-	-	-	-	0.3037	-	-	-	-	-	-	-	-	-
30.2	-	-	-	-	-	0.0183	-	-	-	-	-	-	-	-	-
31	-	-	-	-	-	0.0759	-	-	-	-	-	-	-	-	-
31.2	-	-	-	-	-	0.1021	-	-	-	-	-	-	-	-	-
32	-	-	-	-	-	0.0209	-	-	-	-	-	-	-	-	-
32.2	-	-	-	-	-	0.1152	-	-	-	-	-	-	-	-	-
33.2	-	-	-	-	-	0.0366	-	-	-	-	-	-	-	-	-
34	-	-	-	-	-	0.0026	-	-	-	-	-	-	-	-	-
35	-	-	-	-	-	0.0052	-	-	-	-	-	-	-	-	-

Tabla N°2. Diversidad genética, Análisis descriptivo de cada polimorfismo y estimadores forenses y de paternidad de 15 loci STR's en la población de los Departamentos de Tolima y Huila (Colombia). He: Heterocigosidad Esperada, Ho: Heterocigosidad Observada P: Valores de Probabilidad del test exacto EHW, PM: Probabilidad de "Matching", PD: Poder de Discriminación, CIP: Contenido de Información Polimórfica, PE: Poder de Exclusión, ITP: Índice Típico de Paternidad.

		CSF1PO	D13S317	D16S539	D18S51	D19S433	D21S11	D2S1338	D3S1358	D5S818	D7S820	D8S1179	FGA	TH01	TPOX	vWA	
Diversidad genética	He	0.703	0.823	0.803	0.875	0.844	0.828	0.864	0.712	0.734	0.764	0.778	0.882	0.739	0.676	0.748	
	Ho	0.717	0.827	0.785	0.848	0.806	0.880	0.895	0.681	0.723	0.717	0.801	0.885	0.702	0.728	0.749	
	EHW <i>p</i> -valor	0.635	0.893	0.239	0.164	0.460	0.409	0.026	0.596	0.564	0.905	0.420	0.959	0.706	0.449	0.564	
Análisis descriptivo de cada polimorfismo	Nº alelos	7	7	7	13	13	12	11	7	8	8	9	14	6	7	8	
	Nº genotipos observados	19	27	21	49	48	42	45	17	22	25	27	53	17	17	20	
	Nº alelos reportados	24	16	15	39	25	58	24	32	21	27	20	57	19	15	21	
	Nº teórico de posibles genotipos	300	136	120	780	325	1711	300	528	231	378	210	1653	190	120	231	
	% genotipos observados sobre el total teórico	6.33	19.85	17.50	6.28	14.77	2.45	15.00	3.22	9.52	6.61	12.86	3.21	8.95	14.17	8.66	
	Nº alelos efectivos	3.4	5.6	5	7.9	6.3	5.7	7.2	3.4	3.7	4.2	4.5	8.3	3.8	3.1	3.9	
	% alelos efectivos sobre Nº alelos	48.57	80.00	71.43	60.77	48.46	47.50	65.45	48.57	46.25	52.50	50.00	59.29	63.33	44.29	48.75	
	Nº genotipos posibles con el Nº alelos efectivos	7.48	18.48	15	35.155	22.995	19.095	29.52	7.48	8.695	10.92	12.375	38.595	9.12	6.355	9.555	
	Índices forenses y de Paternidad	PM	0.141	0.060	0.074	0.034	0.045	0.060	0.043	0.127	0.112	0.088	0.081	0.029	0.105	0.161	0.105
		PD	0.859	0.940	0.926	0.966	0.955	0.940	0.957	0.873	0.888	0.912	0.919	0.971	0.895	0.839	0.895
CIP		0.653	0.796	0.771	0.859	0.824	0.805	0.847	0.667	0.692	0.725	0.748	0.868	0.698	0.630	0.708	
PE		0.455	0.651	0.572	0.691	0.611	0.754	0.786	0.399	0.464	0.455	0.601	0.775	0.431	0.472	0.508	
ITP		1.769	2.894	2.329	3.293	2.581	4.152	4.775	1.566	1.802	1.769	2.513	4.548	1.675	1.837	1.990	

Tabla N°3. Análisis molecular de varianza (AMOVA) para la población de los Departamentos de Tolima y Huila (Colombia).

Fuente de variación	Componente de variación	% de variación
Entre las zonas	0.01228	0.20859
Entre municipios dentro de las zonas	-0.01221	-0.20739
Dentro de los municipios	5.88755	99.9988
Total	5.88762	

Tabla N°4. Índices de fijación locus por locus y total producto del análisis molecular de varianza.

Locus	FCT	P-value	FSC	P-value	FST	P-value
CSF1PO	-0.004	0.6059	0.0049	0.5092	0.001	0.3534
D13S317	-0.0068	0.7628	0.008	0.3621	0.0013	0.3118
D16S539	-0.0033	0.7747	-0.0052	0.8725	-0.0085	0.909
D18S51	0.0028	0.2742	-0.002	0.4137	0.0008	0.5331
D19S433	-0.0039	0.8682	0.0005	0.5847	-0.0034	0.759
D21S11	0.0221	0.0063	-0.0143	0.1219	0.0081	0.1188
D2S1338	-0.0028	0.7384	-0.0008	0.7317	-0.0036	0.625
D3S1358	-0.0128	0.6919	0.0258	0.0918	0.0133	0.0538
D5S818	0.0139	0.0876	-0.0064	0.2048	0.0075	0.2575
D7S820	-0.0039	0.653	0.0013	0.6228	-0.0026	0.6681
D8S1179	0.0059	0.294	-0.0123	0.9441	-0.0064	0.842
FGA	0.0031	0.3229	-0.0119	0.9902	-0.0088	0.9798
TH01	0.0018	0.4866	-0.0061	0.7353	-0.0043	0.784
TPOX	0.0099	0.228	0.0046	0.2563	0.0145	0.0318
vWA	0.0094	0.2496	-0.0138	0.6309	-0.0043	0.779
Total	0.00209	0.17004	-0.00208	0.68607	0.00001	0.59444

Tabla N°5. Valores de la probabilidad (LnP) para cada número de agrupaciones (K).

K	LnP
1	-10106.4
2	-10174.6
3	-10208
4	-10348.8
5	-10905.1
6	-10550.3

Figura N°2. Análisis de estructura genética poblacional simulando de una a tres agrupaciones (K).

