

Descripción del posible origen de los patrones trialélicos del locus TPOX en dos familias colombianas y la importancia en la identificación forense

Richard Orlando Salazar Enríquez

Trabajo de investigación para optar al título de Magister en Ciencias Básicas Biomédicas con énfasis en Genética Forense.

Tutor:

**Winston Rojas Montoya, Msc, PhD.
Profesor de la Facultad de Biología, Universidad de Antioquia.**

Miembros del comité tutorial:

**Yeny Cecilia Posada Posada, Msc.
Analista del Laboratorio de Identificación Genética IdentiGEN**

**Leonor Gusmão, PhD.
Profesora de la Universidad Estatal de Rio de Janeiro (Brasil).**

**Adriana Alexandra Ibarra Rodríguez, Msc, PhD.
Directora del Laboratorio De Identificación Genética IdentiGEN.**

**Gonzalo De Jesús Vásquez Palacio, Msc.
Profesor de la Facultad de Medicina, Universidad de Antioquia.**

**Universidad de Antioquia
Corporación de Ciencias Básicas Biomédicas
Medellín
2019**

AGRADECIMIENTOS

Quiero expresar mi gratitud a todos los funcionarios del Laboratorio de Identificación Genética – IdentiGEN, en especial a la analista Yeny Cecilia Posada y a la Doctora Adriana Alexandra Ibarra, por el acompañamiento y el apoyo brindado para el desarrollo del trabajo. A mis asesores por su ayuda a partir de su experiencia, a la profesora Leonor Gusmao por su asesoría en los conceptos y en el desarrollo de las fórmulas para los patrones trialélicos, al profesor Winston Rojas por los momentos de discusión en el desarrollo de la tesis, a los profesores; Juan Fernando Alzate y Felipe Cabarcas por ayudarme con el análisis del genoma, y a los profesores Domingo y Edinson Velásquez por ayudarme en la revisión de las fórmulas del patrón trialélico.

Finalmente, también quiero agradecer al profesor Gonzalo Vásquez por su ayuda incondicional en el desarrollo experimental del trabajo.

Tabla de contenido

1	Resumen	6
2	Introducción.....	8
3	Marco teórico.....	8
3.1	Marcadores usados en genética forense:.....	8
3.1.1	STRs.....	10
3.1.2	Origen molecular de un patrón trialélico:.....	15
3.2	Población de referencia.....	19
3.2.1	Parámetros estadísticos de relevancia en genética forense.....	19
3.2.2	Poblamiento de Colombia.....	23
3.3	Cálculos de filiación: tríos y dúos.....	25
3.4	<i>Next Generation Sequencing (NGS)</i> en búsqueda de variantes estructurales(SV).....	30
4	Planteamiento del problema.....	34
5	Objetivos del proyecto.....	41
5.1	Objetivo general.....	41
5.2	Objetivos específicos.....	41
6	Metodología.....	42
6.1	Localización genómica:.....	45
6.2	Cálculos estadísticos en filiación con patrones trialélicos:.....	46
7	Resultados y discusión.....	77
7.1	Análisis de las genealogías.....	77
7.2	Ancestría con Indels-AIMs.....	83
7.3	Estadísticos forenses y patrones trialélicos tipo 2-B.....	85
7.4	Ubicación genómica del alelo extra.....	87
8	Conclusiones.....	94
9	Referencias bibliográficas.....	95

Lista de figuras

Figura 1 : Clases de STRs	12
Figura 2: Mecanismo Nonallelic homologous recombination.....	18
Figura 3; Probabilidades de compartir alelos IBD Padre-Hijo, suponiendo que el padre y la madre no están emparentados.	30
Figura 4: Patrones de SV	31
Figura 5: Representación esquemática del procedimiento de CGH con arrays.	33
Figura 6: Patrón trialélico tipo 1 y tipo 2 propuesto por Clayton et al. En el año 2004.....	35
Figura 7: Expansión de la mutación del marcador genético TPOX desde África hacia América.....	38
Figura 8: Flujo de trabajo de NGS.....	46
Figura 9: Genealogía de la familia de Frontino-Antioquia.....	78
Figura 10: Genealogía de la familia afrocolombiana estudiada	80
Figura 11: Ejemplo de alineamiento del individuo IV-4 de Frontino	89
Figura 12: Diagrama esquemático de la duplicación y reorganización de la copia adicional de TPOX.	90

Lista de tablas

Tabla 1: Cálculos y fórmulas de prueba de paternidad de un solo locus para 21 diferentes combinaciones de genotipos de trío completo	27
Tabla 2: Cálculos y fórmulas para pruebas de filiación con un solo progenitor	28
Tabla 3: . Probabilidades por pares de genotipos en función del número de alelos compartidos	29
Tabla 4: Reporte de patrones trialélicos para el STR TPOX	37
Tabla 5: Resultados estadísticos de relevancia forense	87
Tabla 6. Resultado de las SV más cercanas al marcador TPOX. Las SV observadas no contienen el marcador TPOX ni parte del gen TPO.	88
Tabla 7. Variantes únicas de cada genoma.....	91
Tabla 8: Resultado del microarreglo.....	93

1 Resumen

El TPOX es un STR ubicado en el intrón 10 del gen de la peroxidasa tiroidea humana (TPO). En este marcador pueden generarse perfiles genéticos con patrones trialélicos y estudios previos sugieren que esta característica tiene origen africano. El presente trabajo se enfoca en describir los patrones trialélicos en dos familias colombianas y su posible origen. Para ello, se estudiaron familias de Chocó y Antioquia, con el fin de observar la segregación de los patrones trialélicos en el marcador TPOX. Además, se obtuvieron estimativos de ancestría con indels-AIMs, ya que el componente ancestral sirve como herramienta para explorar la historia de las familias del presente estudio, aunado a ello se analizó el genoma de un individuo de cada familia en busca del origen molecular. En la genealogía de la familia proveniente de Frontino-Antioquia se observó un patrón trialélico del locus TPOX, donde los alelos 12 y 13 siempre co-segregan a todos los descendientes (hombres y mujeres), lo que es factible en una posible duplicación dentro del cromosoma 2. Por otro lado, la familia de Quibdó-Choco, presentó el alelo 10 en los patrones trialélicos y puede existir una co-segregación con el alelo 6 en presencia de recombinación o una segregación del alelo 10 ligada al cromosoma X. Sumado a las diferencias alélicas y de segregación en las familias, el análisis de ancestría en conjunto con el análisis del genoma completo contribuye a posibles orígenes diferentes.

Palabras clave: patrón trialélico; Segregación; Ancestría; TPOX; mutación.

Abstract

The use of Short Tandem Repeat markers (STRs; or microsatellites) is the most common tool in Forensic Genetics, to perform different tests that are the basis for individual or kinship genetic identifications. The TPOX is a STR marker at intron 10 of human thyroid peroxidase (TPO) gene, located on chromosome 2. The TPOX is included in most multiplex kits that are routinely used in forensic genetics and, as expected for autosomal markers, one or two allelic patterns are usually observed in homozygote and heterozygote individuals, respectively. However, tri-allelic patterns have been reported for TPOX, due to rare events such as mutations or chromosomal aberrations. Since most formulas used to evaluate the statistical weight of the evidences in forensic analysis only account for the expected patterns of one or two alleles for an autosomal locus, it is important to investigate the correct way to solve cases where tri-allelic patterns are present. In addition, ancestry estimates were obtained with indels-AIM, since the ancestral component serves as a tool to explore the history of the families of the present study. Moreover, complete genome sequencing was performed for two individuals with TPOX tri-allelic patterns, to determine the location of the third allele. In the genealogy of the family from Frontino-Antioquia, a tri-allelic pattern of the TPOX locus was observed, where alleles 12 and 13 always segregate together to all descendants (male and female), suggesting a possible duplication within chromosome 2 with a possible co-segregation. These results may differ from the Quibdó-Chocó family, since with the genealogic data it is not yet possible to conclude which is the extra allele or where it is located. Thus, allelic and segregation differences in families, analysis of ancestry in conjunction with analysis of the genome contributes to possible different origins.

Keywords: segregate, tri-allelic pattern, Ancestry, TPOX, mutation.

2 Introducción

En Genética Forense el uso de marcadores moleculares tipo Microsatélite (o STR, del inglés *Short Tandem Repeat*) es la herramienta más común para las diferentes pruebas que se realizan como base en la identificación genética. El marcador genético denominado TPOX se ha incluido en diferentes sistemas utilizados, por lo cual es habitualmente usado en los laboratorios en el mundo. Otro aspecto relevante es que incluye patrones trialélicos en algunos individuos y no se han planteado pautas de análisis, para abordarlos en identificación forense. Por ello, este trabajo pretendió describir el posible origen de los genotipos trialélicos en el locus TPOX y el efecto en la identificación forense.

3 Marco teórico

3.1 Marcadores usados en genética forense

Las repeticiones cortas en tándem (STRs), aunque son frecuentemente usadas en genética forense, presentan dificultades técnicas con las muestras altamente degradadas, muestras en cantidad escasa de ADN o problemas inherentes a la reacción en cadena de la polimerasa (PCR), lo que se evidencia en la obtención de resultados a través de equipos especializados generando inespecificidades. Otras dificultades podrían presentarse por la aparición de alelos extra (*allelic drop-in*) o la pérdida alélica (*allelic drop-out*), además puede presentarse amplificación preferencial (*allelic imbalance*) o resultados con ausencia de amplificación (*locus drop-out*). A pesar de estas desventajas los STRs poseen un gran polimorfismo, y para subsanar el problema descrito con algunas muestras, existen marcadores denominados MiniSTRs que generan

fragmentos de ADN con un rango menor de 150pb (pares de bases), pudiendo amplificar muestras altamente degradadas (1,2).

Igualmente, otro tipo de marcador comúnmente usado son los SNPs (del inglés, *Single Nucleotide Polymorphisms*), estos generalmente son bialélicos y poseen gran variación génica. Los SNPs presentan bajas tasas de mutación alrededor de 10^{-8} , siendo menos informativos comparados con los STRs, pero se puede mejorar la exactitud de los SNPs usando una alta cantidad de éstos y se ha demostrado que el uso de aproximadamente 50 SNPs no ligados, hace que la probabilidad de encontrar dos genotipos iguales tienda a cero (1,3).

Por otra parte, las inserciones y deleciones (Indels) también son ampliamente usadas en identificación genética, estos son polimorfismos de longitud del ADN causados por inserción o deleción de uno o más pares de bases, presentan una baja tasa de mutación aproximada de 10^{-9} , comprenden el 20% de polimorfismos en el genoma humano y se clasifican en cinco tipos (4):

- Inserciones o deleciones de un sólo par de bases.
- Expansiones monoméricas de pares de bases.
- Expansiones Multi-base, de 2-15 unidades de repetición.
- Inserciones de transposones (inserciones de elementos móviles).
- Indels con secuencia aleatoria (random) de ADN con un rango de 2 a 9.989 pb de longitud.

Los STRs, SNPs e Indels, aunque son usados ampliamente alrededor del mundo, se le ha dado prioridad desde hace más de 20 años a los STRs, ya que estos poseen grandes bases de datos. Además, el número de genotipos que se puede obtener con un marcador STR multialélico es mucho mayor a marcadores bialélicos como los SNPs o los Indels. Por ejemplo en un STR de 5 alelos se pueden obtener hasta 15 genotipos distintos, 5 homocigotos y 10 heterocigotos, pero con un SNP bialélico o un Indel solo se obtienen 3 genotipos, un heterocigoto y 2 homocigotos (5,6). También están validados por el *Federal Bureau of Investigation* (FBI) de Estados Unidos, que seleccionó en el año de 1997, 13 loci STRs para usarlos como grupo fundamental en los análisis forenses, denominados CODIS (Combined DNA Index System) y más recientemente en el año 2017 se adicionaron 7 marcadores, para un total de 20. Actualmente destacan varios sistemas de marcadores genéticos validados, que se han convertido en la herramienta de rutina para resolver pruebas forenses y de paternidad en todo el mundo (7–9).

3.1.1 STRs

Los STRs son secuencias de nucleótidos, agrupadas entre 2 y 6 pb, con tándems aproximados entre 6 a 40 veces. Generalmente los fragmentos amplificados de STRs oscilan en un rango aproximado de 100 a 400pb. Además es propio de estos microsatélites que sean ADN no codificante, cuenten con una tasa de mutación alrededor de 1×10^{-3} , y con un alto polimorfismo (7,8,10,11). Estos STRs son de gran importancia en genética forense y se vienen usando desde hace más de 20 años en identificación de personas desaparecidas (MPI), víctimas de desastres masivos (DVI), en filiación, y en el análisis de evidencias biológicas en el campo penal (1,3), aportando a la investigación de casos de carácter criminal o civil. Esto es posible gracias al polimorfismo que existe en el ADN de todos los individuos, generando un perfil único denominado “huella genética”, a excepción de

los gemelos monocigotos que, por su origen univitelino, comparten además de sus características fenotípicas, toda su información genotípica, evidenciada a través de estos marcadores (7).

3.1.1.1 Origen de los STRs

Los STRs pueden originarse como producto de una inserción de 2 o 4 nucleótidos en un locus al azar en el genoma humano y esta inserción puede expandirse en 2 o 5 nuevas repeticiones (12), también pueden generarse por una sustitución como sucede en el locus de n-globin, donde se genera un STR, cambiando la secuencia ATGTGTGT a ATGTATGT, la cual puede encontrarse en muchos primates, así como en el humano con diferente número de repeticiones (13).

De acuerdo a lo anterior, se ha demostrado que las secuencias SINEs (*interspersed nuclear elements*), en particular los elementos Alu, han contribuido en la génesis de los STRs. Para esto se han propuesto tres mecanismos diferentes: 1) el elemento Alu puede insertarse en medio de un STR y crear repeticiones en las secuencias que flanquean dicho elemento; 2) el elemento Alu puede insertarse con una mutación en la cual exista una sustitución o inserción de nucleótidos al retrotranscribirse, pudiendo así generar nuevos STRs y; 3) la acumulación de mutaciones en la cola de poly(A) del elemento Alu, seguido por la expansión por medio del *slippage* o la recombinación, genera nuevos STRs (14). De igual manera se han hallado secuencias de elementos de transposición en las zonas flanqueantes de los STRs hasta un 41%, con relevancia en el origen de estos (15).

3.1.1.2 Clases de STRs

Los STRs pueden ser clasificados como simples, compuestos o complejos. Los STRs simples son secuencias de un único motivo de repetición, los compuestos están conformados por más de un motivo de repetición y los complejos, a diferencia de los compuestos, pueden tener más de un motivo de repetición de diferente tamaño y secuencia (16), como se observa en los ejemplos de la figura 1.

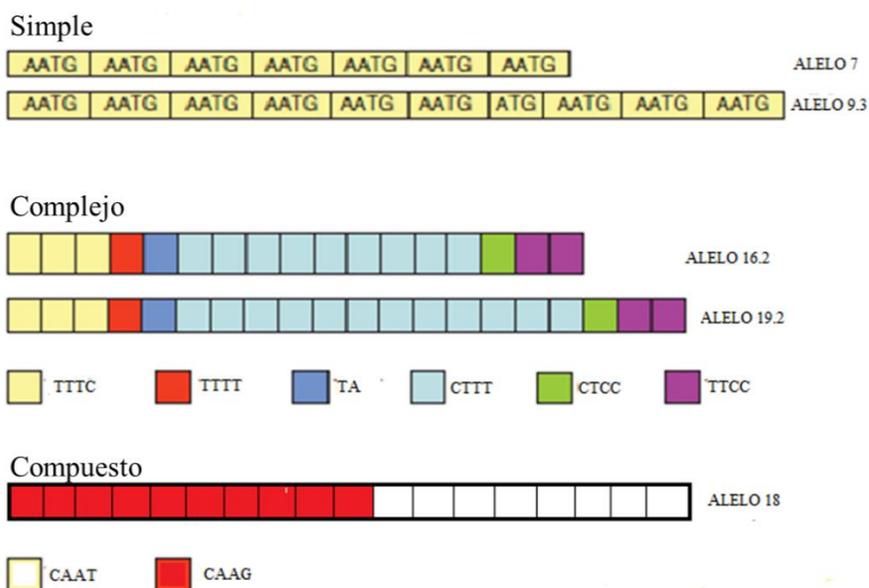


Figura 1 : Clases de STRs, figura modificada de Goodwin et al.(17)

3.1.1.3 Modelos de mutación de los STRs

En la recolección de datos de STRs es importante conocer cual modelo teórico mutacional se debe aplicar, ya que dichos modelos mutacionales se usan para derivar el número esperado de alelos en una población y también en los análisis estadísticos de la variación genética. En general, los modelos usados son (12):

- Modelo mutacional de alelos infinitos. Cada mutación crea aleatoriamente un nuevo alelo, es decir que un alelo con determinado número de repeticiones está estrechamente relacionado con otro sin importar dicho número de repeticiones. En otras palabras, un alelo 10 está relacionado genéticamente con el alelo 15 tanto como con el 9 (12);

- Modelo de mutación por pasos. Al existir una mutación, un STR puede perder o ganar una repetición, implicando que dos alelos que difieren solo en una repetición están más relacionados que aquellos que difieren en varias unidades. Este modelo es usado comúnmente para estructura genética y para estimar la relación entre individuos, excepto en presencia de homoplasia (12);

- Modelo de dos fases. Se introdujo como un complemento del modelo de mutación por pasos. En este se establece que la mayoría de eventos mutacionales resultan en un incremento o disminución de una unidad de repetición, aunque también pueden ocurrir, en menor proporción, alteraciones que involucren grandes números de repeticiones (12);

- Modelo de K-alelos. Si en un locus dado existen K posibles alelos, entonces la probabilidad de que un alelo dado mute es $\mu/k-1$, donde μ es la tasa de mutación (12).

3.1.1.4 Mecanismo de mutación

La tasa de mutación de los STRs oscila entre 10^{-2} hasta 10^{-6} por generación y generalmente se debe al *slippage* de la polimerasa o, en menor medida, a la recombinación desigual durante la replicación o reparación. Sumado a ello está el mecanismo de retrotransposición que también genera cambios en los STRs (18).

Dicha tasa de mutación puede estar influida por diversos factores como el número de repeticiones, la unidad de repetición, la estructura de la repetición, las bases que componen la unidad de repetición, la recombinación, el sexo, la edad y la interrupción de los STRs (18).

El número de repeticiones puede influir en los cambios generados en los STRs, ya que se ha demostrado que una secuencia corta tiende a ganar repeticiones y una secuencia larga tiende a perder unidades de repetición (19).

De acuerdo a la unidad de repetición se ha observado que los tetranucleótidos tienen hasta 4 veces más estabilidad que los dinucleótidos y los trinucleótidos en casos que no involucran la salud tiene una estabilidad mayor que los dinucleótidos hasta 1,5 a 2 veces (20).

Los heterocigotos tienden a aumentar la tasa de mutación de los STR, es decir que un alelo influye sobre el otro y generalmente un alelo corto se ve influenciado por el alelo más grande, a este efecto se le denomina allele-span (21).

Las secuencias con alto contenido de AT suelen aumentar la tasa de mutación del STR, y las secuencias con alto contenido en GC tienden a sufrir menos cambios (21), además no existe evidencia de que la recombinación influya en los cambios de los STRs(18), sin embargo, pueden darse duplicaciones o deleciones en variantes estructurales que pueden conllevar a anomalías en los patrones de lectura de los genotipos(22), esto puede dar origen a patrones trialélicos lo cual se explicará más adelante.

Por otro lado, la mutación de los STRs es más frecuente en los hombres, ya que estos tienen un mayor número de replicaciones en sus células germinales comparado con las mujeres(23), además, cuando los STRs presentan sustituciones o inserciones/deleciones en la secuencia, puede cambiar su tasa de mutación, debido a que interrumpe la conexión entre las unidades de repetición(18).

3.1.2 Origen molecular de un patrón trialélico

Un patrón trialélico en un marcador genético puede deberse a procesos asociados a la recombinación del ADN, la replicación y la reparación, la presencia de elementos de transposición, el mosaicismo y variaciones en el número de cromosomas.

La ganancia o pérdida de uno o más cromosomas dan lugar a las variaciones denominadas aneuploidías. Si hay pérdida puede presentarse una monosomía (ejemplo: síndrome de Turner) o si hay ganancia una trisomía (ejemplo: síndrome de Down). Para estas variaciones en número existe otro término llamado euploidia, el cual define cambios en la totalidad de cromosomas dando origen a organismos haploides, diploides o poliploides, estos últimos se observan generalmente en plantas (6). Los fenómenos de aneuploidía se presentan por la no disyunción cromosómica, en la cual, en el proceso de la meiosis I o II no se separan de forma adecuada los cromosomas homólogos dando lugar a este tipo de alteraciones. Aunque los fenómenos de aneuploidía autosómica se observan con una frecuencia muy baja y la mayoría de los individuos que las presentan son abortados de forma espontánea, como en el caso del cromosoma 2, ya que estos fenómenos se observan mayormente en cromosomas pequeños como el 21, el 18, el 13, y con menor frecuencia en el cromosoma 8 (24).

Ahora bien, en las variaciones de estructura y ordenación de los cromosomas generalmente se deben a roturas causadas por agentes exógenos o endógenos a lo largo del cromosoma, que pueden provocar pérdidas o reordenamiento del material genético, además la exposición a sustancias químicas o radiación aumentan la tasa de daño, y en estos puntos de ruptura se crean segmentos pegajosos que tienden a unirse a otros segmentos, creando duplicaciones, translocaciones y deleciones. Otro de los fenómenos que producen variaciones en el genoma pueden deberse al entrecruzamiento desigual o por error en la replicación, aunque este hecho permite la aparición de conceptos como redundancia génica, alteraciones en el fenotipo y variabilidad génica en la evolución (6).

Así como existen variaciones a nivel cromosómico, también se describen a nivel génico mutaciones que pueden explicar un genotipo trialélico, como las inserciones o deleciones de uno o más pares de nucleótidos (24).

Además del *slippage*, la recombinación desigual también puede generar gran diversidad en los loci, ya que muchos microsatélites y minisatélites se han propuesto como sitios calientes para recombinación. La recombinación como los procesos que impliquen entrecruzamiento desigual de genes puede conducir a grandes contracciones de la escala y expansiones en la matriz de repetición (26). Asimismo, existen mecanismos como *Nonallelic homologous recombination* que pueden generar una duplicación de un segmento del genoma, mecanismo explicado en la figura 2 (27).

La figura 2 representa el mecanismo *Nonallelic homologous recombination*: A | La *Nonallelic homologous recombination* (NAHR) entre repeticiones directamente orientadas genera deleciones

y duplicaciones (parte Aa), mientras que NAHR entre repeticiones invertidas genera la inversión del segmento entre las repeticiones (parte Ab). B | La tasa de NAHR puede variar dentro de un locus particular debido a la presencia de pares de *Low-copy repeats* (LCR) con riesgo de susceptibilidad distintivo. Por ejemplo, la tasa de NAHR se correlaciona positivamente con la longitud de LCR y se correlaciona inversamente con la distancia inter-LCR, que puede conducir a variantes de número de copias recurrentes e infrecuentes (CNV) en un locus determinado identificado en diferentes pacientes con un trastorno genómico (por ejemplo, en las deleciones recurrentes e infrecuentes y las duplicaciones recíprocas en el síndrome de Smith-Magenis y el síndrome de Potocki-Lupski, respectivamente). C | La tasa de NAHR puede variar entre individuos con una arquitectura genómica distinta en un locus determinado o un haplotipo de variante estructural que da como resultado LCR directamente orientado. Esto se ejemplifica mediante haplotipos estructurales invertidos que presentan distintos riesgos de someterse a NAHR debido a la variabilidad estructural en el contenido y orientación de LCR. Por ejemplo, se observan dos haplotipos variantes estructurales divergentes, H1 y H2, en la región 17q21.31, pero las deleciones que causan el síndrome de Koolen-De Vries (OMIM 610443) aparecen en los cromosomas portadores del haplotipo H2. Los genes sensibles a la dosificación flanqueados por LCR están representados por asteriscos. Los LCR están representados por flechas horizontales de colores combinados(27).

