



UNIVERSIDAD NACIONAL DE COLOMBIA

Análisis de marcadores moleculares Y-SNPs para una población bogotana y su aplicación en procesos de identificación humana

Victor Fernando Hidalgo Cerón

Universidad Nacional de Colombia
Facultad de Medicina, Departamento de Morfología,
Bogotá, Colombia

2015

Análisis de marcadores moleculares Y-SNPs para una población bogotana y su aplicación en procesos de identificación humana

Victor Fernando Hidalgo Cerón

Tesis o trabajo de investigación presentada como requisito parcial para optar al título de:

Magister en Genética Humana

Director:

PhD., MSc, Esp, Biólogo. William Usaquén Martínez

Línea de Investigación:

Genética de Poblaciones Humanas

Grupo de Investigación:

Genética de Poblaciones e Identificación

Universidad Nacional de Colombia

Facultad de Medicina, Departamento de Morfología,

Maestría en Genética Humana

Bogotá, Colombia

2015

A Dios por guiar mi camino, por darme salud, vida y concederme la sabiduría para alcanzar este gran paso en mi vida.

A mis padres y hermana por su amor incondicional, sus consejos y confiar en mí en cada momento.

A Angela porque en ti reconozco el verdadero significado del amor, soñando y creyendo en un futuro juntos.

A Marianita, la razón de todos mis actos, mi inspiración y la dueña de mi vida.

A todas las personas, amigos y familiares que de alguna u otra forma me apoyaron y colaboraron en la realización de este sueño.

Agradecimientos

Gracias a Dios por ser el guía en cada uno de mis pasos y darme la oportunidad de alcanzar esta meta.

A mi familia por su gran apoyo incondicional.

A mis profesores y colegas por sus enseñanzas.

Al Grupo de Genética de Poblaciones e Identificación de la Universidad Nacional de Colombia, en especial a mi director de tesis el profesor William Usaquén, que contribuyo con todos sus conocimientos y por la confianza que mostró en mí en todo momento, además de su valiosa paciencia y por demostrar siempre su calidez humana convirtiéndose ante todo en un verdadero amigo.

A la doctora Leonor Gusmao (University of Porto, State University of Rio de Janeiro) por su apoyo, su valiosa colaboración y asesoría en la realización de este proyecto.

Resalto un agradecimiento muy especial a la doctora Martha Roa Bohórquez, Directora Técnica del Grupo Genética de la DNCTI Fiscalía General de la Nación, por su invaluable y desinteresada colaboración en la ejecución de este proyecto.

A Erica Salguero por su importante asesoría durante la fase técnica de este proyecto.

A Andrea Casas por su abnegada hospitalidad y colaboración en sus indicaciones.

También expreso mi más sincera gratitud a todo el Grupo Genética de la DNCTI Fiscalía General de la Nación, por su calidez humana, por brindarme su amistad, colaboración y confianza, además del apoyo técnico y la financiación para la ejecución de este proyecto.

Y a mis amigos y compañeros por su constante motivación y confianza.

Resumen

La investigación en el área de la Genética Forense se ha enfocado en el análisis de marcadores informativos en casos complejos y de identificación humana, no obstante, requiere del apoyo de bases de datos poblacionales robustas que sirvan como referencia en cálculos probabilísticos para toda conclusión forense. Los marcadores Y-SNPs son acertados para el análisis genético y en especial en el campo forense porque expresan diversas ventajas y privilegios respecto a otros marcadores, resultando estratégicos para el estudio del linaje masculino de una muestra poblacional.

La dinámica histórica de Colombia concentra el efecto migratorio en su ciudad capital, lo que ofrece un marco de estudio genético-poblacional conveniente para el establecimiento de registros genéticos como fuente referencial en el análisis con aplicación social, ancestral y forense. En este caso, se discrimina efectivamente la distribución de haplogrupos y subhaplogrupos de 146 muestras provenientes de individuos que habitan en la ciudad de Bogotá, con representación mayoritaria de los linajes europeos (R1b) comunes de la población de península Ibérica, linajes africanos (E1b1) y linajes nativos americanos (Q), lo que permite descifrar el comportamiento de la población masculina bogotana y orientar este conocimiento en el campo de la identificación humana, en aquellos casos donde se pretenda establecer un parentesco o la autoría de un crimen y no se pueda acudir a los marcadores empleados de rutina. Deduciendo de igual manera, que los marcadores seleccionados describen convenientemente el mestizaje colombiano representado en la población bogotana, y pueden resultar útiles si se implementan como herramienta de análisis en casos complejos para laboratorios de genética forense del país.

Palabras clave: Y-SNPs, Haplogrupo, Bogotá, Cromosoma Y, Identificación humana, Linaje.

Abstract

Research in the area of forensic genetics has focused on the analysis of informative markers in complex human identification cases, however, requires the support of robust population data bases that serve as a reference in probabilistic calculations for any forensic conclusion. Markers Y-SNPs are successful for genetic analysis and especially in the forensic field because they express different benefits and privileges over other markers, strategic for the study of the male lineage of a population sample resulting.

The historical dynamics of Colombia accounts for migratory effect in the capital city, which offers a convenient framework for establishing genetic records as a reference source in social analysis, and forensic application ancestral genetic-population study. In this case, the distribution of haplogroups and subhaplogroups of 146 samples from individuals living in the city of Bogotá, with majority representation of common European lineages (R1b) of the population of Iberian peninsula, African lineages (E1b1) effectively discriminates and Native American lineages (Q), allowing decipher the behavior of Bogota males and guide this knowledge in the field of human identification, in cases where it intends to establish a relationship or authorship of a crime and cannot be go to the markers used routinely. Deducing the same way, the markers selected conveniently describe the Colombian miscegenation represented in the Bogota population, and may be useful if implemented as an analytical tool in complex cases for forensic genetics laboratories in the country.

Keywords: Y-SNPs, Haplogroup, Bogotá, Y chromosome, Human Identification, Lineage.

Contenido

Introducción	1
1. La era del Cromosoma -Y en análisis poblacionales y forenses	5
1.1 Polimorfismos en el Cromosoma –Y	6
1.1.1 Marcadores multialélicos	6
1.1.2 Marcadores bialélicos	8
1.2 Historia de los Linajes Masculinos.....	9
1.2.1 Haplogrupos del cromosoma -Y.....	10
1.2.2 Distribución jerárquica de las mutaciones	12
1.3 Asociación entre haplotipos de Y-SNPs y Y-STRs	13
1.4 Ventajas del empleo de marcadores de cromosoma –Y en el campo forense ..	14
2. Desarrollo y adaptación de un método técnico-científico para el estudio de linajes masculinos en la población	18
2.1 Metodología de trabajo	19
2.1.1 Voluntarios y Extracción de ADN.....	20
2.1.2 Cuantificación de ADN por RealTime-PCR	20
2.1.3 Tipificación y Análisis de Y-STRs	20
2.1.4 Selección, Diseño y Validación de Primers para Amplificación	21
2.1.5 Evaluación de los Primers con Gel de Electroforesis	25
2.1.6 Condiciones de la PCR Multiplex.....	25
2.1.7 Purificación de los productos de PCR	27
2.1.8 Selección de Primers para Minisequenciación	28
2.1.9 Reacción de Minisequenciación por SNaPSHOT	29
2.1.10 Purificación del Producto de Minisequenciación.....	30
2.1.11 Electroforesis Capilar	30
2.1.12 Análisis de Datos.....	31
2.1.13 Ensayos de Sensibilidad del Método	32
2.1.14 Pruebas de Especificidad del Método.....	32
2.2 Resultados y Discusión.....	32
2.3 Conclusiones.....	37

3. Estudio de la ancestría y estructura del linaje masculino para una población humana de Bogotá (Colombia)	40
3.1 Materiales y métodos	42
3.1.1 Población analizada	42
3.1.2 Análisis molecular.....	43
3.1.3 Análisis estadístico	43
3.2 Resultados.....	44
3.2.1 Análisis de redes filogenéticas.....	45
3.2.2 Análisis de componentes principales.....	49
3.3 Discusión de resultados.....	51
4. Evaluación del desempeño del método en procesos de identificación humana	55
4.1 Evaluación del método en casos de mezcla de perfiles genéticos.....	58
4.2 Evaluación del método en casos negativos y de degradación de ADN	60
5. Aportes del estudio de marcadores del Cromosoma –Y en el campo forense .	67
5.1 Proyecciones hacia el futuro.....	70
6. Conclusiones y recomendaciones	72
6.1 Conclusiones generales del estudio	72
6.2 Recomendaciones	73
Anexo 1: Encuesta	74
Bibliografía	75

Lista de figuras

Figura 1-1: Distribución de haplogrupos de cromosoma –Y en el mundo.....	12
Figura 1-2: Árbol filogenético del cromosoma –Y.....	13
Figura 2-1: Árbol filogenético con parsimonia máxima para los Y-SNPs seleccionados en el estudio.....	22
Figura 2-2: Geles de Agarosa al 2,2% con los productos de PCR de los 15 amplicones.....	24
Figura 2-3: Reacción de PCR Multiplex.....	25
Figura 2-4: Representación del tratamiento del producto de PCR con ExoSAP-IT.....	27
Figura 2-5: Reacción de Extensión de base simple.....	29
Figura 2-6: Diferenciación de la movilidad electroforética según el tamaño de la cola...	31
Figura 2-7: Haplotipo de Y-STRs para una de las muestras analizadas en el estudio....	32
Figura 2-8: Electroferograma de las dos reacciones multiplex para dos muestras diferentes.....	33
Figura 2-9: Frecuencias de los haplogrupos definidos por los 15 Y-SNPs.....	34
Figura 2-10: Distribución de los principales haplogrupos de Cromosoma –Y encontrados en una población Bogotana. Frecuencias de los sublinajes de los haplogrupos R, E y J.	35
Figura 3-1: Median-joining network del haplogrupo R1b utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.....	45
Figura 3-2: Median-joining network del haplogrupo E utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.....	46
Figura 3-3: Median-joining network del haplogrupo G utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.....	47
Figura 3-4: Median-joining network del haplogrupo J utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.....	47
Figura 3-5: Median-joining network del haplogrupo Q utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.....	48

Figura 3-6: Análisis de componentes principales. A distribución de los haplotipos respecto los haplogrupos. B distribución de los haplotipos respecto las regiones de origen.....	49
Figura 3-7: Estructura genética de la población colombiana.....	50
Figura 4-1: Electroferograma de la tipificación de STRs autosómicos de la mezcla.....	59
Figura 4-2: Electroferograma de los Y-SNPs para la muestra con mezcla de perfiles.....	60
Figura 4-3: Imágenes de algunos de los restos óseos analizados.....	63
Figura 4-4: Electroferogramas de las dos reacciones multiplex para las muestras negativas.....	64

Lista de tablas

Tabla 2-1: SNPs de cromosoma –Y y la secuencia de los primers para amplificación....	23
Tabla 2-2: Condiciones de la reacción de PCR Singleplex.....	24
Tabla 2-3: Condiciones de amplificación para la PCR multiplex.....	26
Tabla 2-4: Secuencias de primers para minisequenciación. Secuencia roja en negrilla corresponde a las colas inespecíficas.....	28
Tabla 2-5: Condiciones de reacción para la minisequenciación con SNaPShot.....	29
Tabla 2-6: Correspondencia entre el color y el nucleótido.....	31
Tabla 3-1: Frecuencias de haplogrupos de cromosoma –Y para la población bogotana.	44

Introducción

Colombia es un país pluricultural, multiétnico, de origen predominantemente mestizo andino o costeño. En el contexto nacional la ampliación de la base de datos referencial en el ámbito genético se encuentra en una fase de crecimiento paulatino, debido a los aportes graduales en el análisis de diferentes marcadores genéticos informativos para el campo poblacional. En general, Colombia presenta una gran diversidad humana dada su ubicación geográfica, un amplio grado de mestizaje de poblaciones autóctonas con poblaciones europeas y africanas a partir de la época de la conquista, así como condiciones ambientales que determinan características genéticas definidas en los diferentes grupos poblacionales de nuestro país.

Si bien, no se cuenta con registros de la variabilidad genotípica-fenotípica- de las distintas regiones colombianas, unas indígenas, otras afrodescendientes aunque de mayoría mestiza caucasoide, se puede describir a la población Bogotana como una sociedad cosmopolita, lo que ofrece un marco de estudio genético-poblacional conveniente para el establecimiento de registros genéticos como fuente referencial en el análisis con aplicación social, ancestral y forense. La capital colombiana es calificada como un centro económico, cultural, académico, social, industrial y turístico; que acoge alrededor de 250 mil personas al año ([Albuja & Ceballos, 2010](#)), lo que la convierte en una candidata acertada para un ambicioso estudio poblacional, y a su vez inevitablemente vincula el adecuado manejo de variables socio-políticas ampliamente influyentes en las tendencias migratorias hacia la capital.

De los más de 8 millones de habitantes de Bogotá, más de 3 millones son de otras regiones ([DANE, 2003](#)), una prueba de la importancia de esta ciudad como centro urbano receptor de mayores y mejores oportunidades de empleo, pues cuenta con la más alta concentración de actividades llamativas. Sin duda alguna, el recrudecimiento del conflicto en el área rural de los últimos años y el asolamiento de la violencia promovida por los grupos ilegales presentes en todo el territorio nacional ha generado desplazamientos

masivos repentinos que culminan en el asentamiento capitalino en búsqueda de oportunidades determinándose como un aspecto de contundente influencia en la representatividad poblacional bogotana. No obstante, estos eventos recientes no son la única causa de aporte al crecimiento demográfico de la capital, diferentes oleadas migratorias a través de la historia también han hecho parte de este incremento, que se unen a momentos históricos a nivel nacional tales como: guerras civiles, luchas y diferencias de índole política, una alta tasa de flujo comercial, la denominación como Distrito Capital del país, e incluso mayor acceso a la educación y finalmente la centralización gubernamental y administrativa del país.

En las últimas décadas se han evidenciado los graves problemas sociales suscitados por el conflicto armado, el terrorismo y la delincuencia común, entre ellos la dramática situación de casi 2 millones de desplazados que con el tiempo incrementan alarmantemente esta cifra, de igual manera, para los últimos veinte años se registran casi 80.000 desaparecidos en todo el territorio nacional ([Registro, Nacional, & Desaparecidos, 2015](#)). Esta particular situación demuestra la necesidad de desarrollar procedimientos y estándares de identificación, tanto para personas vivas como desaparecidas que apoyen la labor de reparación de víctimas de la violencia.

La disponibilidad de estudios genético - poblacionales actual en Colombia se ha visto limitada principalmente por la debilidad del diseño experimental en la fase de muestreo, puesto que estos estudios incluyen el análisis de individuos tomando como criterio de inclusión solamente su lugar de nacimiento, sin tener en cuenta otro tipo de variables que describen de una mejor manera a una población. Entre estos descriptores se deben incluir el tiempo de permanencia de estos individuos en dicho espacio geográfico, que brinde un aporte biológico efectivo garantizando la continuidad de su linaje en la población. Con base en este aspecto, surge la necesidad de hacer énfasis en el análisis retro y prospectivo de cada uno de los individuos a nivel familiar, como criterio de selección.

En los últimos años, la investigación en el área de genética forense se ha enfocado en el análisis de marcadores suplementarios a los marcadores STRs convencionales usados en la casuística forense de casos complejos, con el fin de ser validados en los laboratorios de identificación humana. El avance tecnológico actual ofrece las

herramientas apropiadas para apoyar estudios poblacionales desde la perspectiva genética, ofreciendo técnicas confiables altamente informativas que complementan los análisis de marcadores microsatélites (STRs); la herramienta biológica empleada para cumplir con las necesidades del estudio corresponde a los Polimorfismos de Nucleótido Simple (SNP por sus siglas en inglés), los cuales son considerados como los polimorfismos más frecuentes y por tanto, los más informativos en el genoma porque son muy abundantes y son la base de la individualidad genética, además, su simplicidad hace que sean analizados simultáneamente a partir de escasas cantidades de ADN, haciéndolos prácticos y adecuados para el análisis con tecnologías de alto rendimiento.

En contraste con los STR los SNPs presentan mayor potencia para el análisis simultáneo de distintos marcadores mediante el desarrollo de sistemas multiplex porque son eficaces para ser analizados mediante técnicas de amplificación génica Reacción en Cadena de la Polimerasa (PCR) y minisequenciación ofreciendo una gran sensibilidad, permitiendo el análisis de este tipo de marcadores genéticos a partir de muestras con cantidades muy pequeñas de ADN e incluso cuando este ADN se encuentra en un avanzado estado de degradación, así, el uso de amplicones cortos que aumenta el éxito en la amplificación de ADN degradado.

Teniendo en cuenta que en el campo de la identificación humana los estudios de parentesco juegan un papel crucial en el apoyo de la administración de justicia, se podría hacer referencia a aquellos marcadores que estratégicamente ofrecen ventajas tanto en aspectos analíticos como técnico-científicos, encontramos SNPs en Cromosoma –Y, autosómicos y ADN mitocondrial, sin embargo, el estudio de SNPs en el cromosoma “Y” genera ventajas en la evaluación de la diversidad de haplogrupos para una población, y en este caso permite caracterizar ancestralmente el linaje patrilíneo de la población Bogotana. Así y todo, los Y-SNPs expresan una baja tasa de mutación conveniente para estudios genealógicos (de paternidad), de identificación y ancestría (evolutivos), por ende, su ocurrencia representa un evento único en la evolución de la especie humana resultando estratégicos en la búsqueda de marcadores genéticos informativos con amplio espectro de aplicación. Estas propiedades combinadas con la ausencia de recombinación en gran parte del cromosoma, y la consecuente acumulación de la diversidad genética entre linajes y poblaciones relacionadas, hace a los Y-SNPs una herramienta útil en la predicción del origen étnico y geográfico de un individuo o una muestra apoyándose en la información generada por los Y-STRs, proyectando así una

importante aplicación en el campo forense, teniendo en cuenta que el 98% de los crímenes violentos involucra la participación de hombres.

De esta manera, y en respuesta a la limitada disponibilidad actual de estudios que suplan la carencia de información sobre la caracterización genética de la población Bogotana, se pretende establecer un proyecto que registre el análisis de Y-SNPs informativos que promuevan el desarrollo de la distribución de haplogrupos y haplotipos, así como las correspondientes inferencias ancestrales suscitadas de dichos análisis.

Todos estos argumentos llevan a desarrollar un estudio poblacional confiable, con un diseño experimental lo suficientemente robusto que permita definir fehacientemente el grado de utilidad de estos marcadores y su poder de discriminación empleado en el campo forense además de poder inferir la historia de poblamiento patrilineal, contrastando con las hipótesis existentes sobre la conformación genética de la capital colombiana. Por tanto, este estudio busca emplear herramientas genético-moleculares para describir la conformación de linajes paternos actual de la capital, proyectando su aplicación en el campo de la identificación humana, contemplando el diverso aporte poblacional del panorama capitalino actual.

1. La era del Cromosoma -Y en análisis poblacionales y forenses

El cromosoma -Y contiene alrededor de 60 Mpb. El 60% de este ADN está constituido por secuencias polimórficas, altamente repetidas y está confinado principalmente a la porción heterocromática del brazo largo ([Roewer et al., 1996](#)) y a la región pericentromérica sugiriendo que estas regiones tendrían una funcionalidad limitada.

Se puede afirmar que el cromosoma -Y es un componente importante del genoma humano debido a que lleva a cabo funciones biológicas esenciales como lo son la determinación del sexo y su intervención en la fertilidad masculina, además, debido a la falta de un componente homólogo (que determina una haploidia parcial), la mayor parte del cromosoma -Y no se recombina durante la meiosis. Solo se produce recombinación con el cromosoma -X en dos pequeñas regiones pseudoautosómicas denominadas PAR1 y PAR2 que constituyen solo 3 Mpb de sus 60 Mpb ([M. A. Jobling, Pandya, & Tyler-Smith, 1997](#)). La falta de recombinación determina que todas las secuencias ubicadas se heredan como un bloque constituyendo un grupo de ligamiento o haplotipo.

La secuencia del cromosoma -Y humano fue publicada por primera vez en Nature 423: 825-837 el 19 de junio de 2003. En la actualidad la secuencia del cromosoma -Y está disponible y la forma de acceder a ella es a través de la Universidad de California – Santa Cruz – Genome Bioinformatics website <http://genome.ucsc.edu> o del Center for Biotechnology Information <http://www.ncbi.nlm.nih.gov>.

En el haplotipo de cromosoma -Y se localizan regiones polimórficas que son cedidas intactas de generación en generación por línea paterna y en forma obligada, ellas únicamente cambian por mutación expresando así un registro informativo de su historia. Dichas funciones esenciales le conceden una ventaja para realizar estudios de

migración, y resulta además atractivo para establecer análisis con bases antropológicas, evolutivas y forenses ([Butler, 2003](#)).

Los procesos que implican cambios aleatorios de las frecuencias alélicas que tienden a reducir la variabilidad dentro de una población e incrementan la diferenciación genética entre poblaciones serán más intensos en las regiones ligadas al cromosoma -Y que en cualquier otra región del genoma (exceptuando el ADN mitocondrial), por tanto, al ser su filogenia menos profunda, el cromosoma -Y ha tenido menos tiempo para acumular cambios, lo que se traduce en que contenga una densidad de polimorfismos menor de la de los autosomas. A lo largo de la historia, se encuentra que la dinámica de la población masculina ha sido diferente a la de la población femenina. Actividades como la guerra, la caza, o la poliginia en determinados casos pueden haber reducido el número efectivo de cromosomas -Y en las poblaciones generando haplotipos asociados a grupos poblacionales específicos. De esta manera, se puede inferir que todo familiar por línea paterna tiende a asentar su residencia en el territorio cultural y geográfico de sus ancestros, permitiendo establecer asociaciones para el origen poblacional de un perfil genético dado.

1.1 Polimorfismos en el Cromosoma –Y

Se clasifican en varios tipos según su naturaleza y el interés informativo a la hora de interpretarlos y analizarlos.

1.1.1 Marcadores multialélicos

Para el caso que aquí compete se acude a los polimorfismos de longitud como los Y-STRs o microsatélites, de los cuales se han identificado más de 220 marcadores y que potencialmente ofrecen utilidad en aplicaciones de genética forense ([Gusmao et al., 2006](#)). Cada Y-STR consta de secuencias cortas de entre 2 – 5 nucleótidos de longitud, y minisatélites de los cuales solo se encuentran dos que comprenden secuencias mucho más grandes de entre 10 – 60 pares de bases de longitud ([Gusmao et al., 2005](#)).

Eventualmente pueden ocurrir cambios en las secuencias resultando en un incremento o disminución en el número de repeticiones y por consiguiente individualizando a un alelo específico, el cual testifica la presencia de dicho cambio generación tras generación

convirtiéndolo en un marcador informativo de todo el linaje paterno, teniendo en cuenta además, que obligatoriamente todo individuo vinculado por vía patrilineal comparte por varias generaciones un perfil de Y-STRs idéntico al de sus parientes masculinos.

Es importante hacer referencia a las altas tasas de mutación de estos marcadores, en promedio 6-11% por cada generación por loci minisatelite y ~0.2% por generación por Y-STRs ([Dupuy, Stenersen, Egeland, & Olaisen, 2004](#)), esto hace que un marcador multialélico de esta naturaleza sea extremadamente útil en la discriminación de haplotipos de cromosoma-Y con bastante resolución ([Butler, 2003](#)), tanto así, que se puede hacer una selección de marcadores según su característica polimórfica para incrementar el poder de discriminación del haplotipo.

El minisatélite MSY1 es el marcador bialélico específico del cromosoma -Y más variable, que contiene una secuencia repetida rica en AT de entre 48 a 114 copias. El MSY1 es altamente polimórfico, con variación estructural alta y una considerable tasa de mutación, no obstante, ha demostrado dificultad en la tipificación de muestras degradadas, lo cual, lo convierte en un fuerte candidato para investigación en estudios poblacionales únicamente ([Gusmao, Brion, Gonzalez-Neira, Lareu, & Carracedo, 1999](#)).

Para el caso del MSY2 (DYS440) que es el minisatélite más grande, contiene una secuencia repetida entre 99 y 110 pb mostrando una baja tasa de mutación y niveles detectables de variación en diversos linajes. En la unidad se encuentra una corta secuencia palindrómica CCTAGG y duplicada en la forma CCTAGGCCTAGG ([Bao et al., 2000](#)).

Considerando algunas ventajas que ofrece el uso de marcadores Y-STR como alternativa a otro tipo de análisis, se puede incluir la rápida y eficiente exclusión o inclusión de implicados en casos de parentesco, paternidad o de escena de crimen, empleados también en el análisis de individuos vasectomizados o con semen azoospermico, así mismo, debido a la naturaleza hemicigota de los Y-STR se puede determinar el número de involucrados en delitos de violencia sexual grupal o en casos de asalto sexual donde se puede individualizar el perfil masculino en medio de altas cantidades de ADN femenino, y como ya se ha mencionado, poseen valiosa utilidad en estudios de linajes ancestrales para hacer inferencias históricas, antropológicas y familiares, según la necesidad de cada caso.

Las limitaciones de los Y-STRs son pocas en comparación a otros marcadores, apelando al bajo poder de discriminación en casos donde no se puede excluir a un grupo de individuos que comparten ancestría patrilineal ni mucho menos individualizar al responsable, para esto se hace necesario acudir al análisis de STRs autosómicos, igualmente, no se puede establecer la frecuencia de un solo locus debido a la existencia de múltiples loci específicos que se transmiten a posteriores generaciones sin recombinación, dejando a un lado el empleo de la regla del producto para estimar la frecuencia poblacional de un marcador conduciendo al empleo del término frecuencia haplotípica para este tipo de casos.