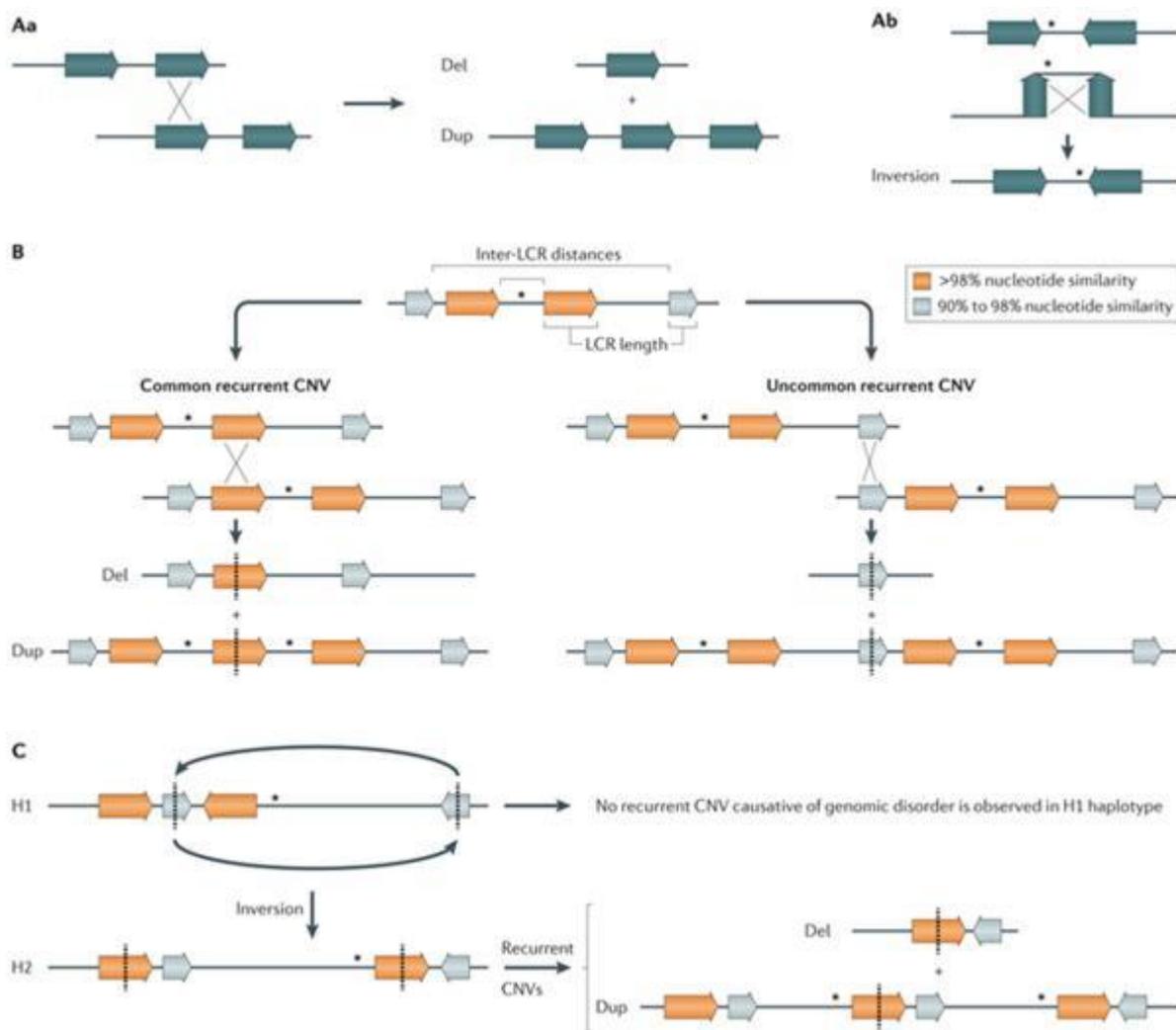


Figura 2: Mecanismo Nonallelic homologous recombination tomada de Carvalho, C. M., & Lupski, J. R.(27)

Como ya se ha mencionado el marcador genético TPOX presenta un tercer alelo en algunos individuos, que pudo haberse originado a través del mecanismo NAHR, característica que en ciertas familias se hereda a sus hijos. Así, al realizar un alineamiento del intrón 10 del gen TPO se pueden observar diversos lugares del genoma con una homología mayor del 90% en los cromosomas 19, 13, 7, 5 y en el X, específicamente X:13665448-13665707(28), que podría dar explicación a la segregación descrita en la familia de población afrodescendiente(29).

Actualmente, el TPOX es un microsatélite incluido en los marcadores más usados en genética forense en todo el mundo, siendo los más comunes; D1S1656, TPOX, D2S441, D2S1338, D3S1358, FGA, CSF1PO, D5S818, SE33, D6S1043, D7S820, D8S1179, D10S1248, TH01, vWA, D12S391, D13S317, Penta E, D16S539, D18S51, D19S433, D21S11, Penta D, y D22S1045(30), estos son generalmente seleccionados para el análisis forense porque son altamente polimórficos, tetraméricos, y no están ligados (31).

3.2 Población de referencia

Frente a las pruebas de ADN generalmente se comparan perfiles genéticos y para ello, se deben tener frecuencias alélicas de la población de referencia, obtenidas a través de un muestreo de diversos individuos no emparentados entre sí (32). Actualmente, los laboratorios pueden emplear las frecuencias alélicas de muestreos genéticos de población colombiana, con un mínimo de 700 individuos (33).

3.2.1 Parámetros estadísticos de relevancia en genética forense

Una población humana es un grupo de individuos que se reproducen entre sí y pueden presentar diferencias génicas respecto a grupos humanos vecinos; es indispensable poseer el soporte estadístico de la variabilidad de cada población mediante la proporción relativa de alelos (34).

Un principio básico en la genética de poblaciones es la Ley del equilibrio de Hardy-Weinberg, que contiene tres principios fundamentales (35):

- 1) En una población las frecuencias alélicas en un locus autosómico no cambian de una generación a la siguiente (equilibrio de las frecuencias alélicas).
- 2) Las frecuencias genotípicas de la población están determinadas de una manera predecible por las frecuencias génicas (equilibrio de las frecuencias genotípicas).
- 3) El equilibrio es neutro, es decir, si se perturba la población el equilibrio se restablecerá en una sola generación de apareamiento al azar, pero con las nuevas frecuencias alélicas y si todos los restantes requisitos se mantienen.

“Como las poblaciones naturales difícilmente se encuentran en equilibrio estricto, hay que determinar si las desviaciones observadas en la muestra pueden atribuirse simplemente al azar o si por el contrario, son tan grandes que lleven, en primer instancia, a investigar una posible ocurrencia de errores sistemáticos en el genotipado o luego permiten afirmar la existencia de algún proceso que esté desviando la población de la situación de equilibrio (mutación, selección, migración, deriva génica, apareamientos preferentes)”(35).

Cuando existe un alto nivel de endogamia o aislamiento en una población puede generarse un exceso de homocigotos y se considera que todo el genoma es afectado y no un locus en concreto, por lo que sería lógico encontrar que la tendencia a la disminución del número de heterocigotos es un proceso general o al menos extendido y sirve como indicio de subestructura poblacional (35).

La subestructura poblacional puede estimarse a través del estadístico F de Wright (1951), ya que una población lleva consigo un efecto similar a la endogamia; y puede utilizarse el "índice de fijación" (simbolizado por F_{ST}), que mide la reducción en la heterocigosis debida a diferenciación genética entre poblaciones (36).

“ F_{ST} es siempre mayor que (o igual a) cero, debido al efecto Wahlund, que asegura que $H_T > H_S$ media. Si todas las subpoblaciones están en equilibrio Hardy-Weinberg con las mismas frecuencias alélicas, entonces $F_{ST} = 0$. A medida que F_{ST} aumenta, aumenta la estructuración, aislamiento y diferencias entre las poblaciones. Donde H_T es el promedio de la heterocigosis esperada en la población total para todos los loci y H_S es el promedio de la heterocigosis esperada dentro de subpoblaciones para todos los loci”(4). Sumado a ello, el F_{ST} puede ser usado como factor de corrección “ θ ” para el cálculo del LR (*Likelihood Ratio*) (37).

Teniendo en cuenta lo anterior a cada marcador STR se le realiza los siguientes parámetros para determinar si son útiles para el uso forense:

- Poder de exclusión (PE): Es la probabilidad de excluir a una persona aleatoria como contribuyente de un alelo en el locus (38). Esta dada por la siguiente formula:

$$PE = H^2(1 - (1-H) H^2)$$

Donde H= Heterocigosidad

- Contenido de información polimórfica (PIC): Refleja la probabilidad de que una descendencia determinada de un padre portador de un alelo en un locus permita la deducción del genotipo parental en el locus y depende del número de alelos de ese marcador y de sus frecuencias relativas en la población (39). La fórmula es la siguiente:

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

Donde n es el número de alelos y p_i es la probabilidad del alelo i

- Probabilidad de coincidencia (P_m): Es la probabilidad de que dos individuos elegidos al azar en una población tengan el mismo perfil genético en uno o varios marcadores y puede simplificarse en la siguiente fórmula (40):

$$P_m = \sum P_i^2$$

Donde P_i es la frecuencia del genotipo i .

- Poder de discriminación (PD): Es la probabilidad de que dos personas elegidas al azar de la población tengan distinto perfil genético y se calcula de la siguiente manera (40):

$$PD = 1 - P_m$$

- Frecuencia alélica mínima (MAF): Algunos alelos ocurren en menor frecuencia que otros por lo que la frecuencia mínima para estos alelos se obtiene de dividir cinco entre $2N$ donde “N” es el número de muestras totales analizadas para la base de datos (41).

$$MAF = \frac{5}{2N}$$

3.2.2 Poblamiento de Colombia

Con el fin de entender las diferencias entre las poblaciones, hay que conocer su historia. Se sabe que entre el Pleistoceno y el Holoceno, llegaban los primeros seres humanos a lo que hoy es Colombia, creando las condiciones físicas y psicológicas para tener éxito en la colonización de los ecosistemas colombianos y en la construcción de territorios humanizados (42). A su llegada al territorio colombiano, las poblaciones europeas se encontraron con una multiplicidad de pueblos, con diversos grados de desarrollo en su actividad productiva y su organización social, y con diferentes referentes simbólicos. Pero su intención de convertir el nuevo mundo en unidad territorial, terminó segregando la diversidad (43). Por otro lado, para los trabajos en minas, haciendas y en la navegación por el río, se trajeron esclavos africanos a partir de 1590. Luego de esto ocurrió un aumento en la población mestiza, indígena y africana, que hacia finales del siglo XVIII ya conformaban la mayoría de la población colonial (44).

Por consiguiente, en Antioquia antes de la llegada de los españoles habitaban diversos grupos indígenas principalmente de dos grandes etnias: los Caribes y los Chibchas (45). La primera hueste de españoles en actitud de conquista que llegó a las tierras de lo que después sería Antioquia

arribaron hacia 1502 al lado oriental del Golfo de Urabá. Con la llegada de estos primeros españoles a Urabá se daba inicio al contacto entre los pueblos americanos, pobladores ancestrales de estas tierras y un conjunto de nuevas realidades que habrían de producir cambios notables y marcarían un nuevo rumbo a la historia de estas comunidades humanas (46). Según los documentos coloniales, el grado de centralización político y la escala territorial de los cacicazgos de la región eran muy limitados (47), además disminuyeron su población notablemente desde el siglo XVI con la represión española (48). La mezcla de las diferentes etnias culturales dio lugar a mestizos entre españoles e indígenas y negros, aunque también existía mezcla entre negros e indígenas (Zambaje), ya que las servidumbres de aquella época convivían y había escasez de mujeres negras (48). En la ubicación del actual municipio de frontino estuvo poblada por indios de Guaca, y como ocurría tan pronto se fundaba una ciudad, sus autoridades procedían a distribuir a los indios entre los conquistadores, y el tipo de economía de expoliación aplicada por los europeos, mantendría a los aborígenes en permanente rebelión durante décadas (49).

La población mestiza siempre fue mayoritaria en los censos realizados a partir del siglo XVII, sin embargo, parecía que “su única posibilidad cultural pasaba por "blanquearse" y esto significaba una especie de condición esquizofrénica desde el punto de vista de sus valores, ideas y símbolos: tratar de ser y parecer "blanco" pero NO ser blanco” (48).

Por otro lado, en el Departamento del Chocó se describe que “hace aproximadamente 2.300 años, existía ya en el Urabá chocoano, conocido con el nombre de Darién, un complejo cultural, conformado por cazadores y pescadores sedentarios que culminó en una fase agrícola. A la llegada de los europeos a América, poblaban el territorio chocoano agrupaciones indígenas conocidas

como los Kunas, Chocoes, Noanamaes (Kunas, Emberas y Waunanas) (50–52). La riqueza en recursos naturales como el oro en el Chocó, estaba ligada a la abundante presencia de esclavos negros en la costa pacífica, aunque los esclavos negros huían de sus amos para construir fortalezas y defenderse, a estos se les conoce como palenques y al acto de rebelión se le daba el nombre de cimarronismo. Todo lo anterior explica porque la población actual del Chocó es mayoritariamente Afrodescendiente (50–54).

De acuerdo a todo lo mencionado, estos hechos de mezcla entre poblaciones, incluyendo la población con origen africano permiten tener una gran diversidad en Colombia, y puede explicar que el patrón trialélico con el alelo extra “10” del marcador TPOX, proviene desde alguna de las embarcaciones de esclavos en algún momento del poblamiento después de 1590.

3.3 Cálculos de filiación: tríos y dúos

Los cálculos en filiación se basan en el IP (índice de paternidad) equivalente al LR (*Likelihood ratio*), donde se obtiene una probabilidad a favor o en contra de la paternidad. Las probabilidades obtenidas a través de la evidencia con el uso de marcadores genéticos, nunca tiene una certeza del 100%, pero existen pautas para catalogar una filiación positiva, como en el artículo 7 de la ley 721 del 2001 de Colombia, la cual designa que para establecer paternidad o maternidad, el índice de probabilidad debe ser superior al 99.9% (55). Aunado a ello la ISFG (*International Society for Forensic Genetics*) recalca que cada laboratorio es responsable de establecer los umbrales de interpretación en los test de paternidad (56).

Las fórmulas usadas en estas pruebas de filiación se basan en la probabilidad de las hipótesis planteadas, como se observa a continuación (56):

$$IP = \frac{\text{Probabilidad(Genotipo observado|H0:el hombre al que se le realiza la prueba sea el padre)}}{\text{Probabilidad(Genotipo observado|H1:cualquier hombre de la poblacion sea el padre)}}$$

El valor IP es obtenido de cada marcador genético utilizado, donde el valor varía según las frecuencias alélicas de la población, a partir de los valores IP de cada marcador es posible tener un índice de paternidad combinado (IPC), como producto final de los IP (56):

$$IPC = \prod_{i=1}^N P_i.$$

Donde $i=1$, hasta N , índices calculados de cada sistema de marcadores genéticos, a partir de esto se puede obtener una probabilidad posterior basada en Bayes (57):

$$Pr(H0|\text{genotipo}) = \frac{Pr(\text{genotipo}|H0)Pr(H0)}{Pr(\text{genotipo}|H0)Pr(H0) + Pr(\text{genotipo}|H1)Pr(H1)}.$$

Generalmente la probabilidad posterior se representa con la letra W una abreviación del alemán “Wahrscheinlichkeit”, introducida por el sueco Essen- Möller. El planteamiento del teorema de Bayes puede simplificarse suponiendo un prior $Pr(H0)=Pr(H1)=0.5$, es decir que la probabilidad de $H0$ es la misma probabilidad de que sea $H1$, obteniendo la siguiente formula (57):

$$W = \frac{Pr(\text{genotipo}|H0)}{Pr(\text{genotipo}|H0) + Pr(\text{genotipo}|H1)} = \frac{IPC}{IPC + 1} = \frac{LR}{LR + 1}$$

De acuerdo a lo anterior se plantea la siguiente tabla con las fórmulas simplificadas para casos de paternidad, donde A, B y R representan alelos diferentes de un locus dado y las frecuencias de los alelos en la población se representan con a, b y r:

	HIJO	MADRE	PRESUNTO PADRE	NUMERADOR	DENOMINADOR	INDICE DE PATERNIDAD (IP). p y q representan las frecuencias del alelo en la población.
1	AA	AA	AA	1	a	1/a
2			AB	1/2	a	1/2a
4		AB	AA	1/2	a/2	1/a
5			AB	1/4	a/2	1/2a
6			AR	1/4	a/2	1/2a
7	AB	AA	BB	1	b	1/b
8			AB	1/2	b	1/2b
9			BR	1/2	b	1/2b
10		AB	AA	1/2	(a+b)/2	1/(a+b)
11			AB	1/2	(a+b)/2	1/(a+b)
12			AR	1/4	(a+b)/2	1/2(a+b)
13			BR	1/4	(a+b)/2	1/2(a+b)
14		BR	AA	1/2	a/2	1/a
15			AB	1/4	a/2	1/2a

Tabla 1: Modificada de cálculos y fórmulas de prueba de paternidad de un solo locus para 21 diferentes combinaciones de genotipos de trío completo del libro *Advanced Topics in Forensic DNA Typing Methodology, chapter 14: Relationship Testing: Kinship Statistics, 2015(58)*

Las fórmulas se fundamentan en el teorema de Bayes; por ejemplo, con el padre AB, madre AB e hijo AB, el numerador generalmente es representado con X y se plantea lo siguiente:

$$X=2ab2ab ((1/2) (1/2)+(1/2) (1/2))=4a^2b^20.5=2a^2b^2$$

Donde 2ab representa la probabilidad de heterocigotos tanto del padre como de la madre, y el 0.5 representa la probabilidad de transmisión del alelo A o B por parte de los padres. Por otro lado, el denominador se representa generalmente con Y obteniendo lo siguiente:

$$Y=2ab \cdot 2ab \cdot (1/2)(a+b) = 2a^2b^2(a+b)$$

Para el denominador hay que tener en cuenta la probabilidad de que cualquier hombre de la población sea el padre, por eso se adiciona la probabilidad de que haya transmitido al hijo el alelo A o el alelo B (representado por $a+b$) y la probabilidad de transmisión de alguno de esos dos alelos es de 0.5.

$$\text{Entonces, } IP = \frac{2a^2b^2}{2a^2b^2(a+b)} = 1/(a+b)$$

Del mismo modo, las pruebas de parentesco con dos individuos se basan en la teoría bayesiana, incluyendo los casos de filiación con el presunto padre y el hijo. Las fórmulas para dúos se observan a continuación:

	HIJO	PRESUNTO PADRE	NUMERADOR	DENOMINADO R	INDICE DE PATERNIDAD (IP)
1	AA	AA	a	a^2	$1/a$
2		AB	$a/2$	a^2	$1/2a$
3		AR	$a/2$	a^2	$1/2a$
4	AB	BB	$1a$	$2ab$	$1/2b$
5		AB	$(a+b)/2$	$2ab$	$(a+b)/4ab$
6		BR	$a/2$	$2ab$	$1/4b$
7		AA	b	$2ab$	$1/2a$
8		AR	$b/2$	$2ab$	$1/4a$

Tabla 2: Cálculos y fórmulas para pruebas de filiación con un solo progenitor, Tabla modificada de Estadística aplicada a la genética forense, Luque J(59).

En los casos donde no se cuenta con la información de la madre, se deben tener en cuenta todas las posibilidades a la hora de plantear las diferentes fórmulas, generalmente se basa en la

probabilidad de compartir alelos idénticos por descendencia (IBD) y se derivan de la siguiente tabla:

	GENOTIPO	I=0	I=1	I=2
1	AA,AA	a^4	a^3	a^2
2	AA,AB	$2a^3b$	a^2b	0
3	AA,BB	a^2b^2	0	0
4	AA,BC	a^22bc	0	0
5	AB,AB	$4a^2b^2$	$ab(a+b)$	$2ab$
6	AB,AC	$4a^2bc$	abc	0
7	AB,CD	$4abcd$	0	0

Tabla 3: . Probabilidades por pares de genotipos en función del número de alelos compartidos, representados con la letra I, es decir cuando I=0 no comparten alelos IBD, si I=1 comparten un alelo IBD y si I=2 comparten 2 alelos IBD. Tabla tomada del libro *Relationship Inference whit Familias and R*(57).

Como ya se ha mencionado, la probabilidad de los alelos IBD es importante a la hora de plantear la solución, por ejemplo, para un padre con genotipo $P_A P_B$ e hijo $P_A P_B$, se forma el numerador de la siguiente manera:

$$X=(0)4a^2b^2 + (1)a b(a + b) + (0)2a b$$

Donde se tiene en cuenta la probabilidad de que compartan 1 o 2 alelos o que no compartan. Así mismo se multiplica cada probabilidad de transmisión observando la genealogía, por ejemplo, la probabilidad de que se transmita un alelo IBD es de 1 para padre e hijo, o de que no transmita es de 0, como puede observarse en la figura 3.

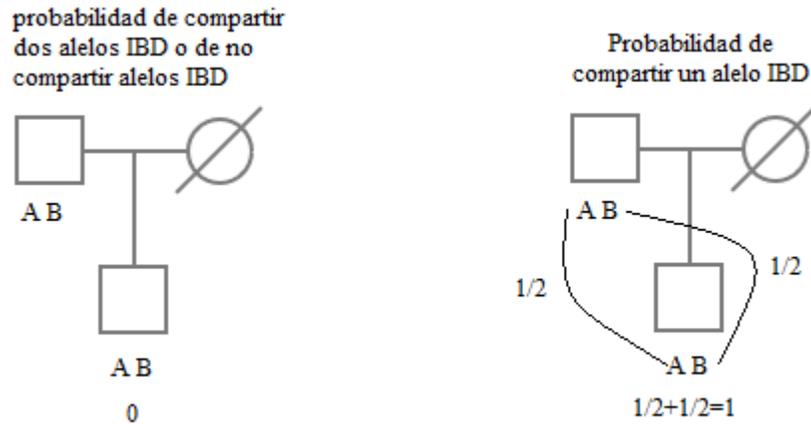


Figura 3; Imagen creada con el programa progeny. Probabilidades de compartir alelos IBD Padre-Hijo, suponiendo que el padre y la madre no están emparentados.

Donde el LR se expresa de la siguiente manera:

$$LR = \frac{(0)4a^2b^2 + (1)ab(a+b) + (0)2ab}{4a^2b^2}$$

$$LR = \frac{ab(a+b)}{4a^2b^2} = \frac{(a+b)}{4ab}$$

3.4 Next Generation Sequencing (NGS) y microarrays en búsqueda de variantes estructurales.

Las variantes estructurales (SV) en el ADN se dan como resultado de mecanismos mutacionales, incluyendo recombinación, replicación y procesos asociados a la reparación. Las SV son inserciones, deleciones, elementos móviles, repeticiones en tándem, translocaciones y variaciones

en el número de copias. Las *SV* pueden ser analizadas mediante secuencia de genoma completo, en búsqueda de puntos de ruptura, diferencias en la profundidad y discordancias en el mapeo con el genoma de referencia (22). En la figura 4 se resumen las estrategias de los diversos softwares usados para el análisis de secuencias que pueden ser usados en el análisis de la posible duplicación del TPOX.