1.1.2 Marcadores bialélicos

Son los polimorfismos más abundantes en el genoma y para el caso del cromosoma –Y incluyen un gran número de Polimorfismos de Nucleótido Simple SNPs y la inserción ALU YAP-DYS287. Los SNPs se generan porque un nucleótido específico padece un cambio en su par de bases mediante una sustitución, transición o transversión el cual se conserva y se hereda. Normalmente, ocurren como resultado de un daño en el ADN, resultando en una deaminación, depurinización u otra alteración de las bases. Entonces, los mecanismos de reparación fallan en la corrección del daño, generando un SNP ([Koblinsky, Liotti, & Oeser-Sweat, 2005](#)). A la fecha se han caracterizado más de 16.300 marcadores bialélicos dentro de haplogrupos según la Sociedad Internacional de Genealogía Genética <http://www.isogg.org/>.

Los SNPs han demostrado ser potencialmente útiles en la construcción de haplogrupos con aplicación en estudios poblacionales y evolutivos; sin embargo, es imperativo que los SNPs más adecuados para cada estudio, deben ser cuidadosamente seleccionados, principalmente cuando el propósito tiene aplicación en la genética forense, en razón a la característica que presentan diferentes SNPs de definir posiblemente el mismo haplogrupo ([Sanchez et al., 2004](#)); igualmente, se debe tener en cuenta la población objeto de estudio para elegir los SNPs que caractericen eficientemente la ancestría masculina esperada para dicha población. Ejemplo de ello son dos SNPs comúnmente usados por su potencial polimórfico, conocidos como P25 y 92R7 usados en estudios poblacionales europeos y los cuales representan diferentes puntos de ramificación en el árbol de haplogrupos de cromosoma –Y, no obstante, su utilidad en el análisis de poblaciones diferentes a las europeas es ciertamente deficiente.

Este tipo de marcadores poseen mecanismos mutacionales que los hace más conservativos y menos variables, esta importante utilidad sobresale en análisis ancestrales y forenses al coleccionar información durante un mayor número de generaciones y poder asociarse a un origen geográfico específico según el momento de aparición de dicha mutación durante el proceso evolutivo del hombre. Adicionalmente, ofrecen una importante aplicación relacionada con el análisis a partir de muestras con cantidades muy pequeñas de ADN e incluso cuando este ADN se encuentra en un avanzado estado de degradación siendo esta una ventaja en casos de identificación de personas cuando se cuenta con muestras biológicas altamente degradadas ([Budowle, 2004](#)), por tanto, su simplicidad hace que sean analizados simultáneamente a partir de escasas cantidades de ADN, haciéndolos prácticos, dóciles y adecuados para el análisis con tecnologías de alto rendimiento.

La principal desventaja al acudir al análisis de SNPs es el bajo poder de discriminación que ofrecen respecto al ofrecido por otros marcadores como los STRs, en razón a su limitada posibilidad polimórfica, así y todo, se debe acudir al estudio simultáneo de un gran número de marcadores para alcanzar un poder de discriminación suficiente para individualizar a una persona. Los análisis computacionales han demostrado que se necesitan entre 25 a 45 loci de SNPs para alcanzar un poder de discriminación cercano a 13 loci de STRs ([Brookes, 1999](#)), otros estudios sugieren que en 50 SNPs con una estimación de frecuencias del 20 al 50% para el menor alelo puede teóricamente resultar un Likelihood ratio (LR) similar a aproximadamente 12 STRs altamente polimórficos ([Chakraborty, Stivers, Su, Zhong, & Budowle, 1999](#)). Para el caso de los Y-SNPs su aplicación se basa en la discriminación del linaje paterno, mas no tienen la capacidad de discriminar entre individuos, mucho menos en casos donde los implicados están relacionados por vía paterna.

1.2 Historia de los Linajes Masculinos

Cuando se analizan los patrones de herencia de las regiones autosómicas respecto a las uniparentales se identifica principalmente entre sus diferencias la recombinación en cada generación presente en las secuencias autosómicas. Consecuentemente, un individuo podrá trazar sus secuencias autosómicas hacia múltiples ancestros pero su ADN mitocondrial y su cromosoma-Y tendrán un único ancestro.

En cualquier población, todos los individuos que la configuran tienen padres, algunos individuos comparten los mismos padres y otros nunca llegan a convertirse en padres. Este hecho explica la convergencia de linajes cuando se analiza la ancestría de un individuo en su filogenia lo que, en genética de poblaciones, se denomina coalescencia. La conclusión lógica de esto es que todos los linajes mitocondriales existentes podrán ser trazados hacia un único ancestro materno o “Eva mitocondrial” y que todos los linajes masculinos existentes a un único ancestro paterno o “Adán”. Sin embargo, a diferencia de sus compañeros bíblicos, nuestros “Eva y Adán” no vivieron solos, sino formando parte de una población, cuyos individuos quizás no contribuyeron con su ADN mitocondrial ni sus cromosomas -Y pero sí con el resto de su ADN genómico. Por lo tanto, aunque los marcadores uniparentales son útiles para estudiar la historia genética de los linajes masculinos y femeninos, es importante recordar que solo comprenden el 5% del material genético.

Para el caso del cromosoma -Y con frecuencia se utiliza el término haplogrupo para designar el conjunto de cromosomas que comparten un mismo haplotipo. Dada la naturaleza única de la mutación en los polimorfismos que los caracteriza, se espera que los haplogrupos sean grupos de cromosomas relacionados por la descendencia, es decir, grupos de cromosomas con un antepasado común; por el contrario los haplotipos definidos con marcadores con una mayor tasa de mutación, como los microsatélites pueden ser idénticos por descendencia pero no tan conservativos generación tras generación.

1.2.1 Haplogrupos del cromosoma -Y

En genética humana, los haplogrupos más comúnmente estudiados del cromosoma -Y (ADN-Y) pueden ser usados para definir poblaciones genéticas atribuyéndole un origen geográfico específico. Hay dos tipos de variaciones dentro de la composición del cromosoma -Y. La primera es una variación que cambia de manera rápida y parcialmente permite distinguir a hombres diferentes dentro de una población funcionando como una especie de marca personal.

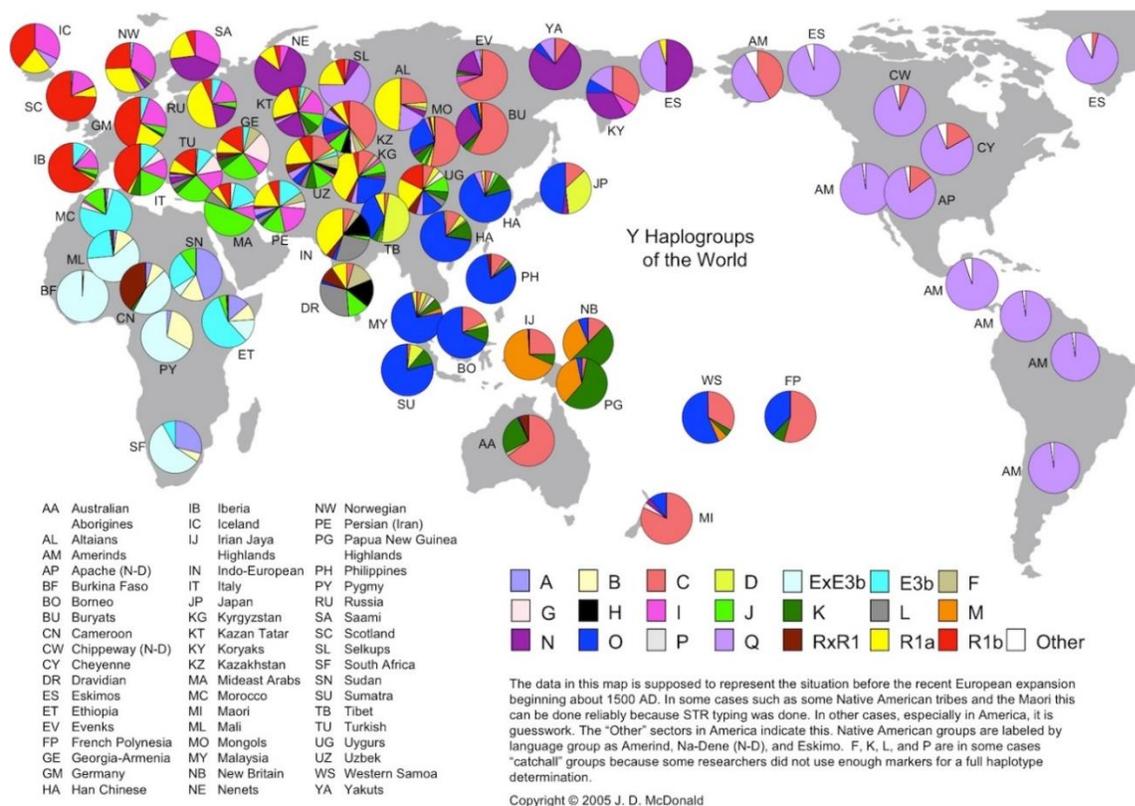
El otro tipo de variación cambia lentamente así que podrían haber grandes números de hombres con un tipo similar de cromosoma -Y. Estas secuencias permiten agrupar a los tipos de cromosomas -Y en diferentes familias denominadas Haplogrupos (HG). El

análisis del tipo de variación dependerá de la elección de los marcadores y el grado de variación que expresan para cada población.

Para designarlos, se ha utilizado una nomenclatura que hasta antes del año 2002 había sido variada. Por lo menos existían siete diferentes sistemas para nombrar los diferentes haplogrupos lo cual dificultaba la comparación e interpretación de resultados, ya que en algunas ocasiones se encontraban diferentes nomenclaturas para un mismo haplogrupo ([M. A. Jobling & Tyler-Smith, 2000](#)), ([P. A. Underhill et al., 2000](#)), ([Hammer et al., 2001](#)), ([T. Karafet et al., 2001](#)), ([Semino et al., 2000](#)), ([Su et al., 1999](#)), ([Capelli et al., 2001](#)).

La nomenclatura de los haplogrupos del cromosoma -Y está representada en un árbol filogenético binario el cual se identifica como YCC2003 y se encuentra disponible como referencia en la página web del Consorcio Cromosoma -Y (<http://ycc.biosci.arizona.edu>). En este árbol filogenético, el clado de -Y contiene los haplogrupos principales que se identifican con letras mayúsculas de la A hasta la R. Dentro de estos se encuentran los paragrupos que se definen como un grupo de haplotipos que contiene alguno pero no a todos los descendientes de un linaje ancestral y se indican con un asterisco. El haplogrupo al cual pertenece un tipo de cromosoma -Y, proporciona en muchos casos una evidencia clara sobre su origen lo que hace que sean útiles para ser empleados en una reconstrucción genealógica. En la figura 1-1 se encuentra el mapa de la distribución mundial de los haplogrupos, identificados para cada uno de los continentes con la nomenclatura actual para designar cada uno de los haplogrupos.

Una comparación entre polimorfismos autosómicos, ADNmt y de cromosoma -Y, demuestra que este último tiene mayor diferenciación proporcionalmente a la distancia geográfica relacionando un fenómeno de patrilocalidad ([Seielstad, Minch, & Cavalli-Sforza, 1998](#)). Dicha diferenciación geográfica de haplotipos -Y confirma que estos marcadores pueden ser útiles en la predicción del origen de una población.

Figura 1-1: Distribución de haplogrupos de cromosoma –Y en el mundo.

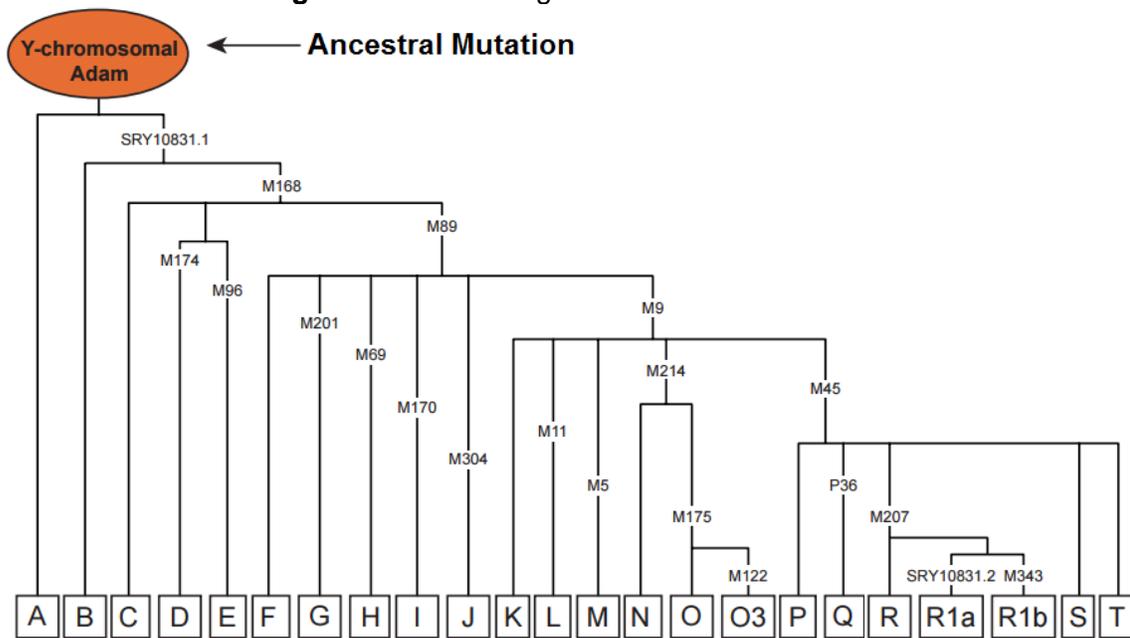
Tomado de: <http://www.transpacificproject.com/index.php/transpacific-migrations/>

Los perfiles de haplogrupos para diferentes regiones del mundo son considerablemente diferentes, otorgando a esta especificidad regional un beneficio potencialmente útil en el campo de la identificación humana. Ejemplo de ello es el haplogrupo Q definido por el polimorfismo en el locus M242 y descrito inicialmente por Seielstad y colaboradores en el 2003, el cual se cree que apareció en el momento de la primera migración hacia el continente americano, argumentado por qué tiene presencia en grupos étnicos del este asiático, de igual manera, es un polimorfismo predominante en grupos étnicos nativo-americanos, lo que permite asociarlo a un origen geográfico del continente americano. Para el caso de haplogrupos que no muestran mucha especificidad, se asume que su frecuencia es de similar magnitud en diversas poblaciones, lo que puede reducir su utilidad en el análisis de una muestra de origen desconocido y de la cual se pretenda buscar información orientadora.

1.2.2 Distribución jerárquica de las mutaciones

El estudio de marcadores uniparentales como los del cromosoma –Y, que no se someten a recombinación, permite comprender la aparición de cada polimorfismo en un momento evolutivo dado, para esto, se puede establecer un polimorfismo ancestral que evidentemente todos los cromosomas –Y actuales poseen. A su vez han surgido nuevos polimorfismos asociados siempre al proceso migratorio del hombre a lo largo del planeta, generando una gran diversidad de linajes masculinos que testifican momentos claves de la historia y de igual manera, se pueden atribuir a la colonización geográfica de un sitio específico. La interpretación de una serie de marcadores asociados a la definición de un haplogrupo debe hacerse teniendo en cuenta que su distribución obedece a un proceso jerárquico, consecutivo y relacionado con el momento de aparición de cada polimorfismo, partiendo del más ancestral y el cual es frecuente en toda la población, hasta el polimorfismo más reciente y por tanto, menos frecuente en la población, lo que permite discriminar muy bien el linaje que define, formando el conocido árbol genealógico del cromosoma –Y.

Figura 1-2: Árbol filogenético del cromosoma –Y.



Adaptado de: Which markers are tested in the Y-DNA SNP Haplogroup Backbone Panel? Genebase Tutorials. May 2015, from <http://www.genebase.com/learning/article/55>

En la figura 1-2 se observa la expresión gráfica de la distribución jerárquica de haplogrupos reconociendo los haplogrupos más jóvenes o recientes ubicados hacia el lado derecho de la imagen, por tanto, si un individuo pertenece a un haplogrupo como el

Q o el R, obligatoriamente debe presentar las mutaciones ancestrales que dieron paso a la aparición de dicho polimorfismo, sirviendo esto además, como un control a la hora de asignar un haplogrupo para una muestra, descartando así una posible mutación *de novo* que dé lugar a una falsa asignación pasándose por alto las mutaciones ancestrales que debe poseer dicha muestra.

1.3 Asociación entre haplotipos de Y-SNPs y Y-STRs

Inicialmente aquel individuo que expresaba un Y-SNP ancestral que definía un haplogrupo específico, obligatoriamente portaba un haplotipo en loci multialélicos o Y-STRs característico y asociado a dicho haplogrupo, una vez los descendientes de aquel individuo experimentaban mutaciones en su haplotipo, se generaba una diversidad haplotípica de Y-STRs generación tras generación, sin embargo, la definición del haplogrupo por el Y-SNP ancestral se conservaba sin mayor variación llegando a acumular un número considerable de diversos haplotipos de Y-STRs congregados en un haplogrupo específico. De esta manera, se puede asociar que un SNP que defina un haplogrupo principal es considerablemente atribuido a un origen poblacional particular, y así mismo los haplotipos asignados a dicho haplogrupo ([Forster et al., 2000](#)).

Dado el origen y la naturaleza de un haplotipo de Y-STRs, este puede eventualmente usarse para la predicción de un haplogrupo y por consiguiente proveer información sobre el origen poblacional de la muestra aportante del haplotipo. El proceso de generación de nuevos haplogrupos asociados a un origen poblacional específico, paulatinamente permite hacer un seguimiento de linajes ancestrales según la magnitud de la diversidad de los alelos y el grado de polimorfismo de cada marcador, para evaluar el nivel de divergencia respecto al haplotipo ancestral o modal y así poder estimar un aproximado del tiempo de aparición y la capacidad poblacional de distribución geográfica.

Resulta interesante también asociar fenómenos sociales a la dispersión poblacional de los haplogrupos a lo largo de las diferentes regiones geográficas propias de cada uno, se observan entonces conceptos como la patrilocalidad característica de la mayoría de las poblaciones humanas donde las mujeres tienden a radicarse en los lugares de nacimiento de sus esposos luego de casarse, promoviendo una diferenciación geográfica de los linajes masculinos. Si se acude además al análisis de marcadores uniparentales y autosómicos que evidentemente se comportan de manera diferente, se puede apreciar

una distribución genético-poblacional diferente para cada marcador, e incluso entre los marcadores uniparentales, el cromosoma –Y suele expresar diferenciación genética proporcional a la distancia geográfica ([Seielstad et al., 1998](#)), infiriendo que la distribución de haplogrupos de cromosoma –Y no se realiza de manera aleatoria dentro de las poblaciones.

1.4 Ventajas del empleo de marcadores de cromosoma – Y en el campo forense

Si se parte de la premisa que el genoma de cada individuo es único (con la excepción de los gemelos idénticos) y a su vez, es heredado proporcionalmente de padres a hijos, cuando se acude a un análisis de tipo forense bien sea para establecer identificación, parentesco o uniprocedencia, se debe planear estratégicamente el empleo de marcadores según su capacidad informativa, las necesidades del caso y la disponibilidad de muestras objeto de estudio; no obstante, si se tiene en cuenta que en más del 90% de los delitos al menos uno de los implicados corresponde a una persona de sexo masculino, resulta de gran importancia estudiar el cromosoma –Y porque permite obtener una amplificación diferencial en casos de asaltos sexuales separando el perfil de la víctima. Así mismo, puede establecer un linaje masculino en casos de paternidad cuando los marcadores autosómicos no llegan a ser lo suficientemente informativos o cuando se establezca un parentesco en casos de identificación de desaparecidos respecto a muestras de referencia con familiares en vía patrilineal.

La pérdida de recombinación hace que al analizar un grupo de marcadores polimórficos dentro del cromosoma –Y, permita hacer el estudio comparativo de individuos masculinos separados por largos períodos de tiempo, e incluso después de varias generaciones familiares y de igual manera, se pueden plantear estudios de ancestría masculina y la estimación del tiempo de aparición de ancestros comunes para una población determinada. A pesar de ello, dicha ventaja puede llegar a convertirse en un factor adverso debido a que no se puede distinguir o individualizar familiares que lleven el mismo linaje paterno, lo que reduce considerablemente su poder de discriminación.

Aunque se analicen muchos marcadores los loci no son independientes, por tanto, no se puede generar una probabilidad de coincidencia con la regla del producto (multiplicación de probabilidades), en lugar de ello se analiza el haplotipo o la combinación de alelos observados. El método de análisis haplotípico ha evidenciado utilidad en la

reconstrucción de relación paterna, por ejemplo, investigaciones de desastres masivos y la atribución de un origen étnico a una muestra de ADN de la cual se desconoce su procedencia.

De igual manera, la ocurrencia de cada SNP representa un evento único en la evolución de la especie humana resultando estratégicos en la búsqueda de marcadores genéticos informativos con amplio espectro de aplicación. Dichas propiedades combinadas con la ausencia de recombinación en gran parte del cromosoma, y la consecuente acumulación de la diversidad genética entre linajes y poblaciones relacionadas, proyecta una importante aplicación de los Y-SNPs en el campo forense.

([Roewer, 2009](#)) describió las aplicaciones del análisis de marcadores de cromosoma –Y en casos criminales, de manera acertada:

- En casos de asalto sexual donde la evidencia en cuestión es amilasa-positiva y por tanto, se espera una mezcla de perfiles masculino y femenino (trazas de besos o mordiscos).
- En casos en donde se encuentren manchas antiguas de semen, donde se sospecha que la mayoría de células espermáticas están degradadas y la lisis diferencial no tiene éxito o es riesgosa.
- En casos de asalto sexual donde se encuentra un gran número de manchas de semen.
- En casos donde se necesita determinar el número de donantes masculinos presentes en una mancha de semen.
- En casos donde la evidencia en cuestión se espera incluya células del perpetrador (uñas de una víctima que demostró defenderse durante el ataque).
- En casos donde se necesita determinar la relación patrilínea del donante de una mancha de semen.
- En casos donde se necesite inferir el posible origen poblacional del donante.

En la mayoría de los casos donde se analizan marcadores de cromosoma –Y relacionados en procesos judiciales, se deben plantear inferencias del posible origen del haplotipo encontrado respecto a la información poblacional disponible, para ello se acude a bases de datos de referencia como lo es la de haplotipos de cromosoma –Y disponible en línea URL: www.yhrd.org, la cual representa el depósito más grande de haplotipos a la

fecha. La versión actual de la base de datos contiene 365.578 haplotipos de 2.415 poblaciones muestreadas a escala mundial.

2.Desarrollo y adaptación de un método técnico-científico para el estudio de linajes masculinos en la población

Los marcadores SNPs aparecen en el genoma en una alta densidad de 1 a 10 por cada 1000 nucleótidos, convirtiéndolos en la fuente de variación más abundante en el genoma humano ([Sobrino, Brion, & Carracedo, 2005](#)). Consecuentemente, a lo largo de los últimos años se han desarrollado diversas tecnologías para el análisis de marcadores moleculares como los SNPs. Usualmente la elección del método más apropiado depende de variables como el número de marcadores y de individuos a analizar; sin embargo, para los propósitos forenses no todas las técnicas resultan convenientes porque en su mayoría establecen requerimientos como el análisis a gran escala o el acervo de grandes cantidades de ADN en óptimo estado de integridad.

Para el caso de estudios forenses, por lo general se analizan muestras individuales o uniprocedentes con disponibilidad de cantidades mínimas de ADN, razón por la cual, el método de minisequenciación por SNaPshot® es una de las técnicas más aceptadas en el campo forense para tipificar varios SNPs de interés en una reacción multiplex, muy probablemente debido a su universalidad porque los equipos necesarios para el método, son de rutina y empleados en cualquier laboratorio de genética forense.

El método describe un procedimiento para realizar un análisis filogenético de varios Y-SNPs en una reacción multiplex jerárquica, es decir, con un orden específico que va del más ancestral al polimorfismo más reciente. Dicho método exige la apropiada y estratégica selección de marcadores SNPs acorde a la población objeto de estudio; así mismo, se requiere de un eficiente diseño de primers, optimización de las reacciones de PCR y la correspondiente extensión de base simple, así como el respectivo análisis de los resultados. El principio del método SNaPshot consiste en una minisequenciación por medio de una extensión alelo-específica de un primer que se alinea en posición

adyacente al SNP de interés. Dicho proceso se adelanta con el uso de didesoxinucleótidos (ddNTPs) fluoromarcados como los de secuenciación por el método Sanger, los cuales inhiben la continuidad de la elongación por la pérdida del grupo 3'-hidroxilo en el azúcar ([Carracedo, 2005](#)). Los primers con la base simple extendida en su extremo 3' se diseñan con diferente longitud adicionando una cola de nucleótidos inespecíficos en su extremo 5' para poder analizarlos individualmente durante la migración en la electroforesis capilar.

Entonces, resulta estratégico poder implementar un método sensible de minisequenciación para el estudio de SNPs que permita analizar tanto muestras con material biológico en buen estado de integridad, como aquellas que ostentan potencial degradación de ADN, apoyándose además en bases de datos poblacionales que puedan mapear la diferenciación genética entre poblaciones a escala local o incluso mundial, tal y como lo permite la Base de Datos de Referencia de Haplotipos de cromosoma -Y (YHRD), la cual permite obtener información orientadora asignando la ancestría geográfica más probable para casos de personas desaparecidas o sospechosos desconocidos.

Se pretende por tanto, plantear el diseño y validación de reacciones PCR multiplex de marcadores Y-SNPs, que permitan genotipificar jerárquicamente el árbol filogenético del cromosoma -Y para una población masculina colombiana. Las reacciones PCR multiplex propuestas en este estudio, se han desarrollado de tal manera, que permitan explorar los marcadores que probablemente sean capaces de determinar el genotipo de haplogrupos -Y colombianos más frecuentes. Para la selección de marcadores se tuvo en cuenta entre otras cosas la posición en el árbol filogenético reportado en la actualización más reciente del Consorcio del Cromosoma -Y ([Genealogy, 2014](#)) y el reporte en estudios previos con poblaciones latinoamericanas o de mestizaje similar ([Acosta, Blanco-Verea, Lareu, Brion, & Carracedo, 2009](#); [L A Alonso, 2013](#); [Bisso-Machado et al., 2011](#); [A. Blanco-Verea, Jaime, Brion, & Carracedo, 2010](#); [Caratti, Gino, Torre, & Robino, 2009](#); [Kidd et al., 2011](#); [Lovo-Gomez, Blanco-Verea, Lareu, Brion, & Carracedo, 2007](#); [Nunez, Geppert, Baeta, Roewer, & Martinez-Jarreta, 2012](#); [Palha et al., 2012](#); [Sanchez-Diz et al., 2009](#)). El principal aporte forense que ofrece el método aquí propuesto consiste en asignar la ancestría más probable a una muestra de ADN masculino con origen desconocido, mediante el análisis de la distribución geográfica del haplogrupo de Y-SNPs establecido.

2.1 Metodología de trabajo

La metodología aplicada consistió de un procedimiento sistemático que involucró el diseño de un componente teórico que sirviera como soporte para el componente experimental.

2.1.1 Voluntarios y Extracción de ADN

Se analizaron 144 muestras de individuos no emparentados, obtenidas del banco de muestras de referencia del Grupo Genética del Cuerpo Técnico de Investigación de la Fiscalía General de la Nación, de las cuales se disponían muestras de sangre líquida, y algunas en tarjetas FTA® Whatman. Como criterio de inclusión se tenía en cuenta que fuesen mayores de edad y residentes en la ciudad de Bogotá D.C. y municipios circundantes, con un tiempo mínimo de residencia de 5 años.

Quienes cumplían con los criterios, además de revisar el proceso documental estándar para todo muestradante del banco de muestras, se verificó la aprobación del consentimiento informado para estudios poblacionales y de igual manera, se contactaron telefónicamente para diligenciar una breve encuesta con información básica para realizar el análisis y las inferencias poblacionales correspondientes, garantizando siempre la confidencialidad de los datos suministrados (Ver anexo 1).

Para la extracción de ADN se empleó el DNA IQ™ Casework Pro Kit for Maxwell® (Promega), que utiliza la resina DNA IQ™ para purificar el ADN. El procedimiento se realizó en el instrumento Maxwell® 16 Instrument (Nº de cat. AS3060), que es un instrumento de manipulación mediante partículas paramagnéticas que transporta eficazmente la resina DNA IQ™ a través de reactivos de purificación en cartuchos prellenados, y así combina la resina con los reactivos de lisis y lavado durante todo el procedimiento ([Bierke et al., 2006](#)).

2.1.2 Cuantificación de ADN por RealTime-PCR

La PCR en tiempo real o PCR cuantitativa es una variación de la PCR estándar y es utilizada para la cuantificación de ADN de una muestra. Utilizando primers específicos de secuencia y sondas TaqMan, es posible determinar el número de copias o la cantidad relativa de una determinada secuencia de ADN, monitoreando el progreso de la

amplificación mientras ocurre ([Swango, Timken, Chong, & Buoncristiani, 2006](#)). Para ello se empleó el Kit de cuantificación Quantifiler® Human DNA (Life Technologies), siguiendo las instrucciones del fabricante y usando el equipo Applied Biosystems 7500 Real-Time PCR System, con el fin de determinar la cantidad total de ADN humano amplificable en unidades (ng/μL) y estimar la presencia potencial de inhibidores de PCR.

2.1.3 Tipificación y Análisis de Y-STRs

El ADN extraído se sometió a amplificación por medio de la Reacción en Cadena de la Polimerasa (PCR), en este caso mediante una PCR multiplex para amplificar los diecisiete (17) loci microsatélites nucleares de cromosoma -Y (DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385 a/b, DYS393, DYS391, DYS439, DYS635, DYS392, Y GATA H4, DYS437, DYS438, DYS448) incluidos en el kit Yfiler® (Applied Biosystems), siguiendo las directrices del fabricante. Posteriormente, las muestras se tipificaron empleando un analizador genético de electroforesis capilar ABI 3130XL y con ayuda del software *GeneMapper® ID-X* se realizó el análisis y la asignación alélica correspondiente. Con los haplotipos obtenidos se utilizó la aplicación Haplogroup Predictor (www.hprg.com/hapest5/) con el fin de hacer una clasificación preliminar de los haplogrupos de cromosoma -Y en cada muestra. Este programa se basa en una aproximación bayesiana que utiliza frecuencias alélicas que fueron calculadas a partir de colecciones de haplotipos tomados de artículos publicados y bases de datos ([Athey, 2005](#)). Este enfoque determina la probabilidad de que un haplotipo Y-STR se encuentre dentro de un determinado haplogrupo. El haplogrupo que se asigna es aquel que tiene la mayor probabilidad de pertenecer dicho haplotipo.

2.1.4 Selección, Diseño y Validación de Primers para Amplificación

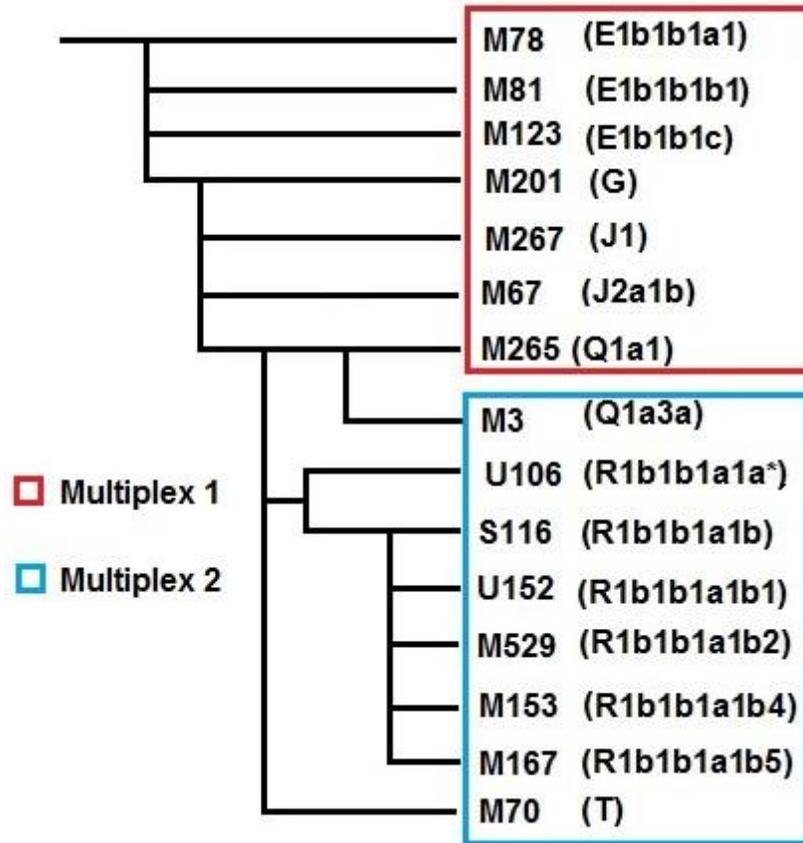
El proceso inició con una exhaustiva revisión bibliográfica en búsqueda de aquellas frecuencias alélicas de SNPs que hubieran sido reportados previamente, específicamente para poblaciones latinoamericanas caracterizadas por su alto grado de mestizaje, así entonces, se seleccionaron 15 Y-SNPs que potencialmente pudieran caracterizar y determinar los haplogrupos más frecuentes presentes en la población colombiana.

De igual manera, se tuvo en cuenta la estructura jerárquica de los Y-SNPs, dado que su análisis debe realizarse en dos pasos, iniciando con el marcador más basal en el árbol filogenético para sentar bases en la asignación de marcadores posteriores y que permitieran la asignación de nuevos clados ([Brion, Sobrino, Blanco-Verea, Lareu, & Carracedo, 2005](#)). Esta es la principal estrategia que permite formular el análisis de varios marcadores simultáneamente, maximizando la eficiencia en tiempo y en costos en el análisis del linaje masculino para una muestra. De manera que se pueden estudiar varios loci paralelamente en una reacción multiplex, empleando una mínima cantidad de ADN lo cual es común en estudios de naturaleza forense.

Los SNPs y el árbol de haplogrupos definidos en el presente estudio se expresan en la figura 2-1. Por consiguiente, los marcadores seleccionados corresponden al M78, M81 y M123 que discriminan el haplogrupo E1b1b1, el cual expresa predominancia en África y el este de Europa ([Adams et al., 2008](#)). Se estudiaron también los marcadores M201, M267 y M67 característicos de poblaciones semitas en el medio oriente ([Nunez et al., 2012](#)). Análogamente, era importante incluir marcadores que definieran el origen Nativo-Americano en dos de sus polimorfismos más representativos para el haplogrupo Q ([Bortolini et al., 2003](#)). Para el caso de los marcadores U106, S116, U152, M529, M153 y M167 son los que diversifican el haplogrupo R1b, el cual es el más frecuente en estudios de población mestiza latinoamericana ([Acosta et al., 2009](#); [L A Alonso, 2013](#); [Bisso-Machado et al., 2011](#); [Kidd et al., 2011](#); [Lovo-Gomez et al., 2007](#); [Nunez et al., 2012](#); [Palha et al., 2012](#)), por ende se hizo necesario mejorar la resolución en el análisis discriminatorio de este haplogrupo. El marcador M70 se implementó para clasificar de manera general haplogrupos poco frecuentes.

Inicialmente, las secuencias de algunos primers se obtuvieron de la Dra. Leonor Gusmao (Institute of Molecular Pathology and Immunology of the University of Porto IPATIMUP), sin embargo, se prescindió de su uso en razón a la longitud de los amplicones que oscilaba entre 200 y 400pb, lo cual resultaba contraproducente para los intereses forenses de este estudio. De ahí que se acudió a obtener la información de la secuencia de cada par de primers usados para amplificar los 15 segmentos de ADN de cromosoma –Y que incluían los correspondientes 15 loci Y-SNPs.

Figura 2-1: Árbol filogenético con parsimonia máxima para los Y-SNPs seleccionados en el estudio.



Nota: Se describen los SNPs incluidos en cada multiplex y el respectivo haplogrupo. La longitud de los brazos no tiene significancia.

Para ello, se accedió a la información 200 pb corriente arriba y abajo del sitio polimórfico, disponible en el GenBank (www.ncbi.nlm.nih.gov). Dichas secuencias se llevaron a la interfase Primer3Plus del software Primer3 <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plusAbout.cgi> para elegir el par de primers más adecuados. En el proceso de elección de primers se tuvieron en cuenta aspectos como una longitud del amplicón no superior a 230pb, valores de Tm para todos los primers entre 58 y 62 °C, con diferencia máxima de ± 1 °C entre cada par de primers y ± 3 °C entre el set completo de primers como se observa en la tabla 2-1. Para proveer una buena especificidad, la longitud de los primers oscila entre 20 y 28 nucleótidos. El porcentaje de GC se mantuvo entre 20 y 40%. También se revisó la autocomplementaridad para evitar la formación de dímeros de primer o estructuras secundarias entre los mismos.

Tabla 2-1: SNPs de cromosoma –Y y la secuencia de los primers para amplificación.

Locus	ID en GeneBank o dbSNPs	Polimorfismo	Primer Forward (5'→3')	Primer Reverse (5'→3')	μM*	Tamaño del producto (pb)
MULTIPLEX 1						
M78	AC010889	C>T	TGCATTACTCCGTATGTTGAC	TGGAAGCTTACCATCTTTTATGA	0,25	132
M81	Rs2032640	C>T	CATCTCTTAACAAAAGAGGTAAATTTG	CATTGTGTTACATGGCCTATAATATTC	0,57	179
M123	AC010889	C>T	GAATTTGCATTCTAGGTATTCAGG	GTTGCCTTTTGAAATGAATAAATC	0,65	175
M201	Rs2032636	G>T	TCAGATCTAATAATCCAGTATCAACTG	GTTCAAATCCCATATCCAGCAT	0,4	93
M267	Rs9341313	T>G	CCAAGTCTGGATAGCGGATT	CCAGGCTAGTGTCAAACTCGTA	0,32	156
M67	Rs2032628	A>T	TCCTAAAAGAAAAATGAGATTGTGAAT	AGTCTTTTCACTTGTTCTGTTGGA	0,55	114
M265	Rs3212294	C>A	GGGGTTGGGAGAACATCTTAAA	TACTGATACATTCCACCTCATCAT	0,32	157
MULTIPLEX 2						
M3	Rs3894	C>T	CTGCCAGGGCTTTCAAATAG	TTAAGGGCATCTTTTCATTTAGGTA	0,6	95
U106	Rs16981293	C>T	TCCTGAATAGCAAATCCCAAAG	AATGGCAGAGGTAGGAGGAAAT	0,57	192
S116	Rs34276300	C>A	GGTCTGCTAATGTATCTGCTG	ATGAGCAGAAGACACAAGATGC	0,11	84
M153	21706360	A>T	TTCTCAGACACCAATGGTCCTA	TCTGACTTGAAAGGGGAAA	0,37	76
M167	Rs1800865	T>C	GAGGCTGGCCAAGTTAAG	CTTCCTCGGAACCACTACCA	0,53	131
U152	Rs1236440	C>T	GAAACATTCCACGCTTGAGG	AGCCTCTTTTGGCTTCCAT	0,02	163
M529	Rs11799226	C>G	GCCCCCAAACAACAGAATA	GGAAGCATTGAGAAGCAGGT	0,59	228
M70	Rs2032672	A>C	TCATAGCCCACTATACTTTGGAC	GTGATCTGAGGGCTGGACTAT	0,18	86

*Concentración final en la reacción.

Cada primer individualmente se chequeó mediante la herramienta de alineamiento BLAT <https://genome.ucsc.edu/cgi-bin/hgBlat> con el fin de evaluar su especificidad en el genoma completo; posteriormente, con el fin de confirmar la especificidad conjunta de cada par de primers, estos se evaluaron en la aplicación Primer-BLAST http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome, teniendo en cuenta los parámetros ya mencionados. Finalmente, se realizó una valoración del desempeño de los primers mediante una PCR-*in silico* mediante la herramienta UCSC In-Silico PCR <https://genome.ucsc.edu/cgi-bin/hgPcr?command=start> y así determinar la eficiencia de los primers en el proceso de amplificación.

Los primers fueron sintetizados por Eurofins MWG/Operon y se obtuvieron purificados por desalinización y liofilizados para ser resuspendidos en agua miliQ ultrapura asegurando siempre un pH de 7.0 y llevados a una concentración final de 100 μM, conservados a -20°C. Así mismo, para el caso de los primers de minisequenciación se purificaron por cromatografía líquida de alta eficacia HPLC para optimizar su desempeño.

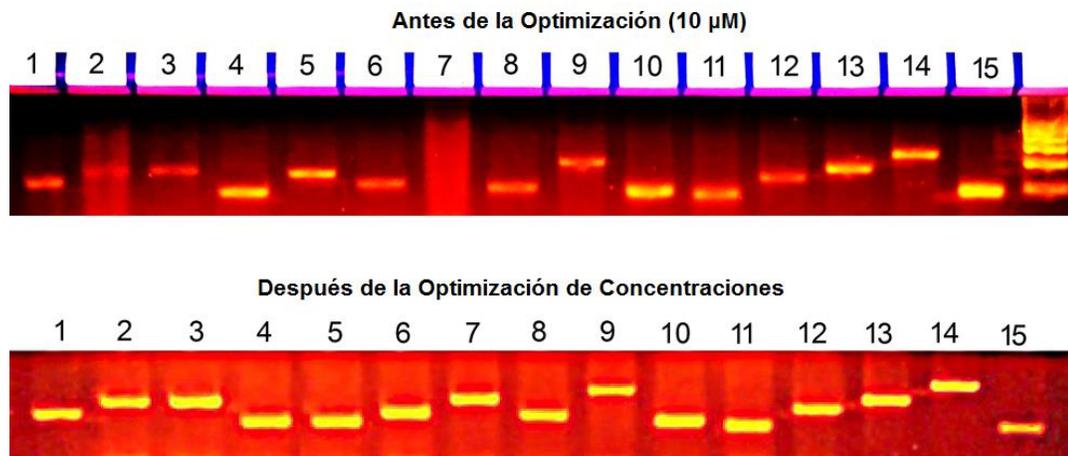
2.1.5 Evaluación de los Primers con Gel de Electroforesis

Con el fin de verificar la eficiencia en el diseño de cada par de primers, se evaluaron por amplificación singleplex con un control de ADN masculino de 0,25 ng/μL, en un volumen final de 12,5 μL como se muestra en la tabla 2-2:

Tabla 2-2: Condiciones de la reacción de PCR Singleplex.

Reactivos			Condiciones de PCR
Volumen	Descripción	Origen	
1,5 μL	Buffer 10X PCR	Roche – Applied Biosystems	95°C → 11 Min
0,75 μL	MgCl ₂ Solution 25 mM	Roche – Applied Biosystems	
0,25 μL	Bovine Serum Albumin	Promega	95°C → 10 Seg
	BSA Acetylated 10 ng/μL		
0,25 μL	dNTP Mix 10mM	Applied Biosystems	36 Ciclos
0,25 μL	AmpliTaq Gold [®] 5U/μL	Applied Biosystems	
5 μL	H ₂ O dd	----	
4 μL	ADN 0,25 ng/ μL	----	72°C → 7 Min
0,25 μL	Primer Forward 10 μM	Eurofins MWG/Operon	4°C → ∞
0,25 μL	Primer Reverse 10 μM	Eurofins MWG/Operon	
12,5 μL	VOLUMEN TOTAL		

Figura 2-2: Geles de Agarosa al 2,2% con los productos de PCR de los 15 amplicones.

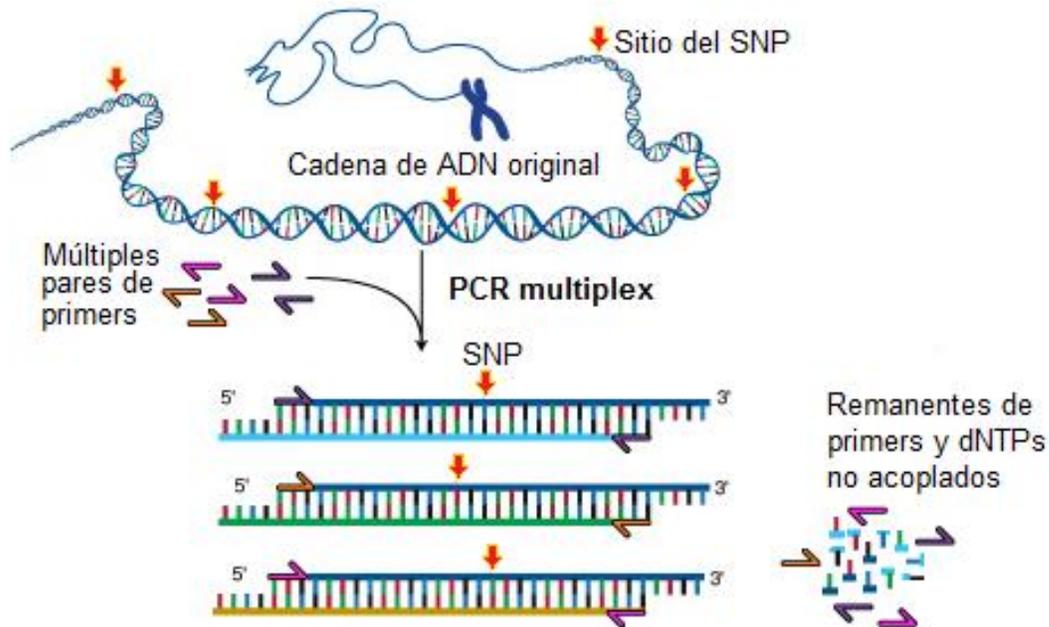


El producto de amplificación de cada pareja de primers se verificó en un gel de agarosa al 2,2% mediante un Cassette FlashGel[®] System (Lonza), y se usó un marcador de peso molecular 100 pb DNA Step Ladder (Promega) para verificar el tamaño esperado de los amplicones. Mediante un proceso de optimización de concentraciones con sucesivas amplificaciones, se buscó encontrar una homogeneidad en la intensidad de la banda y así conservar la proporción de las concentraciones en la PCR multiplex, tal y como se observa en la figura 2-2.

2.1.6 Condiciones de la PCR Multiplex

Básicamente la reacción de PCR multiplex busca amplificar simultáneamente más de una secuencia. Para ello, se combinan varios pares de primers y se suministran los reactivos en cantidades suficientes para abastecer el proceso de amplificación simultánea de varios segmentos de ADN a partir de un único molde, como se observa en la figura 2-3.

Figura 2-3: Reacción de PCR Multiplex.



Para el caso de las dos reacciones multiplex las condiciones fueron similares diferenciando su contenido de primers (concentraciones se detallan en la tabla 2-1), para ello se usaron 5 μL de AmpliTaq Gold® 360 Master Mix (Applied Biosystems), 1 μL del primer mix, ddH₂O y 0,5 ng de ADN, resultando un volumen final de 10 μL , los parámetros de la amplificación se describen en la tabla 2-3.

Tabla 2-3: Condiciones de amplificación para la PCR multiplex.

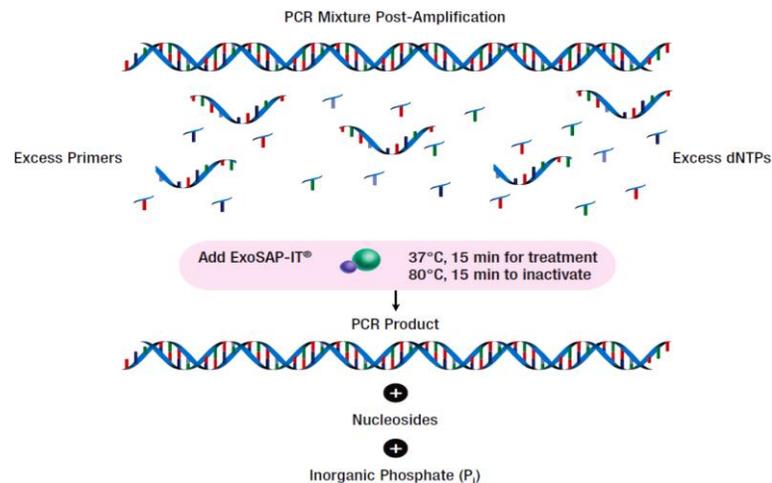
Estado	Paso	Temperatura °C	Tiempo
Hold	Desnaturalización inicial	95	10 Min
28 Ciclos	Desnaturalización	95	30 Seg
	Alineamiento	60	30 Seg
	Extensión	72	60 Seg
Hold	Extensión final	72	7 Min
Hold	Hold final	4	∞

En el proceso de validación del método se tuvieron en cuenta diversas variables que potencialmente influían en la obtención de resultados exitosos, entre ellas se estimó el número de ciclos para la amplificación de la multiplex y la cantidad de ADN por muestra, pensando siempre en la aplicación forense de la metodología. Evidentemente, se ajustaron nuevamente las concentraciones de los primers en cada reacción multiplex conservando las proporciones de las concentraciones evaluadas en los geles, esta vez asegurándose de obtener picos homogéneos y de similar cantidad de Unidades Relativas de Fluorescencia RFU en todos los electroferogramas analizados.

2.1.7 Purificación de los productos de PCR

Los productos de la amplificación por PCR multiplex se sometieron a purificación con el compuesto ExoSAP-IT (Affymetrix), que está constituido básicamente por dos enzimas hidrolíticas, una exonucleasa y una fosfatasa alcalina de camarón, la primera rompe los enlaces fosfodiéster del ADN de una sola hebra y la segunda elimina los grupos fosfato de los dNTPs (Dugan, Lawrence, Hares, Fisher, & Budowle, 2002). Como se puede observar en la figura 2-4, para llevar a cabo el proceso de purificación se mezcla 1 μ L del producto de PCR con 0,5 μ L del reactivo ExoSAP-IT, conservando siempre la baja temperatura con la ayuda del hielo en todo momento. Posteriormente, se incuba la mezcla a 37°C durante 15 minutos para que las enzimas actúen degradando los remanentes de primers y de nucleótidos y finalmente, se incuba nuevamente la mezcla a 80°C durante 15 minutos para inactivar la enzima. El producto queda ya disponible para el proceso de minisequenciación.

Figura 2-4: Representación del tratamiento del producto de PCR con ExoSAP-IT.



(tomado de http://www.affymetrix.com/catalog/131310/USB/ExoSAP-IT+For+PCR+Product+Cleanup#1_2)

2.1.8 Selección de Primers para Minisequenciación

Los primers para la minisequenciación o extensión de base simple SBE fueron diseñados manualmente, de tal manera que su terminal 3' quedaba inmediatamente adyacente al SNP, y el extremo 5' contenía una secuencia inespecífica a manera de cola, que difiere en número de nucleótidos entre cada primer con el fin de ajustar la diferenciación durante la movilidad electroforética. Del mismo modo, cada primer se estudió en la interfase Primer3Plus del software Primer3 <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plusAbout.cgi> para valorar la proporción de Guaninas, Citocinas y la Tm. En todos los casos se evaluó la viabilidad de cada primer, tanto en orientación forward o reverse, para elegir aquella que satisficiera eficazmente los requisitos y así considerar si el primer se hibrida con la cadena principal o lo hace con la complementaria, aspecto que se debe tener en cuenta a la hora del análisis de los resultados.