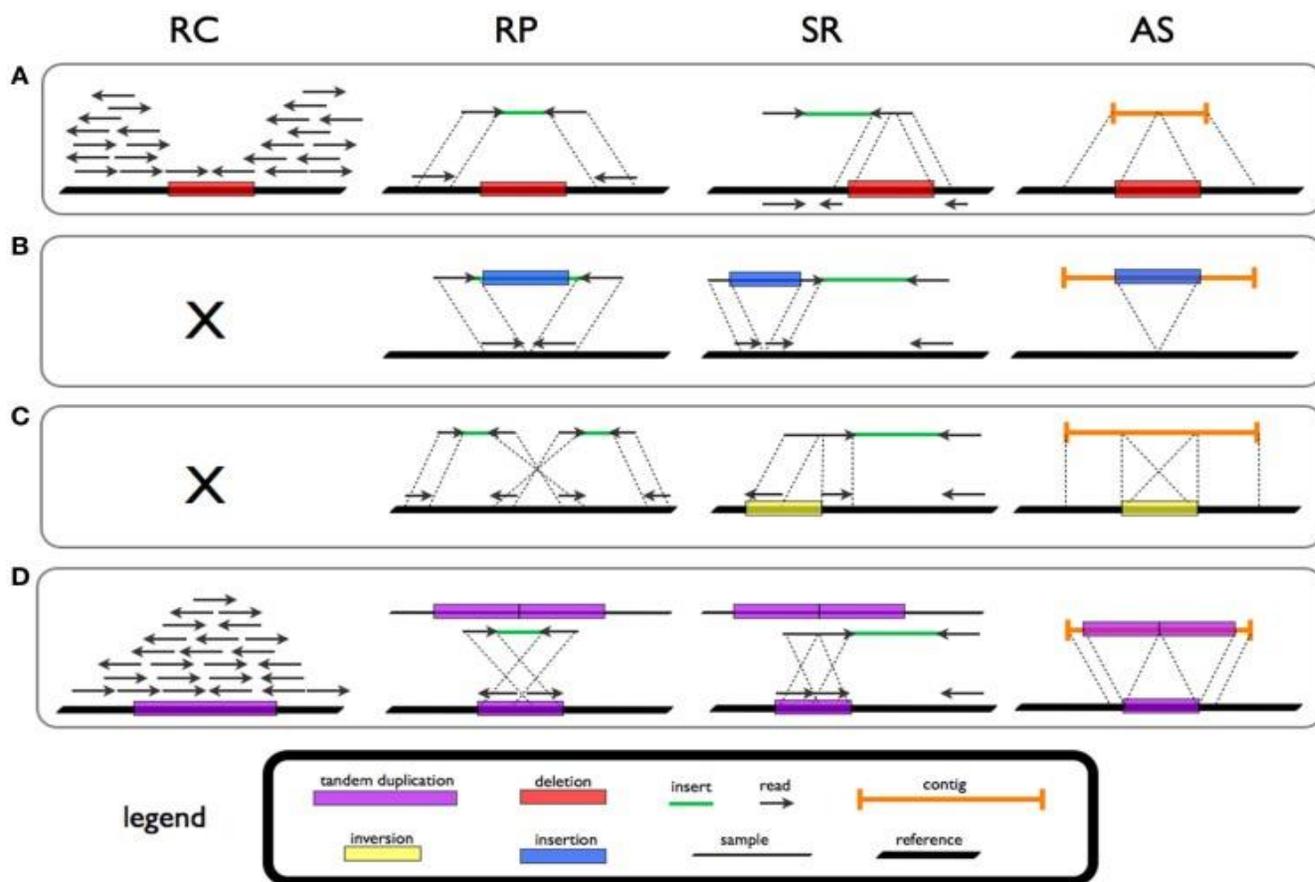


Figura 4: tomada de Tattini, L., D'Aurizio, R., y Magi, A. Patrones de SV para deleciones (A), nueva secuencia de inserción (B), inversión (C) y duplicación en tándem (D), en métodos como; read-count (RC), read-pair (RP), split-read (SR), y de novo assembly (AS)(60).

Los métodos de *read-pair* (RP) se basan en la evaluación del intervalo y la orientación de las lecturas (*reads*) de extremos pareados. Se recopilan los pares discordantes en los que el intervalo de mapeo y / o la orientación de los pares de *reads* son inconsistentes con el tamaño de inserción

esperado. Se pueden investigar varias clases de SV mediante este enfoque. El mapeo de pares de *reads* demasiado alejado está asociado a deleciones, mientras que los que se encuentran más cerca de lo esperado son indicativos de inserciones. Además, las inconsistencias de orientación pueden representar inversiones y una clase específica de duplicaciones en tándem(60).

Los enfoques de *read-count (RC)* asumen una distribución aleatoria (Poisson o Poisson modificada) en profundidad de mapeo e investigan la divergencia de esta distribución para resaltar las duplicaciones y deleciones. La secuencia de las regiones duplicadas da como resultado una mayor profundidad de *reads*, mientras que las regiones eliminadas muestran una profundidad de *reads* reducida en comparación con las regiones normales(61).

Los métodos de *split-read (SR)* permiten la detección de SV con una resolución de par de base única. La presencia de un punto de ruptura de SV se investiga sobre la base de una lectura de secuencia dividida que rompe la alineación con la referencia (60).

Teóricamente, todas las formas de variación estructural podrían investigarse por medio de métodos de *novu assembly (AS)*. El AS se refiere a fusionar y ordenar fragmentos cortos para volver a ensamblar la secuencia original a partir de la cual se tomaron muestras de los fragmentos cortos, aunque las características intrínsecas de los datos de NGS, como la longitud de los *reads* (corta), limitan el uso de enfoques AS para la investigación de variantes (62).

Además, una clase específica de SV, inserciones de elementos móviles (*ME*), se puede investigar explotando la información de *reads* discordantes y recortados (*DC*) (60). Los *ME* son secuencias de ADN repetitivas que pueden cambiar de posición dentro del genoma, debido a esta característica su detección es desafiante. Las últimas estimaciones sugieren que más de la mitad del genoma

humano es repetitivo o derivado de repetición (63). Si bien el enfoque de *DC* se puede atribuir a los métodos *RP* y *SR*, las parejas de *reads* se asignan a una biblioteca personalizada y configurable de secuencias consenso activas de *ME* conocidas (64).

Por otro lado, los microarreglos o *arrays* de hibridación genómica comparada (aCGH), muestran la cantidad de ADN en diferentes loci del genoma al confrontar las concentraciones relativas de ADN de 2 genomas (control y paciente), marcados con fluorocromos diferentes, que se empalman con fragmentos de ADN de secuencia conocida adheridas a un soporte de cristal. El matiz de la fluorescencia en los diferentes sitios del aCGH indica la concentración relativa de cada secuencia de ADN y permite observar la existencia de ganancias o de pérdidas en zonas específicas del genoma (65,66). La principal ventaja de aCGH es la capacidad de detectar simultáneamente aneuploidías, deleciones, duplicaciones y / o amplificaciones de cualquier locus representado en una matriz, también posee ventaja sobre la hibridación fluorescente in situ (FISH), ya que la matriz CGH es un experimento FISH coordinado y concurrente en cientos o miles de loci (67), como se observa en el ejemplo de la figura 5.

Arrays de CGH



Figura 5: Tomado de Rawe V, et Cuzzi j. Representación esquemática del procedimiento de CGH con arrays (68).

4 Planteamiento del problema

Actualmente uno de los marcadores STR más usados en genética forense es el TPOX que está ligado al gen de la peroxidasa tiroidea humana (TPO), ubicado en el cromosoma 2:1489617-1489848, específicamente en el intron 10. El gen TPO tiene un tamaño aproximado de 150Kb, compuesto por 17 exones, donde cuatro exones únicos en el gen (exones 13-16) abarcan un módulo de proteínas diferentes. Tres de estos módulos, que representan los exones 13, 14 y 15, presentan similitudes significativas con la C4b / beta-2-glicoproteína, el receptor EGF / LDL, y un dominio transmembrana típico, respectivamente. Los genes que codifican estos módulos probablemente se fusionaron a un gen de la peroxidasa ancestral para generar el presente gen TPO. Por lo tanto, la pérdida o ganancia de intrones, y el arrastramiento del exón han desempeñado un papel importante en la evolución del gen TPO (69).

Aunado a ello, el gen TPO codifica para una glicoproteína de membrana, que ayuda a la yodación de residuos de tirosina en la formación de tiroglobulina y fenoxiéster entre pares de tirosinas yodadas para generar las hormonas tiroideas, tiroxina y triyodotironina. Las mutaciones en este gen pueden producir varias alteraciones de hormogenesis, como el hipotiroidismo congénito, el bocio congénito y el defecto de la organización de la hormona tiroidea IIA(70). Al igual que los otros loci STR de uso forense, el TPOX está constituido por repeticiones de tetranucleótidos, particularmente por la secuencia AATG, que se repite aproximadamente entre 6 a 14 veces (71,72); aunque a nivel de la población en general, el 90% de los individuos presentan los alelos 8, 9 u 11 (73). En general, en Colombia, se ha observado que los alelos 8,9 y 11 tienen una representación del 82% (74), mientras que en la población antioqueña y chocoana se encuentran con una frecuencia del 79% (11) y 77% respectivamente (75). También, dicho marcador TPOX puede

presentar perfiles genéticos con patrones trialélicos, debido a eventos inusuales como mutaciones, que pueden causar un número irregular de alelos en algunos individuos (7,76).

Asimismo, los perfiles genéticos con patrones trialélicos se caracterizaron hacia el año 2004, clasificados por el *Forensic Science Service* (FSS) de acuerdo a la intensidad que presentaba cada alelo. Se establecieron dos modelos a saber, el Tipo 1, en el cual los tres alelos exhibían diferente intensidad, indicando que se trataba de una mutación somática en un locus heterocigoto, mientras que en el Tipo 2, se observaban tres alelos de igual intensidad, sugiriendo una duplicación (77). Esto se puede observar en la Figura 6.

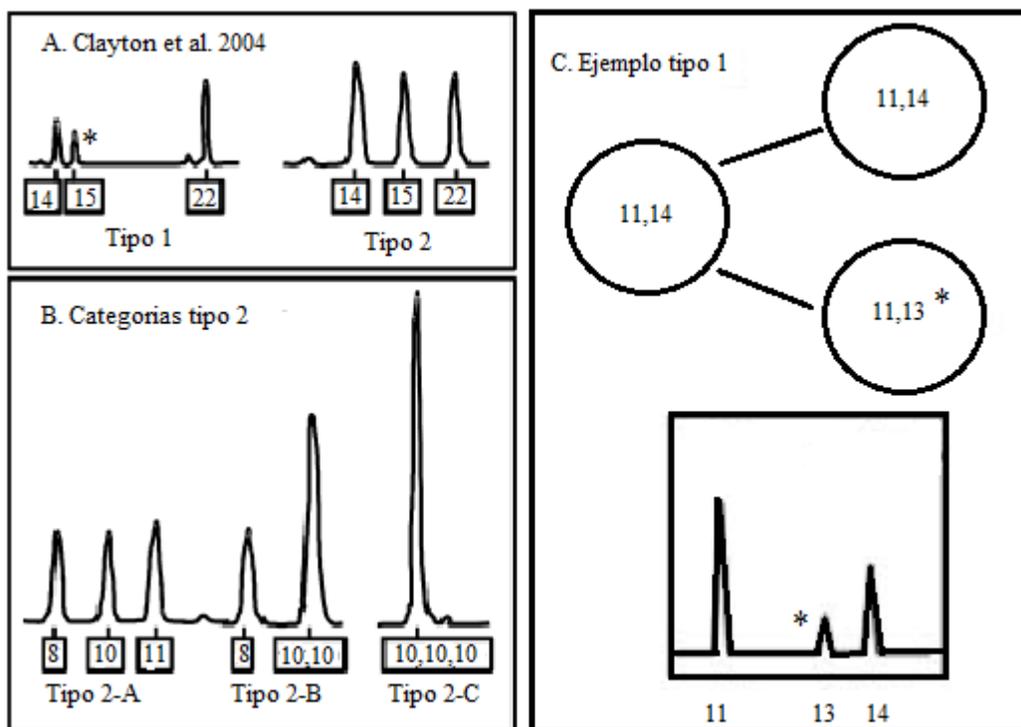


Figura 6: Imagen modificada de Clayton et al. En la figura (A), se observa un patrón trialélico tipo 1 y tipo 2 propuesto por Clayton et al. En el año 2004, donde el alelo mutante putativo se denomina por *. El patrón tipo 1 es debido a dos líneas celulares, el cual una de ellas contiene la variación y el tipo 2 puede deberse a una duplicación del marcador. En la figura (B), se subclasifica el tipo 2 según Picanço et al. En el año 2015, Tipo 2-A (8-10-11): tres picos regulares de tres alelos diferentes; Tipo 2-B (8-10-10): una dosis doble pico y un pico regular; Tipo 2-C: un solo pico triple dosis. En la figura (C) se muestra un ejemplo de dos líneas celulares, una línea celular con la mutación presente y la otra línea celular normal, además incluye el electroferograma del amplificado.

Además, los patrones trialélicos tipo 2-B y tipo 2-C pueden omitirse en la observación de un electroferograma, dado que el alelo extra contiene el mismo número de repeticiones que el ancestral, pero el tipo 2-A se puede observar claramente el alelo extra pudiendo deberse a los mecanismos de mutación que alteran el número de repeticiones en el alelo (16).

Estudios previos de los patrones trialélicos del marcador TPOX sugieren que esta característica tiene origen africano, específicamente del sur del continente, dado que aproximadamente el 2,4% de los sudafricanos poseían un genotipo de tres alelos para este marcador (29), que puede deberse a una duplicación dentro del mismo cromosoma 2 o una duplicación de la secuencia de este marcador en el cromosoma X (29,76,77). Además, hay evidencia que la mutación no se quedó en territorio africano, testimonio de ello, son varios estudios realizados en América. Por ejemplo en Brasil, Fridman y colaboradores encontraron un TPOX trialélico en 1 de 561 (0,00178) muestras que componían su estudio (78). Allí mismo, años después, hallaron igual patrón en 88 muestras de 150.226 (0,00059) (79). También Díaz y colaboradores observaron la presencia de la variante en República Dominicana (80). Así pues, aunque los hallazgos son pocos y las frecuencias bajas, dan prueba de la llegada de tal genotipo por parte de la población sudafricana, como se aprecia en la figura 7.

Dados los hallazgos de estas variantes trialélicas en el marcador TPOX a nivel mundial, el *National Institute of Standards and Technology* (NIST), que cuenta con una base de datos en línea de los marcadores STR utilizados en la práctica forense, tiene un apartado exclusivo alusivo a los patrones trialélicos de este marcador, construida a partir de los reportes realizados por diversas instituciones alrededor del mundo. Actualmente existe una recopilación de 120 reportes de

genotipos trialélicos, no obstante, entre estos 120 datos, solamente se encuentran 19 patrones diferentes, como se puede ver en la tabla 1, que es una adaptación de la base de datos original, que puede ser encontrada en la página web http://www.cstl.nist.gov/strbase/var_TPOX.htm#Tri.

Genotipo	Kit	Instrumento	Frecuencia	Confirmación	Contribución
6, 8, 10	CO	ABI377	1/39000	SI	Chris Piwonka
6, 9, 10	CO	ABI377	1/39000	SI	Kelly Solis
6, 10, 11	CO	ABI377	1/39000	SI	Chris Piwonka
6, 10, 12	CO	ABI310	N/A	SI	Diane Hanson
7, 8, 10	PP16	ABI310	1/1100	SI	Angelo Abati
7, 9, 10	CO	ABI3100	1/20000	SI	Neil Quigley
7, 10, 11	CO	ABI377	1/20000	SI	Karl Morton
8, 9, 9	ID	ABI3500	N/A	SI	Brittany Porter
8, 9, 10	CO	ABI377	1/20000	SI	Karl Morton
8, 9, 11	PR-ID	ABI3100	N/A	SI	Chun Byun-Won
8, 10, 11	CO	ABI377	1/39000	SI	Chris Piwonka
8, 10, 12	CO	ABI377	1/39000	SI	Chris Piwonka
8, 11, 12	PS-CO	ABI3100	1/10613	SI	Misty Marra
8, 11, 14.3	ID	ABI3130;ABI3130xl	N/A	SI	Valerie Huebner
8, 10, 11	CO	ABI377	N/A	SI	Karl Morton
9, 10, 12	CO	ABI377	1/20000	SI	Shantel Kaster
9, 11, 12	PP16	ABI3100	1/28252	SI	Sun Hongyu <i>et al</i>
10, 10, 11	CO	ABI3100	N/A	SI	Garon Foster
10, 11, 12	CO	ABI377;ABI310	1/30000	SI	William Stubbs

Tabla 4: adaptada de STRbase del NIST. Reporte de patrones trialélicos para el STR TPOX: CO: Applied Biosystems COfiler, PR: Applied Biosystems Profiler, PS: Applied Biosystems Profiler Plus, ID: Applied Biosystems Identifier, PP16: Promega PowerPlex 16.

De los 19 modelos trialélicos que han sido informados, 14 presentan el alelo 10 en su genotipo. Sin embargo, en estos perfiles, se puede notar la existencia de un alelo 9 u 11, pudiendo así derivarse de un alelo 10 ancestral (71). Esto puede deberse a cambios en la longitud de un STR, bien sea por ganancia o pérdida de sus unidades de repetición (81). Este tipo de mutaciones, pueden generarse por *slippage* (82).

En Colombia los afrodescendientes ocupan alrededor del 10,6% del total de la población, aunque el mayor asentamiento de individuos con ascendencia africana están en el departamento de Chocó, localizado en la región del Pacífico Colombiano (83). Justamente en individuos de esa población se observó un caso de patrón trialélico del marcador TPOX, en una prueba de paternidad de rutina llevada a cabo en el Laboratorio de Identificación Genética – IdentiGEN de la Universidad de Antioquia.



Figura 7: Mapa tomado de Vélez et al(84), y modificado de <http://www.lahistoriaconmapas.com/atlas/inglaterra-mapas/el-croquis-del-mundo-mapa.htm>. Expansión de la mutación del marcador genético TPOX desde África hacia América. En verde oscuro, se representa la población africana estudiada por el Dr. Lane. En azul, los países americanos donde se ha encontrado la mutación (Brasil, República Dominicana y Alabama (EEUU)). En verde claro, este estudio llevado a cabo en una familia de Chocó, Colombia. Las flechas indican la expansión desde África. Crouse et al, no indican en su estudio el posible origen del genotipo TPOX trialélico.

Asimismo, en un estudio previo en el laboratorio IdentiGEN se analizaron 17 individuos pertenecientes a un grupo familiar procedente del departamento de Chocó, Colombia. Después de genotipificar cada una de las muestras, se encontró que 5 de ellos eran portadores del patrón trialélico del locus TPOX, concluyendo que el alelo 10 segregaba de manera independiente posiblemente en el cromosoma X (84).

Es importante analizar los patrones trialélicos debido a que pueden afectar el producto total del *LR* y de esa forma influir en el resultado final de la prueba de ADN (79). Además, en el cálculo del *LR* solo se describen marcadores diploides, por lo cual no existen pautas para el cálculo de los marcadores trialélicos, representando un problema a la hora de resolver los casos en identificación genética, incluso en algunos casos se recomienda no tener en cuenta el marcador trialélico para el análisis final (85). A raíz de esto se hace necesario crear pautas en base a las recomendaciones de organizaciones internacionales (*Federal Bureau of Investigation, National Institute of Standards and Technology, International Society for Forensic Genetics* y el Grupo de Habla Española y portuguesa de la *International Society for Forensic Genetics*), que permitan un adecuado desarrollo de los casos que presentan perfiles genéticos trialélicos. En resumen, un marcador trialélico permite aumentar el valor final del *LR* y por ende aumenta el poder de exclusión, discriminación e identificación, y si la transmisión de padre a hijo es clara, la interpretación estadística podría mejorarse durante la práctica forense (79).

Por otra parte, la investigación de los perfiles de ancestría genética de las poblaciones humanas es una herramienta valiosa para comprender la dinámica de las migraciones y los eventos de colonización, así como para determinar los patrones de mezcla dentro de las poblaciones (86),

aunado a ello, se sugiere una estrecha relación entre la distancia geográfica y la diferenciación genética (87), por consiguiente, el componente ancestral sirve como herramienta para explorar la historia de las familias del presente estudio.

Finalmente, en el campo de la genética forense, tener un conjunto de marcadores que puedan proporcionar estimaciones del componente ancestral, puede ayudar a identificar la población de origen del donante de cierta evidencia y ayudar a dirigir la investigación criminal (86).

5 Objetivos del proyecto

5.1 Objetivo general

Describir el posible origen de los patrones trialélicos en el locus TPOX en dos familias colombianas y su importancia en la identificación forense.

5.2 Objetivos específicos

- Analizar la segregación del locus TPOX en dos familias colombianas
- Describir la ancestría de las dos familias por medio de 46 Indels-AIMs
- Calcular los parámetros estadísticos de relevancia en genética forense con datos recolectados desde el año 2011 hasta el año 2016 en el locus TPOX y buscar patrones trialélicos tipo 2-B.
- Evaluar la importancia de los patrones trialélicos del locus TPOX en la identificación forense.
- Explorar la posible localización genómica del tercer alelo observado en el marcador TPOX.

6 Metodología

Población: Familias portadoras del genotipo trialélico del marcador TPOX observadas en el laboratorio IdentiGEN.

Muestreo y observación: Se toman los resultados de las bases de datos del laboratorio IdentiGEN desde el año 2011 hasta el año 2016 con el fin de buscar patrones trialélicos tipo 2-B en el locus TPOX.

Las dos familias sobre las que se hace el análisis de segregación del locus TPOX también fueron obtenidas a partir de un individuo portador que dio el consentimiento para el análisis familiar.

Unidad de observación: Individuos portadores del patrón trialélico del locus TPOX dentro de las familias a analizar.

Unidad de análisis: Las familias portadoras del patrón trialélico del locus TPOX observadas en el laboratorio IdentiGEN.

Criterios de inclusión: Individuos con patrones trialélicos incluidos en la base de datos del laboratorio IdentiGEN. Familias que denotan el patrón trialélico del locus TPOX obtenidas a partir de un individuo portador en los estudios de rutina del laboratorio IdentiGEN.

Criterios de exclusión: Toda población cuya muestra no haya sido obtenida dentro del laboratorio de Identificación Genética y las familias no portadoras del patrón trialélico del marcador TPOX.

Muestreo: Se realiza el análisis de toda la población entre los años 2011 y 2017 en busca de patrones trialélicos tipo 2-B.

Se estudian individuos provenientes de familias colombianas, con el fin de establecer en ellos la presencia del genotipo trialélico en el marcador TPOX. De cada individuo se toma una muestra de sangre sobre tarjetas FTA® de Whatman. Se amplifican STRs autosómicos con el kit Identifiler de la casa comercial Applied Biosystems, que contiene el marcador de interés y también se confirman con el kit PowerPlex 16, de la casa comercial promega..

Los resultados obtenidos en la amplificación del ADN de interés se someten a electroforesis capilar en el Analizador Genético AB3130 de Applied Biosystems, y leídas con el software GeneMapper versión 3.2.

El LOD score de cada familia se calcula de acuerdo a la siguiente formula:

$$\text{LOD score} = Z = \log_{10} \left(\frac{\theta^{(\# \text{ OF RECOMBINANTS})} \cdot (1-\theta)^{(\# \text{ OF NONRECOMBINANTS})}}{(1/2)^{(\# \text{ OF RECOMBINANTS})} (1/2)^{(\# \text{ OF NONRECOMBINANTS})}} \right)$$

Y se realiza un cálculo de Chi cuadrado basado en la siguiente formula:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

Donde o=genotipos observados y, e=genotipos esperados.

Por otra parte, la potencia de cada genealogía se calcula con el programa SPLINK.