La migración electroforética según el tamaño de cada primer fue considerada usando el SNaPshot® Primer Focus Kit (Applied Biosystems), y los datos se analizaron con el software GeneMapper ID-X versión 1.1.1 (Applied Biosystems) y de esta manera establecer los sitios donde se espera encontrar cada SNP.

Tabla 2-4: Secuencias de primers para minisequenciación. Secuencia roja en negrilla corresponde a las colas inespecíficas.

Mutación	Secuencia (5'→3')	Orientación	μM*
MULTIPLEX 1			
M78	TTATTTTGAATATTTGGAAGGGC	Reverse	0,2
M81	AACA GTTTGTGTGAGTATACTCTATGAC	Reverse	0,2
M123	TCTGTCCC GCATTCTAGGTATTCAGGCGATG	Forward	0,2
M201	AAAGTCCACCCC CTAATAATCCAGTATCAACTGAGG	Forward	0,2
M267	TACACTTCCCAAATCC TGCATGGAAGCATTTTTGTAAATA	Forward	0,2
M67	AACCCACCGTCCCCTCCTGT AAAATATGTGTAATTCAAAAAACA	Forward	0,2
M265	CACATACGGCCCCCTTACCTAC TACTAAATTAAGTTCTGACTTGCT	Forward	0,2
MULTIPLEX 2			
M3	AATGGGTCACCTCTGGGACTGA	Forward	0,425
U106	TGAC AATAGCAAATCCCAAAGCTCCA	Forward	0,5
S116	AAATTAAG GGAGTTGGGGCTAAAGTGAAAAG	Reverse	0,425
M153	CCCCCCCCC AAAGCTCAAAGGGTATGTGAACA	Reverse	0,325
M167	ACCCCTTTCCAAATT CCAAGGAAGCCCCACAGGGTGC	Forward	0,3
U152	CCCTACTTTCAGTCCCA AAAAATGAGTATTGTGAAAATA	Forward	0,025
M529	CCCCCTTACCCACAGTCCGCC ATAACAACCGCTCTCTCAGACA	Forward	0,475
M70	ACAGCTTCGCGAAATCCTTACCCCA GATTCTGTTGTGGTAGTCTTAG	Reverse	0,25

*Concentración final en la reacción.

2.1.9 Reacción de Minisequenciación por SNaPShot

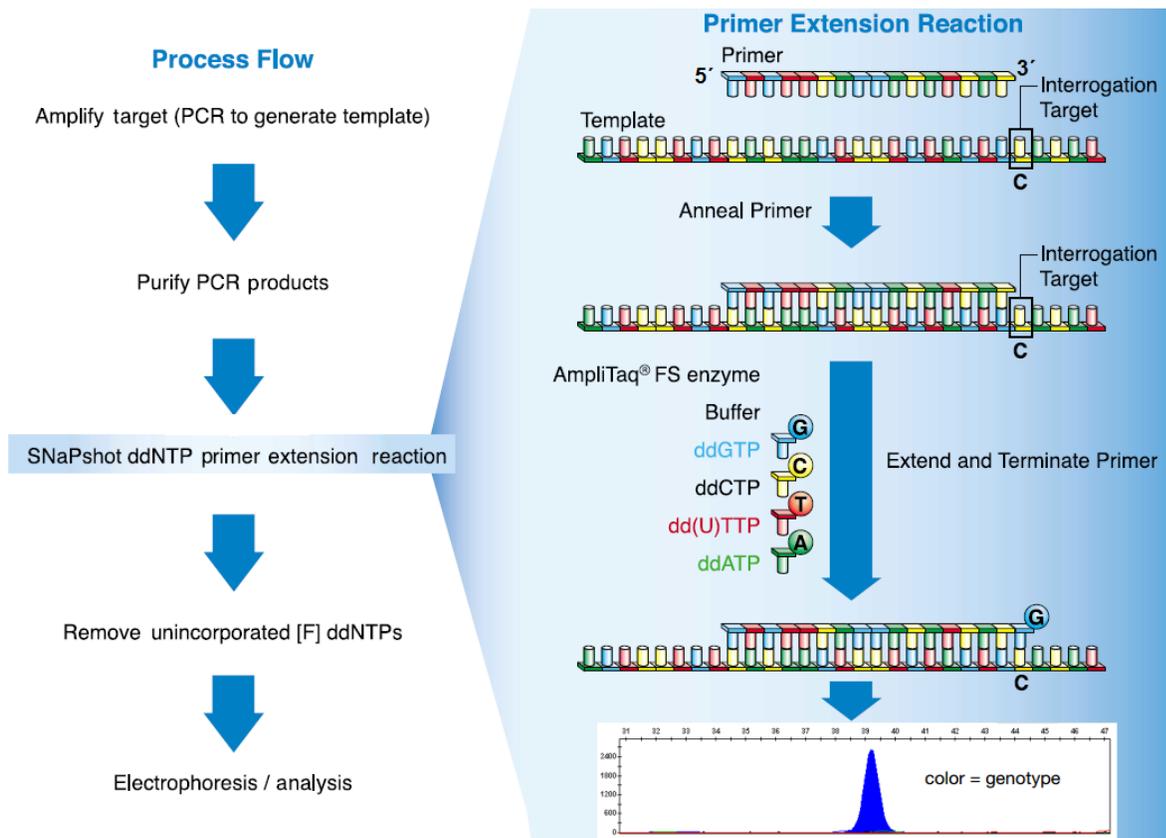
En continuidad al proceso de validación de la técnica, la reacción de minisequenciación se realizó inicialmente para cada primer de manera individual con el fin de confirmar su especificidad y asegurar la ausencia de artefactos. Posteriormente se realizaron ensayos con el mix de primers para cada multiplex con la finalidad de ajustar y optimizar las concentraciones, evaluando la altura en unidades relativas de fluorescencia RFU y el balanceo de los picos de cada SNP.

La metodología se desarrolla empleando el kit comercial de minisequenciación SNaPshot® Multiplex, para lo cual se adiciona a 1,5 μ L del producto de PCR purificado aproximadamente 10ng, 2,5 μ L del SNaPshot Multiplex Ready Reaction Mix y 1 μ L del mix de primers, para un volumen final de 5 μ L. En la tabla 2-5 se describen las condiciones de reacción.

La reacción de minisequenciación o extensión de base simple SBE, es básicamente una reacción de PCR que cuenta con los mismos reactivos convencionales, la diferencia es que no emplea desoxinucleótidos dNTPs, sino que cuenta con cuatro didesoxinucleótidos ddNTPs terminadores y marcados cada uno con un fluorocromo diferente. En la figura 2-5 se puede observar que una vez el primer hibrida sobre el molde de ADN previamente amplificado y purificado (pueden ser uno o varios los moldes a analizar en una sola reacción de SNaPshot), la polimerasa extiende el ddNTP correspondiente, el cual queda incorporado al primer y listo para la detección de la fluorescencia.

Tabla 2-5: Condiciones de reacción para la minisequenciación con SNaPShot.

Estado	Paso	Temperatura °C	Tiempo
25 Ciclos	Desnaturalización	96	10 Seg
	Alineamiento	50	5 Seg
	Extensión	60	30 Seg
Hold	Hold final	4	∞

Figura 2-5: Reacción de Extensión de base simple.

Adaptado de: Product Bulletin 2001. ABI Prism® SNaPshot™ Multiplex System. Applied Biosystems

2.1.10 Purificación del Producto de Minisequenciación

A pesar de que existen varias posibilidades para el proceso de remoción de remanentes de la minisequenciación, se implementó el método de purificación con el kit BigDye® XTerminator™ Purification (Applied Biosystems), debido a que está diseñado para eliminar componentes como iones de sal y ddNTPs que puedan coinyectarse con los productos fluoromarcados, interfiriendo con la captura de señal durante la electroforesis ([XTerminator™, 2007](#)).

El procedimiento consiste en la adición de 25 µL de la mezcla solución SAM+XTerminator a los 5 µL del producto de la minisequenciación en un plato de 96 pozos, se sella cada pozo y se somete a vortex durante 45 minutos, en este paso la solución purificadora captura e inmoviliza los componentes indeseados. A continuación, el producto de la reacción es brevemente centrifugado y en la fracción sobrenadante se conservan los productos purificados que pueden ser llevados directamente al análisis.

2.1.11 Electroforesis Capilar

Para el análisis del producto purificado, se empleó el equipo ABI PRISM® 3130 Genetic Analyzer (Applied Biosystems) con un capilar de 36 cm. Polímero POP4 (Applied Biosystems), módulo E5 y el GeneScan 120 LIZ como size standart internal. Se usaron 10 µL del sobrenadante de la purificación de la minisequenciación, mezclado con 0,1 µL del LIZ 120 (Applied Biosystems).

La electroforesis se realizó con un tiempo de inyección de 2 segundos, voltaje de 15kV y temperatura de 60 °C. Se empleó una matriz DS-02 (Applied Biosystems) con cinco colores dR110, dR6G, dTAMRA y dROX en el filtro E5 para la detección de la fluorescencia en la base incorporada que revela la presencia del SNP.

2.1.12 Análisis de Datos

Los datos obtenidos fueron analizados usando el software GeneMapper™ *ID-X* (Ver. 1.1.1, Applied Biosystems) con el umbral de detección mínimo fijado en 100 RFU. Para el análisis obligatoriamente se deben tener en cuenta el diseño de los primers de minisequenciación, porque la posición de cada SNP dependerá del tamaño de la cola inespecífica como se puede observar en la figura 2-6. Cada SNP es representado por un pico del color correspondiente al nucleótido según la tabla 2-6.

Figura 2-6: Diferenciación de la movilidad electroforética según el tamaño de la cola.

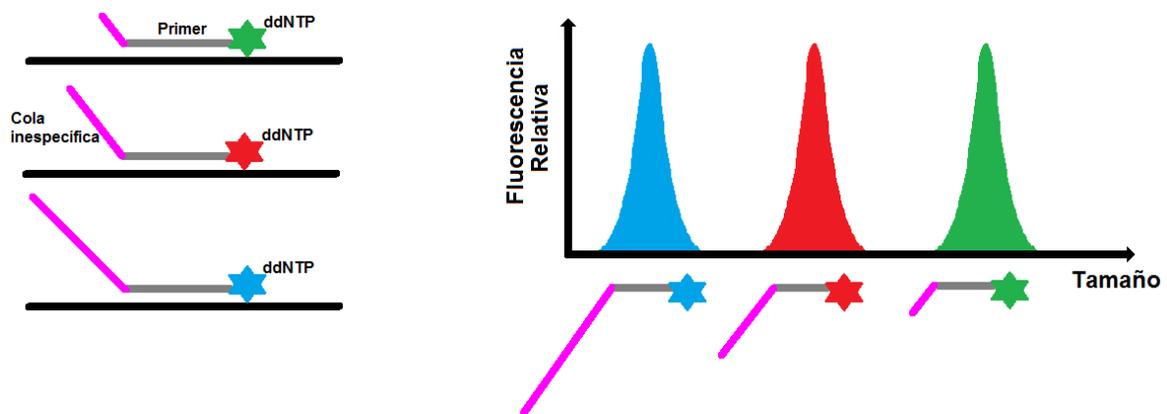


Tabla 2-6: Correspondencia entre el color y el nucleótido.

Color	Nucleótido
Azul	Guanina
Verde	Adenina
Negro	Citosina
Rojo	Timina
Naranja	LIZ 120

2.1.13 Ensayos de Sensibilidad del Método

La sensibilidad de los 15 Y-SNPs fue evaluada mediante diluciones seriadas de dos muestras de ADN masculino previamente cuantificadas. Las cantidades analizadas fueron 1 ng/ μL y 500; 250; 125; 62,5; 31,25; 15,75 y 7,125 pg/ μL .

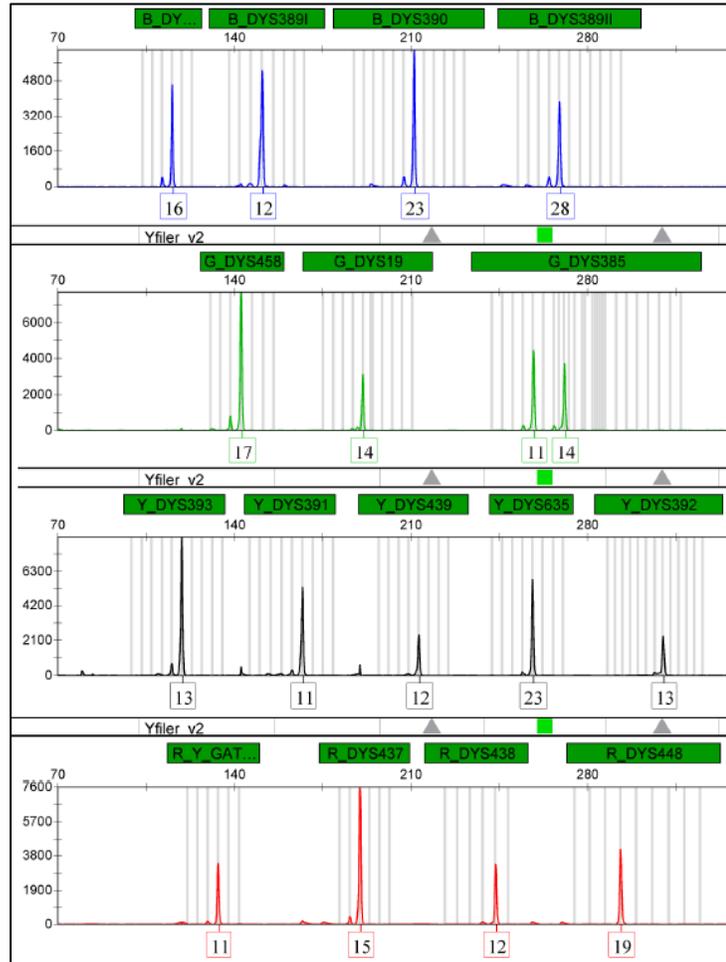
2.1.14 Pruebas de Especificidad del Método

Para confirmar la especificidad del ensayo, se analizaron dos muestras de ADN femenino y una muestra de ADN animal, obtenida de un perro y disponible en el laboratorio del Grupo Genética del CTI de la Fiscalía empleándose como control en las pruebas preliminares de rutina.

2.2 Resultados y Discusión

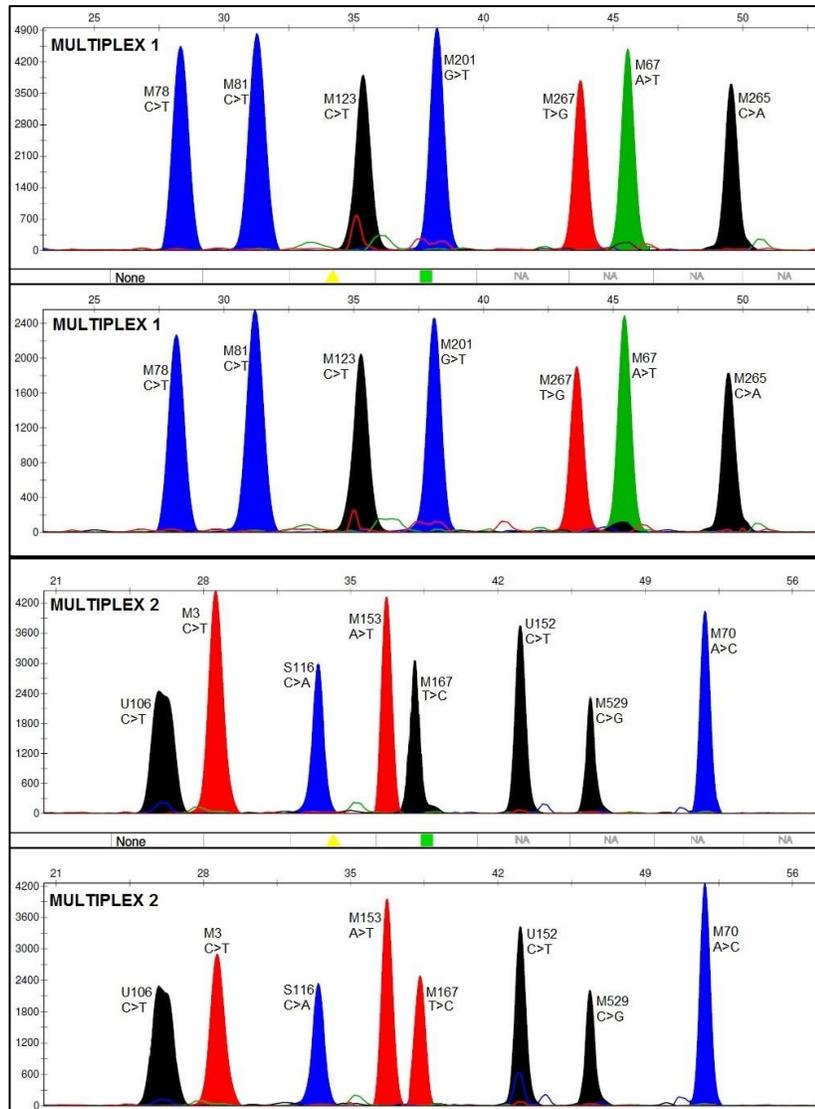
El procedimiento de análisis de Y-STRs es un método ya estandarizado y validado en el laboratorio del Grupo Genética del CTI, por esta razón, su desarrollo estuvo sujeto a las pautas establecidas en el protocolo generado para tal fin, obteniendo el haplotipo de los 17 Y-STRs de manera satisfactoria tal y como se observa en la figura 2-7. Cabe resaltar, que se recomienda realizar el análisis de STRs cuando se establece una validación de SNPs, en razón de conformar un control de calidad de los resultados y fortalecer el análisis dual de los marcadores.

Figura 2-7: Haplotipo de Y-STRs para una de las muestras analizadas en el estudio.



Las estrategias empleadas en la implementación y validación del método resultaron eficaces en la medida en que se observaron productos de las PCRs claros y sin artefactos, en los pasos iniciales se intentó la minimización de la presencia de productos no específicos controlados en su mayoría por la actividad eficiente de los métodos de purificación y el empleo de primers purificados por HPLC. Adicionalmente, se puede observar en la figura 3-8 el resultado de la optimización de las concentraciones de los primers y el manejo de las diversas variables de tipo técnico para alcanzar la validación del método, resultando en una homología lo más satisfactoria posible en la altura en RFU de los picos de cada SNP para las dos reacciones multiplex. En algunos corridos se observaron leves variaciones en la movilidad de los picos de las variantes alélicas y por tanto, en las posiciones de los picos, lo que en ningún caso afectó la asignación alélica de los SNPs.

Figura 2-8: Electroferograma de las dos reacciones multiplex para dos muestras diferentes.

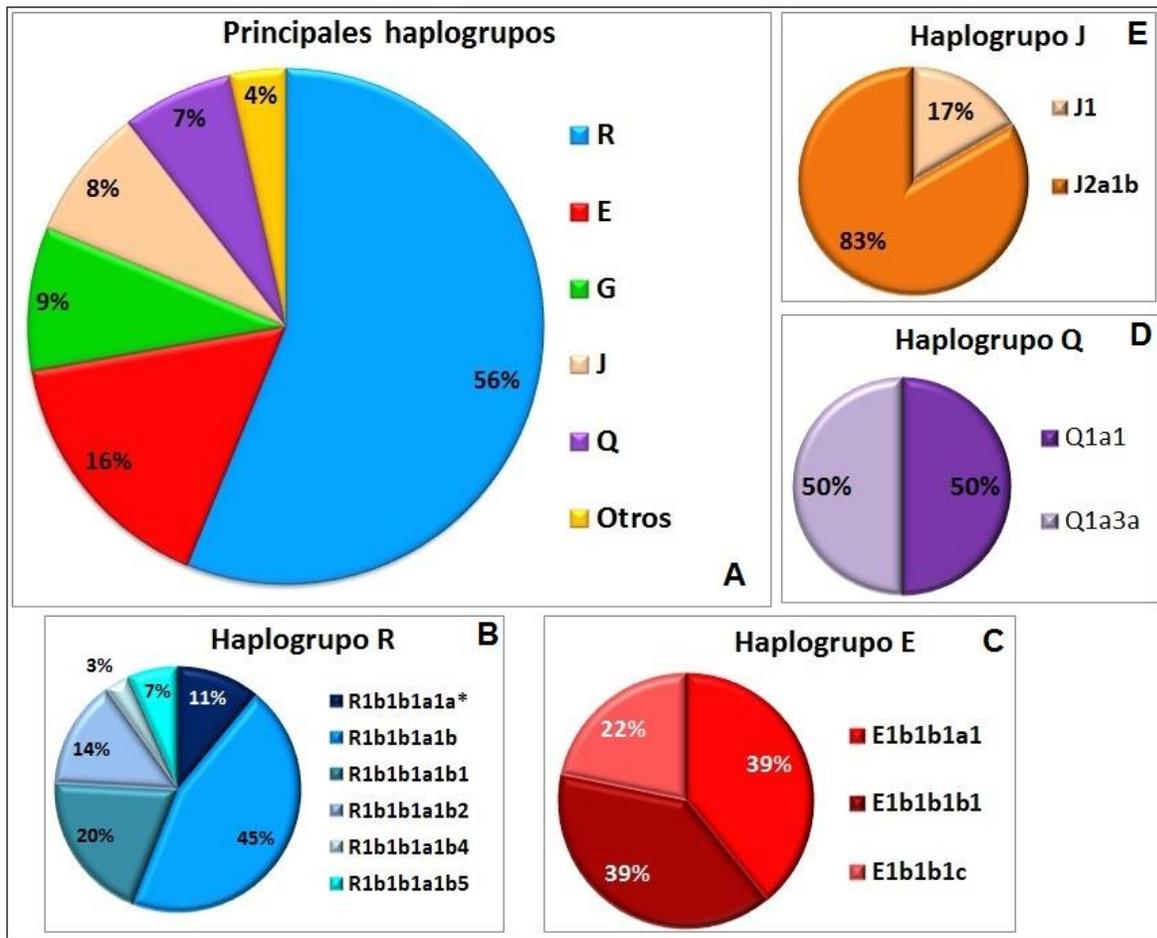


La asignación de haplogrupos en este estudio está basada en la actualización más reciente del árbol filogenético del Consorcio del Cromosoma –Y (<http://ycc.biosci.arizona.edu>). Los resultados de cada muestra se confirmaron replicando cada ensayo en días diferentes, obteniendo resultados similares.

La mayoría de las muestras analizadas se pueden clasificar en el pool genético Euroasiático como se observa en la figura 2-9. El haplogrupo más frecuente fue el R (56%), particularmente el linaje R1b1b1a1b-S116 (27%). Seguido por los haplogrupos E

subhaplogrupos. A pesar de no representar un porcentaje considerable en el total de muestras, se podría contemplar la idea de incluir un par de marcadores SNPs que definan linajes del haplogrupo I y G, discriminando óptimamente este agrupamiento.

Figura 2-10: Distribución de los principales haplogrupos de Cromosoma –Y encontrados en una población Bogotana. Frecuencias de los sublinajes de los haplogrupos R, E y J.



Es importante mencionar que para este estudio no se encontraron incoherencias entre los resultados arrojados con los Y-STRs en el Haplogroup Predictor y los resultados obtenidos en el análisis de los Y-SNPs, obteniendo prácticamente resultados confirmatorios bilateralmente.

La predicción de haplogrupos basados en haplotipos de Y-STRs es una opción eficiente y conveniente cuando se dispone de dicha información y puede considerarse como un buen estimador de este tipo de categorizaciones, sin embargo, el método disponible y mayormente utilizado, el Haplogroup Predictor, expresa una tasa de error bastante

considerable (4.8%) ([Athey, 2005](#)), argumentada en la limitación de una adecuada base de datos de haplotipos de Y-STRs. Dicha justificación no se podría generalizar para todas las poblaciones, pues es bien sabido que existen haplogrupos bien caracterizados en poblaciones como la europea y la africana, no obstante, existen haplogrupos poco frecuentes y por consecuencia son escasamente estudiados.

Es importante entonces, acudir a marcadores Y-SNPs para hacer la determinación de haplogrupos de manera más acertada y preferiblemente en conjunto con los Y-STRs para obtener resultados confirmatorios entre sí. Incluso aun, la determinación de haplogrupos poco frecuentes debe ser cautelosa si se acude exclusivamente a Y-STRs o en poblaciones que no reporten la presencia de esos haplogrupos raros.

Por otro lado, los resultados de este estudio confirman hallazgos en estudios anteriores ([Acosta et al., 2009](#); [L A Alonso, 2013](#); [Bailliet et al., 2011](#); [Bertoni, Jin, Chakraborty, & Sans, 2005](#); [Bolnick, Bolnick, & Smith, 2006](#); [Carvajal-Carmona et al., 2000](#); [Collins-Schramm et al., 2002](#); [Mao et al., 2007](#); [Price et al., 2007](#); [Rodas, Gelvez, & Keyeux, 2003](#); [Sans, 2000](#); [Seldin et al., 2007](#); [Wang et al., 2008](#)) para poblaciones colombianas pero de otras regiones de origen, donde se encontró una predominancia de la ancestría Europea, representada por el linaje R1b y asociada principalmente al proceso de conquista del continente americano por parte de colonias europeas. Con todo y que a Colombia se le considera un país de población mestiza, la predominancia de la ancestría europea es evidente.

Resulta interesante observar que otro haplogrupo frecuente es el E (16%), el cual es muy frecuente en África, pero también se reporta en el medio oriente y el sur de Eurasia ([T. M. Karafet et al., 2008](#); [Semino et al., 2004](#); [M. van Oven, Toscani, van den Tempel, Ralf, & Kayser, 2013](#)), y de igual manera, es muy probablemente asociado con el proceso de conquista y el mercado de esclavos provenientes principalmente de África. Caso contrario, para el haplogrupo Q (7%) se esperaba mayor prevalencia en la población y está asociado a las comunidades nativo-americanas específicamente, por tanto, en teoría se pensaba que enriquecía el pool genético mestizo de Colombia y potencialmente podía tener una considerable presencia en la población, pero como en los estudios previos su frecuencia es realmente muy baja.

Respecto a la prueba de sensibilidad, ésta permitió ubicar el límite de detección de haplotipos completos con 15 Y-SNPs y con picos de altura mayor a 100 RFU, a 125 pg

de ADN. Con cantidades de ADN entre 125 y 31 pg, se obtuvieron haplotipos incompletos y por debajo de los 31 pg no se obtuvo amplificación de Y-SNPs.

La evaluación de especificidad permitió confirmar la ausencia de amplificación en las muestras femeninas y la del perro. Esta especificidad resulta ventajosa para análisis de mezclas de perfiles masculino y femenino, comúnmente encontrados en casos de asalto sexual.

2.3 Conclusiones

De manera general, se puede inferir que la cuidadosa selección de los 15 Y-SNPs fue acertada, porque se logró abarcar en cantidad mínima de marcadores, la aproximación más informativa posible acerca de los principales haplogrupos de cromosoma –Y presentes en la población bogotana analizada. Esto se evidencia porque los SNPs definieron bien los linajes masculinos de la población, lo que es observable en la homogénea distribución de frecuencias de los haplogrupos en la población en general.

Sin apartarse del enfoque forense de este estudio, se dice que la genotipificación de Y-SNPs realizada por minisequenciación, es una técnica simple y eficiente. Por esta razón, el esfuerzo se enfocó en optimizar un método que genere resultados en casos extremos de análisis de ADN. Una ventaja importante de este ensayo es el reducido número de fragmentos coamplificados que facilitan la optimización del método en un laboratorio forense y minimiza los efectos de competición de reactivos durante la amplificación de muestras con escasas cantidades de ADN.

Así pues, las pruebas de evaluación mostraron que el ensayo con los 15 Y-SNPs es un método sensible que provee resultados confiables, robustos y específicos para el estudio del ADN masculino humano.

Aunque el número de haplogrupos es relativamente pequeño, ellos pueden proveer una herramienta simple pero poderosa de exclusión en casos de análisis forenses. De modo similar, la especificidad poblacional atribuida a cada Y-SNP puede determinar el origen biogeográfico de un linaje masculino en particular, teniendo en cuenta eso sí, que podrían existir salvedades en aquellos haplogrupos ampliamente ubicuos en toda población.

El patrón poblacional encontrado en los linajes masculinos de la muestra bogotana, probablemente refleja el impacto de la conquista española durante la época colonial y asociando además el tráfico de esclavos africanos acorde a la época, dados los múltiples movimientos migratorios que se han documentado de Europa hacia América principalmente desde el siglo XIX.

En conclusión, este trabajo deja en disposición dos reacciones PCR multiplex de 15 Y-SNPs seleccionados estratégicamente a lo largo del árbol filogenético y adecuado para estudios con proyección evolutiva o forense.

3. Estudio de la ancestría y estructura del linaje masculino para una población humana de Bogotá (Colombia)

La sucesiva acumulación de marcadores genéticos binarios como los SNPs revela una cascada de diferenciación genética que coincide aleatoriamente con episodios originarios de varias poblaciones en un contexto geográfico y temporal específico ([Carracedo, 2005](#)). De esta manera, los Y-SNPs pueden considerarse como marcadores informativos de ancestría (AIMs) ofreciendo información de la ancestría geográfica de un ADN en particular.

Para el análisis de la paternidad heredada y distribuida geográficamente de manera no aleatoria, se acude a los Y-SNPs ubicados dentro de la región masculino-específica del cromosoma $-Y$, convirtiéndose en una poderosa herramienta para la evaluación del origen biogeográfico de linajes paternos. La diferente combinación de esos marcadores Y-SNPs definen los haplogrupos de cromosoma $-Y$, lo cual comúnmente se asocia con ubicaciones biogeográficas bien definidas, en razón a su profunda caracterización a lo largo de los años en diferentes poblaciones humanas ([Consortium, 2002](#); [Mark A Jobling & Tyler-Smith, 2003](#); [T. M. Karafet et al., 2008](#); [Peter A Underhill & Kivisild, 2007](#); [Mannis van Oven, Ralf, & Kayser, 2011](#)).

Para el caso colombiano, desde la perspectiva genética y cultural, es una región con marcada diversidad poblacional, a pesar de ello, se reconocen tres etnias principales pertenecientes a grupos europeos, africanos y nativo americanos, los cuales probablemente caracterizan la región como una consecuencia de la ubicación geográfica del país. Asentándose al extremo norte de Sudamérica, Colombia es un paso obligado para el acceso al resto de países. Además, las condiciones topográficas y geográficas del territorio, representadas por las regiones naturales, exhiben características únicas en cuanto a clima, suelo y relieve ([IGAC, 2010](#)). Estas diferencias potencialmente

permitieron que los diferentes grupos humanos establezcan asentamiento a lo largo del tiempo en las diferentes regiones asociándose con la heterogeneidad poblacional.

Remitiéndose específicamente a la ciudad de Bogotá, es de extenso conocimiento la diversidad y los amplios orígenes geográficos de los habitantes de la ciudad. Claro es, desde la perspectiva histórica, que los primeros habitantes de Bogotá fueron los muisca (de la familia lingüística chibcha), quienes ocupaban la altiplanicie que incluía parte actual de Boyacá y Santander. Para la época de la conquista, la consigna era establecer un centro urbano organizado y bajo un gobierno estable ([Palacio Castañeda, 2008](#); [Soler](#)). La influencia migratoria fue evidente, pues llegaron nuevos grupos europeos y africanos producto de la esclavitud. Por ello, a lo largo de los siglos el paisaje urbano de Bogotá ha cambiado notablemente, para mediados del siglo XX hubo una oleada migratoria intensa hacia la capital promovida por los actos de violencia de la época, persecución política y por el asesinato de Jorge Eliecer Gaitán (El bogotazo). Igualmente, el crecimiento de la población bogotana se debió a la industrialización comercial que amplió las fuentes de trabajo, el desarrollo de vías de comunicación y la centralización política y administrativa del país ([Soler, 2001](#)).

Para la actualidad, la diversidad poblacional es innegable, desde todas las regiones del país el destino migratorio casi siempre es la ciudad capital, todo ello en busca de nuevas oportunidades laborales, académicas, económicas e incluso como resultado del desplazamiento generado por el conflicto armado del país.

Actualmente, para las poblaciones sudamericanas, los datos obtenidos a partir del análisis de marcadores uniparentales están en concordancia con la evidencia histórica del proceso de emparejamiento asimétrico en relación con el sexo y el origen étnico que predominó en la formación de las poblaciones de América ([Suarez-Kurtz, 2005](#)). Uno de los hallazgos del estudio, es la introducción de genes de origen europeo en la población americana a través de los hombres; ya que históricamente existió un número desigual de uniones entre hombres Europeos y mujeres nativas y africanas. En consecuencia, el ADN mitocondrial muestra una contribución americana predominantemente nativa mientras que el cromosoma -Y indica una contribución principalmente europea ([L A Alonso, 2013](#); [Bortolini et al., 2002](#); [Gonzalez-Andrade, Roewer, Willuweit, Sanchez, & Martinez-Jarreta, 2009](#); [Mesa et al., 2000](#); [Tarazona-Santos et al., 2001](#)).

En este sentido, no se dispone de muchos estudios que busquen evaluar la diversidad y la estructura genética de la población masculina de Bogotá. En la escasa disponibilidad de estudios se abordan otro tipo de poblaciones o lo hacen de manera muy general donde se incluye un número muy pequeño de representantes de cada población o lo que es peor, muchos de los estudios aún no se han publicado. Por tanto, se puede decir que no existe una caracterización exhaustiva de la diversidad de los linajes de cromosoma –Y para la población Bogotana. Alternativamente, ciertos estudios han procurado realizar la reconstrucción de la ancestría patrilineal utilizando como base de análisis la dinámica de poblamiento de las comunidades latinoamericanas en general ([Luz Angela Alonso & Usaquén, 2013](#); [Gonçalves, Spinola, & Brehm, 2007](#); [Gonzalez-Andrade et al., 2009](#); [Mendizabal et al., 2008](#); [Nunez et al., 2012](#); [Nuñez et al., 2010](#); [Simms et al., 2011](#); [Simms et al., 2012](#); [Jada Benn Torres, Doura, Keita, & Kittles, 2012](#); [J Benn Torres, Kittles, & Stone, 2007](#)).

Con todo esto, la población Bogotana ofrece un marco de estudio genético-poblacional atractivo y conveniente para el establecimiento de registros genéticos como fuente referencial para análisis con aplicación social, ancestral y forense, teniendo en cuenta que el registro de frecuencias genéticas en el contexto biosocial o población de referencia representa la base del proceso de identificación. Así y todo y en respuesta a la limitada disponibilidad actual de estudios que suplan la carencia de información sobre la caracterización genética de la población masculina Bogotana, se establece un análisis de 15 haplogrupos definidos por marcadores tipo SNPs y 17 microsatélites de cromosoma –Y informativos que promuevan el análisis de la distribución de haplotipos, así como las correspondientes inferencias ancestrales suscitadas de dichos análisis.

3.1 Materiales y métodos

A partir de la disponibilidad de Y-SNPs en el árbol filogenético reportado en la actualización más reciente del Consorcio del Cromosoma –Y, se eligieron los marcadores que se esperaba puedan ajustarse y caracterizar la distribución de linajes masculinos para la población bogotana.

3.1.1 Población analizada

Los criterios a tener en cuenta para la elección de muestras, fueron un tiempo de residencia mínima de 5 años en la ciudad de Bogotá o municipios cercanos, descartando

además, cualquier posibilidad de parentesco familiar entre los muestradantes. Básicamente, las 144 muestras de sangre líquida y algunas en tarjetas FTA® Whatman, se eligieron del banco de muestras de referencia del Grupo Genética del CTI de la Fiscalía. Las cuales contaban con un consentimiento informado que incluyó una autorización para usar las muestras en estudios poblacionales, con fines académicos y científicos. Además, se diligenció una encuesta que incluía información sobre sus familiares por vía paterna, entre otros datos. Esta investigación contó con el aval del comité de ética de la facultad de Medicina de la Universidad Nacional de Colombia (Acta No. 006-036-15).

Para la extracción de ADN se empleó el DNA IQ™ Casework Pro Kit for Maxwell® (Promega).

3.1.2 Análisis molecular

Las muestras se cuantificaron mediante PCR en tiempo real, se analizaron los 17 marcadores Y-STRs del kit Y-filer® (Applied biosystems), la detección y separación de los productos amplificados fue realizada en un analizador genético ABI PRISM 3130xl y con ayuda del software *GeneMapper® ID-X* se realizó el análisis de datos.

Para el análisis de los Y-SNPs, se diseñaron tanto los primers de amplificación de los 15 segmentos de ADN, como los de minisequenciación, se optimizaron las concentraciones y se validaron las condiciones técnicas del procedimiento. El panel de marcadores filogenéticos tipo Y-SNPs, seleccionados porque se esperaba que definan los haplogrupos más frecuentes para la población bogotana son: M78, M81, M123, M201, M267, M67, M265, M3, U106, S116, U152, M529, M153, M167 y M70, analizados en forma de dos reacciones multiplex descritas en el capítulo 2.

3.1.3 Análisis estadístico

Se determinaron las frecuencias de los haplogrupos encontrados y se construyeron las redes Median-Joining para determinar las relaciones existentes entre los principales haplogrupos de cromosoma -Y, utilizando el programa Network 4.6.1.1 (www.fluxustechology.com), el cuál asume un modelo de mutación Stepwise para los microsatélites. Las redes construidas para los haplogrupos utilizaron 15 loci

microsatélites (DYS19, DYS389I y DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, y GATA H4). En todos los casos el marcador DYS385 fue excluido, dado que este corresponde a dos loci diferentes que no pueden ser discriminados por el método de tipificación utilizado ([Toscanini et al., 2008](#)). Adicionalmente como el locus DYS389II contiene el locus DYS389I, se substrajo el tamaño de DYS389I a partir de DYS389II ([Bolnick et al., 2006](#)). Cada una de las redes fue resuelta mediante la aplicación median joining ([Bandelt, Forster, & Rohl, 1999](#)). Los pesos de cada uno de los loci en la red fueron asignados de 3 a 10 siguiendo la metodología propuesta por ([Muzzio, Muzzio, Bravi, & Bailliet, 2010](#)) y las tasas mutacionales reportadas en el YHRD.

Se realizó un análisis de componentes principales (PCA) para visualizar si las regiones de ancestría comparten distribuciones de frecuencias y determinar qué haplogrupos son los que generan una mayor separación dada su frecuencia, esto se realizó mediante el programa MVSP (Multivariate Statistical Package) versión 3.1 ([Kovach, 1998](#)).

3.2 Resultados

En este estudio se identificaron 15 haplogrupos y las frecuencias de cada uno de ellos se reporta en la tabla 3-1. En el análisis de las frecuencias se observó que el haplogrupo R1b1b1a1b-S116 tuvo la mayor representación en la muestra con un 21,5%. Junto a este alto porcentaje se encontró una gran diversidad de linajes derivados de este haplogrupo: R1b1b1a1b1-U152 (n=18), R1b1b1a1b2-M529 (n=13), R1b1b1a1b4-M153 (n=3) y R1b1b1a1b5-M167 (n=6), los cuales sumados en total dan una representación de linajes R1b1b1a1b del 49,31%. Adicionalmente en una frecuencia de 6,9% (n=10) se observó el haplogrupo R1b1b1a1a*-U106 el cual también representa otra de las múltiples ramas que caracteriza el haplogrupo R. En general, se encontró que 81 cromosomas -Y de la muestra bogotana pertenecen al haplogrupo R1b-M343 con el 56,25%.

De igual manera, el haplogrupo que siguió en orden de frecuencias fue el E, con los linajes E1b1b1a1-M78 (n=9), E1b1b1b1-M81 (n=9) y E1b1b1c-M123 (n=5), los cuales sumados en total dan una representación de linajes E del 15,98%, con un total de 23 cromosomas -Y representados con este linaje.

En consecuencia, el haplogrupo G es el que siguió en el orden de las frecuencias, con $n=13$ y un porcentaje total de 9%, teniendo en cuenta que fue un haplogrupo principal, el cual incluye varios linajes.

Tabla 3-1: Frecuencias de haplogrupos de cromosoma –Y para la población bogotana.

haplogrupo	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
E1b1b1a1	9	6,3	6,3	6,3
E1b1b1b1	9	6,3	6,3	12,5
E1b1b1c	5	3,5	3,5	16,0
G	13	9,0	9,0	25,0
J1	2	1,4	1,4	26,4
J2a1b	10	6,9	6,9	33,3
Q1a1	5	3,5	3,5	36,8
Q1a3a	5	3,5	3,5	40,3
R1b1b1a1a*	10	6,9	6,9	47,2
R1b1b1a1b	31	21,5	21,5	68,8
R1b1b1a1b1	18	12,5	12,5	81,3
R1b1b1a1b2	13	9,0	9,0	90,3
R1b1b1a1b4	3	2,1	2,1	92,4
R1b1b1a1b5	6	4,2	4,2	96,5
T	5	3,5	3,5	100,0
Total	144	100,0	100,0	

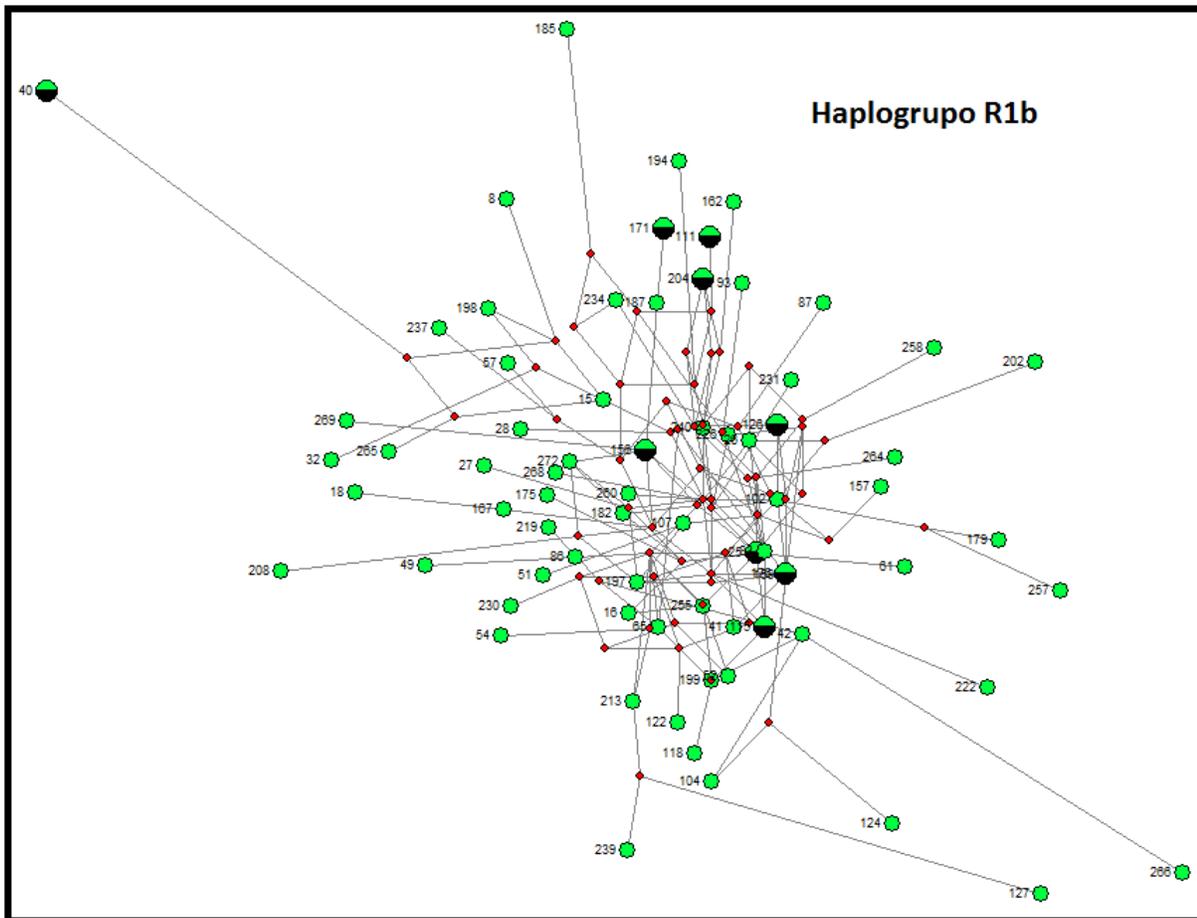
Los 27 cromosomas –Y restantes se distribuyeron en los siguientes haplogrupos: los linajes del haplogrupo J con J1 ($n=2$) y J2a1b ($n=10$) para un total de 8,34%. Se encontraron para el haplogrupo principal Q al linaje Q1a1 ($n=5$) y Q1a3a ($n=5$), encontrando que el 6,95 % de los cromosomas –Y analizados, corresponden a este haplogrupo. Y de igual manera, el haplogrupo T reportó $n=5$ para un 3,47% del total de muestras analizadas.

3.2.1 Análisis de redes filogenéticas

Con el propósito de tener un mejor conocimiento de las bases moleculares de la variabilidad encontrada en la muestra, se construyeron redes median-joining con todos los haplotipos pertenecientes a los haplogrupos R1b, E, G, J y Q. En el primer caso se utilizaron 15 Y-STRs salvo el locus DYS385a/b. Se observó que esta red carece de estructura, es decir, todos sus haplotipos se encuentran distribuidos al azar sin ningún tipo de patrón que les agrupe como se aprecia en la figura 3-1, lo cual es consistente

desde el punto de vista genético dados los niveles de variabilidad que se encontraron para ese haplogrupo y desde el punto de vista histórico, dados los múltiples movimientos migratorios que se han documentado de Europa hacia América principalmente desde la colonia siglo XIX. No obstante, se observa el aislamiento de algunos haplotipos ubicándose en los extremos de la gráfica, así mismo, se observa que algunos haplotipos se repiten, lo que se refleja en los círculos de dos colores.

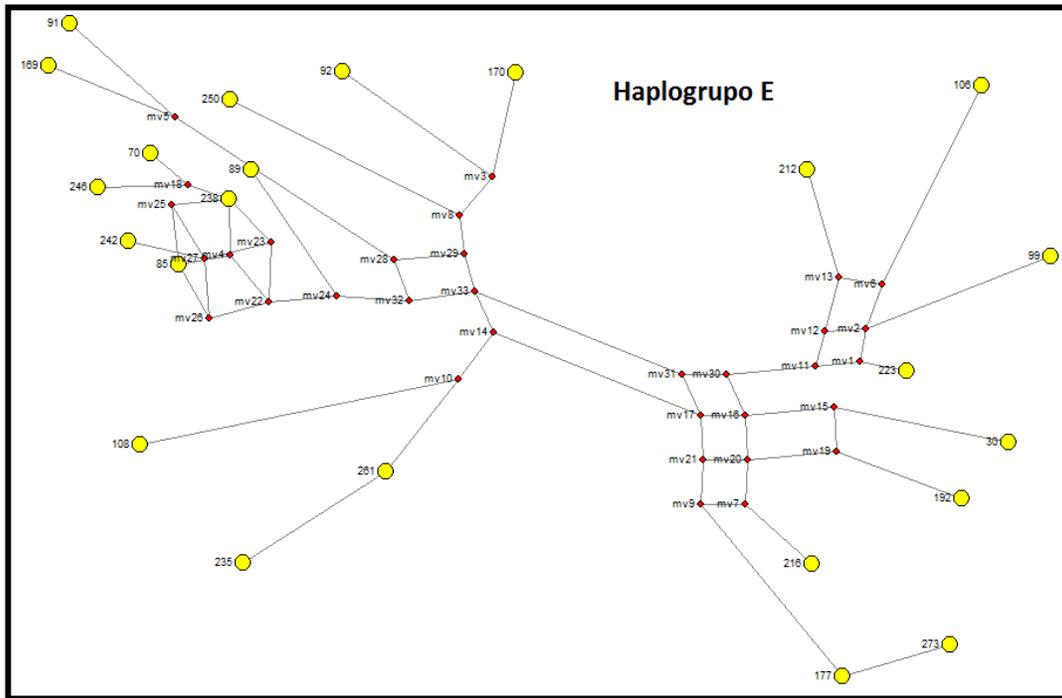
Figura 3-1: Median-joining network del haplogrupo R1b utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.



Para el haplogrupo E se observa una distribución diferente en la red, la figura 3-2 muestra que no hay repetición de haplotipos, además se observa que la distribución se divide en dos grandes grupos, a pesar de ser tres los SNPs analizados para este haplogrupo, al parecer el clado E1b1b1c-M123 con $n=5$, se acopló en su distribución al

clado E1b1b1b1-M81, el cual es su ancestro directo y por tanto, es con el que conserva marcadores en común. Así y todo, se observa una distribución generalizada, pues el agrupamiento no es evidente para este haplogrupo.

Figura 3-2: Median-joining network del haplogrupo E utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.



Para el caso del haplogrupo G, su distribución es bastante dispersa, en la figura 3-3 se observa claramente la abundancia de haplotipos intermedios o hipotéticos en la muestra, lo cual explica la amplia separación de los haplotipos y de igual manera se destaca que el haplogrupo más extendido y por tanto, el más aislado, está repetido, correspondiendo probablemente a las muestras pertenecientes al haplogrupo I, las cuales se clasificaron dentro de este haplogrupo principal, en razón a la falta de un marcador SNP que defina este haplogrupo, por lo cual, se incluyen en el haplogrupo más cercano ancestralmente, lo que en este caso le corresponde al haplogrupo G.

Algo similar se obtuvo para el haplogrupo J, en la figura 3-4 se observa una distribución de haplotipos dispersa, destacando que en este caso se observan haplotipos únicos, los cuales en su mayoría distan entre sí mediante varios pasos mutacionales expresados en

el amplio número de haplotipos hipotéticos que enlazan los haplotipos identificados en este estudio para este haplogrupo en particular.

Figura 3-3: Median-joining network del haplogrupo G utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.