También se tomaron muestras de suero de la mayoría de los individuos para determinar los valores de TSH y T4 libre, con el fin de observar presencia de alguna patología y finalmente, se utiliza la plataforma oligo-SNP's de Affymetrix CytoScan 750K para el análisis por microarreglos.

Todos los procedimientos se realizan bajo los protocolos estandarizados y validados por el Laboratorio de Identificación Genética – IdentiGEN de la Universidad de Antioquia.

Para cada individuo estudiado, se cuenta con el respectivo consentimiento informado, es decir que a los participantes se les explica de forma verbal el carácter confidencial del procesamiento, divulgación y publicación de los resultados y el carácter voluntario de su participación, así como los riesgos a los que están expuestos, tras la toma de la muestra. Posteriormente, se obtendrá el consentimiento informado del padre o responsable de los menores de edad participantes, con la capacidad de libre elección y sin coacción alguna.

En esta investigación prevalece el respeto a la dignidad humana y se protege la privacidad del individuo, identificándolo solo cuando los resultados lo requieran y éste lo autorice.

6.1 Localización genómica

La secuencia del genoma completo de un individuo IV-4 de Frontino y el II-3 de Quibdó se hace a través de Illumina con NGS (del inglés *Next Generation Sequencing*). Esta tecnología tiene cuatro pasos básicos (88):

- Preparación de la muestra: el ADN se extrae de una muestra. Después de realizar el control de calidad (QC), la muestra aprobada se procesa con la construcción de la biblioteca.
- Construcción de la biblioteca: la biblioteca de secuenciación se prepara por fragmentación aleatoria de la muestra de ADN, seguida por la ligación de adaptadores en los extremos 5' y 3'. Alternativamente, la "etiquetación" combina la fragmentación y las reacciones de ligación en un solo paso que aumenta en gran medida la eficacia del proceso de preparación de la biblioteca. Los fragmentos ligados al adaptador se amplifican por PCR y se purifican en gel. En este estudio se tiene en cuenta el kit *Truseq PCR free library*.
- Secuenciación: para la generación de clúster, la biblioteca se carga en una celda de flujo donde se capturan fragmentos en un césped de oligos unidos a la superficie, complementarios a los adaptadores de la biblioteca. Cada fragmento luego se amplifica en clusters clonales distintos a través de la amplificación. Cuando se completa la generación de clúster, las plantillas están listas para la secuencia. La tecnología Illumina SBS utiliza un método patentado basado en un terminador reversible que actúa en el extremo 3' y detecta bases individuales a medida que se incorporan en cadenas de plantillas de ADN. El resultado es una secuencia base por base altamente precisa que prácticamente elimina los

errores específicos del contexto de la secuencia, incluso dentro de las regiones secuenciales repetitivas y los homopolímeros.

- Datos crudos: La secuencia de datos se convierte en datos crudos para el análisis.

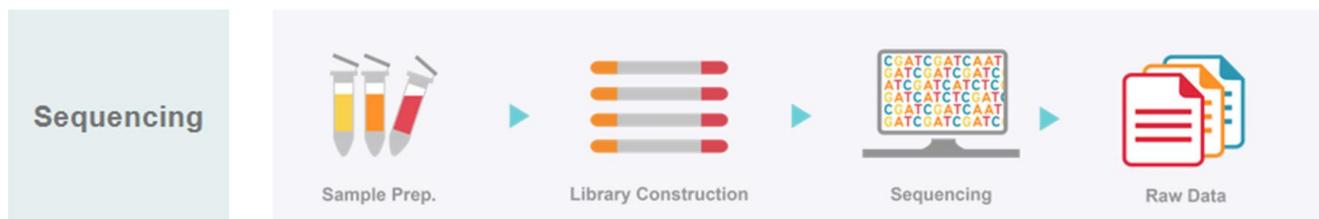


Figura 8: Flujo de trabajo de NGS

Una vez obtenidos los datos crudos se realiza el análisis para la ubicación del tercer alelo y para ello se tiene en cuenta la profundidad de las secuencias, y aquellas secuencias con una mayor profundidad pueden deberse a una duplicación. Para una ubicación más precisa en el genoma se usan las secuencias flanqueantes, y la respectiva descripción con los genes implicados si se observan. El análisis de mapeo se lleva a cabo a través del software *Speedseq align* y *LUMPY*.

Además, se realiza un estudio por aCGH al individuo II-3 de Quibdó, en busca de la duplicación del marcador TPOX. Solo se le realiza dicho estudio a un individuo por dificultades con la muestra del individuo IV-4 de Frontino.

6.2 Cálculos estadísticos en filiación con patrones trialélicos

Basado en la probabilidad Bayesiana, donde el numerador en pruebas de parentesco, representado por X , es la probabilidad de que el individuo “x” sea el padre del niño y el denominador representa

la hipótesis de que cualquier individuo al azar en la población (diferente al individuo “x”) sea el padre (89). De acuerdo a lo anterior, se plantean las posibles hipótesis para un patrón trialélico tanto para el numerador como para el denominador de la siguiente manera:

$$LR = \frac{Pi+Pj+Pk+Pl+Pm}{Ri+Rj+Rk+Rl+Rm}$$

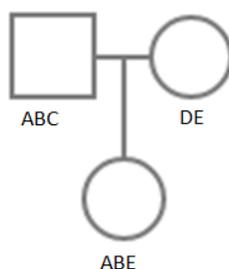
Donde “P” representa la hipótesis del presunto padre, R representa la hipótesis de que cualquier individuo al azar en la población sea el padre, “i” representa la probabilidad de que el alelo extra sea autosómico, “j” representa la probabilidad de que el alelo extra \neq “i” sea autosómico, “k” representa la probabilidad de un haplotipo, “l” representa la probabilidad del alelo extra ubicado en el cromosoma X, y “m” representa la probabilidad del alelo extra \neq “l” ubicado en el cromosoma X. Lo anterior se puede simplificar de la siguiente manera:

$$LR = \frac{\sum_{n=i}^m(Pn)}{\sum_{n=i}^m(Rn)}$$

Donde “P” representa la hipótesis del presunto padre, R representa la hipótesis de que cualquier individuo al azar en la población sea el padre, y “n” representa las diferentes hipótesis. Dicha fórmula será representada con la letra “u” a partir de ahora.

A partir de lo anterior, existen dos formas de plantear las hipótesis: 1) La inserción del alelo extra “z” está en un locus diferente del marcador genético usado en filiación que presente el patrón trialélico, y al tener una ubicación diferente tiene una frecuencia diferente a la de los alelos ya establecidos en dicho marcador, por ende, puede usarse la frecuencia del patrón trialélico para “z”

y la ausencia de la inserción puede representarse como “1-z”; 2) El “z” puede tomarse a partir del planteamiento de una mutación, el cual daría resultados más conservadores a la hora de abordar dicho patrón. Sobre la base de las consideraciones anteriores, se plantea la solución de diferentes casos de filiación de la siguiente manera:



Caso 1: Donde A, B, C, D y E representan alelos diferentes, a, b, c, d y e representan sus frecuencias respectivas, “trans” representa la transferencia de un alelo dado y los alelos entre corchetes como [ab], representan un haplotipo. Al aplicar la primera forma para plantear las diferentes hipótesis se obtiene lo siguiente:

$$\begin{aligned}
 X = & 2ac \ 2de \ (1/2)(1/2) \ (1) \ (\text{trans } b) \\
 & + 2bc \ 2de \ (1/2)(1/2) \ (1) \ (\text{trans } a) \\
 & + 2c \ [ab] \ 2de \ (1/2)(1/2) \\
 & + 2ac \ 2de \ 2(\text{trans } b) \ (1-\text{trans } de \ b) \ (1/2)(1/2) \ (1/2) \\
 & + 2bc \ 2de \ 2(\text{trans } a) \ (1-\text{trans } de \ a) \ (1/2)(1/2) \ (1/2)
 \end{aligned}$$

The first two terms are grouped by a bracket pointing to a box: **Hipótesis de que el alelo A o B estén insertados en el cromosoma X**
 The third term is grouped by a bracket pointing to a box: **Hipótesis de que el alelo A o B estén ligados**
 The last two terms are grouped by a bracket pointing to a box: **Hipótesis de que el alelo A o B estén insertados en cualquier cromosoma autosómico.**

$$\begin{aligned}
 X &= acde (\text{tran } b) (1 + (1 - \text{trans } b)) \\
 &\quad + bcde (\text{trans } a) (1 + (1 - \text{trans } a)) \\
 &\quad + cde[ab]
 \end{aligned}$$

$$\begin{aligned}
 Y &= 2ac \ 2de (\text{tans. } b) (1/2)a (1) (\text{tans. } b) \\
 &\quad + 2bc \ 2de (\text{trans. } a) (1/2)b (1) (\text{trans. } a) \\
 &\quad + 2c[ab] \ 2de (1/2)[ab] \\
 &\quad + 2ac \ 2de \ 2(\text{trans.}b)(1-\text{trans.}b) (1/2)a (1/2) (\text{trans. } b) \\
 &\quad + 2bc \ 2de \ 2(\text{trans. } a)(1-\text{trans } a) (1/2)b (1/2)(\text{trans. } a)
 \end{aligned}$$

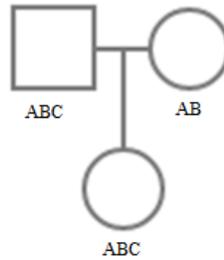
$$\begin{aligned}
 Y &= 2a^2 \ cde (\text{tran } b)^2 (1 + (1 - \text{trans } b)) \\
 &\quad + 2b^2 \ cde (\text{trans } a)^2 (1 + (1 - \text{trans } a)) \\
 &\quad + 2c[ab]^2 de
 \end{aligned}$$

$$LR = \frac{cde(a(\text{trans } b)(2 - \text{trans } b) + b(\text{trans } a)(2 - \text{trans } a) + [ab])}{2cde(a^2 (\text{trans } b)^2 (2 - \text{trans } b) + b^2 (\text{trans } a)^2 (2 - \text{trans } a) + [ab]^2)}$$

El desarrollo anterior se puede simplificar más, solo si se usa la frecuencia del patrón trialélico para la probabilidad de transferencia, entonces $\text{trans } b = \text{trans } a = [ab]$ y para simplificar se representaría con la letra “z”. Además, si $(1 - \text{trans } a \text{ ó } b) \approx 1$, entonces el desarrollo sería el siguiente:

$$LR = \frac{2a + 2b + 1}{z(2a^2 + 2b^2 + 1)}$$

El caso anterior está basado en la primera forma de plantear las hipótesis. A partir de acá se desarrollan varios casos a partir de la segunda forma, que es más conservadora y se basa en las fórmulas de mutación.



Caso 2: Donde A, B y C representan alelos diferentes, y teniendo en cuenta la probabilidad de translocación, además de la probabilidad de que cualquiera de los alelos sea el alelo extra.

- Planteamiento de la fórmula, donde Y representa el denominador y X el numerador:

$$X = 2ab2ab((1/2)(1/2) + (1/2)(1/2))Pr(\text{transl de C}) Pr(1-\text{trasl de C}).$$

$$Y = 2ab2ab(1/2)(a+b)Pr(1-\text{traslocacion de C}) 1c$$

$$LR = \frac{Pr(\text{translocación de C})}{c(a+b)}$$

- Definiciones:

Los alelos se representan con letras mayúsculas como “A”, “B”, “C” o “D” y las frecuencias alélicas se representan con letras minúsculas como “a”, “b”, “c” o “d”.

Pr(trans de C): probabilidad de transferencia

Pr(1-trans de C): probabilidad de que no haya transferencia

$2ab(1/2)(1/2) + (1/2)(1/2)$: probabilidad de que el padre transmita el alelo A o el alelo B y la probabilidad de que la madre transmita el alelo A o el alelo B.

$2ab(1/2)(a+b)$: probabilidad de transmisión de la madre y probabilidad de transmisión de cualquier individuo en la población.

$1c$: probabilidad de que cualquier individuo en la población transmita el alelo C.

Al añadirle la probabilidad de que esté ligado al alelo B o al A, además de la probabilidad de estar ubicado en el cromosoma X, y la probabilidad de que sea cualquiera de los tres alelos el alelo extra, queda de la siguiente manera:

$$X = 2ab \left[2ab \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) + \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) \right] \text{Pr}(\text{trans C}) \text{Pr}(1\text{-transl C})$$

$$+ 2bc \left[2ab \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) \right] \text{Pr}(\text{trans A}) \text{Pr}(1\text{-trans A})$$

$$+ 2ac \left[2ab \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) \right] \text{Pr}(\text{trans B}) \text{Pr}(1\text{-trans B})$$

$$+ 2a[bc] \left[2ab \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) + 2b(ac) \right] 2ab \left(\frac{1}{2} \right) \left(\frac{1}{2} \right)$$

Hipótesis de que el alelo A, B o C estén transferidos en el cromosoma X

Hipótesis de que el alelo A o B estén ligados

$$\begin{aligned}
 &+ 2ab \cdot 2ab \left(\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \right) \cdot \frac{1}{2} \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 2bc \cdot 2ab \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ 2ac \cdot 2ab \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \Pr(\text{trans B}) \Pr(1\text{-trans B})
 \end{aligned}$$

Hipótesis de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

La fila que contiene este resultado $2a(bc) \cdot 2ab \cdot \frac{1}{2} \cdot \frac{1}{2} + 2b(ac) \cdot 2ab \cdot \frac{1}{2} \cdot \frac{1}{2}$, representa las posibilidades de que [bc] o [ac] segreguen como haplotipos.

$$\begin{aligned}
 X = & 2a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B}) \\
 & + a^2 [bc] b + b^2 [ac] a \\
 & + a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + \frac{1}{2} b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + \frac{1}{2} a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B})
 \end{aligned}$$

$$\begin{aligned}
 X = & 3a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + \frac{3}{2} b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + \frac{3}{2} a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B}) \\
 & + a^2 [bc] b + b^2 [ac] a
 \end{aligned}$$

$$\begin{aligned}
 X = & (ab) \left(3ab \Pr(\text{trans C}) \Pr(1\text{-trans C}) + \frac{3}{2} b c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + \frac{3}{2} a c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + a [bc] + b [ac] \right)
 \end{aligned}$$

$$\begin{aligned}
 Y = & 2ab \cdot 2ab \cdot \frac{1}{2} \cdot (a+b) \Pr(1\text{-trans de C}) \cdot 1c \\
 & + 2bc \cdot 2ab \cdot \frac{1}{2} \cdot (a+b) \Pr(1\text{-trans de A}) \cdot 1a
 \end{aligned}$$

$$\begin{aligned}
&+ 2ac2ab (1/2) (a+b) \Pr(1\text{-trans de B}) 1b \\
&+ 2a[bc] 2ab (1/2) [bc] \\
&+ 2b[ac] 2ab (1/2) [ac] \\
&+ 2ab2ab(1/2) (a+b) \Pr(1\text{-trans de C}) (1/2)c \\
&+ 2bc2ab(1/2) (a+b) \Pr(1\text{-trans de A}) (1/2)a \\
&+ 2ac2ab (1/2) (a+b) \Pr(1\text{-trans de B}) (1/2)b
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y} = & 2a^2 b^2 (a+b) \Pr(1\text{-trans de C}) c \\
& + 2b^2 c a^2 (a+b) \Pr(1\text{-trans de A}) \\
& + 2a^2 c b^2 (a+b) \Pr(1\text{-trans de B}) \\
& + 2a^2 [bc]^2 b \\
& + 2b^2 [ac]^2 a \\
& + a^2 b^2 (a+b) \Pr(1\text{-trans de C}) c \\
& + b^2 c a^2 (a+b) \Pr(1\text{-trans de A}) \\
& + a^2 c b^2 (a+b) \Pr(1\text{-trans de B})
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y} = & 3a^2 b^2 (a+b) \Pr(1\text{-trans de C}) c \\
& + 3b^2 c a^2 (a+b) \Pr(1\text{-trans de A}) \\
& + 3a^2 c b^2 (a+b) \Pr(1\text{-trans de B}) \\
& + 2a^2 [bc]^2 b \\
& + 2b^2 [ac]^2 a
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y} = & (ab) (3a b (a+b) \Pr(1\text{-trans de C}) c + 3b c a (a+b) \Pr(1\text{-trans de A}) + 3a c b (a+b) \Pr(1\text{-trans de B}) + \\
& 2a [bc]^2 + 2b [ac]^2)
\end{aligned}$$

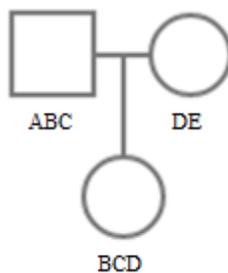
Simplificando la fórmula:

LR=

$$\frac{3ab \Pr(\text{trans C})\Pr(1-\text{trans C}) + (3/2) b c \Pr(\text{trans A})\Pr(1-\text{trans A}) + (3/2) a c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + a [bc] + b [ac]}{(3abc)(a+b)(\Pr(1-\text{trasl. de C}) + \Pr(1-\text{trasl. de A}) + \Pr(1-\text{trans de B})) + 2a [bc]^2 + 2b [ac]^2}$$

Igual al caso anterior, el desarrollo se puede simplificar aún más, solo si se usa la frecuencia del patrón trialélico para la probabilidad de transferencia, entonces $\text{trans c} = \text{trans b} = \text{trans a} = [bc] = [ac]$ y para simplificar se representaría con la letra “z”. Además, si $(1-\text{trans a,b ó c}) \approx 1$, entonces el desarrollo sería el siguiente:

$$LR = \frac{(z)(3ab + (\frac{3}{2})bc + (\frac{3}{2})ac + a + b)}{3abc(a+b)(3 + (z^2)(2a+2b)}$$



Caso 3: A partir de este caso en adelante no se simplifica usando la misma frecuencia del patrón trialélico para la probabilidad de transferencia, entonces se obtiene lo siguiente:

$$\begin{aligned} X &= 2ab \cdot 2de \cdot (1/2)(1/2) \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 2ac \cdot 2de \cdot (1/2)(1/2) \Pr(\text{trans B}) \Pr(1-\text{trans B}) \end{aligned}$$

Probabilidad de que el alelo A, B o C estén transferidos en el cromosoma X
--

$$+ 2a[bc] \text{ de } (1/2)(1/2)$$

$$+ 2ab \text{ de } (1/2) (1/2) (1/2) \text{Pr(trans C) Pr(1-trans C)}$$

$$+ 2ac \text{ de } (1/2) (1/2) (1/2) \text{Pr(trans B) Pr(1-trans B)}$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico
--

$$X = \text{de } (3/2) ab \text{ de Pr(trans C) Pr(1-trans C)}$$

$$+ (3/2) ac \text{ de Pr(trans B) Pr(1-trans B)}$$

$$+ a[bc] \text{ de}$$

$$X = \text{de } ((3/2) ab \text{ Pr(trans C) Pr(1-trans C)} + (3/2) ac \text{ Pr(trans B) Pr(1-trans B)} + a[bc])$$

$$Y = 2ab \text{ de } (1/2) (b) \text{ Pr(1-trans de C) } 1c$$

$$+ 2ac \text{ de } (1/2) (c) \text{ Pr(1-trans de B) } 1b$$

$$+ 2a[bc] \text{ de } (1/2) [bc]$$

$$+ 2ab \text{ de } (1/2) (b) \text{ Pr(1-trans de C) } (1/2)c$$

$$+ 2ac \text{ de } (1/2) (c) \text{ Pr(1-trans de B) } (1/2)b$$

$$Y = 3ab \text{ de } (b) \text{ Pr(1-trans de C) } c$$

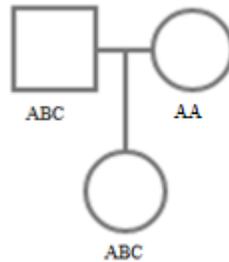
$$+ 2a[bc]^2 \text{ de}$$

$$+ 3ac \text{ de } (c) \text{ Pr(1-trans de B) } b$$

$$Y = \text{de } ((3ab^2) \text{ Pr(1-trans de C) } c + (3ac^2) \text{ Pr(1-trans de B) } b + 2[bc]^2)$$

Simplificando la fórmula:

$$LR = \frac{\frac{3}{2} ab \Pr(\text{trans C}) \Pr(1-\text{trans C}) + 3/2 ac \Pr(\text{trans B}) \Pr(1-\text{trans B}) + [bc]}{(3ab^2) \Pr(1-\text{trans de C})c + (3ac^2) \Pr(1-\text{trans de B})b + 2[bc]^2}$$



Caso 4. En este caso existen dos posibles alelos como alelo extra y se desarrolla de la siguiente manera:

$$\begin{aligned} X &= 2ab a^2 (1/2)(1) \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 2ac a^2 (1/2)(1) \Pr(\text{trans B}) \Pr(1-\text{trans B}) \\ &+ 2a[bc] a^2 (1/2)(1) \\ &+ 2ab a^2 (1/2) (1) (1/2) \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 2ac a^2 (1/2) (1) (1/2) \Pr(\text{trans B}) \Pr(1-\text{trans B}) \end{aligned}$$

$$\begin{aligned} X &= 3/2 ab a^2 \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 3/2 ac a^2 \Pr(\text{trans B}) \Pr(1-\text{trans B}) \\ &+ a[bc] a^2 \end{aligned}$$

$$X = a^3 (3/2 b \Pr(\text{trans C}) \Pr(1-\text{trans C}) + 3/2 c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + [bc])$$

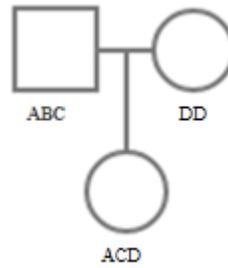
$$\begin{aligned}
Y &= 2ab a^2 (1) (b) \Pr(1\text{-trans de C}) 1c \\
&+ 2ac a^2 (1) (c) \Pr(1\text{-trans de B}) 1b \\
&+ 2a[bc] a^2 (1) [bc] \\
&+ 2ab a^2 (1) (b) \Pr(1\text{-trans de C}) (1/2)c \\
&+ 2ac a^2 (1) (c) \Pr(1\text{-trans de B}) (1/2)b
\end{aligned}$$

$$\begin{aligned}
Y &= 3ab a^2 (b) \Pr(1\text{-trans de C}) c \\
&+ 2a[bc]^2 a^2 \\
&+ 3ac a^2 (c) \Pr(1\text{-trans de B}) b
\end{aligned}$$

$$Y = a^3 ((3b^2)c \Pr(1\text{-trans de C}) + (3c^2)b \Pr(1\text{-trans de B}) + 2[bc]^2)$$

Simplificando la fórmula:

$$LR = \frac{3/2 b \Pr(\text{trans C}) \Pr(1\text{-trans C}) + 3/2 c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + [bc]}{(3b^2)c \Pr(1\text{-trans de C}) + (3c^2)b \Pr(1\text{-trans de B}) + 2[bc]^2}$$



Caso 4. En este caso el alelo extra puede ser el A o el C.

$$\begin{aligned}
 X &= 2ab d^2 (1/2)(1) \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 2bc d^2 (1/2)(1) \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ 2b[ac] d^2 (1/2)(1) \\
 &+ 2ab d^2 (1/2) (1) (1/2) \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 2bc d^2 (1/2) (1) (1/2) \Pr(\text{trans A}) \Pr(1\text{-trans A})
 \end{aligned}$$

$$\begin{aligned}
 X &= ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ b[ac] d^2 \\
 &+ (1/2) ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ (1/2) bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A})
 \end{aligned}$$

$$\begin{aligned}
 X &= 3/2 ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 3/2 bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ b[ac] d^2
 \end{aligned}$$

$$X = bd^2 (3/2 a \Pr(\text{trans C}) \Pr(1\text{-trans C}) + 3/2 c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + [ac])$$