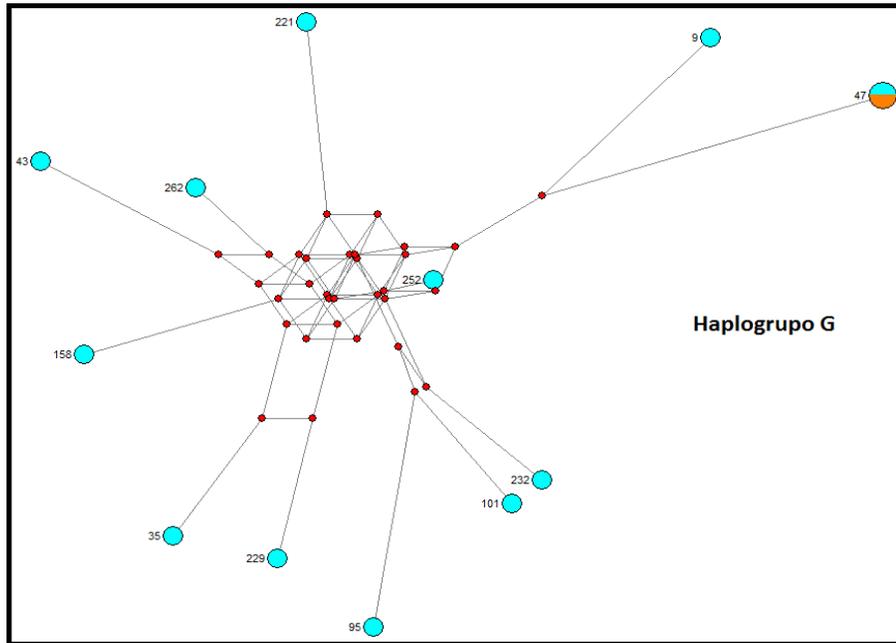
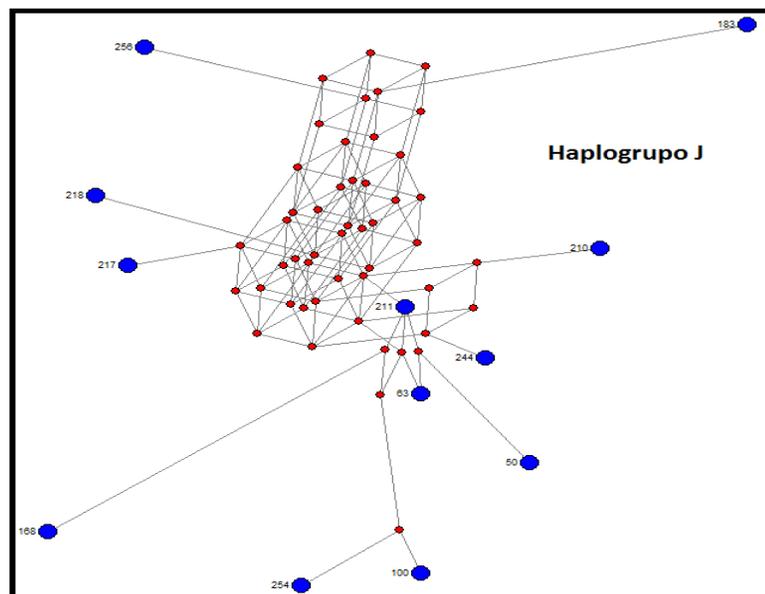
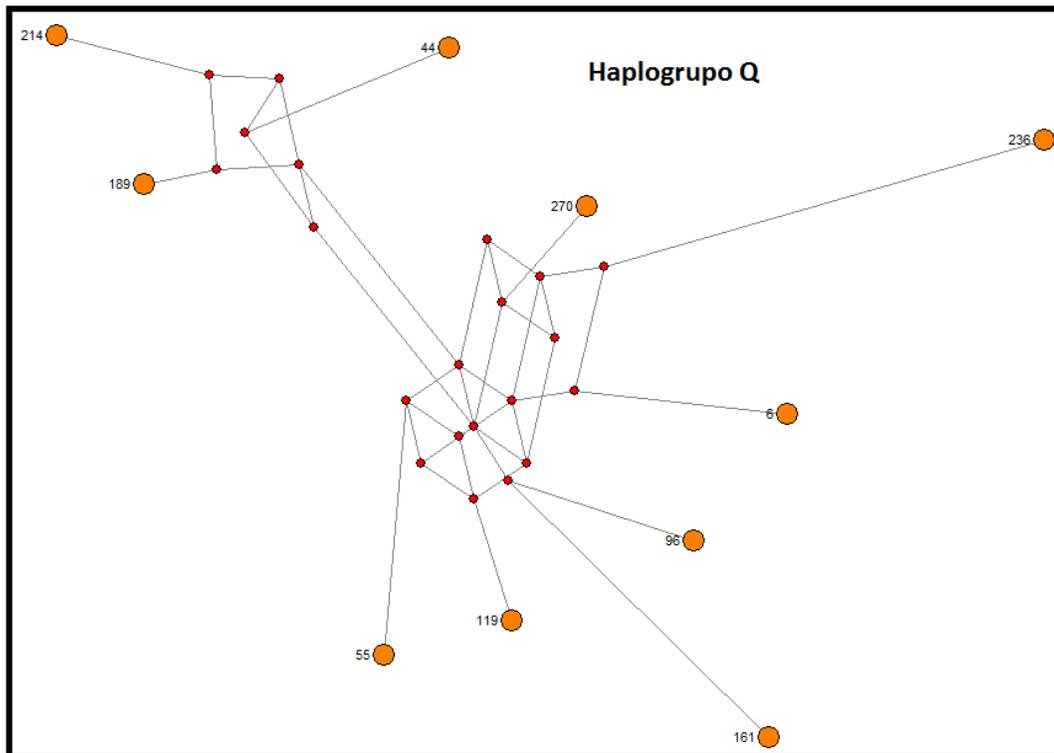


Figura 3-4: Median-joining network del haplogrupo J utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.



El análisis del haplogrupo Q, permite observar a través de la figura 3-5, que la dispersión de los haplotipos es muy amplia, por tanto, se puede decir que los haplotipos tienen pocos marcadores en común, e incluso, la expresión alélica es muy diversa. No se observan haplotipos repetidos. Se debe tener en cuenta además, que este haplogrupo se esperaba en mayor frecuencia, debido a que caracteriza el linaje nativo americano propio de la región.

Figura 3-5: Median-joining network del haplogrupo Q utilizando 15 sistemas microsatélites. Los círculos representan haplotipos con área proporcional a su frecuencia. Los median vectors (haplotipos ausentes o extintos) se encuentran en rojo.

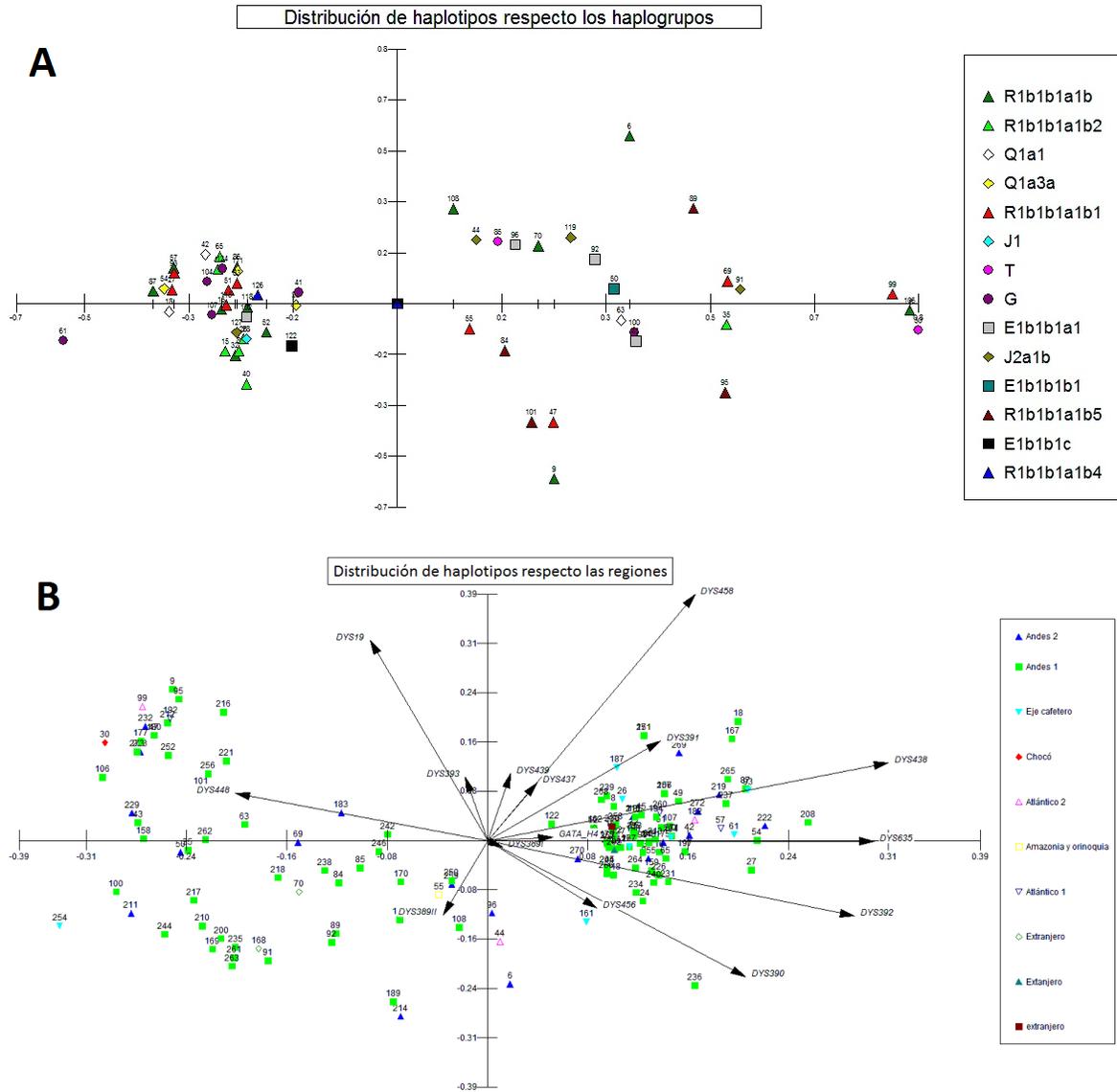


3.2.2 Análisis de componentes principales

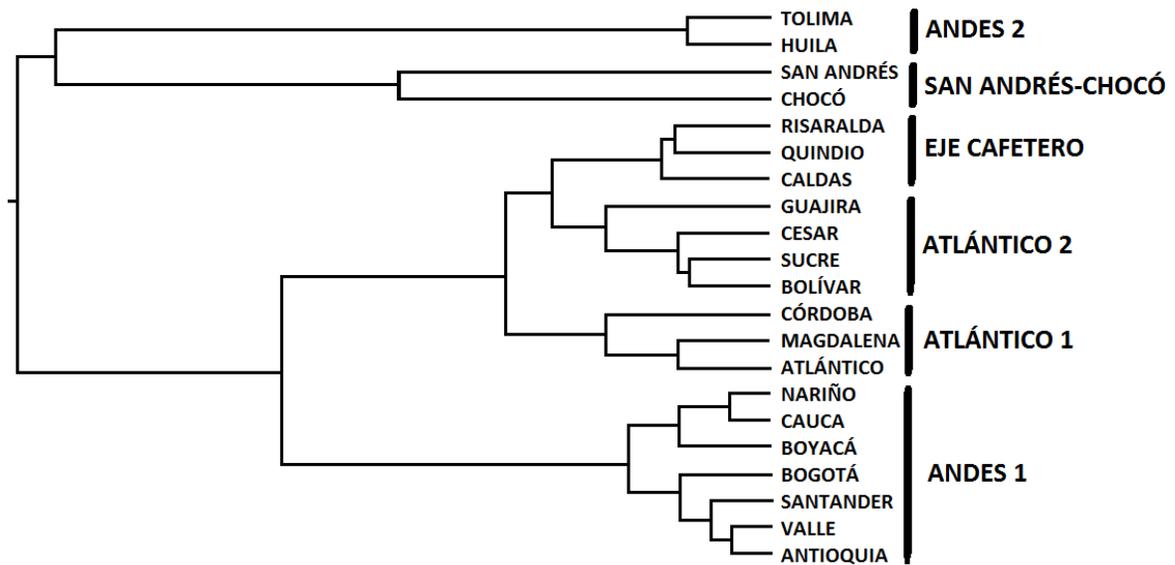
Los haplotipos de Y-STRs se sometieron a análisis multivariado de componentes principales (PCA), para visualizar si la distribución espacial de los haplotipos sigue un patrón definido y acorde a variables como la ancestría o las regiones de origen. En la figura 3-6 se observan dos tipos de agrupamientos para las dos opciones de análisis, bien sea la distribución de haplotipos respecto a los haplogrupos o respecto a las regiones de origen. Las agrupaciones en los cuadrantes de ambos lados, se observa que

una de ellas es más dispersa que la otra, sin embargo, no se observa que obedezcan a algún patrón de distribución en particular asociado a una de las variables.

Figura 3-6: Análisis de componentes principales. **A** distribución de los haplotipos respecto los haplogrupos. **B** distribución de los haplotipos respecto las regiones de origen.



Para este análisis se contó con la distribución de regiones planteada por (Usaquén, 2006), donde se establecen agrupamientos de 23 poblaciones diferentes, provenientes de diferentes departamentos de Colombia, según la estructura genética demostrada en los análisis multivariados por el método Ward, que permite la formación de seis regiones claramente agrupadas como se observa en la figura 3-7.

Figura 3-7: Estructura genética de la población colombiana.

3.3 Discusión de resultados

En la última década se han publicado diversos estudios que han caracterizado la gran heterogeneidad genética de las poblaciones latinoamericanas ([L A Alonso, 2013](#); [Bailliet et al., 2011](#); [Bertoni et al., 2005](#); [Bolnick et al., 2006](#); [Carvajal-Carmona et al., 2000](#); [Collins-Schramm et al., 2002](#); [Mao et al., 2007](#); [Price et al., 2007](#); [Rodas et al., 2003](#); [Sans, 2000](#); [Seldin et al., 2007](#); [Wang et al., 2008](#)).

Recientemente se han realizado una gran cantidad de esfuerzos por caracterizar la diversidad biológica y cultural de las poblaciones humanas en Colombia, abordándola desde diferentes enfoques metodológicos, sin embargo, los estudios de cromosoma -Y desde la perspectiva filogenética y genético-poblacional empleando microsatélites y polimorfismos de nucleótido simple, ciertamente son escasos. Del mismo modo, la mayoría de los mismos se han concentrado en poblaciones pequeñas o en comunidades nativas, no obstante, en este estudio la estrategia de análisis con marcadores informativos estuvo enfocada en caracterizar los principales haplogrupos de cromosoma -Y, en relación con la importancia histórica y social de la ciudad de Bogotá y de esta manera, poder describir de una forma más profunda y robusta la diversidad genética de los linajes masculinos de la zona.

Colombia como a varias poblaciones latinas se le considera de origen mestizo o tri-étnico, que comúnmente se describe en términos de linajes característicos de

poblaciones Europeas, Indígenas y Africanas como producto de la conquista a partir de 1500 y más fuertemente para el siglo XIX ([Bolnick et al., 2006](#); [Carvajal-Carmona et al., 2000](#); [Geppert et al., 2015](#); [Mark A Jobling & Tyler-Smith, 2003](#)).

El estudio publicado por Carvajal-Carmona et al. (2000) fue uno de los primeros en describir este proceso en una población Colombiana, en el departamento de Antioquia. En dicho estudio las principales conclusiones mencionan que “aproximadamente el 94% de los cromosomas -Y es de origen Europeo, el 5% Africano y el 1% Amerindio”. Por el contrario, “la proporción de linajes maternos de DNA mitocondrial fue aproximadamente 90% Amerindia”. De esta diferencia proviene esa denominada asimetría entre las frecuencias de linajes paternos y maternos en población colombiana, confirmada por estudios posteriores, incluyendo el aquí planteado.

En los 144 cromosomas -Y analizados en este estudio, 134 fueron únicos (93,05%) y se distribuyeron en 6 linajes principales de cromosoma -Y: haplogrupos: R, Q, J, G, T y E. De estos haplogrupos se encontró una representación mayoritaria del linaje específico R1b (56,25%), donde 73 de sus haplotipos se caracterizan por ser únicos, de igual manera, los SNPs corriente abajo de la mutación discriminaron muy bien el haplogrupo principal, por consiguiente la diversidad encontrada en este haplogrupo puede ser el reflejo del proceso de colonización en la región, que introdujo hombres provenientes principalmente de la Península Ibérica; lugar donde el haplogrupo R1b presenta frecuencias mayores al 80% y se considera originario de dicha zona ([Adams et al., 2008](#)).

Era muy importante profundizar en el haplogrupo R1b, pues en todos los estudios previos con poblaciones mestizas latinoamericanas, se observó una prevalencia superior al 60%, razón por la cual se pretendía discriminar mejor este haplogrupo en pro de la perspectiva forense de este estudio. Dicho objetivo se alcanzó de manera eficiente, no obstante, se recomendaría implementar un par de marcadores tipo SNPs que caractericen aún más este clado, dejando como precedente la caracterización determinada en este estudio.

Por otro lado, el haplogrupo E mostró una frecuencia moderada en la muestra. Este haplogrupo es el más frecuente en África, pero también se reporta en el medio oriente, el sur de Europa y Asia ([Nunez et al., 2012](#)). Al parecer su influencia en la población colombiana es considerable, por tanto, resultaría acertado prestar mayor atención al estudio de esta ancestría, pues se observa que la distribución está fragmentada en dos

grandes agrupamientos asociados a los marcadores analizados, no obstante, los haplotipos del marcador más reciente y menos frecuente, el M123 (n=5) se asocian a los haplotipos del marcador M81 que es el inmediatamente ancestral, esto probablemente se deba a que la distancia genética y ancestral entre estos marcadores es reducida.

Debido a que la Península Ibérica y la región norteafricana se encuentran cercanas desde la perspectiva geográfica e histórica, esta última caracterizada por los diferentes y múltiples movimientos poblacionales a lo largo del tiempo a través del estrecho de Gibraltar ([Bosch et al., 2003](#)). En esta investigación se encontraron miembros del linaje E1b1b1 con un 6,3% de contribución (N=9) en la muestra. Este haplogrupo es el más frecuente en poblaciones que hablan lenguas afroasiáticas africanas (camito-semíticas), así como en muchas poblaciones de lenguas Nilo-saharianas, y su mayor diversidad se encuentra en Etiopía. E1b1b1 se encuentra ampliamente extendido en el cuerno de África, Norte de África, Sudán, árabes y judíos askenazíes y sefardíes; también en el Sur de Europa, los Balcanes y gran parte de África subsahariana ([Wood et al., 2005](#)).

En la muestra también se encuentran otros haplogrupos frecuentes en Europa Occidental; como es el caso del linaje J con frecuencia del 8,3% (n=12) que se encuentra muy difundido en Oriente Próximo, el Cáucaso y África del norte, pero en frecuencias mucho menores en el sur de Europa, cuerno de África, Asia central y Sur de Asia, así como el haplogrupo T con frecuencia del 3,5% (n=5) y el haplogrupo G con el 9% (n=13).

El conjunto de cromosomas –Y pertenecientes al linaje Q1a3a-M3 y al Q1a1-M265 al ser analizados filogenéticamente en conjunto, fueron diferentes, ya que se ubican de manera indistinta en la red genética (Figura 3-5), así mismo, la muy baja frecuencia observada, contrario a lo esperado, podría representar la presencia de haplotipos casi extintos o en muy baja frecuencia en las comunidades nativas actuales. Todo esto se argumenta como resultado de la dinámica de expansión, contracción y extinción que han vivido las comunidades indígenas en América sumado al fuerte efecto que tiene el fenómeno de deriva y la selección natural en el cromosoma –Y ([Wang et al., 2008](#)).

Respecto al análisis de componentes principales, se analizaron los haplotipos de Y-STRs respecto a las regiones de origen de los muestradantes, clasificados así según el lugar de nacimiento de la persona o de sus familiares por vía paterna, para ello se empleó la clasificación regional establecida por Usaquén 2006, donde se analizó la estructura genética de 23 poblaciones colombianas, obteniendo un cladograma de las relaciones

filogenéticas de todas las poblaciones, esto permitió generar agrupaciones regionales como se observa en la figura 3-7, las cuales se emplearon para clasificar la ancestría de las muestras de este estudio. No obstante, el patrón de distribución ciertamente es aleatorio, a pesar de observar dos grandes agrupaciones a ambos lados de los cuadrantes, pero dicha distribución no se asocia con las regiones. Del mismo modo, se observa un panorama similar cuando se analizan los haplotipos de Y-STRs respecto los haplogrupos definidos por los marcadores SNPs, en la figura 3-6A, se aprecia nuevamente las dos grandes agrupaciones a ambos lados de los cuadrantes, en este caso hay una agrupación más concentrada, sin embargo, no se observa claramente que dicha agrupación esté relacionada directamente con la distribución de haplogrupos.

En consecuencia, el análisis de componentes principales demuestra que la distribución de haplotipos se relaciona con algunos marcadores Y-STRs en especial; los cuales son más conservativos y no tan polimórficos, permitiendo la agrupación más concentrada de haplotipos, a diferencia de los marcadores polimórficos que dispersan más las muestras con una agrupación más extendida, argumentándose en las pequeñas diferencias en las tasas de mutación de algunos marcadores Y-STRs, que incluso son de gran importancia en el campo forense debido a su capacidad polimórfica y poder discriminatorio, llegando incluso en algunos casos a individualizar personas separadas generacionalmente, pero vinculadas por vía patrilineal.

Estos hallazgos ponen en manifiesto la importancia de hacer una exhaustiva caracterización de los linajes de cromosoma -Y en poblaciones latinoamericanas indígenas como no indígenas y de esta manera evaluar y proponer mejores modelos que expliquen y aporten una mayor resolución al proceso de poblamiento americano, pues la pérdida de linajes nativos es evidente.

De manera general se puede decir que se observó una población con una gran diversidad genética >90% que contiene haplotipos únicos para los linajes característicos nativos americanos y los euro asiáticos. El hecho de profundizar en la caracterización genética de linajes con una resolución tan eficiente, es un importante logro para establecer estudios genético-forenses que definan acertadamente la ancestría masculina de una muestra desconocida.

4. Evaluación del desempeño del método en procesos de identificación humana

El análisis forense de una muestra de ADN pretende establecer una asociación entre la evidencia biológica y su fuente, bien sea, el sospechoso, la víctima o una tercera persona. Dicho análisis tiene como principio asociativo, los alelos presentes en la evidencia comparado con los alelos presentes en las referencias.

Se debe tener en cuenta además, que la base del análisis genético en el campo forense es el polimorfismo, el cual, dependiendo de su naturaleza, es decir si es bialélico o polialélico, permitirá individualizar a una persona dada su variabilidad interindividual que difícilmente podría repetirse de manera idéntica entre dos individuos no emparentados. Dicho argumento aplica perfectamente para los polimorfismos autosómicos, no obstante, para el caso de los polimorfismos uniparentales, su variabilidad permite individualizar el linaje, mas no a la persona; dicho en otras palabras, mediante el análisis de marcadores uniparentales pueden establecer un parentesco familiar por vía paterna o materna exclusivamente, siendo causal de exclusión de cualquier individuo no relacionado familiarmente por cualquiera de estas vías, mas no podría especificar la procedencia del polimorfismo dentro de un linaje familiar.

Entre las consideraciones relevantes del estudio aquí planteado, se encuentra la aplicación de la técnica en procesos de identificación humana. Entendiendo que el método de minisequenciación expresa flexibilidad para diversos fines, adaptándolo acorde a las necesidades investigativas, es decir, si el análisis está enfocado en estudios antropológicos o poblacionales, el tamaño de los amplicones no es relevante, porque se dispone en su mayoría de fuentes biológicas que aún conservan el ADN en buen estado. Si el enfoque es clínico, no habría necesidad de usar muchos marcadores simultáneamente, simplificando la técnica del método hacia uno o dos marcadores de interés. No obstante, si la perspectiva de análisis es forense, como lo planteado en este

estudio, las condiciones técnico-científicas deben adaptarse a los intereses de dicha ciencia. Tal es el caso de las fuentes biológicas de análisis que se pueden encontrar, tratándose de restos óseos sometidos a un sin fin de escenarios ambientales adversos, muestras trazas que incluyen cabellos, semen, dientes, orina, saliva, tejidos, uñas y pequeñas manchas con fluidos biológicos encontrados en escenas de crimen, así mismo, pueden haber manchas de sangre diluidas, sometidas a agentes dañinos o fuentes de contaminación. Por tanto, para tratar este tipo de muestras, se debe desarrollar un sistema de análisis muy sensible, procurando controlar la mayoría de variables involucradas. Para el aspecto técnico, el principal objetivo es contar con amplicones con el menor tamaño posible y optimizar el número de ciclos de amplificación de las reacciones de PCR. El aspecto analítico también debe tenerse en cuenta, pues la elección de los marcadores SNPs debe ser estratégica, pensando en la utilidad informativa en el campo forense, ya que se dispone de una gran variedad de marcadores, sin embargo, en el campo forense se pretende establecer un posible origen biogeográfico poblacional y una ancestría asociada al linaje masculino de la muestra; y que sean representativos de la población objeto de estudio.

Entonces, para la ciencia forense en general, es muy importante el apoyo del análisis de marcadores genéticos, disponiendo de una variedad de ellos y de diferente naturaleza. Lo importante ahora, es la significativa aplicación forense de los polimorfismos genéticos, empleados rutinariamente al estudio de los vestigios biológicos presente en las escenas donde ocurrió un delito; y muchas veces, es la principal herramienta para la demostración de la participación de una persona en un hecho delictivo. Otro valor agregado es la posibilidad de aplicación a cualquier tipo de vestigio biológico, bien sea sangre, semen, cabellos, saliva, tejidos, etc y sobre aquellos objetos que puedan contener o ser soporte para este tipo de vestigios.

Por otro lado, los marcadores localizados en el cromosoma –Y se convierten día a día en una herramienta de gran valor en diversas situaciones en el campo forense, pueden ser de gran interés para detectar específicamente los restos celulares de un varón en casos en los que se realiza una identificación genética de una mezcla de fluidos biológicos de varón y de mujer en los que la fracción de células procedentes del varón sea minoritaria con respecto a la fracción de células de la mujer; situación que ocurre en una proporción de casos de violación en los que en la toma de fluidos biológicos por las circunstancias

de los hechos o de la toma de muestras, se detecta una pequeña cantidad de espermatozoides del agresor que se encuentran mezclados con una gran cantidad de células de descamación del epitelio vaginal de la víctima.

Otra aplicación importante es la detección de células epiteliales masculinas procedentes de individuos vasectomizados y de permitir una mejor aproximación al número de contribuyentes de semen cuando se trata de un caso de violación con agresores múltiples, obteniendo en este caso una mezcla de haplotipos de los agresores.

Sin embargo, es necesario recordar las limitaciones de los resultados obtenidos a partir del cromosoma $-Y$ como prueba suficiente de inclusión en casos forenses dada la posibilidad que cada haplotipo concreto de polimorfismos del cromosoma $-Y$ sea compartido por todos los miembros varones de una misma familia, los cuales no podrán ser excluidos solo a partir del cromosoma $-Y$.

En general, los marcadores del cromosoma $-Y$ son de gran interés en la investigación biológica de la paternidad de hijos varones, cuando el presunto padre falleció y no se pueda disponer de una muestra biológica del mismo, ya que permite investigar a cualquier pariente del presunto padre y comprobar si su haplotipo de marcadores $-Y$ es coincidente.

El hecho de encontrar que distintas regiones geográficas presentan una distribución de haplogrupos marcadamente diferente y que en algunos se pueda hablar incluso de haplogrupos específicos de población, sirve como herramienta para predecir la población de origen de una muestra de DNA (masculino) desconocido, como ya se ha mencionado antes ([Wetton, Tsang, & Khan, 2005](#)). Ahora bien, acudiendo al análisis de Y-SNPs se puede apoyar en el proceso de identificación de víctimas en desastres a gran escala, donde la información acerca del origen biogeográfico puede ayudar a orientar, facilitar o acelerar las investigaciones ([Brion, Sanchez, et al., 2005](#); [M. A. Jobling, 2001](#)).

También se ha visto que el uso de marcadores de cromosoma $-Y$ presenta ventajas no solamente en el consumo de menor tiempo, sino también en minimizar la pérdida de ADN espermático y reducir el riesgo de contaminación especialmente cuando se trata de muestras escasas ([Hanson & Ballantyne, 2004](#)). Además la investigación del cromosoma $-Y$ en muestras de cadáveres o restos cadavéricos amplía la posibilidad de conocer la identidad y sexo en donde se analizan pequeños fragmentos humanos.

4.1 Evaluación del método en casos de mezcla de perfiles genéticos

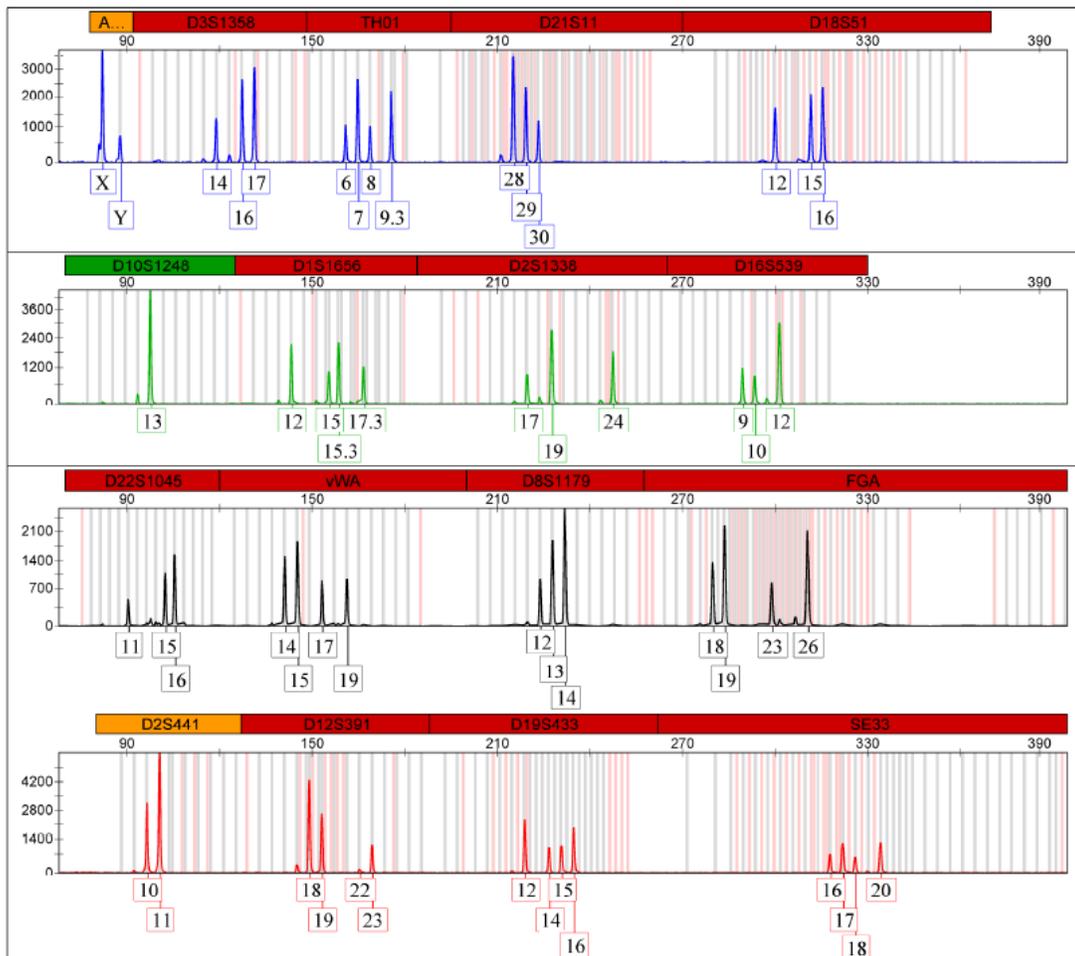
Por lo general, las evidencias obtenidas en casos de asalto sexual tales como frotis vaginales o ropa interior de la víctima, poseen componentes tanto masculino como femenino. Ahora bien, con la intención de aislar el componente masculino, se realiza un procedimiento conocido como “*lisis diferencial*”, pues en teoría, las células espermáticas y las epiteliales femeninas serían completamente separadas y el ADN de cada uno puede ser aislado. Sin embargo, en la práctica, el método lítico usualmente falla, sin alcanzar el 100% de la separación, resultando en la presencia de ADN femenino en cantidades significativas dentro de la fracción de ADN masculino. Cuando la proporción de ADN femenino versus el masculino es 10 veces mayor, el componente femenino sería preferencialmente amplificado y el componente masculino prácticamente sería enmascarado ([Koblinsky et al., 2005](#)).

Por tanto, el empleo de marcadores de cromosoma -Y resulta estratégico en este tipo de casos, la amplificación va dirigida específicamente a marcadores de linaje masculino, de hecho, hay estudios que demuestran la especificidad del análisis incluso cuando la proporción de ADN femenino versus el masculino es 4000 a 1, sin observar que los polimorfismos de cromosoma -Y sean enmascarados, esto en razón a que el ADN femenino carece totalmente de genes -Y. Asumiendo que a pesar de la reducida proporción de ADN, la obtención de producto amplificado proviene del o de los asaltantes. Consecuentemente, este tipo de métodos también resulta útil cuando más de un hombre participa en el asalto sexual, porque al analizar patrón mezclado en la evidencia, se puede deducir la cantidad de individuos participantes.

Según esto, es importante entonces, evaluar la eficiencia del método propuesto respecto al desempeño en casos de mezclas de perfiles genéticos, para ello, se analizó una muestra de uno de los casos trabajados previamente por el Grupo Genética de la DNCTI, correspondiente a una mezcla de ADN masculino y femenino obtenida de un fragmento de tela que hacía parte de la ropa interior de la víctima de un caso de agresión sexual, allegada al Grupo como uno de los Elementos Materiales Probatorios del caso. A pesar de que se realizó la correspondiente extracción diferencial, en este caso se empleó una muestra con extracción directa, la cual, no diferencia la proveniencia del material genético. Posteriormente, se realizó una cuantificación de ADN de la muestra extraída a

través de una PCR en tiempo real. El proceso de cuantificación indicó que había 0,456 ng/ μ L de ADN humano, con un IPC de 28,634, lo que indica que no hay presencia de agentes inhibidores de la reacción de PCR, si este valor es ≥ 30 , se puede interpretar como una presencia de inhibidores. En la figura 4-1 se observa la tipificación de marcadores STRs autosómicos para la muestra cuantificada, donde se puede evidenciar la presencia de una típica mezcla de perfiles genéticos.

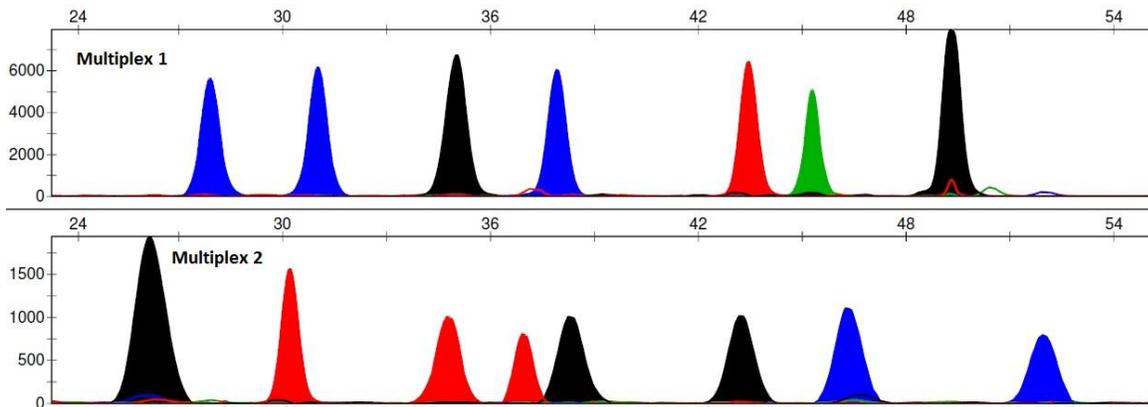
4-1: Electroferograma de la tipificación de STRs autosómicos de la mezcla.



Con todo esto, se procedió a analizar la muestra con el set de Y-SNPs a través de las dos reacciones multiplex, teniendo en cuenta el valor de la cuantificación, el cual incluía la suma de las cantidades de ADN de las dos muestras, por tanto, la cantidad se manejó asumiendo que del valor arrojado por la cuantificación, la mitad corresponde al ADN masculino. Las condiciones del método aplicadas, fueron las mismas empleadas para

muestras de buena calidad, sin hacer alguna modificación en el procedimiento. En la figura 4-2 se puede observar el resultado para las dos reacciones multiplex.

4-2: Electroferograma de los Y-SNPs para la muestra con mezcla de perfiles.



Se obtuvo entonces, para la mezcla de perfiles analizada, un perfil único en las dos reacciones multiplex, atribuida evidentemente a la fracción de ADN masculino presente en dicha mezcla. Se observa además, la ausencia de artefactos o picos inespecíficos que puedan suscitar falsos positivos o interpretaciones erradas. Por consiguiente, se confirma la especificidad del método y la importante aplicación en este tipo de casos, permitiendo dilucidar procesos judiciales complejos con un argumento científico. E incluso, los marcadores del cromosoma -Y también pueden servir como herramienta en la identificación de otro tipo de mezclas como por ejemplo sangre-sangre, sangre-saliva o mezcla de tejidos similares, en las que no se puede aplicar la lisis celular diferencial.

4.2 Evaluación del método en casos negativos y de degradación de ADN

El proceso de degradación del ADN es un elemento permanentemente ineludible en las muestras forenses, si se trata de muestras de escena de crimen, la contaminación es una adversidad que siempre recae sobre este tipo de muestras, pues se deben adoptar la mayor cantidad de medidas posibles para controlarlo. La exposición a radiación ultravioleta, aunque no se crea, es un riesgo para la integridad del ADN, porque se sabe que dicha radiación genera dímeros de Timina y rupturas de cadena simple y doble en el ADN; además, el manejo inadecuado de la muestra al momento de la toma y del transporte, pueden exponer la muestra a daño por radiación ultravioleta directa. La

degradación por agentes abrasivos también es un riesgo inherente a este tipo de muestras, por lo general en la búsqueda por eliminar toda evidencia y borrar todo rastro de cualquier delito, la escena del crimen se somete a limpieza con detergentes, blanqueadores y demás compuestos químicos que diluyen y en muchos casos degradan la molécula de ADN.

Si la evidencia biológica corresponde a un resto óseo o estructura dental producto de una exhumación, la exposición a efectos ambientales, a contaminantes biológicos y a agentes inhibidores de la amplificación de ADN, es inevitable. Se consideran muestras difíciles por la exposición multifactorial, donde se incluyen además las condiciones del terreno, la carga orgánica o la presencia de minerales en el suelo.

Cualquier método que pretenda usar este tipo de muestras como base de su análisis, debe contemplar el control de todas estas variables, además debe expresar la sensibilidad suficiente para amplificar muestras de ADN altamente degradadas. En el mercado comercial existen kits para la extracción y amplificación de ADN en muestras forenses, que intentan controlar estas adversidades, en muchos casos demostrando eficiencia en la obtención de resultados positivos. No obstante, en algunos casos no es posible obtener resultados a partir de este tipo de muestras; a pesar de que se implementan estrategias técnicas adicionales, como el incremento del ciclaje de amplificación, aumento de la cantidad de muestra para la extracción de ADN o el aumento de la concentración del producto de ADN extraído luego de la lisis celular. Aun así, no se consiguen resultados concluyentes, probablemente porque el ADN ya es muy escaso o su integridad está gravemente alterada.

La cuestión aquí, radica en la imposibilidad de identificar la fuente de la evidencia biológica, recordando que en la mayoría de los casos, el análisis genético es la última y/o única posibilidad de alcanzar un resultado en el proceso de identificación de un individuo. Lo cual implica que la persona de quien proviene la muestra analizada, probablemente nunca pueda ser identificada y por tanto, desde la perspectiva social, se puede inferir que toda una familia seguirá padeciendo la abnegada incertidumbre de no lograr la identificación, bien sea de su familiar desaparecido o el autor del delito del cual han sido víctimas, según sea el contexto. Evidentemente, como genetistas forenses, este tipo de casos de ninguna manera son deseables, pero lastimosamente su frecuencia es

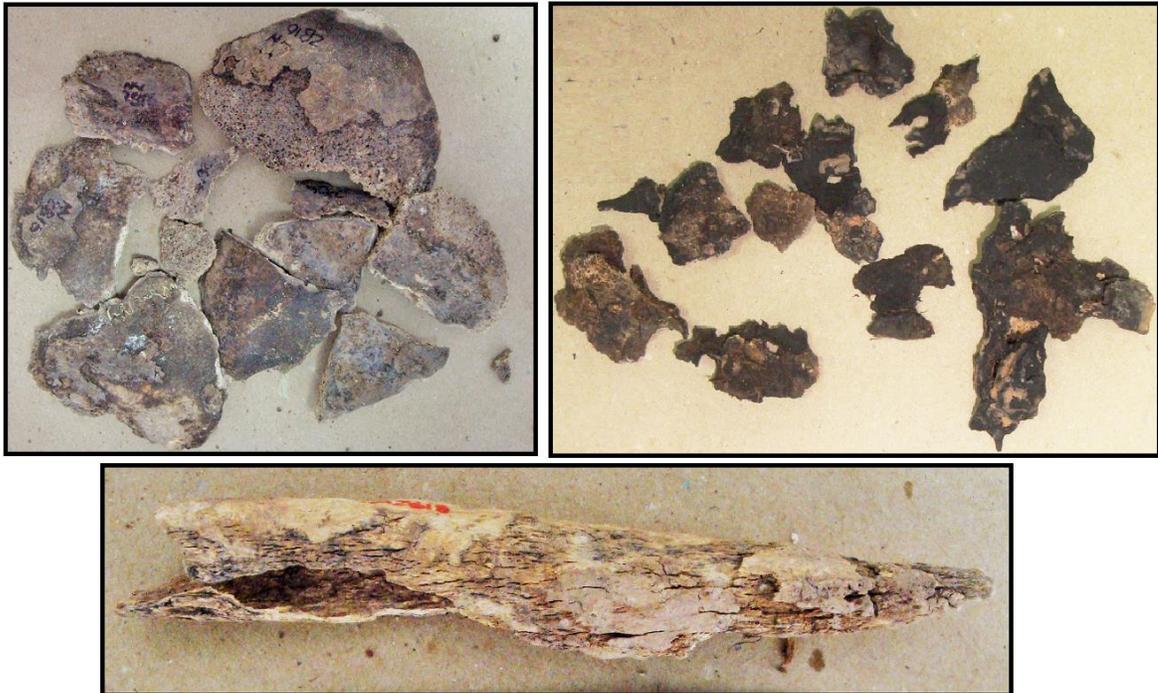
ocasional y no queda más remedio que informar la carencia de resultados concluyentes, a pesar de ser conscientes de las implicaciones negativas que ello conlleva.

Dada esta situación, en este estudio se pretende que el método propuesto logre resultados satisfactorios en este tipo de muestras, pues en otros estudios de referencia intentan evaluar la sensibilidad del método ([Muro et al., 2011](#); [Onofri et al., 2006](#); [Valverde, Kohnemann, Cardoso, Pfeiffer, & de Pancorbo, 2013](#)), lo hacen induciendo una degradación artificial de ADN, intentando simular la baja calidad en el ADN, e incluso comparan la eficiencia del método respecto a la amplificación de kits comerciales de amplificación de STRs, lo cual resulta ilógico, pues el hecho de tener resultados así sean parciales en la amplificación de STRs, significa que existe ADN analizable, y no resultaría un estudio comparativo convincente, pues no es informativo de la capacidad de análisis en muestras degradadas de la técnica.

Por esta razón, en este estudio se planteó la evaluación de la sensibilidad y eficiencia del método, pero en muestras reales, adoptadas a partir de casos negativos del Grupo Genética de la DNCTI, las cuales a pesar de múltiples intentos y la adopción de estrategias técnicas, no se obtuvieron resultados concluyentes y por tanto, la conclusión en cada caso fue negativa. Es importante aclarar que los casos elegidos tienen en común que los resultados con los análisis de marcadores STRs fueron totalmente negativos, sin la amplificación de ningún sistema ni la obtención parcial de ningún alelo. En la figura 4-3 se puede observar el estado de algunas de las muestras analizadas, donde se evidencia la baja calidad en la integridad desde la misma muestra; y como valor agregado se analizó una muestra de un resto óseo incinerado el cual expresa muy mal estado de integridad. Realmente, estas condiciones de análisis resultan desafiantes para el método y vale mencionar que cualquier resultado que se obtenga por más mínimo que sea, es considerado como una ganancia en el proceso de identificación de dichas muestras.

De esta manera, se analizaron los 15 Y-SNPs seleccionados en este estudio, mediante las dos reacciones multiplex, adicionando la cantidad máxima de ADN a cada reacción. Cabe mencionar además, que estas muestras se sometieron a cuantificación por PCR en tiempo real, obteniendo un resultado indeterminado, lo que se interpreta como ausencia de ADN en buen estado. Así mismo, se intentó desarrollar un proceso de validación para este tipo de muestras, teniendo en cuenta que no se disponía de mucho producto de ADN extraído, así y todo, se lograron optimizar algunas variables.

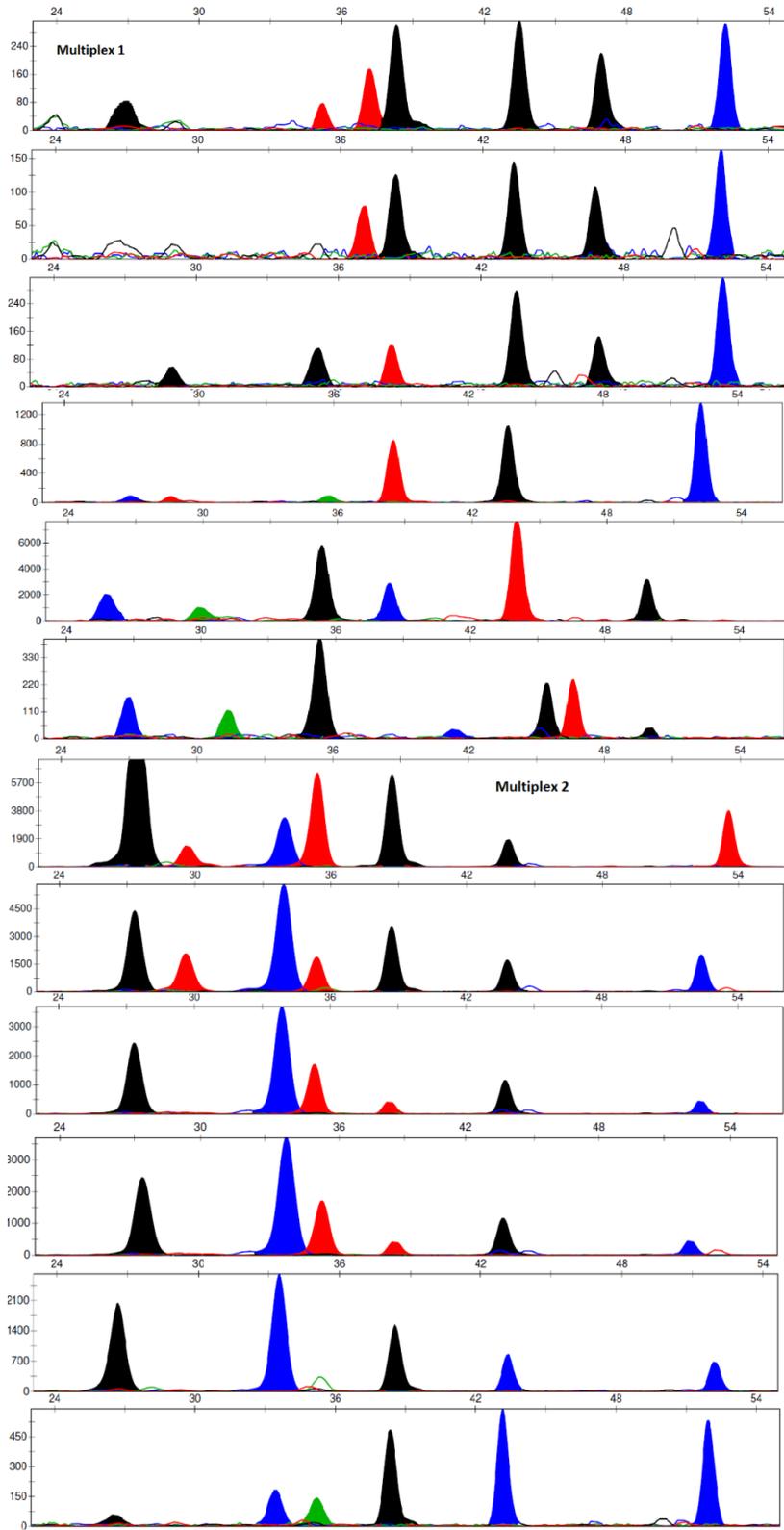
Figura 4-3: Imágenes de algunos de los restos óseos analizados.



Por el estado que presentaban las muestras y según lo que habían demostrado en análisis anteriores, se esperaban resultados particulares para todas las muestras analizadas. En la figura 4-4 se observa que los resultados proveen haplotipos con notable desbalance, así y todo, la amplificación en algunos SNPs es preferencial, debido a la competencia durante la reacción; sin embargo, para los demás casos, se observó una proporción entre los niveles de RFU y los amplicones con tamaño mayor a 150 pb, los cuales fueron más bajos respecto a los demás.

Se observa además para el caso de la multiplex 1, que dos muestras presentan un haplotipo completo. Para el caso de la multiplex 2, a todos los haplotipos les falta un SNP, no obstante, de manera general, se obtuvieron resultados para todas las muestras negativas analizadas.

Figura 4-4: Electroferogramas de las dos reacciones multiplex para las muestras negativas.



Respecto a la fracción metodológica, y después de la respectiva validación, se puede deducir que la amplificación del ADN degradado mejoró cuando el número de ciclos en la PCR multiplex fue incrementado de 28 a 35. El control del desbalance de picos en este caso es muy difícil porque no se puede establecer si la región del ADN objeto de amplificación de cada SNP está en buen estado, por tanto, en aquellos amplicones donde el ADN estaba en mal estado, expresan una amplificación de baja calidad respecto a las regiones de otros SNPs que puedan estar con buena integridad, eso puede explicar la variabilidad en el desbalance.

Según la literatura disponible sobre el tema, es aconsejable hacer un esfuerzo especial en los procesos de purificación, tanto de la multiplex como de la minisequenciación y así obtener claridad en los electroferogramas, e incluso se aconseja incrementar las cantidades de EXOSAP-IT y BigDye Xterminator, así como los tiempos de tratamiento ([Valverde et al., 2013](#)), para controlar el ruido de fondo y los artefactos, aunque este no fue el caso, cabe mencionar que este problema también está relacionado con la pobreza en la calidad del ADN.

Los efectos ambientales sobre las muestras forenses pueden resultar en degradación del ADN o daño en sitios aleatorios, por actividad bien sea de DNAsas o por factores similares. Por consiguiente, la habilidad de la tipificación de Y-SNPs en este estudio, para amplificar ADN degradado no fue ideal, pero se rescata la obtención de resultados mínimos. Con esto se apoyó además, el proceso de evaluación de la sensibilidad del método con muestras que resultaron siendo todo un desafío para el método. Así, los 15 Y-SNPs elegidos en el método pueden resultar valiosos para obtener una idea preliminar de la ancestría biogeográfica de la muestra, en razón de que no se tiene información alguna de cada una de estas.