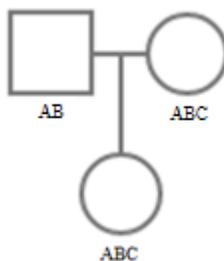
$$\begin{aligned}
 Y &= 2ab d^2 (1) (a) \Pr(1\text{-trans de C}) 1c \\
 &+ 2bc d^2 (1) (c) \Pr(1\text{-trans de A}) 1a \\
 &+ 2b[ac] d^2 (1) [bc] \\
 &+ 2ab d^2 (1) (a) \Pr(1\text{-trans de C}) (1/2)c \\
 &+ 2bc d^2 (1) (c) \Pr(1\text{-trans de A}) (1/2)a
 \end{aligned}$$

$$\begin{aligned}
 Y &= 3ab d^2 (1) (a) \Pr(1\text{-trans de C}) c \\
 &+ 3bc d^2 (1) (c) \Pr(1\text{-trans de A}) 1a \\
 &+ 2b[ac] d^2 [ac]
 \end{aligned}$$

$$Y = bd^2 (3ac \Pr(1\text{-trans de C}) + 3ac \Pr(1\text{-trans de A}) + 2[ac]^2)$$

Simplificando la fórmula:

$$LR = \frac{3/2 a \Pr(\text{trans C}) \Pr(1\text{-trans C}) + 3/2 c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + [ac]}{3ac (\Pr(1\text{-trans de C}) + \Pr(1\text{-trans de A})) + 2[ac]^2}$$



Caso 5: Donde A, B y C representan alelos diferentes, y teniendo en cuenta la probabilidad de translocación, además de la probabilidad de que cualquiera de los alelos sea el alelo extra.

$$\begin{aligned}
 \mathbf{X} = & 2ab \left[\frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} \right] \frac{1}{2} \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + 2bc \frac{1}{2} \frac{1}{2} \Pr(\text{trans A}) \frac{1}{2} \Pr(1\text{-trans A}) \\
 & + 2ac \frac{1}{2} \frac{1}{2} \Pr(\text{trans B}) \frac{1}{2} \Pr(1\text{-trans B}) \\
 & + 2a[bc] \frac{1}{2} \frac{1}{2} + 2b[ac] \frac{1}{2} \frac{1}{2} \\
 & + 2ab \left[\frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} \right] \frac{1}{2} \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + 2bc \frac{1}{2} \frac{1}{2} \frac{1}{2} \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + 2ac \frac{1}{2} \frac{1}{2} \frac{1}{2} \Pr(\text{trans B}) \Pr(1\text{-trans B})
 \end{aligned}$$

Probabilidad de que el alelo A, B o C estén en el cromosoma X

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

$$\begin{aligned}
 \mathbf{X} = & a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + \frac{1}{2} b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + \frac{1}{2} a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B}) \\
 & + a^2 [bc] b + b^2 [ac] a \\
 & + a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + \frac{1}{2} b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + \frac{1}{2} a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B})
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{X} = & 2a^2 b^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 & + b^2 c a \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 & + a^2 c b \Pr(\text{trans B}) \Pr(1\text{-trans B}) \\
 & + a^2 [bc] b + b^2 [ac] a
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{X} = & [ab] \left(2ab \Pr(\text{trans C}) \Pr(1\text{-trans C}) + b c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + a c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + a \right. \\
 & \left. [bc] + b [ac] \right)
 \end{aligned}$$

$$\mathbf{Y} = 2ab \frac{1}{2} (a+b) \Pr(1\text{-trans de C}) \frac{1}{2}$$

$$\begin{aligned}
&+ 2bc2ab(1/2) (b+c) \Pr(1\text{-trans de A}) (1/2) \\
&+ 2ac2ab (1/2) (a+c) \Pr(1\text{-trans de B}) (1/2) \\
&+ 2a[bc] 2ab (1/2) [bc] \\
&+ 2b[ac] 2ab (1/2) [ac] \\
&+ 2ab2ab(1/2) (a+b) \Pr(1\text{-trans de C}) (1/2) \\
&+ 2bc2ab(1/2) (b+c) \Pr(1\text{-trans de A}) (1/2) \\
&+ 2ac2ab (1/2) (a+c) \Pr(1\text{-trans de B}) (1/2)
\end{aligned}$$

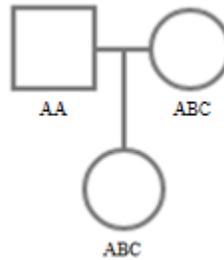
$$\begin{aligned}
\mathbf{Y} &= a^2 b^2 (a+b) \Pr(1\text{-trans de C}) \\
&+ b^2 c a (b+c) \Pr(1\text{-trans de A}) \\
&+ a^2 c b (a+c) \Pr(1\text{-trans de B}) \\
&+ 2a^2 [bc]^2 b \\
&+ 2b^2 [ac]^2 a \\
&+ a^2 b^2 (a+b) \Pr(1\text{-trans de C}) \\
&+ b^2 c a (b+c) \Pr(1\text{-trans de A}) \\
&+ a^2 c b (a+c) \Pr(1\text{-trans de B})
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y} &= 2a^2 b^2 (a+b) \Pr(1\text{-trans de C}) \\
&+ 2b^2 c a (b+c) \Pr(1\text{-trans de A}) \\
&+ 2a^2 c b (a+c) \Pr(1\text{-trans de B}) \\
&+ 2a^2 [bc]^2 b \\
&+ 2b^2 [ac]^2 a
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y} &= [ab] (2a b (a+b) \Pr(1\text{-trans de C}) + 2b c (b+c) \Pr(1\text{-trans de A}) + 2a c (a+c) \Pr(1\text{-trans de B}) + 2a \\
&[bc]^2 + 2b [ac]^2)
\end{aligned}$$

Simplificando la fórmula:

$$LR = \frac{2ab \Pr(\text{trans C})\Pr(1-\text{trans C}) + b c \Pr(\text{trans A})\Pr(1-\text{trans A}) + a c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + a [bc] + b [ac]}{2a b (a+b) \Pr(1-\text{trans de C}) + 2b c (b+c) \Pr(1-\text{trans de A}) + 2a c (a+c) \Pr(1-\text{trans de B}) + 2a [bc]^2 + 2b [ac]^2}$$



Caso 6. Manteniendo las mismas definiciones de los casos anteriores se desarrolla este problema, en el cual existen dos posibles alelos como alelo extra y se desarrolla de la siguiente manera:

$$\begin{aligned} X &= 2ab a^2 (1/2)(1) (1/2) \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 2ac a^2 (1/2)(1) (1/2) \Pr(\text{trans B}) \Pr(1-\text{trans B}) \\ &+ 2a[bc] a^2 (1/2)(1) \\ &+ 2ab a^2 (1/2) (1) (1/2)\Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ 2ac a^2 (1/2) (1) (1/2)\Pr(\text{trans B}) \Pr(1-\text{trans B}) \end{aligned}$$

$$\begin{aligned} X &= ab a^2 \Pr(\text{trans C}) \Pr(1-\text{trans C}) \\ &+ ac a^2 \Pr(\text{trans B}) \Pr(1-\text{trans B}) \\ &+ a[bc] a^2 \end{aligned}$$

$$X = a^3 (b \Pr(\text{trans C}) \Pr(1\text{-trans C}) + c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + [bc])$$

$$Y = 2ab a^2 (1) (a+b) \Pr(1\text{-trans de C}) (1/2)$$

$$+ 2ac a^2 (1) (a+c) \Pr(1\text{-trans de B}) (1/2)$$

$$+ 2a[bc] a^2 (1) [bc]$$

$$+ 2ab a^2 (1) (a+b) \Pr(1\text{-trans de C}) (1/2)$$

$$+ 2ac a^2 (1) (a+c) \Pr(1\text{-trans de B}) (1/2)$$

$$Y = 3ab a^2 (b) \Pr(1\text{-trans de C}) c$$

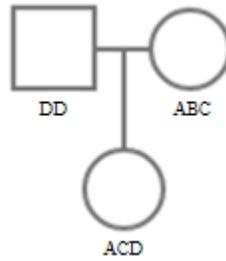
$$+ 2a[bc]^2 a^2$$

$$+ 3ac a^2 (c) \Pr(1\text{-trans de B}) b$$

$$Y = a^3 ((3b^2) \Pr(1\text{-trans de C}) + (3c^2) \Pr(1\text{-trans de B}) + 2[bc]^2)$$

Simplificando la fórmula:

$$LR = \frac{3/2 b \Pr(\text{trans C}) \Pr(1\text{-trans C}) + 3/2 c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + [bc]}{(3b^2) \Pr(1\text{-trans de C}) + (3c^2) \Pr(1\text{-trans de B}) + 2[bc]^2}$$



Caso 7. En este caso el alelo extra puede ser el A o el C.

$$\begin{aligned}
 X &= 2ab d^2 (1/2)(1) (1/2) \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 2bc d^2 (1/2)(1) (1/2) \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ 2b[ac] d^2 (1/2)(1) \\
 &+ 2ab d^2 (1/2) (1) (1/2) \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ 2bc d^2 (1/2) (1) (1/2) \Pr(\text{trans A}) \Pr(1\text{-trans A})
 \end{aligned}$$

$$\begin{aligned}
 X &= (1/2)ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ (1/2) bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ b[ac] d^2 \\
 &+ (1/2) ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ (1/2) bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A})
 \end{aligned}$$

$$\begin{aligned}
 X &= ab d^2 \Pr(\text{trans C}) \Pr(1\text{-trans C}) \\
 &+ bc d^2 \Pr(\text{trans A}) \Pr(1\text{-trans A}) \\
 &+ b[ac] d^2
 \end{aligned}$$

$$X = bd^2 (a \Pr(\text{trans C}) \Pr(1\text{-trans C}) + c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + [ac])$$

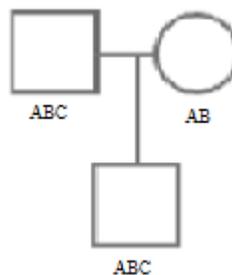
$$\begin{aligned}
 Y &= 2ab d^2 (1/2) (a+b) \Pr(1\text{-trans de C}) \\
 &+ 2bc d^2 (1/2) (b+c) \Pr(1\text{-trans de A}) \\
 &+ 2b[ac] d^2 (1/2) [bc] \\
 &+ 2ab d^2 (1/2) (a+b) \Pr(1\text{-trans de C}) \\
 &+ 2bc d^2 (1/2) (b+c) \Pr(1\text{-trans de A})
 \end{aligned}$$

$$\begin{aligned}
 Y &= 2ab d^2 (a+b) \Pr(1\text{-trans de C}) \\
 &+ 2bc d^2 (b+c) \Pr(1\text{-trans de A}) \\
 &+ b[ac] d^2 [ac]
 \end{aligned}$$

$$Y = bd^2 (2a(a+b) \Pr(1\text{-trans de C}) + 2c(b+c) \Pr(1\text{-trans de A}) + 2[ac]^2)$$

Simplificando la fórmula:

$$LR = \frac{a \Pr(\text{trans C}) \Pr(1\text{-trans C}) + c \Pr(\text{trans A}) \Pr(1\text{-trans A}) + [ac]}{2a(a+b) \Pr(1\text{-trans de C}) + 2c(b+c) \Pr(1\text{-trans de A}) + 2[ac]^2}$$



Caso 8: Donde A, B y C representan alelos diferentes, y teniendo en cuenta la probabilidad de translocación, además de la probabilidad de que cualquiera de los alelos sea el alelo extra.

$$X = 2ab \left[\frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} \right] \Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ 2bc \frac{1}{2} \frac{1}{2} \Pr(\text{trans A}) \Pr(1-\text{trans A})$$

$$+ 2ac \frac{1}{2} \frac{1}{2} \Pr(\text{trans B}) \Pr(1-\text{trans B})$$

$$+ 2a[bc] \frac{1}{2} \frac{1}{2} + 2b[ac] \frac{1}{2} \frac{1}{2}$$

$$+ 2ab \left[\frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} \right] \frac{1}{2} \Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ 2bc \frac{1}{2} \frac{1}{2} \frac{1}{2} \Pr(\text{trans A}) \Pr(1-\text{trans A})$$

$$+ 2ac \frac{1}{2} \frac{1}{2} \frac{1}{2} \Pr(\text{trans B}) \Pr(1-\text{trans B})$$

Probabilidad de que el alelo A, B o C estén transferidos en el cromosoma X. En este caso la probabilidad de padre a hijo sería cero.

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

Como la probabilidad de que el padre le transmita el cromosoma X a su hijo varón es cero, la fórmula queda de la siguiente manera:

$$X = a^2 [bc] b + b^2 [ac] a$$

$$+ a^2 b^2 \Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ \frac{1}{2} b^2 c a \Pr(\text{trans A}) \Pr(1-\text{trans A})$$

$$+ \frac{1}{2} a^2 c b \Pr(\text{trans B}) \Pr(1-\text{trans B})$$

$$X = a^2 b^2 \Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ \frac{1}{2} b^2 c a \Pr(\text{trans A}) \Pr(1-\text{trans A})$$

$$+ \frac{1}{2} a^2 c b \Pr(\text{trans B}) \Pr(1-\text{trans B})$$

$$+ a^2 [bc] b + b^2 [ac] a$$

$$X = [ab] (ab \Pr(\text{trans C})\Pr(1-\text{trans C}) + (1/2) b c \Pr(\text{trans A})\Pr(1-\text{trans A}) + (1/2) a c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + a [bc] + b [ac])$$

$$Y = 2a[bc] 2ab (1/2) [bc]$$

$$+ 2b[ac] 2ab (1/2) [ac]$$

$$+ 2ab2ab(1/2) (a+b) \Pr(1-\text{trans de C}) (1/2)c$$

$$+ 2bc2ab(1/2) (a+b) \Pr(1-\text{trans de A}) (1/2)a$$

$$+ 2ac2ab (1/2) (a+b) \Pr(1-\text{trans de B}) (1/2)b$$

$$Y = 2a^2 [bc]^2 b$$

$$+ 2b^2 [ac]^2 a$$

$$+ a^2 b^2 (a+b) \Pr(1-\text{trans de C}) c$$

$$+ b^2 c a^2 (a+b) \Pr(1-\text{trans de A})$$

$$+ a^2 c b^2 (a+b) \Pr(1-\text{trans de B})$$

$$Y = a^2 b^2 c (a+b) \Pr(1-\text{trans de C})$$

$$+ b^2 c a^2 (a+b) \Pr(1-\text{trans de A})$$

$$+ a^2 c b^2 (a+b) \Pr(1-\text{trans de B})$$

$$+ 2a^2 [bc]^2 b$$

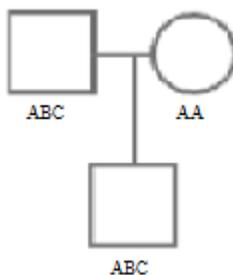
$$+ 2b^2 [ac]^2 a$$

$$Y = [ab] (a b c (a+b) \Pr(1-\text{trans de C}) + b c a (a+b) \Pr(1-\text{trans de A}) + a c b (a+b) \Pr(1-\text{trans de B}) + 2a [bc]^2 + 2b [ac]^2)$$

Simplificando la fórmula:

LR=

$$\frac{ab \Pr(\text{trans C})\Pr(1-\text{trans C}) + (1/2) b c \Pr(\text{trans A})\Pr(1-\text{trans A}) + (1/2) a c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + a [bc] + b [ac]}{a b c (a+b) \Pr(1-\text{trans de C}) + b c a (a+b) \Pr(1-\text{trans de A}) + a c b (a+b) \Pr(1-\text{trans de B}) + 2a [bc]^2 + 2b [ac]^2}$$



Caso 9. Manteniendo las mismas definiciones de los casos anteriores se desarrolla este problema, en el cual existen dos posibles alelos como alelo extra y se desarrolla de la siguiente manera:

$$X = 2a[bc] a^2 (1/2)(1)$$

$$+ 2ab a^2 (1/2) (1) (1/2)\Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ 2ac a^2 (1/2) (1) (1/2)\Pr(\text{trans B}) \Pr(1-\text{trans B})$$

$$X = 1/2 ab a^2 \Pr(\text{trans C}) \Pr(1-\text{trans C})$$

$$+ 1/2 ac a^2 \Pr(\text{trans B}) \Pr(1-\text{trans B})$$

$$+ a[bc] a^2$$

$$X = a^3 (1/2 b \Pr(\text{trans C}) \Pr(1-\text{trans C}) + 1/2 c \Pr(\text{trans B}) \Pr(1-\text{trans B}) + [bc])$$

$$Y = 2a[bc] a^2 (1) [bc]$$

$$+ 2ab a^2 (1) (b) \Pr(1\text{-trans de C}) (1/2)c$$

$$+ 2ac a^2 (1) (c) \Pr(1\text{-trans de B}) (1/2)b$$

$$Y = ab a^2 (b) \Pr(1\text{-trans de C}) c$$

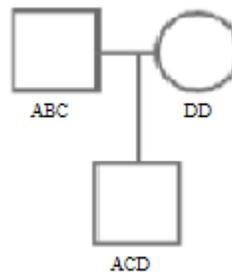
$$+ 2a[bc]^2 a^2$$

$$+ ac a^2 (c) \Pr(1\text{-trans de B}) b$$

$$Y = a^3 ((b^2 c) \Pr(1\text{-trans de C}) + (c^2 b) \Pr(1\text{-trans de B}) + 2[bc]^2)$$

Simplificando la fórmula:

$$LR = \frac{1/2 b \Pr(\text{trans C}) \Pr(1\text{-trans C}) + 1/2 c \Pr(\text{trans B}) \Pr(1\text{-trans B}) + [bc]}{(c b^2) \Pr(1\text{-trans de C}) + (b c^2) \Pr(1\text{-trans de B}) + 2[bc]^2}$$



Caso 10. En este caso el alelo extra puede ser el A o el C.

$$\begin{aligned}
 X &= 2b[ac] d^2 (1/2)(1) \\
 &+ 2ab d^2 (1/2) (1) (1/2)\text{Pr}(\text{trans C}) \text{Pr}(1\text{-trans C}) \\
 &+ 2bc d^2 (1/2) (1) (1/2)\text{Pr}(\text{trans A}) \text{Pr}(1\text{-trans A})
 \end{aligned}$$

$$\begin{aligned}
 X &= b[ac] d^2 \\
 &+ (1/2) ab d^2 \text{Pr}(\text{trans C}) \text{Pr}(1\text{-trans C}) \\
 &+ (1/2) bc d^2 \text{Pr}(\text{trans A}) \text{Pr}(1\text{-trans A})
 \end{aligned}$$

$$X = bd^2 (1/2 a \text{Pr}(\text{trans C}) \text{Pr}(1\text{-trans C}) + 1/2 c \text{Pr}(\text{trans A}) \text{Pr}(1\text{-trans A}) + [ac])$$

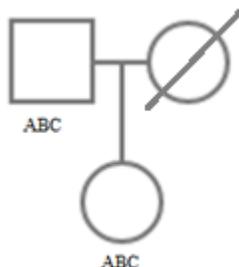
$$\begin{aligned}
 Y &= 2b[ac] d^2 (1) [ac] \\
 &+ 2ab d^2 (1) (a) \text{Pr}(1\text{-trans de C}) (1/2)c \\
 &+ 2bc d^2 (1) (c) \text{Pr}(1\text{-trans de A}) (1/2)a
 \end{aligned}$$

$$\begin{aligned}
 Y &= ab d^2 (1) (a) \text{Pr}(1\text{-trans de C}) c \\
 &+ bc d^2 (1) (c) \text{Pr}(1\text{-trans de A}) 1a \\
 &+ 2b[ac] d^2 [ac]
 \end{aligned}$$

$$Y = bd^2 (ac \text{Pr}(1\text{-trans de C}) + ac \text{Pr}(1\text{-trans de A}) + 2[ac]^2)$$

Simplificando la fórmula:

$$LR = \frac{1/2 a \Pr(\text{trans C}) \Pr(1 - \text{trans C}) + 1/2 c \Pr(\text{trans A}) \Pr(1 - \text{trans A}) + [ac]}{ac \Pr(1 - \text{trans de C}) + ac \Pr(1 - \text{trans de A}) + 2[ac]^2}$$



Caso 11: a partir de este punto se toman los casos con dúos, donde se obtiene el perfil del Padre y del hijo.

$$X = ab(a+b) (1/2) \Pr(\text{trans C}) \Pr(1 - \text{trans C}) \\ + ac(a+c) (1/2) \Pr(\text{trans B}) \Pr(1 - \text{trans B}) \\ + bc(b+c) (1/2) \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

$$+ [ab]c ([ab]+c) \\ + [ac]b ([ac]+b)$$

Probabilidad de que cosegreden los alelos (ab), (ac) o (bc).

$$+ [bc]a ([bc]+a)$$

$$+ ab(a+b) \Pr(\text{trans C}) \Pr(1 - \text{trans C}) \\ + ac(a+c) \Pr(\text{trans B}) \Pr(1 - \text{trans B}) \\ + bc(b+c) \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

Probabilidad de que el alelo A, B o C estén transferidos en el cromosoma X.

$$X = (3/2) ab(a+b) \Pr(\text{trans C}) \Pr(1 - \text{trans C}) \\ + (3/2) ac(a+c) \Pr(\text{trans B}) \Pr(1 - \text{trans B}) \\ + (3/2) bc(b+c) \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

$$+ [ab]c ([ab]+c)$$

$$+ [ac]b ([ac]+b)$$

$$+ [bc]a ([bc]+a)$$

$$Y = 4a^2 b^2 \Pr(1 - \text{trans C}) (1/2) c$$

$$+ 4a^2 c^2 \Pr(1 - \text{trans B}) (1/2) b$$

$$+ 4b^2 c^2 \Pr(1 - \text{trans A}) (1/2) a$$

$$+ 4[ab]^2 c^2$$

$$+ 4[ac]^2 b^2$$

$$+ 4[bc]^2 a^2$$

$$+ 4a^2 b^2 \Pr(1 - \text{trans C}) c$$

$$+ 4a^2 c^2 \Pr(1 - \text{trans B}) b$$

$$+ 4b^2 c^2 \Pr(1 - \text{trans A}) a$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

Probabilidad de que cosegreden los alelos (ab), (ac) o (bc).

Probabilidad de que el alelo A, B o C estén transferidos en el cromosoma X.

$$Y = 6a^2 b^2 \Pr(1 - \text{trans C}) c$$

$$+ 6a^2 c^2 \Pr(1 - \text{trans B}) b$$

$$+ 6b^2 c^2 \Pr(1 - \text{trans A}) a$$

$$+ 4[ab]^2 c^2$$

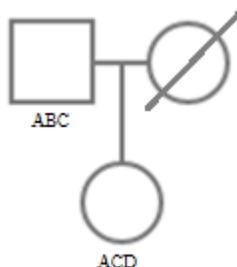
$$+ 4[ac]^2 b^2$$

$$+ 4[bc]^2 a^2$$

Simplificando:

LR=

$$\frac{\left(\frac{3}{2}\right)ab(a+b)\Pr(\text{trans C})\Pr(1 - \text{trans C}) + \left(\frac{3}{2}\right)ac(a+c)\Pr(\text{trans b})\Pr(1 - \text{trans b}) + \left(\frac{3}{2}\right)bc(b+c)\Pr(\text{trans A})\Pr(1 - \text{trans A}) + [ab]c ([ab]+c) + [bc]a ([bc]+[bc]a ([bc]+a))}{6^{a^2} b^2 \Pr(1 - \text{trans C})c + 6^{a^2} c^2 \Pr(1 - \text{trans B})b + 6b^2 c^2 \Pr(1 - \text{trans A})a + 4[ab]^2 c^2 + 4[ac]^2 b^2 + 4[bc]^2 a^2}$$

**Caso 12:**

$$X = abd \left(\frac{1}{2}\right) \Pr(\text{trans C}) \Pr(1 - \text{trans C})$$

$$+ bcd \left(\frac{1}{2}\right) \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

$$+ [ac]b ([ac]+b)$$

$$+ abd \Pr(\text{trans C})\Pr(1 - \text{trans C})$$

$$+ bcd \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

Probabilidad de que el alelo A o C estén transferidos en cualquier cromosoma autosómico

Probabilidad de que coseguren los alelos (ac).

Probabilidad de que el alelo A o C estén transferidos en el cromosoma X.

$$X = \left(\frac{3}{2}\right) abd \Pr(\text{trans C}) \Pr(1 - \text{trans C})$$

$$+ \left(\frac{3}{2}\right) bcd \Pr(\text{trans A}) \Pr(1 - \text{trans A})$$

$$+ [ac]b ([ac]+b)$$

$$Y = 4a^2 b^2 \Pr(1 - \text{trans C}) \left(\frac{1}{2}\right) c$$

$$+ 4b^2 c^2 \Pr(1 - \text{trans A}) \left(\frac{1}{2}\right) a$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

$$+ 4[ac]^2 b^2$$

Probabilidad de que coseguren los alelos (ab), (ac) o (bc).

$$+ 4a^2 b^2 \Pr(1 - \text{trans } C) c$$

$$+ 4b^2 c^2 \Pr(1 - \text{trans } A) a$$



Probabilidad de que el alelo A, B o C estén transferidos en el cromosoma X.

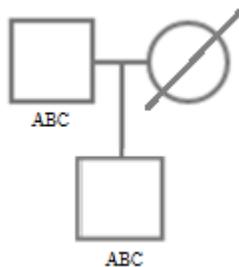
$$Y = 6a^2 b^2 \Pr(1 - \text{trans } C) c$$

$$+ 6b^2 c^2 \Pr(1 - \text{trans } A) a$$

$$+ 4[ac]^2 b^2$$

Simplificando la fórmula:

$$LR = \frac{(3/2) ad \Pr(\text{transl } C) \Pr(1 - \text{transl } C) + (3/2) cd \Pr(\text{transl } A) \Pr(1 - \text{transl } A) + [ac] ([ac] + b)}{6a^2 b \Pr(1 - \text{trans } C) c + 6b c^2 \Pr(1 - \text{trans } A) a + 4[ac]^2 b}$$



Caso 13:

$$X = ab(a+b) (1/2) \Pr(\text{trans } C) \Pr(1 - \text{trans } C)$$

$$+ ac(a+c) (1/2) \Pr(\text{trans } B) \Pr(1 - \text{trans } B)$$

$$+ bc(b+c) (1/2) \Pr(\text{trans } A) \Pr(1 - \text{trans } A)$$



Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

$$+ [ab]c ([ab]+c)$$

$$+ [ac]b ([ac]+b)$$

$$+ [bc]a ([bc]+a)$$

Probabilidad de que cosegreden los alelos (ab), (ac) o (bc).

$$Y = 4a^2 b^2 \Pr(1 - \text{trans C}) (1/2) c$$

$$+ 4a^2 c^2 \Pr(1 - \text{trans B}) (1/2) b$$

$$+ 4b^2 c^2 \Pr(1 - \text{trans A}) (1/2) a$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

$$+ 4[ab]^2 c^2$$

$$+ 4[ac]^2 b^2$$

Probabilidad de que cosegreden los alelos (ab), (ac) o (bc).

$$+ 4[bc]^2 a^2$$

$$Y = 2a^2 b^2 \Pr(1 - \text{trans C}) c$$

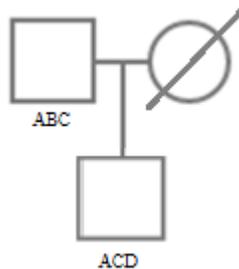
$$+ 2a^2 c^2 \Pr(1 - \text{trans B}) b$$

$$+ 2b^2 c^2 \Pr(1 - \text{trans A}) a$$

$$+ 4[ab]^2 c^2$$

$$+ 4[ac]^2 b^2$$

$$+ 4[bc]^2 a^2$$



Caso 14:

$$\begin{aligned}
 X = & abd (1/2) \Pr(\text{trans } C) \Pr(1 - \text{trans } C) \\
 & + bcd (1/2) \Pr(\text{trans } A) \Pr(1 - \text{trans } A) \\
 & + [ac]b ([ac]+b)
 \end{aligned}$$

Probabilidad de que el alelo A o C estén transferidos en cualquier cromosoma autosómico

Probabilidad de que cosegreden los alelos (ac).

$$\begin{aligned}
 Y = & 4a^2 b^2 \Pr(1 - \text{trans } C) (1/2) c \\
 & + 4b^2 c^2 \Pr(1 - \text{trans } A) (1/2) a \\
 & + 4[ac]^2 b^2
 \end{aligned}$$

Probabilidad de que el alelo A, B o C estén transferidos en cualquier cromosoma autosómico

Probabilidad de que cosegreden los alelos (ab), (ac) o (bc).

$$\begin{aligned}
 Y = & 2a^2 b^2 \Pr(1 - \text{trans } C) c \\
 & + 2b^2 c^2 \Pr(1 - \text{trans } A) a \\
 & + 4[ac]^2 b^2
 \end{aligned}$$

Simplificando la fórmula:

$$LR = \frac{(1/2) ad \Pr(\text{transl } C) \Pr(1 - \text{transl } C) + (1/2) cd \Pr(\text{transl } A) \Pr(1 - \text{transl } A) + [ac] ([ac]+b)}{2a^2 b^2 \Pr(1 - \text{trans } C)c + 2b^2 c^2 \Pr(1 - \text{trans } A)a + 4[ac]^2 b}$$

Caso 15: para los casos donde se coteja un elemento material probatorio (EMP) con un sospechoso pueden plantearse las fórmulas de la siguiente manera:

Marcador	Genotipo hallado en EMP	Genotipo del sospechoso
Marcador 1	ABC	ABC

Probabilidad de coincidencia: $2ab \Pr(\text{trans } C)$

$$LR = \frac{1}{2ab \Pr(\text{trans } C) \Pr(1 - \text{trans } C)}$$

7 Resultados y discusión

7.1 Análisis de las genealogías

En la familia de Frontino-Antioquia se observa un patrón trialélico del marcador TPOX con los alelos 12 y 13 siempre presentes en la descendencia, sugiriendo una segregación dependiente en el cromosoma 2 (Figura 9). Al realizar un Chi cuadrado bajo la hipótesis nula de co-segregación del alelo 12 y 13 se obtiene un resultado de 9,5 ($p \sim 0,1$) con 5 grados de libertad, dicho resultado acepta la hipótesis nula, pero no es concluyente, asimismo al realizar un LOD Score, el resultado no es > 3 , por el tamaño de la muestra. Además los genotipos observados en esta genealogía (Figura 9) no están descritos en el reporte del NIST (90), y tampoco posee el alelo 10, ni presenta una segregación compatible con la teoría del cromosoma X descrita por Lane et al. (29). A partir de esto se proponen dos hipótesis: 1. El origen molecular del patrón trialélico puede ser diferente en las familias analizadas o; 2. El alelo 10 puede presentar una mutación de dos o más pasos.

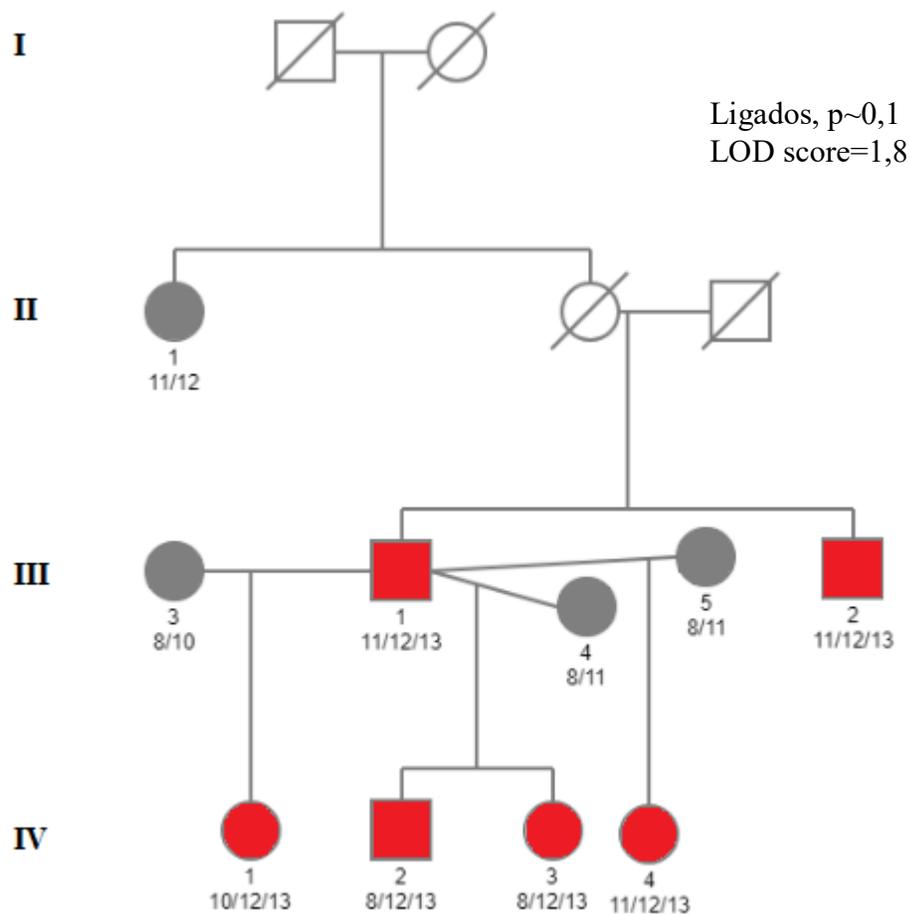


Figura 9: Genealogía de la familia de Frontino-Antioquia: En color rojo, se representan los individuos analizados, que portan el genotipo TPOX trialélico. La numeración de los individuos se encuentra encima de cada representación (cuadro o círculo) y los alelos observados están debajo de cada representación.

En la figura 9 se representa la familia de Frontino-Antioquia y se observa que todos los descendientes de la generación IV heredan la condición del patrón trialélico y probablemente los alelos 12 y 13 co-segreden, concordando con un estudio realizado con población de China y Corea que también se observa la transmisión del patrón trialélico del padre al hijo, pero no contienen el alelo 12 o el 13 (91), sin embargo, pueden tener el mismo origen molecular.

También, la influencia estadística del marcador trialélico, brinda mayor certeza a los resultados obtenidos en los laboratorios de genética forense y se propone en caso de ligamiento, se calcule la frecuencia de ese haplotipo, usar la frecuencia del patrón trialélico o se utilice la frecuencia alélica mínima (MAF) y se apliquen las diferentes posibilidades descritas anteriormente. En efecto, para una prueba de paternidad entre los individuos III-3, III-1 y IV-2, sin tener en cuenta la segregación familiar observada para aplicar las dos formas descritas de desarrollo.

$$\text{Primera forma: } LR = \frac{F_{12} + F_{13} + 1}{2Z(2F_{12}^2 + 2F_{13}^2 + 1)} = \frac{0,0895 + 0,0007 + 1}{2(0,0004)(2 \times 0,0895^2 + 2 \times 0,0007^2 + 1)} = \frac{1,0902}{0,000812800784} = 1341,29$$

Donde “F” representa la frecuencia de un alelo dado.

$$W = 0,99925$$

No obstante, se puede aplicar un factor de corrección “ $\theta = 0,03$ ” recomendado por el NRC de 1996 (37) de la siguiente manera para un resultado conservador:

$$LR = \frac{0,0902(1 - 0,03) + 1}{(0,0008)(0,0160 \times 0,97 + 0,00000098 \times 0,97 + 1)} = \frac{1,087494}{0,0012416008} = 875,88$$

$$W = 0,99885$$

Por consiguiente, el resultado del LR disminuye su valor, pero aun así aporta al producto final en conjunto con los demás marcadores. Ahora, usando la segunda forma, la cual se basa en la fórmula de mutación, se obtiene el siguiente resultado:

$$LR = \frac{z\left(\frac{1}{2}\right)F_{12} + \left(\frac{1}{2}\right)F_{13} + 1}{F_{12}^2 F_{13} + F_{13}^2 F_{12} + z^2} = \frac{(0,0004)(0,04475 + 0,00035 + 1)}{0,00801025(0,0007) + 0,00000049(0,0895) + 0,00000016} = \frac{(1,0451)(0,0004)}{0,00881074}$$

$$LR = \frac{0,00041804}{0,000005811} = 71,93$$

$$W = 0,9862$$

Siempre que se utilice la segunda forma se obtienen resultados más conservadores, pero la primera forma es la más correcta de abordar para un patrón trialélico, ya que tiene en cuenta la presencia del tercer alelo representada como “z” y su ausencia como “1-z”, abordando a un heterocigoto como $2(z)(1-z)$, donde “z” es igual a la frecuencia del patrón trialélico observado en Colombia en los ejemplos de Frontino y Quibdó.

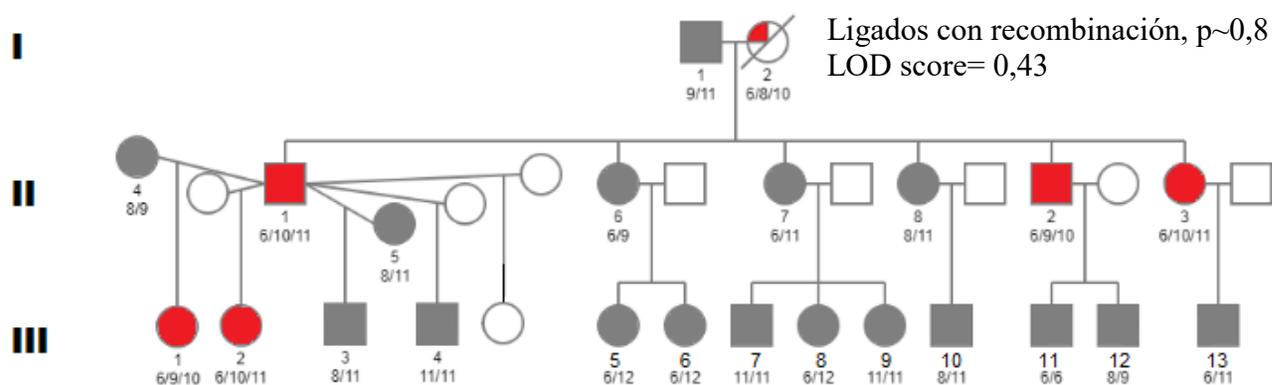


Figura 10: Genealogía de la familia afrocolombiana estudiada. En color rojo, se representan los individuos analizados, que portan el genotipo TPOX trialélico. En gris, los individuos estudiados, cuyo genotipo TPOX fue diploide. En blanco, los individuos que no participaron de la investigación. La numeración de los individuos se encuentra encima de cada representación (cuadro o círculo) y los alelos observados están debajo de cada representación.

Del mismo modo se analiza la familia de Quibdó- Chocó representada en la figura 10. Aquí, bajo la hipótesis nula de co-segregación de los alelos 6 y 10 en presencia de recombinación, se obtiene

un Chi cuadrado de 3,1414 ($p \sim 0,8$) con 6 grados de libertad, dando a entender la posible co-segregación de dichos alelos, aunque el LOD Score muestra una tendencia de ligamiento entre el alelo 6 y el 10, no es concluyente debido al tamaño muestral. Por consiguiente, no es posible descartar ni afirmar que el alelo extra esté ligado al cromosoma X como se sugiere en la literatura (29,71,76,79,80). Asimismo, cabe aclarar que ambas familias tienen una potencia de 0 para realizar estimaciones precisas.

En la figura 10 se representa a la familia de Quibdó-Chocó, donde los alelos 6 y 10 se encuentran en los individuos con patrones trialélicos, de los cuales se observa que tres individuos de la segunda generación, dos hombres y una mujer, son trialélicos para el STR TPOX (II-1, II-2 y II-3); sin embargo, el único elemento que se estudia de la primera generación, presenta un genotipo diploide para este marcador, aludiendo entonces que la herencia provino de la madre (I-2). Adicionalmente, el individuo II-1, les transmite tal característica a sus dos hijas mujeres (III-1 y III-2), pero no a los hijos varones, también se observa que el patrón trialélico del sujeto II-2 no lo hereda su hijo varón III-11 y III-12. Dichas observaciones en la segregación permiten mantener los indicios encontrados en la literatura sobre el alelo 10 ubicado en el cromosoma X (29,71,76,79,80), aunque también puede existir recombinación en los individuos II-6, II-7, III-11 y III-13, en consecuencia no se puede descartar la hipótesis de que el alelo 10 sea sintético con el alelo 6. Además, se han podido describir individuos varones con el alelo 10, transmitirlo al hijo varón, aunque no propiamente en población africana (91).

A fin de realizar la aplicación de las fórmulas propuestas para patrones trialélicos se toman los genotipos de los individuos II-1, II-4 y III-1. Entonces el desarrollo de la primera forma queda de la siguiente manera:

$$LR = \frac{2F_6 + 2F_{10} + 1}{z(2F_6^2 + 2F_{10}^2 + 1)} = \frac{2(0,0517) + 2(0,0881) + 1}{(0,0004)(2(0,00267289) + 2(0,00776161) + 1)} = \frac{1,2796}{0,0004083476} = 3133,6$$

$$W = 0,9997$$

De manera que se puede aplicar el factor de corrección “ $\theta=0,03$ ” recomendado por el NRC de 1996 (37) de la siguiente manera para un resultado conservador:

$$LR = \frac{(1,97)(0,0517) + (1,97)(0,0881) + 1}{(0,0004)(1,97(0,00267289) + 1,97(0,00776161) + 1)} = \frac{1,275406}{0,0004082224} = 3124,3$$

$$W = 0,9997$$

De manera que el LR disminuye su valor, pero en este caso no representa diferencia en el W y tampoco se afectaría el producto final del LR con los diferentes marcadores. En comparación con la segunda forma que es mucho más conservadora, y se observa una gran diferencia en el LR:

$$LR = \frac{(z)\left(\frac{3}{2}F_6 + \frac{3}{2}F_{10} + 1\right)}{(3F_6^2F_{10}) + (3F_{10}^2F_6) + 2z^2}$$

$$LR = \frac{(0,0004)\left(\frac{3}{2}(0,0517) + \frac{3}{2}(0,0881) + 1\right)}{3(0,00267289 \times 0,0881) + 3(0,00776161 \times 0,0517) + 2(0,00000016)}$$

$$LR = \frac{(0,0004)(0,07755+0,13215+1)}{0,0007064448+0,0012038242+0,00000032}$$

$$LR = \frac{0,00048388}{0,001910589} = 0,25$$

$$W = 0,3125$$

Dado que el resultado del LR es de 0,25, reduce el valor del producto de los diferentes marcadores, de esta manera se sugiere abordar los patrones trialélicos a partir de la primera forma y de ser necesario un resultado un poco conservador se puede usar un valor de corrección “ θ ” (92).

7.2 Ancestría con Indels-AIMs.

Una descripción de la ancestría en estas familias puede contribuir a inferir sobre el origen del patrón trialélico, considerando el alto componente Indígena de los individuos III-1 y III-2 de la familia de Frontino - Antioquia, que puede provenir de las etnias indígenas de Caquetá (Andakíes, Huitotos y los Coreguajes)(93,94), ya que el padre de dichos individuos (III-1 y III-2), nació en Florencia-Caquetá y migró a Frontino-Antioquia siendo el posible portador del patrón trialélico descrito. Asimismo, la herencia europea la pudo recibir por las incursiones españolas en dicha región desde el año 1542, aunque la ola de colonos surgió en el siglo XIX por la explotación del caucho(95), sin embargo también presentó incursiones efectuadas por comerciantes portugueses, criollos y negros que escapaban del sistema esclavista (96). Estos últimos quizás pudieron aportar el componente africano que se observa en los resultados de ancestría.

Sumado a la segregación observada y las diferencias en el origen de ambas familias, se sugiere que el alelo extra tiene una ancestría diferente en ambas familias, aunque debe realizarse una ancestría local para estudios futuros. En la siguiente grafica está presente el porcentaje de cada individuo:

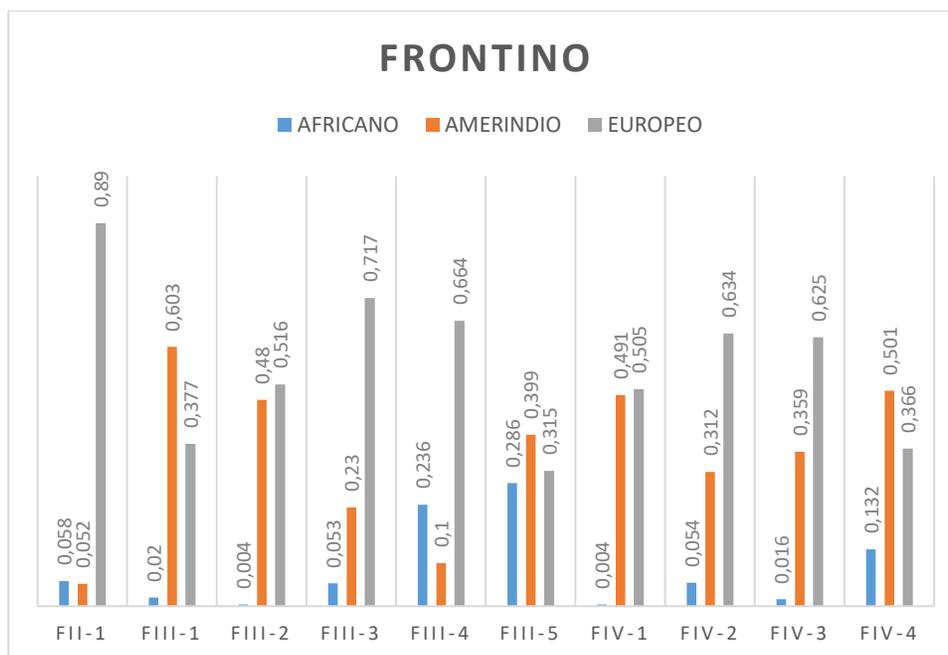


Grafico 1: Porcentaje de ancestría de la familia de Frontino-Antioquia.

De acuerdo a la gráfica 1, la familia de Quibdó-Chocó presenta una mayor ancestría africana, acorde a su ubicación geográfica, fenotipo e historia (97), además de coincidir con la segregación descrita en la literatura donde el alelo extra puede estar ubicado en el cromosoma X, y por ende su origen puede ser africano como lo sugiere *Lane et al.* (29).

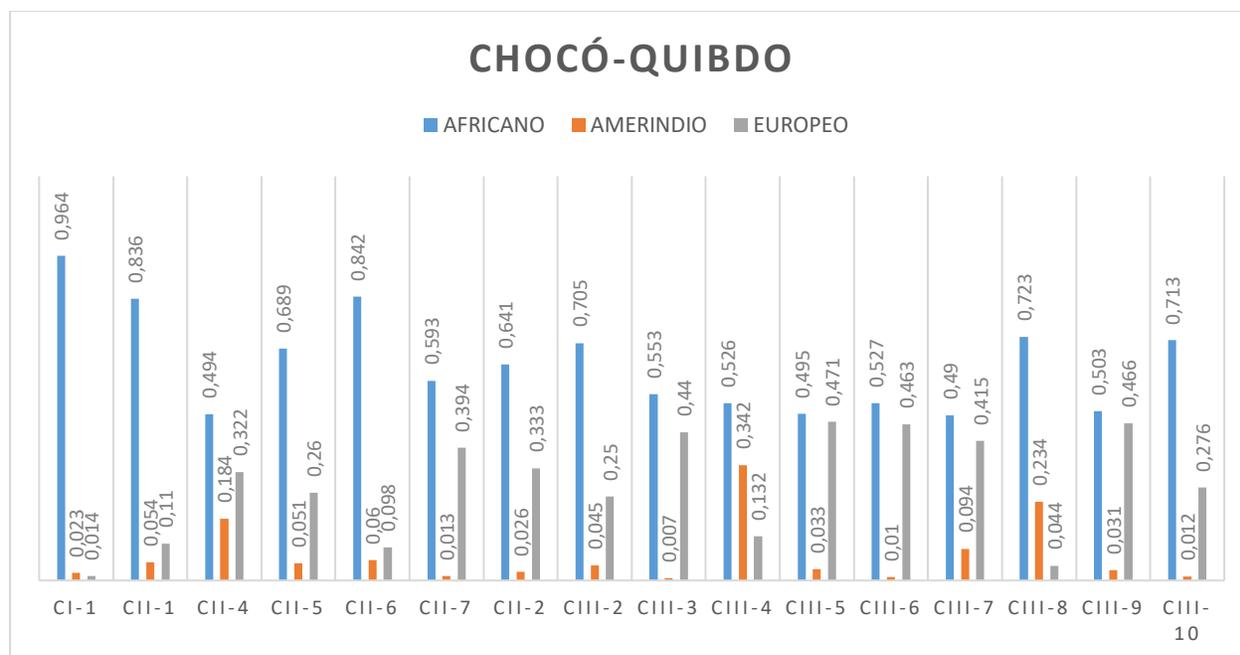


Grafico 2: Porcentaje de ancestría de la familia de Quibdó-Chocó.

7.3 Estadísticos forenses y patrones trialélicos tipo 2-B.

Las frecuencias alélicas obtenidas presentan al alelo 8 como el más frecuente seguido del alelo 11. Esto concuerda con estudios realizados en Colombia (11,98) y en otras poblaciones de Latinoamérica como Perú y México (99,100). Además esta tendencia también se observa en Afrodescendientes y Europeos (101). Sin embargo, en poblaciones del sur de Iraq los alelos más frecuentes son el 11,12 y 13 (102), aunado a ello, otras poblaciones de Serbia y aborígenes de Australia no tienen el alelo 8 como el más frecuente (103).

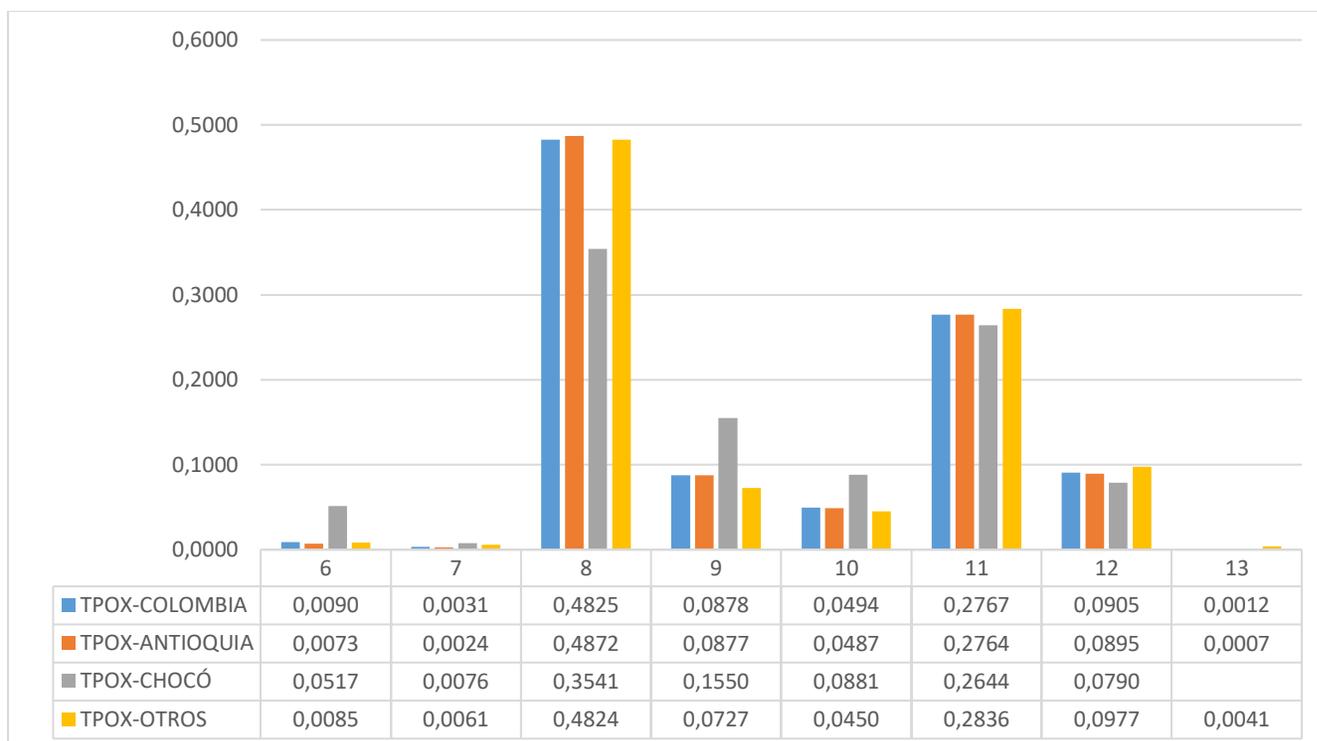


Grafico 3: Frecuencias alélicas del locus TPOX

Otras características del marcador TPOX, indican que es informativo ($PIC > 0,5$) (104,105), y posee un índice de paternidad típico ($PI_{typical}$) de 1,5 sin tener en cuenta los patrones trialélicos. Sumado a ello, los resultados de los diversos parámetros forenses son muy similares a estudios realizados en afrodescendientes, europeos y Latinos (99–101), aunque difieren con poblaciones de Iraq, serbia y aborígenes australianos que presentan un mayor polimorfismo (102,103). Aunque este marcador en general es poco polimórfico, se observan parámetros aceptables para ser usado en identificación forense y la presencia de patrones trialélicos aumenta el valor del LR .

	N	Ho	He	PIC	PE	PD	PM	PI_{typical}	MAF
TPOX-COLOMBIA	9999	0,6690	0,6723	0,6250	0,3820	0,6722	0,3278	1,5107	0,0003
TPOX-ANTIOQUIA	8191	0,6681	0,6681	0,6202	0,3806	0,6681	0,3319	1,5063	0,0003
TPOX-CHOCÓ	329	0,7356	0,7651	0,7295	0,4854	0,7639	0,2361	1,8908	0,0075
TPOX-OTROS	1479	0,6599	0,6701	0,6216	0,3690	0,6698	0,3302	1,4702	0,0017

Tabla 5: Resultados estadísticos de relevancia forense

Así mismo, los parámetros forenses entre la población colombiana no presentan mayores diferencias, aunque el $PI_{typical}$ en la población de Chocó es más alto debido a que el marcador TPOX es un poco más polimórfico, pero no se observa el alelo 13, pudiendo deberse a la baja cantidad de individuos incluidos en el cálculo de frecuencias (106).

A su vez el patrón trialélico tipo 2-B no se observa en la base de datos, quizás por la baja frecuencia que se presenta en la población y al tamaño de la muestra (79).

7.4 Ubicación genómica del alelo extra

En la base de datos de *1000 genomes* se pueden identificar hasta 132 *SV* (107), en los cuales se involucra la ubicación del TPOX en el cromosoma 2:1489617-1489848(108). En consecuencia existen varias alternativas que pueden explicar la presencia del alelo extra sinténico en el cromosoma 2, pero no se reporta evidencia de transferencia de dicha región a otro lugar del genoma(107). Aunque a través de un alineamiento in silico, usando una región del intrón 10, se encontraron similitudes en diferentes lugares del genoma, que pueden explicar la transferencia de dicho alelo extra, incluyendo el cromosoma X. Como ya se ha mencionado, puede presentarse

recombinación con diversos lugares del genoma con una homología mayor del 90% como se observa en los cromosomas 19, 13, 7, 5 y en el X:13665448-13665707(28), que podría dar explicación a la segregación descrita en población afrodescendiente (29).

Con el objetivo de ubicar el alelo extra, se analizó el genoma completo de un individuo de cada familia, el individuo IV-4 de Frontino y el individuo II-3 de Quibdó. Para el ensamblaje del genoma se seleccionó la herramienta *Speedseq* y para la búsqueda de *SV* se usó *Lumpy express*, ya que estas herramientas incluyen el uso de distintas formas de analizar el alineamiento de los *reads* como *read-pair*, *split-read and read-count* (109), y se observa que para el cromosoma 2 en la familia de Frontino hay 458 *SV* y en la familia de Quibdó 474 *SV*, pero solo se observan *SV* cerca al gen TPO.

SV DE FRONTINO			SV DE QUIBDÓ		
TIPO	UBICACIÓN CROMOSOMA 2	TAMAÑO(pb)	TIPO	UBICACIÓN CROMOSOMA 2	TAMAÑO(pb)
DELECIÓN	1094252-1094900	648	DELECIÓN	1094240-1094900	660
DELECIÓN	1596108-1597169	1061	DUPLICACIÓN	1565664-1566368	704
*Ubicación del Gen TPO: Cromosoma 2: 1,374,066-1,543,711					
*Ubicación del marcador TPOX: Cromosoma 2:1489617-1489848					

Tabla 6. Resultado de las *SV* más cercanas al marcador TPOX. Las *SV* observadas no contienen el marcador TPOX ni parte del gen TPO.

De acuerdo a lo anterior, el análisis se centra en el gen TPO, pero no se evidencian cambios en la profundidad de los *reads* (*read-count*), como se observa en la figura 11 y tampoco se observan variantes estructurales que contengan la ubicación del TPOX, esto puede deberse a que la

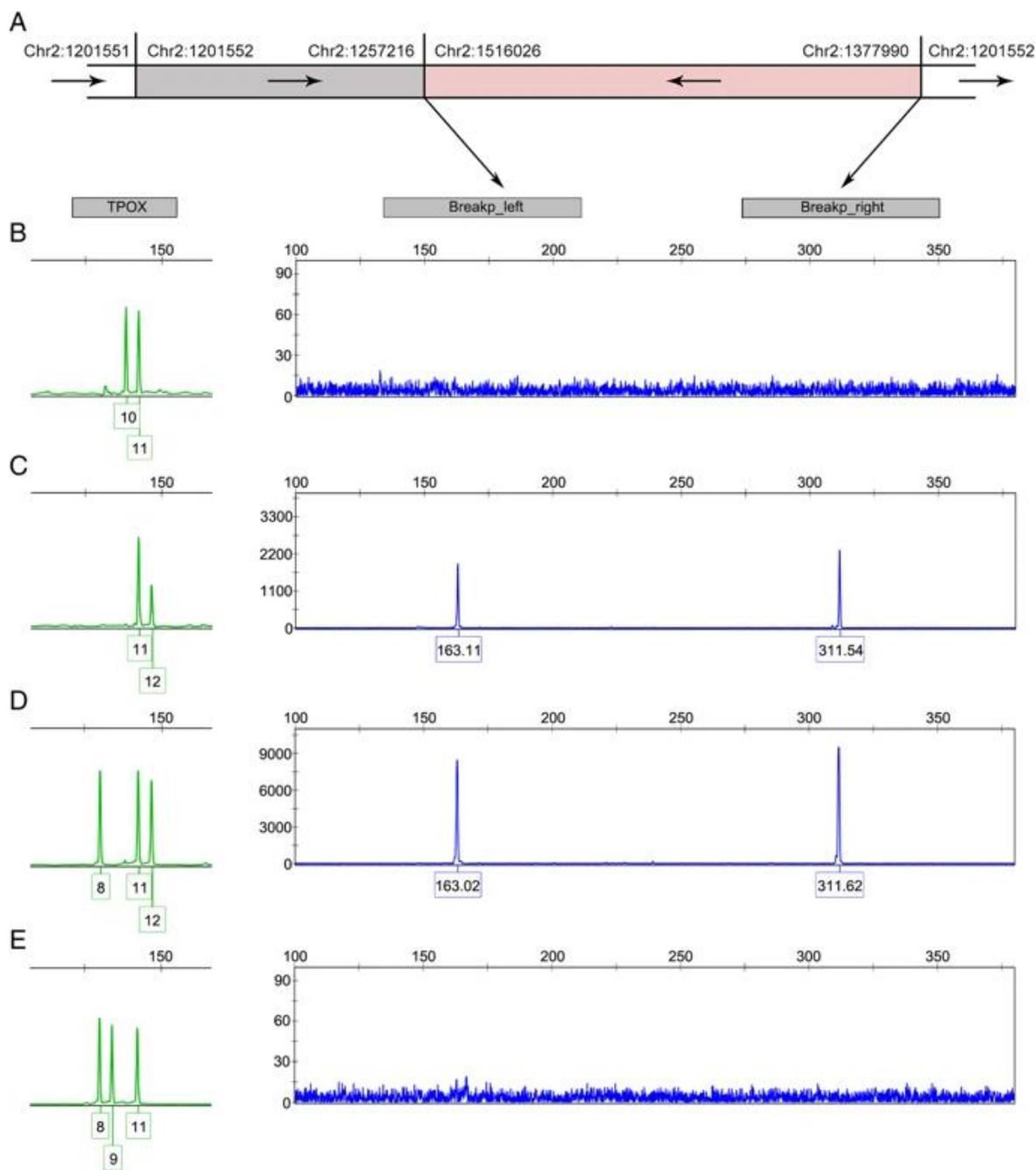


Figura 12: Tomada de Yang et al. La detección de dos puntos de ruptura en individuos no relacionados con patrón trialélico en TPOX. A, diagrama esquemático de la duplicación y reorganización de la copia adicional de TPOX. La región gris representa una duplicación de SNTG2 y la región rosa representa una duplicación de TPO. Las flechas indican la dirección de los genes en el cromosoma 2. B, La detección de dos uniones de punto de ruptura por Electroforesis Capilar en un individuo normal y, C, D Y E, tres individuos no relacionados con patrón tri-alelico en TPOX(91).

Los resultados obtenidos por Yang et al, indican que no todos los individuos presentan el mismo punto de ruptura y por ende los patrones trialélicos del locus TPOX pueden tener orígenes

moleculares diferentes(91), y por la segregación observada en el presente estudio el origen molecular puede ser distinto en ambas familias.

Sumado a lo anterior, también se observa que los SNPs cercanos al gen TPO tienen 4 variantes en la familia de Quibdó que solo se observan en poblaciones africanas y están ausentes en la familia de Frontino. Además, se observan 2 variantes en la familia de Frontino que no se han descrito en población africana y están ausentes en la familia de Quibdó. Debido a lo anterior, ambos patrones trialélicos pueden tener un origen diferente como lo propuesto por Yang et al(91). Aunado a ello, al patrón trialélico de la familia de Quibdó se le sugiere un origen Africano, mientras que a la Familia de Frontino se le propone un origen Europeo o Amerindio, pero se debe realizar una ancestría local que confirme lo que se observa en ambos genomas. A pesar de que se observan muchas variantes de SNPs e Indels, se toman las más relevantes para realizar la comparación en las regiones cercanas al gen TPO, tabuladas a continuación:

VARIANTES GENOMA DE FRONTINO		VARIANTES GENOMA DE QUIBDÓ	
rs28915689*	G/T	rs558315966	T/A
rs4641984	C/T	rs556166876	A/T
*Frecuencia < 1% en población africana.		rs142501081	C/T
		rs114651797	C/T

Tabla 7. Variantes únicas de cada genoma: las 4 variantes observadas en Quibdó solo se han descrito en población africana, mientras que las 2 variantes de Frontino no se han descrito en población africana o están con una frecuencia inferior al 1%.

Por otro lado, el resultado del aCGH del individuo II-3 de Quibdó no muestra signos de duplicación en el locus del marcador TPOX, pudiendo deberse a las limitaciones de la técnica, ya que la

plataforma oligo-SNP's de Affymetrix CytoScan 750K usa un microarreglo que contiene 750.436 sondas, incluidos 550.000 sondas no polimórficos y 200.436 de marcadores de SNP's, con una resolución de hasta 750Kbp (114), y si el tamaño de la duplicación es cercana a la descrita por Yang et al, no podría observarse la duplicación mediante esta técnica (91), por otro lado, en un estudio anterior se describe una duplicación de alrededor de 1,59Mpb (76), pero se descarta por el resultado del microarreglo.

Además de lo anterior, no siempre los microarreglos pueden detectar la presencia de aberraciones cromosómicas y tampoco se puede visualizar el tipo de reordenamiento en el caso de que la matriz detecte la deleción o la duplicación (115). También, en un estudio anterior realizado con cariotipo sobre los individuos de Chocó con patrón trialélico, no se observa ninguna alteración que pueda dar explicación a dicho patrón (84), a diferencia del microarreglo usado en el presente estudio, el cariotipo descarta grandes alteraciones del genoma, de 4 a 10 Mpb de tamaño, descartando translocaciones o eventos que involucren translocaciones-inversiones (116).

TIPO	UBICACIÓN	TAMAÑO (pb)
DUPLICACION	2q13(112,192,266-112,566,199)	373,933
DELECIÓN	5q21.1(99,412,589-99,659,949)	247,360
DELECIÓN	8p11.22(39,226,335-39,379,654)	153,319
DUPLICACIÓN	14q32.33(106,167,581- 106,863,507)x3	695,926
LOH (Perdida de heterocigocidad)	Xq22.2q23(103,672,013- 112,259,594)	8,587,581
DELECIÓN	Xq13.3(75,410,627-75,587,764)	177,137
DELECIÓN	Xq11.2q12(64,573,287-64,717,718)	144,431

Tabla 8: Resultado del microarreglo.

Con relación a las duplicaciones descritas que involucran al marcador TPOX, se han reportado 30 casos en la base de datos de variantes genómicas (DGV) (117), 11 casos en la base de datos de Normas Internacionales para Matrices Citogenómicas (ISCA) (118), y 14 casos reportados por Repnikova et al. (119). Aunque en varios casos de los ya mencionados se relaciona con enfermedades, en el presente estudio no se observa ninguna relación aparente con algún fenotipo patológico y se descartan trastornos tiroideos, ya que a cada individuo se le realiza una medición de TSH y T4 libre con resultados dentro del intervalo biológico permitido.

8 Conclusiones

En este estudio, se discutieron las principales características de los patrones trialélicos en TPOX, y se observó que los alelos 8 y 11 son los más frecuentes en la población colombiana, además de que el patrón trialélico no está siempre ligado al alelo 10 o a una segregación ligada al cromosoma X, sino que también puede tener una segregación diferente con otros alelos como el 12 y el 13, aparentemente sinténicos. También se realizó la secuenciación del genoma completo, revelando posibles orígenes diferentes en el patrón trialélico de ambas familias, pero se debe confirmar con ancestría local en futuros estudios. Además, se sugiere un origen molecular diferente por las diferencias observadas en la segregación y por las diversas variantes observadas en los genomas. Por otro lado, se determinó que la influencia del marcador trialélico brinda mayor poder estadístico a los resultados obtenidos en los laboratorios de genética forense y se propone en caso de ligamiento, se calcule la frecuencia de ese haplotipo, usar la frecuencia del patrón trialélico o se utilice la frecuencia alélica mínima y se apliquen las diferentes posibilidades descritas anteriormente. Aunque no se determinó la ubicación del alelo extra debido a las limitaciones de la técnica, se sugiere reanalizar los genomas con el algoritmo propuesto por Yang et al(91). Finalmente, la técnica por microarreglos es limitada al no poder detectar ganancias genómicas o pérdidas menores a su resolución (750Kpb), sugiriendo que la duplicación presente requiere un método de mayor resolución si se desea estimar el tamaño.

9 Referencias bibliográficas

1. Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet.* marzo de 2011;12(3):179-92.
2. Ibarra A. Estudio de marcadores genéticos bialélicos para aplicaciones forenses: SNPs autosómicos e Indels específicos de cromosoma X. [Medellín]: Universidad de Antioquia; 2015.
3. Senge T, Madea B, Junge A, Rothschild MA, Schneider PM. STRs, mini STRs and SNPs – A comparative study for typing degraded DNA. *Leg Med.* marzo de 2011;13(2):68-74.
4. Posada Y. Polimorfismos Inserción/Delección – Indels: Una herramienta genética en la identificación plena de individuos NN en Colombia. [Bogotá, Colombia]: Escuela de Investigación Criminal - ESINC; 2014.
5. ADN en Medicina Legal: Problemas [Internet]. [citado 27 de julio de 2017]. Disponible en: http://www.biologia.arizona.edu/human/problem_sets/DNA_forensics_2/01t.html
6. Klug W, Cummings S M, Charlotte S, Palladino M. Conceptos de genética. 10.^a ed. España: Pearson Educación, S.A., Madrid; 2013.
7. Klug W, Cummings M, Charlotte S, Palladino M. Conceptos de genética. En: 10.^a ed. España: Pearson Educación, S.A., Madrid; 2013. p. 544-55.
8. Martínez-Sevilla VM, Aguilar-Velázquez JA, Inclán-Sánchez A, Romero-Rentería O, Díaz-Navarro XX, Saiz M, et al. Parámetros forenses del sistema Powerplex® 21 (Promega Corp.) en población mestiza del occidente de México. *Rev Esp Med Leg.* :10-6.
9. University of Arizona. Overview of Blackett Family DNA Activity [Internet]. 2000 [citado 24 de marzo de 2017]. Disponible en: http://www.biology.arizona.edu/human_bio/activities/blackett2/overview.html
10. Bravo M. La verdad genética de la paternidad. Universidad de Antioquia; 2009. 166 p.
11. Reyes HO, Peñaranda ET, Moreno H, Horta J, Moreno G. Frecuencias alélicas de 14 STR's autosómicos en una población de Antioquia, Colombia. *NOVA- PUBLICACION Cient EN Cienc Biomed.* 6 de enero de 2010;8(13):25-9.
12. Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC. Origin, evolution and genome distribution of microsatellites. *Genet Mol Biol* [Internet]. 2006 [citado 17 de abril de 2018];29(2):294-307. Disponible en: http://www.scielo.br/scielo.php?script=sci_abstract&pid=S1415-47572006000200018&lng=en&nrm=iso&tlng=en

13. Messier W, Li S-H, Stewart C-B. The birth of microsatellites. *Nature* [Internet]. junio de 1996 [citado 17 de abril de 2018];381(6582):483. Disponible en: <https://www.nature.com/articles/381483a0>
14. Arcot SS, Wang Z, Weber JL, Deininger PL, Batzer MA. Alu Repeats: A Source for the Genesis of Primate Microsatellites. *Genomics* [Internet]. 1 de septiembre de 1995 [citado 17 de abril de 2018];29(1):136-44. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0888754385712244>
15. Ramsay L, Macaulay M, Cardle L, Morgante M, Ivanissevich S degli, Maestri E, et al. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *Plant J* [Internet]. 1 de febrero de 1999 [citado 17 de abril de 2018];17(4):415-25. Disponible en: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-313X.1999.00392.x>
16. Paredes M. Análisis mutacional de microsatélites humanos implicaciones evolutivas, poblacionales y forenses [Internet]. [Bogota, Colombia]: Universidad Nacional De Colombia; 2014. Disponible en: <http://www.bdigital.unal.edu.co/39690/1/manuelparedes.2014.pdf>
17. Goodwin W, Linacre A, Hadi S. *An Introduction to Forensic Genetics*. John Wiley & Sons; 2011. 232 p.
18. Fan H, Chu J-Y. A Brief Review of Short Tandem Repeat Mutation. *Genomics Proteomics Bioinformatics* [Internet]. 2007 [citado 18 de abril de 2018];5(1):7-14. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5054066/>
19. Lai Y, Sun F. The Relationship Between Microsatellite Slippage Mutation Rate and the Number of Repeat Units. *Mol Biol Evol* [Internet]. 1 de diciembre de 2003 [citado 19 de abril de 2018];20(12):2123-31. Disponible en: <https://academic.oup.com/mbe/article/20/12/2123/978562>
20. Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc Natl Acad Sci U S A* [Internet]. 4 de febrero de 1997 [citado 19 de abril de 2018];94(3):1041-6. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC19636/>
21. Amos W, Sawcer SJ, Feakes RW, Rubinsztein DC. Microsatellites show mutational bias and heterozygote instability. *Nat Genet*. agosto de 1996;13(4):390-1.
22. Wechselblatt B, Rudd MK. Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet TIG* [Internet]. octubre de 2015 [citado 19 de abril de 2018];31(10):587-99. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4600437/>
23. Shimmin LC, Chang BH-J, Li W-H. Male-driven evolution of DNA sequences. *Nature* [Internet]. abril de 1993 [citado 19 de abril de 2018];362(6422):745-7. Disponible en: <https://www.nature.com/articles/362745a0>

24. Pierce BA. *Genética: Un enfoque conceptual*. Ed. Médica Panamericana; 2009. 836 p.
25. Huang Q-Y, Xu F-H, Shen H, Deng H-Y, Liu Y-J, Liu Y-Z, et al. Mutation Patterns at Dinucleotide Microsatellite Loci in Humans. *Am J Hum Genet*. marzo de 2002;70(3):625-34.
26. Bhargava A, Fuentes FF. Mutational dynamics of microsatellites. *Mol Biotechnol*. marzo de 2010;44(3):250-66.
27. Carvalho CMB, Lupski JR. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet*. abril de 2016;17(4):224-38.
28. Homo sapiens thyroid peroxidase (TPO), RefSeqGene on chromosome 2. 24 de febrero de 2019 [citado 23 de abril de 2019]; Disponible en: http://www.ncbi.nlm.nih.gov/nucore/NG_011581.1
29. Lane AB. The nature of tri-allelic TPOX genotypes in African populations. *Forensic Sci Int Genet*. marzo de 2008;2(2):134-7.
30. Gettings KB, Aponte RA, Vallone PM, Butler JM. STR allele sequence variation: Current knowledge and future issues. *Forensic Sci Int Genet*. septiembre de 2015;18:118-30.
31. CODIS and NDIS Fact Sheet [Internet]. Federal Bureau of Investigation. [citado 27 de julio de 2017]. Disponible en: <https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet>
32. Espuny MJF. Introducción a la tecnología del ADN aplicada en el laboratorio forense. *Estud Juríd* [Internet]. 2004 [citado 19 de septiembre de 2018];(2004). Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=4050840>
33. Usaquén Martínez W. Validación y consistencia de información en estudios de diversidad genética humana a partir de marcadores microsatélites [Internet] [phd]. Universidad Nacional de Colombia; 2012 [citado 12 de agosto de 2019]. Disponible en: <http://www.bdigital.unal.edu.co/8963/>
34. Jarreta MBM. La prueba del ADN en medicina forense: la genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad [Internet]. Masson; 1999 [citado 20 de septiembre de 2018]. Disponible en: <https://dialnet.unirioja.es/servlet/libro?codigo=71666>
35. González-Andrade F. Análisis molecular de variación de polimorfismos STR autosómicos y de cromosoma «Y» en grupos étnicos de Ecuador [Internet] [<http://purl.org/dc/dcmitype/Text>]. Universidad de Zaragoza; 2006 [citado 23 de octubre de 2018]. Disponible en: <https://dialnet.unirioja.es/servlet/tesis?codigo=13742>
36. Holsinger KE, Weir BS. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nat Rev Genet*. septiembre de 2009;10(9):639-50.

37. Update NRC (US) C on DFSA. Overview [Internet]. National Academies Press (US); 1996 [citado 1 de mayo de 2019]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK232606/>
38. Cifuentes LO, Martínez EH, Acuña MP, Jonquera HG. Probability of exclusion in paternity testing: time to reassess. *J Forensic Sci.* marzo de 2006;51(2):349-50.
39. Butler JM. Chapter 10 - STR Population Data Analysis. En: Butler JM, editor. *Advanced Topics in Forensic DNA Typing: Interpretation* [Internet]. San Diego: Academic Press; 2015 [citado 22 de abril de 2019]. p. 239-79. Disponible en: <http://www.sciencedirect.com/science/article/pii/B9780124052130000105>
40. Yepes JC. Matemáticas aplicadas a la genética forense. En: *Matemáticas aplicadas a la genética forense*. Ministerio del Interior, Secretaría General Técnica; 2006. p. 33-5.
41. Rodríguez AWH, Medinilla F de MT. Estudio Genético Poblacional de Frecuencias Alélicas para 15 marcadores STR presentes en la Población del Estado de Zacatecas Aplicado a la Práctica Forense. *Arch Med* [Internet]. 2014 [citado 13 de noviembre de 2019];10(1):1. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=4635959>
42. Ardilla G. Colombia: migraciones transnacionalismo y desplazamiento [Internet]. CES; 2006. Disponible en: <https://www.ugr.es/~cuadgeo/docs/articulos/041/041-013.pdf>
43. Novoa E. La metamorfosis de la cuestión espacial en Colombia. BogotáUniversidad Nac Colomb. 2010;
44. Palacios M, Safford F. Historia de Colombia: país fragmentado, sociedad dividida [Internet]. 2012 [citado 19 de septiembre de 2018]. Disponible en: <http://babel.banrepcultural.org/cdm/pageflip/collection/p17054coll10/id/2581/type/compoundobject/show/2579/cpdtype/document/pftype/pdf>
45. Antioquiatic. Culturas Indígenas de Antioquia [Internet]. [citado 21 de octubre de 2018]. Disponible en: <http://www.antioquiatic.edu.co/noticias-general/item/226-culturas-indigenas-de-antioquia>
46. Alvarez V. HISTORIA DE ANTIOQUIA. Un ensayo de Historia Regional. LA FORMACION DE LA SOCIEDAD COLONIAL EN ANTIOQUIA.
47. Carl Henrik L. ANTECEDENTES INDÍGENAS DEL URBANISMO COLONIAL EN DOS REGIONES DE COLOMBIA: LOS ANDES ORIENTALES Y EL VALLE DE ABURRÁ. UNA VISIÓN DESDE LA ARQUEOLOGÍA. *Rev Estud Soc* [Internet]. febrero de 2002 [citado 21 de octubre de 2018];(11):47-56. Disponible en: <https://revistas.uniandes.edu.co/doi/10.7440/res11.2002.05>
48. Victor Alvarez. La sociedad colonial en Antioquia: Mestizaje y exclusión. En: *Globalización, cultura y poder en Colombia : una mirada interdisciplinaria*. 1.^a ed. Medellín; p. 121-46.

49. La conquista de Antioquia, 1500-1580 [Internet]. [citado 23 de abril de 2019]. Disponible en: http://www.jorgeorlandomelo.com/conquista_de_antioquia.htm
50. Nuestro Departamento [Internet]. 2015 [citado 22 de octubre de 2018]. Disponible en: https://web.archive.org/web/20150923202930/http://www.choco.gov.co/informacion_general.shtml#historia
51. HISTORIA DEL CHOCÓ [Internet]. [citado 22 de octubre de 2018]. Disponible en: <http://www.angelfire.com/or/choco/HISTORIA.html>
52. Historia del Chocó ~ Departamento del Chocó [Internet]. [citado 22 de octubre de 2018]. Disponible en: <http://magazinchoco.blogspot.com/2015/04/historia-del-choco.html>
53. Generalidades y Símbolos del Choco [Internet]. [citado 22 de octubre de 2018]. Disponible en: <http://www.lavozdelaselva.galeon.com/productos965456.html>
54. Chocó en la cartografía histórica: de territorio incierto a departamento de un país llamado Colombia | banrepcultural.org [Internet]. 2011 [citado 22 de octubre de 2018]. Disponible en: <https://web.archive.org/web/20111104154955/http://www.banrepcultural.org/blaavirtual/publicacionesbanrep/boletin/boleti1/bol43/b43o.htm>
55. Consulta de la Norma: [Internet]. [citado 18 de octubre de 2017]. Disponible en: <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=9565>
56. Gjertson DW, Brenner CH, Baur MP, Carracedo A, Guidet F, Luque JA, et al. ISFG: Recommendations on biostatistics in paternity testing. *Forensic Sci Int Genet.* 1 de diciembre de 2007;1(3):223-31.
57. Egeland T, Kling D, Mostad P. Chapter 2 - Basics. En: *Relationship Inference with Families and R* [Internet]. San Diego: Academic Press; 2016 [citado 17 de octubre de 2017]. p. 11-50. Disponible en: <http://www.sciencedirect.com/science/article/pii/B9780128024027000023>
58. Butler JM. Chapter 11 - DNA Profile Frequency Estimates and Match Probabilities. En: *Advanced Topics in Forensic DNA Typing: Interpretation* [Internet]. San Diego: Academic Press; 2015 [citado 6 de octubre de 2017]. p. 281-308. Disponible en: <http://www.sciencedirect.com/science/article/pii/B9780124052130000117>
59. Luque J. Estadística aplicada a la genética forense.
60. Tattini L, D'Aurizio R, Magi A. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. *Front Bioeng Biotechnol* [Internet]. 25 de junio de 2015 [citado 23 de abril de 2019];3. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4479793/>
61. Magi A, Tattini L, Pippucci T, Torricelli F, Benelli M. Read count approach for DNA copy number variants detection. *Bioinforma Oxf Engl.* 15 de febrero de 2012;28(4):470-8.

62. Earl D, Bradnam K, St John J, Darling A, Lin D, Fass J, et al. Assemblathon 1: a competitive assessment of de novo short read assembly methods. *Genome Res.* diciembre de 2011;21(12):2224-41.
63. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* diciembre de 2011;7(12):e1002384.
64. Thung DT, de Ligt J, Vissers LEM, Steehouwer M, Kroon M, de Vries P, et al. Mobster: accurate detection of mobile element insertions in next generation sequencing data. *Genome Biol.* 2014;15(10):488.
65. Baldwin EL, Lee J-Y, Blake DM, Bunke BP, Alexander CR, Kogan AL, et al. Enhanced detection of clinically relevant genomic imbalances using a targeted plus whole genome oligonucleotide microarray. *Genet Med Off J Am Coll Med Genet.* junio de 2008;10(6):415-29.
66. Edelmann L, Hirschhorn K. Clinical utility of array CGH for the detection of chromosomal imbalances associated with mental retardation and multiple congenital anomalies. *Ann N Y Acad Sci.* enero de 2009;1151:157-66.
67. Bejjani BA, Shaffer LG. Application of Array-Based Comparative Genomic Hybridization to Clinical Diagnostics. *J Mol Diagn JMD [Internet].* noviembre de 2006 [citado 14 de noviembre de 2019];8(5):528-33. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1876176/>
68. Rawe V, Cuzzi J. La CGH con arrays (aCGH) para todos los cromosomas: Un concepto avanzado para el diagnóstico genético preimplantatorio. [citado 12 de noviembre de 2019];27(2). Disponible en: <http://www.samer.org.ar/revista/volumen27.asp?numero=2>
69. Kimura S, Hong YS, Kotani T, Ohtaki S, Kikkawa F. Structure of the human thyroid peroxidase gene: comparison and relationship to the human myeloperoxidase gene. *Biochemistry.* 16 de mayo de 1989;28(10):4481-9.
70. TPO thyroid peroxidase [Homo sapiens (human)] - Gene - NCBI [Internet]. [citado 13 de julio de 2017]. Disponible en: <https://www.ncbi.nlm.nih.gov/gene/7173>
71. Picanço JB, Raimann PE, Paskulin GA, Alvarez L, Amorim A, Batista dos Santos SE, et al. Tri-allelic pattern at the TPOX locus: A familial study. *Gene.* 10 de febrero de 2014;535(2):353-8.
72. AmpFISTR® Identifiler® PCR Amplification [Internet]. [citado 24 de marzo de 2017]. Disponible en: http://tools.thermofisher.com/content/sfs/manuals/cms_039944.pdf
73. Nakagawa T, Maruyama K, Takeyama H, Matsunaga T. Determination of microsatellite repeats in the human thyroid peroxidase (TPOX) gene using an automated gene analysis system with nanoscale engineered biomagnetite. *Biosens Bioelectron.* 15 de abril de 2007;22(9–10):2276-81.

74. Paredes M, Galindo A, Bernal M, Avila S, Andrade D, Vergara C, et al. Analysis of the CODIS autosomal STR loci in four main Colombian regions. *Forensic Sci Int.* 14 de octubre de 2003;137(1):67-73.
75. Laboratorio de Identificación Genética - IdentiGEN. Frecuencias UdeA 2014 - Chocó. 2014.
76. Lukka M, Tasa G, Ellonen P, Moilanen K, Vassiljev V, Ulmanen I. Triallelic patterns in STR loci used for paternity analysis: Evidence for a duplication in chromosome 2 containing the TPOX STR locus. *Forensic Sci Int.* 1 de diciembre de 2006;164(1):3-9.
77. Clayton TM, Guest JL, Urquhart AJ, Gill PD. A Genetic Basis for Anomalous Band Patterns Encountered During DNA STR Profiling. *J Forensic Sci.* 2004;49(6):1-8.
78. Fridman C, dos Santos PCC, Kohler P, Garcia CF, Lopez LF, Massad E, et al. Brazilian population profile of 15 STR markers. *Forensic Sci Int Genet.* marzo de 2008;2(2):e1-4.
79. Picanço JB, Raimann PE, Motta CHAS da, Rodenbusch R, Gusmão L, Alho CS. Identification of the third/extra allele for forensic application in cases with TPOX tri-allelic pattern. *Forensic Sci Int Genet.* mayo de 2015;16:88-93.
80. Díaz V, Rivas P, Carracedo A. The presence of tri-allelic TPOX genotypes in Dominican Population. *Forensic Sci Int Genet Suppl Ser.* diciembre de 2009;2(1):371-2.
81. Castañeda Fernández M. Estudio de los microsatélites y miniSTRs del cromosoma X de aplicación forense. 14 de junio de 2013 [citado 24 de marzo de 2017]; Disponible en: <https://repositorio.unican.es/xmlui/bitstream/handle/10902/3053/TesisMCF.pdf?sequence=7>
82. Watson J, Baker T, Bell S, Gann A, Levine M, Losick R. The mutability and repair of DNA. *Molecular Biology of the gene.* En: *Molecular Biology of the gene.* 7.^a ed. PEARSON EDUCACIÓN, Inc; 2014. p. 911.
83. Hernández A. La visibilización estadística de los grupos étnicos colombianos. Colombia: DANE; 2005.
84. Velez A, Posada Y. Caracterización del genotipo trialélico del microsatelite (STR)TPOX en una familia afrocolombiana. En Bucaramanga;
85. Butler JM. Chapter 5 - STR Profiles: Multiplex PCR, Tri-Alleles, Amelogenin, and Partial Profiles. En: *Advanced Topics in Forensic DNA Typing: Interpretation* [Internet]. San Diego: Academic Press; 2015 [citado 6 de octubre de 2017]. p. 109-27. Disponible en: <http://www.sciencedirect.com/science/article/pii/B9780124052130000051>
86. Manta FS de N, Pereira R, Vianna R, Araújo ARB de, Gitaí DLG, Silva DA da, et al. Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLOS ONE* [Internet]. 20 de septiembre de 2013 [citado 14 de noviembre de 2019];8(9):e75145. Disponible en: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0075145>

87. Rojas MY, Alonso LA, Sarmiento VA, Eljach LY, Usaquén W. Structure analysis of the La Guajira-Colombia population: a genetic, demographic and genealogical overview. *Ann Hum Biol.* marzo de 2013;40(2):119-31.
88. illumina. Sequencing Technology | Sequencing by Synthesis [Internet]. Sequencing by Synthesis (SBS) Technology. [citado 27 de abril de 2017]. Disponible en: <https://www.illumina.com/technology/next-generation-sequencing/sequencing-technology.html>
89. Yepes JC. Capitulo 5. En: *Matemáticas aplicadas a la genética forense*. Ministerio del Interior, Secretaría General Técnica; 2006. p. 109-78.
90. TPOX Variants [Internet]. [citado 16 de septiembre de 2018]. Disponible en: https://strbase.nist.gov/var_TPOX.htm#Tri
91. Yang Q, Liu B, Shao C, Zhou Y, Yao Y, Pan Y, et al. Characterization of the extra copy of TPOX locus with tri-allelic pattern. *BMC Genet* [Internet]. 14 de febrero de 2019 [citado 12 de noviembre de 2019];20(1):18. Disponible en: <https://doi.org/10.1186/s12863-019-0723-2>
92. Update NRC (US) C on DFSA. Statistical Issues [Internet]. National Academies Press (US); 1996 [citado 2 de mayo de 2019]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK232615/>
93. Silva PM. Introducción histórica a la etnia Coreguaje: Cultura Tucano occidental de Colombia. *Rev Bras Linguística Antropológica* [Internet]. 25 de octubre de 2013 [citado 18 de noviembre de 2019];5(1):193-215. Disponible en: <https://periodicos.unb.br/index.php/ling/article/view/16547>
94. Chaves M. Florencia: conectora de regiones [Internet]. *Revista Credencial*. 2011 [citado 17 de noviembre de 2019]. Disponible en: <http://www.revistacredencial.com/credencial/historia/temas/florencia-conectora-de-regiones>
95. HISTORIA DEL CAQUETÁ - Gobernación de Caquetá [Internet]. [citado 18 de noviembre de 2019]. Disponible en: <http://www.caqueta.gov.co/departamento/historia-del-caqueta>
96. Herrera y Tordesillas A de. Historia General de los hechos de los castellanos en las Islas i tierra firme del Mar oceano. En *Ma[drid]: en la Emplenta Real*; 1601.
97. Información general - Gobernación del Chocó [Internet]. [citado 17 de noviembre de 2019]. Disponible en: <http://www.choco.gov.co/departamento/informacion-general>
98. Ocampo1 2*, Martinez1 RA, Rocha JF, Cardona H, Ocampo1 2*, Martinez1 RA, et al. Genetic characterization of Colombian indigenous sheep. *Rev Colomb Cienc Pecu* [Internet]. junio de 2017 [citado 22 de abril de 2019];30(2):116-25. Disponible en: http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S0120-06902017000200116&lng=en&nrm=iso&tlng=en

99. Alej, WendolyHernandez-Rodriguez R, TrejoMedinilla F de M. Estudio Genético Poblacional de Frecuencias Alélicas para 15 marcadores STR presentes en la Población del Estado de Zacatecas Aplicado a la Práctica Forense. Arch Med [Internet]. 2014 [citado 5 de febrero de 2019];10(1). Disponible en:
<http://www.archivosdemedicina.com/abstract/estudio-gentico-poblacionalde-frecuencias-allicas-para-15marcadores-str-presentes-en-lapoblacin-del-estado-de-zacatecasaplicado-a-501.html>
100. Delgado E, Neyra CD. Frecuencias alélicas de 21 marcadores STR autosómicos en una población mestiza peruana aplicado a la práctica forense. Rev Esp Med Leg [Internet]. octubre de 2018 [citado 22 de abril de 2019];S0377473218300622. Disponible en:
<https://linkinghub.elsevier.com/retrieve/pii/S0377473218300622>
101. Butler JM, Schoske R, Vallone PM, Redman JW, Kline MC. Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. J Forensic Sci. julio de 2003;48(4):908-11.
102. Hameed IH, Ommer AJ, Murad AF, Mohammed GJ. Allele frequency data of 21 autosomal short tandem repeat loci in Mesan and Basra provinces in South Iraq. Egypt J Forensic Sci [Internet]. 1 de diciembre de 2015 [citado 22 de abril de 2019];5(4):150-6. Disponible en:
<http://www.sciencedirect.com/science/article/pii/S2090536X1400080X>
103. ALLSTR*^R Autosomal Database for short tandem repeats [Internet]. [citado 22 de abril de 2019]. Disponible en:
<http://allstr.de/allstr/showMarker.seam?markerName=TPOX&populationId=ccc3d0b1-a214-448c-82eb-642a1826afac&accessionNumber=ASTRTP283312&cid=461440>
104. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. mayo de 1980;32(3):314-31.
105. Molecular Genetics, Gene Transfer, and Therapy. Elsevier; 1997. 244 p.
106. Bravo MLJ, Builes JJ, de Pancorbo MM, Moreno MA. Analysis of 12 STR loci in Antioquia (Colombia) population sample. Int Congr Ser [Internet]. 1 de abril de 2004 [citado 6 de mayo de 2019];1261:151-3. Disponible en:
<http://www.sciencedirect.com/science/article/pii/S0531513103015322>
107. Gene: TPO (ENSG00000115705) - Structural variants - Homo sapiens - Ensembl genome browser 96 [Internet]. [citado 2 de mayo de 2019]. Disponible en:
http://www.ensembl.org/Homo_sapiens/Gene/StructuralVariation_Gene?db=core;g=ENSG00000115705;r=2:1374066-1543711
108. Marker: TPOX - Details - Homo sapiens - Ensembl genome browser 96 [Internet]. [citado 2 de mayo de 2019]. Disponible en:
http://www.ensembl.org/Homo_sapiens/Marker/Details?m=TPOX

109. Layer RM, Chiang C, Quinlan AR, Hall IM. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol* [Internet]. 26 de junio de 2014 [citado 5 de noviembre de 2018];15(6):R84. Disponible en: <https://doi.org/10.1186/gb-2014-15-6-r84>
110. HiSeq X Series | Ultra-high-throughput 1000 dollar genome sequencing [Internet]. [citado 2 de mayo de 2019]. Disponible en: <https://www.illumina.com/systems/sequencing-platforms/hiseq-x.html>
111. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* [Internet]. julio de 2013 [citado 2 de mayo de 2019];15(7):565-74. Disponible en: <https://www.nature.com/articles/gim201373>
112. Clark BE, Shooter C, Smith F, Brawand D, Thein SL. Next-generation sequencing as a tool for breakpoint analysis in rearrangements of the globin gene clusters. *Int J Lab Hematol*. mayo de 2017;39 Suppl 1:111-20.
113. Chiang C, Layer RM, Faust GG, Lindberg MR, Rose DB, Garrison EP, et al. SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nat Methods* [Internet]. 10 de agosto de 2015;12:966. Disponible en: <https://doi.org/10.1038/nmeth.3505>
114. Affymetrix Support by Product for The CytoScan 750K Array Kit and Reagent Kit Bundle - Available Outside US/Canada Only [Internet]. [citado 16 de diciembre de 2019]. Disponible en: http://www.affymetrix.com/support/technical/byproduct.affx?product=CytoScan750K_Array
115. Evangelidou P, Alexandrou A, Moutafi M, Ioannides M, Antoniou P, Koumbaris G, et al. Implementation of High Resolution Whole Genome Array CGH in the Prenatal Clinical Setting: Advantages, Challenges, and Review of the Literature [Internet]. *BioMed Research International*. 2013 [citado 18 de noviembre de 2019]. Disponible en: <https://www.hindawi.com/journals/bmri/2013/346762/>
116. Albertson DG, Pinkel D. Genomic microarrays in human genetic disease and cancer. *Hum Mol Genet*. 15 de octubre de 2003;12 Spec No 2:R145-152.
117. Stankiewicz P, Lupski JR. The genomic basis of disease, mechanisms and assays for genomic disorders. *Genome Dyn*. 2006;1:1-16.
118. Riggs ER, Jackson L, Miller DT, Van Vooren S. Phenotypic information in genomic variant databases enhances clinical care and research: the International Standards for Cytogenomic Arrays Consortium experience. *Hum Mutat*. mayo de 2012;33(5):787-96.
119. Repnikova EA, Rosenfeld JA, Bailes A, Weber C, Erdman L, McKinney A, et al. Characterization of copy number variation in genomic regions containing STR loci using array comparative genomic hybridization. *Forensic Sci Int Genet*. septiembre de 2013;7(5):475-81.