5. Aportes del estudio de marcadores del Cromosoma –Y en el campo forense

Los marcadores ubicados en la región no recombinante del cromosoma –Y (NRY), la cual escapa de la recombinación y es heredada de generación en generación como un bloque de haplotipo conservado, pueden proveer ventajas importantes en toda investigación forense, casos de mezclas de perfiles y pruebas de paternidad compleja. Por esta razón, se puede relacionar que los Y-SNPs son especialmente útiles en estudios de poblaciones humanas porque las mutaciones son conservadas y transferidas generación tras generación como una huella genética de la historia evolutiva del locus en particular ([Serdar & Sema, 2009](#)).

Todo análisis poblacional debe hacerse teniendo en cuenta aspectos multidisciplinarios, capturando perspectivas culturales, antropológicas y sociales. Esto permite reconocer comportamientos poblacionales diferentes, ejemplo de ello es la ancestría masculina versus la femenina, donde se evidencia el dimorfismo estructural de la población al analizar marcadores uniparentales. Si se tiene en cuenta que en muchas culturas los parientes por vía paterna tienden a vivir en el territorio geográfico y cultural de sus ancestros, de ahí que el análisis del cromosoma –Y ofrece un valioso potencial en el establecimiento de inferencias sobre el origen de una población de un perfil masculino dado.

Para el caso de la población suramericana, que expresa un reconocido mestizaje poblacional y cultural, resulta cautivante para los estudiosos de la genética de poblaciones, establecer análisis multidisciplinarios que comprendan y caractericen de manera meta analítica el comportamiento de la población, pero resulta aún más interesante plantear la valoración del pool genético de una población con estas características. Con todo esto, la estimación genética de este tipo de poblaciones, resulta todo un reto bien sea desde la perspectiva técnica como de la analítica, pues conlleva la

acertada elección de marcadores, el adecuado manejo del muestreo y la población objeto de estudio y la implementación de métodos simples, eficientes e informativos, para cumplir con los objetivos que plantea cualquier estudio poblacional. Aun así, la ancestría masculina de estudios poblaciones en países sudamericanos revela una marcada prevalencia europea, muy superior a la nativo-americana esperada para esta población ([A Blanco-Verea et al., 2006](#); [A. Blanco-Verea et al., 2010](#); [Geppert et al., 2015](#); [Martinez-Cortes et al., 2012](#)). Caso contrario ocurre con el ADN mitocondrial, donde se evidencia un patrón de ancestría predominantemente nativo-americano, lo que es acorde a la teoría del aporte masculino europeo durante la época de la conquista ([Bedoya et al., 2006](#); [Fuselli et al., 2003](#); [Mendizabal et al., 2008](#)), confirmando la divergencia en patrón ancestral de los linajes.

Muchas veces, cuando se dispone de muestras biológicas de origen desconocido o no se dispone de información alguna acerca del donante de la muestra, cualquier indicio, orientación o dato acerca de su procedencia puede ser valioso en cualquier investigación judicial. Determinar y atribuirle un haplogrupo de cromosoma –Y y por consiguiente, asociarle un posible origen biogeográfico a una muestra desconocida, se convierte en un factor relevante en la generación del *a priori* de un caso, pues reduce las posibilidades de imputar la autoría de un delito a un universo de posibles individuos. De manera que, hasta cierto punto es un invaluable aporte para la labor judicial en la aclaración de casos complejos y en la identificación personal, no obstante, dicha asignación del posible origen biogeográfico para una muestra obtenida en una población altamente mestiza, debe hacerse con cautela, pues la frecuencia de ciertos haplogrupos con extensa distribución poblacional, no permite asignar un origen en especial, justificado principalmente por las considerables tasas migratorias de algunos focos poblacionales relacionados a factores sociales como el comercio o la violencia ([Serdar & Sema, 2009](#)). De modo similar, se pueden encontrar casos donde las características biogeográficas del sospechoso no coincidan con las que se espera comúnmente encontrar para el haplogrupo –Y asignado, puede ocurrir eventualmente en individuos mestizos o en grupos poblacionales que en generaciones pasadas emigraron a una región geográfica diferente ([Valverde et al., 2013](#)).

Ahora bien, si resulta arriesgado establecer una afiliación biogeográfica de una muestra de origen poblacional mestizo, se puede acudir entonces a una estrategia más conveniente y acertada, analizando la ancestría del linaje masculino de la muestra, en este caso no se asocia a un origen geográfico específico, sino a un patrón de haplotipo de Y-SNPs asociado a un linaje discriminado en una caracterización genético-poblacional previa. Este aspecto funcionaría equivalentemente al *a priori* del origen geográfico, reduciendo el universo muestral de posibles donantes de la muestra analizada, a aquellos que expresen el mismo linaje masculino para la población. Por ello, es importante caracterizar adecuadamente la distribución poblacional de linajes determinados, mediante marcadores lo suficientemente informativos para discriminar cada sublinaje asociado a un haplogrupo principal y que sirva como referente en cualquier análisis de tipo poblacional que acuda a marcadores uniparentales como los Y-SNPs.

Los STRs son marcadores que por excelencia son comúnmente empleados en el campo forense, adoptados como ensayos de rutina para análisis que busquen establecer identificación o parentesco de una o varias muestras. El surgimiento de nuevos marcadores de ninguna manera pretende delegar o remplazar el estudio de STRs, porque difícilmente podrían igualar la capacidad polimórfica y el poder de discriminación que ostentan, si precisamente esta característica es la principal herramienta usada en el proceso de individualización de una muestra. Lo que se pretende al proponer análisis de nuevos marcadores es complementar el volumen informativo de los métodos de rutina, e incrementar el valor probabilístico de las conclusiones en casos complejos. La intención además, es formular alternativas que expresen sensibilidad, confiabilidad y eficiencia cuando el análisis de STRs resulte negativo, conociendo de antemano que dicho método requiere de ADN en buen estado de integridad, lo cual en muchos casos es un factor limitante. Como valor agregado se citan los casos de mezclas de perfiles masculino y femenino, donde los Y-STRs no alcancen un perfil concluyente, puede acudir al análisis de Y-SNPs para establecer el linaje masculino de la muestra y asociarlo al de los sospechosos. Razón por la cual, se dice que el análisis en conjunto tanto de marcadores Y-STRs y Y-SNPs se convierte en una poderosa estrategia de análisis en cualquier caso de naturaleza forense, siempre y cuando se disponga de la posibilidad de analizarlos dualmente, caso contrario, se acudiría al método que esté a la mano, teniendo en cuenta el valor informativo de cada uno ellos, según sea el caso.

5.1 Proyecciones hacia el futuro

Sin duda alguna, el siguiente paso en el análisis de marcadores moleculares como los SNPs, está dirigido hacia la secuenciación de nueva generación, donde se pueden realizar estudios en masa, con un volumen analítico de miles marcadores simultáneamente, planteando la detección de SNPs específicos con un método directo, e incluso es una buena herramienta para el mapeo de nuevos polimorfismos. Sin embargo, sistemas convencionales de análisis no se descartan del todo, en casos donde se busca realizar la detección específica de un polimorfismo conocido, evitando hacer un análisis masivo innecesario.

Recientemente, se viene fortaleciendo el análisis genético de caracteres visibles externamente conocido como fenotipificación, un método que permite valorar diversos marcadores genéticos asociados a caracteres poligénicos como el color de la piel, del cabello, de los ojos o incluso precisar los rasgos faciales de un individuo. No obstante, grandes controversias éticas han surgido relacionadas con este tipo de análisis, considerando que el uso indebido de esta información puede promover la discriminación contra miembros de grupos poblacionales minoritarios. A este respecto, se debería poner mayor atención al manejo de la información genética, con ello, a las responsabilidades adquiridas por el administrador de las muestras, a los compromisos de confidencialidad de la información y las sanciones de aquellos que desvíen la aplicación de este tipo de información. Con todo esto, se considera importante continuar mejorando esta rama de la genética forense, porque promete llegar a ser una herramienta contundente en la investigación criminal, generando invaluable información del fenotipo completo de un individuo a partir del análisis de una muestra de origen desconocido.

Haciendo un aparte de lo ya mencionado, cabe realizar un análisis desde otra perspectiva, tratando de dejar a un lado la acepción inquisitiva del potencial aplicativo de la técnica, pues evidentemente puede emplearse como herramienta de exoneración en falsas imputaciones. Tanto así, que en Estados Unidos se creó en 1992 el “Proyecto Inocencia”, donde un grupo de abogados se dedicaron a retomar casos condenados que involucraban homicidios y violencia sexual, para así confirmar o desmentir la culpabilidad, apoyándose de la tecnología en el análisis del cromosoma –Y; hasta el año 2006 se habían exonerado 138 personas a partir del análisis de ADN ([Koblinsky et al., 2005](#)). Se

registran incluso casos donde la evidencia de ADN prueba la inocencia de una persona, a pesar de que había confesado la autoría del delito ([Wilgoren, 2002](#)). Definitivamente, este propósito adicional impulsa aún más la aplicabilidad y utilidad para la metodología planteada, ampliando su campo de acción y otorgándole un valor agregado desde la perspectiva social.

Para el caso que aquí compete, sería una opción muy conveniente, plantear una caracterización poblacional más ambiciosa en cuanto al tamaño muestral para la ciudad de Bogotá, contemplando los marcadores aquí propuestos y que demostraron una eficiente discriminación de los linajes masculinos. En consecuencia, se podría hacer una evaluación de la hipótesis muestral que incluye a la ciudad de Bogotá como el foco poblacional representativo del pool genético de todo el país, teniendo en cuenta que la ciudad capital es el destino migratorio de la mayoría de colombianos en búsqueda de oportunidades económicas, académicas, sociales y de desarrollo para mejorar la calidad de vida de las personas. Con esto, se optimizaría el proceso analítico de las bases de datos poblacionales disponibles a nivel nacional y clasificadas por regiones como si existiese un proceso de subestructura poblacional para el país. Y mejor aún, se unificaría una base de datos de referencia a partir del análisis bogotano, adoptado finalmente como representativo de la distribución genética de Colombia, empleándose en los cálculos probabilísticos forenses como un todo, en lugar de hacerlo por sectores, lo que al final podría estar sesgando los análisis. Para ello, se debe hacer un estudio comparativo entre lo encontrado en la población bogotana y lo correspondiente a las regiones del país, aunque pueda sonar algo ambicioso y resultar un tanto complejo, sería una muy buena opción para el desarrollo de la genética forense colombiana.

6. Conclusiones y recomendaciones

6.1 Conclusiones generales del estudio

El método propuesto para el análisis de 15 marcadores tipo Y-SNPs demostró capacidad de determinar eficientemente los haplogrupos informativos de cromosoma –Y, y por consiguiente, se considera una herramienta invaluable para asignar la ascendencia patrilínea biogeográfica de una muestra masculina. El ensayo tiene la ventaja de poder inferir la expresión de los principales macrohaplogrupos de cromosoma –Y, a través del análisis de un par de reacciones multiplex, sin necesidad de acudir a estudios complejos o de mayor labor técnica con resultados similares.

En este sentido, se destaca la acertada elección de marcadores tipo SNP, porque a pesar de la amplia disponibilidad de marcadores en el árbol filogenético del cromosoma –Y, era muy importante realizar una estratégica selección de marcadores, pensando siempre en la capacidad informativa para los intereses forenses, así como en la población objeto de estudio. Esta aplicabilidad se confirma con la uniforme distribución de frecuencias observada en los resultados, lo que implica que los marcadores caracterizaron homogéneamente los linajes masculinos de la población bogotana analizada.

Desde la perspectiva forense, este aporte es importante para conocer el poder discriminatorio de los marcadores elegidos, evitando caer en el error de analizar un marcador muy general que describa la presencia de un haplogrupo muy frecuente, como el caso del R1b y que por consiguiente no demuestre capacidad de diferenciar entre individuos de una misma población.

Mediante este estudio, se confirman las hipótesis de la ascendencia tri-étnica presente en la población colombiana, pues las frecuencias demuestran que los linajes masculinos para

Bogotá son principalmente europeos, africanos y nativos americanos, con una amplia diversidad en la distribución alélica de los haplotipos.

Sin lugar a dudas, la sensibilidad expresada por el método lo convierte en una herramienta valiosa para el análisis de casos complejos en el marco judicial; de esta manera, se acomete uno de los aspectos indeseables y de considerables secuelas negativas en el campo de la genética forense, y es el hecho de no obtener resultados de un análisis como consecuencia de la degradación del material genético. Por ello, siempre serán bienvenidas las propuestas que mediante un fundamento científico, permitan apoyar la investigación judicial en el proceso de orientación e individualización de personas.

No se puede dejar a un lado el valioso aporte social de este proyecto, pues el hecho de permitir retomar casos negativos, con la esperanza de tener un resultado al menos orientador, reabre la posibilidad de aclarar una deuda judicial ante la sociedad, y lo que eso conlleva, como la pérdida de incertidumbre de personas en espera de decisiones judiciales o la posible identificación de un familiar del cual no se sabe nada.

Es importante mencionar que lo propuesto en este estudio puede abrir las puertas para establecer un meta-análisis poblacional de la ancestría masculina colombiana, evaluando la posibilidad de convertir a la ciudad de Bogotá en el referente poblacional del país, sin la necesidad de acudir a análisis específicos para cada región, como actualmente se hace en los cálculos estadísticos de identificación forense.

6.2 Recomendaciones

Se pueden contemplar aspectos a futuro, con el fin de fortalecer la investigación realizada, para ello recomendaría adoptar un par de marcadores SNPs complementarios que promuevan aún más la discriminación del haplogrupo R1b, de amplia prevalencia en la población analizada. La elección de esos nuevos marcadores debe ser acorde a la distribución jerárquica de los SNPs e igualmente informativos, en este caso recomendaría implementar marcadores que correspondan a subclados del linaje R1b1b1a1b-S116, el cual resultó ser el más frecuente en este estudio, y por tanto sería conveniente discriminarlo aún más, para los intereses forenses.

Anexo 1: Encuesta

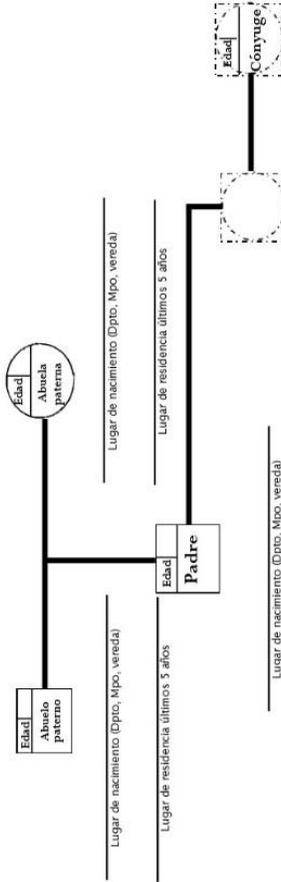
FECHA	A	M	D
Lugar de toma	Lab-Gen	UNJYP	jornada otro
Quien toma la muestra			
Código			

Caracterización genética de la población Bogotana

Encuesta para diligenciar a personas residentes (últimos 5 años) en Bogotá y municipios cercanos



1er Apellido	2do Apellido	Nombres	Documento
Municipio de Residencia	Tiempo de residencia en el lugar		
Municipio de nacimiento	Edad		
Edad del hijo mayor	Lugar de nacimiento del hijo mayor		



Motivo por el cual su familia se radicó en este Departamento	Económico	Violencia	Trabajo	Estudio	Político	Otro
Generación en la que su familia se radicó en este Departamento	ACTUAL		PADRES		ABUELOS	Tiempo aproximado

Observaciones:

Municipios cercanos incluidos en el estudio:
 - Soacha
 - Chia
 - Funza

Anexo 2: Consentimiento informado



FISCALIA
GENERAL DE LA NACION

PROGRAMA DE BANCO DE MUESTRAS DE SANGRE DE FUNCIONARIOS
LABORATORIO DE GENÉTICA FORENSE – FORMATO DE RECOLECCIÓN DE MUESTRAS

DATOS PERSONALES	
APELLIDOS _____	NOMBRES _____ Doc. Identidad _____ N° _____
Lugar de Nacimiento (Municipio): _____ (Dpto): _____	EDAD <input type="text"/> a GÉNERO F <input type="checkbox"/> M <input type="checkbox"/>
Dependencia: _____	Teléfono: _____ Ext: _____
Durante los últimos 3 meses ha recibido transfusión sanguínea:	SI: <input type="checkbox"/> NO: <input type="checkbox"/>
Ha recibido trasplante de médula ósea:	SI: <input type="checkbox"/> NO: <input type="checkbox"/>
INFORMACIÓN FAMILIAR	
MADRE	
APELLIDOS Y NOMBRES: _____	Lugar de Nacimiento (Ciudad, Depto): _____
PADRE	
APELLIDOS Y NOMBRES: _____	Lugar de Nacimiento (Ciudad, Depto): _____
ABUELA MATERNA	
APELLIDOS Y NOMBRES: _____	Lugar de Nacimiento (Ciudad, Depto): _____
ABUELO PATERNO	
APELLIDOS Y NOMBRES: _____	Lugar de Nacimiento (Ciudad, Depto): _____
Con el fin de evitar duplicidad en la base de datos es necesario analizar únicamente 1 individuo de cada familia.	
¿TRABAJAN OTROS FAMILIARES RELACIONADOS POR VIA MATERNA EN LA FGN?	SI <input type="checkbox"/> NO <input type="checkbox"/>

En la ciudad de _____, el _____ de _____ del _____. Firma: _____

CONSENTIMIENTO INFORMADO

Yo, _____ identificado con cédula de ciudadanía No. _____ de _____, autorizo la toma de muestra de sangre, como muestra se preservará y almacenará en el Laboratorio de Genética Forense de F.G.N. respetando mi derecho a la intimidad y no será usada con fines diferentes a los de Identificación, estudios genéticos poblacionales y creación de bases de datos, las cuales serán utilizadas como fundamento estadístico para la interpretación de pruebas realizadas por el Laboratorio de Genética Forense de la F.G.N.

FIRMA
C.C.

TESTIGO
C.C.

Bibliografía

Acosta, M. A., Blanco-Verea, A., Lareu, M. V., Brion, M., & Carracedo, A. (2009). The genetic male component of two South-Western Colombian populations. *Forensic Sci Int Genet*, 3(2), e59-61. doi: 10.1016/j.fsigen.2008.06.006

Adams, S. M., Bosch, E., Balaesque, P. L., Ballereau, S. J., Lee, A. C., Arroyo, E., . . . Jobling, M. A. (2008). The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet*, 83(6), 725-736. doi: 10.1016/j.ajhg.2008.11.007

Albuja, L., & Ceballos, M. (2010). Urban displacement and migration in Colombia. *Forced Migration Review*, 34, 10-12.

Alonso, L. A. (2013). Caracterización de la Población Humana de los Departamentos de Tolima y Huila. Perspectivas: Demográficas, Genéticas y Socioculturales. *Trabajo de Grado para optar al título de Magíster*.

Alonso, L. A., & Usaquén, W. (2013). Y-chromosome and surname analysis of the native islanders of San Andres and Providencia (Colombia). *HOMO-Journal of Comparative Human Biology*, 64(1), 71-84.

Athey, T. W. (2005). Haplogroup prediction from Y-STR values using an allele-frequency approach. *J Genet Geneal*, 1, 1-7.

Bailliet, G., Ramallo, V., Muzzio, M., Santos, M. R., Motti, J., Bianchi, N. O., & Bravi, C. M. (2011). Antecedentes y nuevos aportes en el estudio del Cromosoma Y en poblaciones humanas sudamericanas. *BAG. Journal of basic and applied genetics*, 22(1), 0-0.

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, 16(1), 37-48.

Bao, W., Zhu, S., Pandya, A., Zerjal, T., Xu, J., Shu, Q., . . . Tyler-Smith, C. (2000). MSY2: a slowly evolving minisatellite on the human Y chromosome which provides a useful polymorphic marker in Chinese populations. *Gene*, 244(1-2), 29-33.

Bedoya, G., Montoya, P., García, J., Soto, I., Bourgeois, S., Carvajal, L., . . . Hedrick, P. W. (2006). Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. *Proceedings of the National Academy of Sciences*, 103(19), 7234-7239.

Bertoni, B., Jin, L., Chakraborty, R., & Sans, M. (2005). Directional mating and a rapid male population expansion in a hybrid Uruguayan population. *American Journal of Human Biology*, 17(6), 801-808.

- Bisso-Machado, R., Jota, M. S., Ramallo, V., Paixao-Cortes, V. R., Lacerda, D. R., Salzano, F. M., . . . Bortolini, M. C. (2011). Distribution of Y-chromosome Q lineages in Native Americans. *Am J Hum Biol*, 23(4), 563-566. doi: 10.1002/ajhb.21173
- Bjerke, M., Kephart, D., Knox, C., Stencel, E., Krueger, S., & Lindsay, C. (2006). Forensic application of the Maxwell™ 16 Instrument. *Profiles in DNA*, 9(1), 3-5.
- Blanco-Verea, A., Brión, M., Sanchez-Diz, P., Jaime, J., Lareu, M., & Carracedo, A. (2006). *Analysis of Y chromosome lineages in native South American population*. Paper presented at the International Congress Series.
- Blanco-Verea, A., Jaime, J. C., Brion, M., & Carracedo, A. (2010). Y-chromosome lineages in native South American population. *Forensic Sci Int Genet*, 4(3), 187-193. doi: 10.1016/j.fsigen.2009.08.008
- Bolnick, D. A., Bolnick, D. I., & Smith, D. G. (2006). Asymmetric male and female genetic histories among Native Americans from Eastern North America. *Molecular biology and evolution*, 23(11), 2161-2174.
- Bortolini, M. C., Salzano, F. M., Bau, C. H., Layrisse, Z., Petzl-Erler, M. L., Tsuneto, L. T., . . . Ruiz-Linares, A. (2002). Y-chromosome biallelic polymorphisms and Native American population structure. *Ann Hum Genet*, 66(Pt 4), 255-259. doi: doi:10.1017/S0003480002001148
- Bortolini, M. C., Salzano, F. M., Thomas, M. G., Stuart, S., Nasanen, S. P., Bau, C. H., . . . Ruiz-Linares, A. (2003). Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet*, 73(3), 524-539. doi: 10.1086/377588
- Bosch, E., Calafell, F., Plaza, S., Pérez-Lezaun, A., Comas, D., & Bertranpetit, J. (2003). Genética e historia de las poblaciones del norte de Africa y la península Ibérica. *Investigación y Ciencia*, 63.
- Brion, M., Sanchez, J. J., Balogh, K., Thacker, C., Blanco-Verea, A., Borsting, C., . . . Morling, N. (2005). Introduction of a single nucleotide polymorphism-based "Major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. *Electrophoresis*, 26(23), 4411-4420. doi: 10.1002/elps.200500293
- Brion, M., Sobrino, B., Blanco-Verea, A., Lareu, M. V., & Carracedo, A. (2005). Hierarchical analysis of 30 Y-chromosome SNPs in European populations. *Int J Legal Med*, 119(1), 10-15. doi: 10.1007/s00414-004-0439-2
- Brookes, A. J. (1999). The essence of SNPs. *Gene*, 234(2), 177-186.
- Budowle, B. (2004). SNP typing strategies. *Forensic Sci Int*, 146 Suppl, S139-142.
- Butler, J. M. (2003). Recent Developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism Analysis. *Forensic Science Review*, 15(2), 92-111.
- Capelli, C., Wilson, J. F., Richards, M., Stumpf, M. P., Gratrix, F., Oppenheimer, S., . . . Goldstein, D. B. (2001). A predominantly indigenous paternal heritage for the

Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am J Hum Genet*, 68(2), 432-443. doi: 10.1086/318205

Caratti, S., Gino, S., Torre, C., & Robino, C. (2009). Subtyping of Y-chromosomal haplogroup E-M78 (E1b1b1a) by SNP assay and its forensic application. *Int J Legal Med*, 123(4), 357-360. doi: 10.1007/s00414-009-0350-y

Carracedo, A. (2005). *Forensic DNA typing protocols* (Vol. 297): Springer Science & Business Media.

Carvajal-Carmona, L. G., Soto, I. D., Pineda, N., Ortíz-Barrientos, D., Duque, C., Ospina-Duque, J., . . . Bedoya, G. (2000). Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *The American Journal of Human Genetics*, 67(5), 1287-1295.

Collins-Schramm, H. E., Phillips, C. M., Operario, D. J., Lee, J. S., Weber, J. L., Hanson, R. L., . . . Seldin, M. F. (2002). Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *The American Journal of Human Genetics*, 70(3), 737-750.

Consortium, Y. C. (2002). A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome research*, 12(2), 339-348.

Chakraborty, R., Stivers, D. N., Su, B., Zhong, Y., & Budowle, B. (1999). The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis*, 20(8), 1682-1696. doi: 10.1002/(SICI)1522-2683(19990101)20:8<1682::AID-ELPS1682>3.0.CO;2-Z

DANE. (2003). Evidencia reciente del comportamiento de la migración interna en Colombia a partir de la Encuesta Continua de Hogares. *Dirección de Metodología y Producción Estadística*.

Dugan, K. A., Lawrence, H. S., Hares, D. R., Fisher, C. L., & Budowle, B. (2002). An improved method for post-PCR purification for mtDNA sequence analysis. *J Forensic Sci*, 47(4), 811-818.

Dupuy, B. M., Stenersen, M., Egeland, T., & Olaisen, B. (2004). Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. *Hum Mutat*, 23(2), 117-124. doi: 10.1002/humu.10294

Forster, P., Rohl, A., Lunnemann, P., Brinkmann, C., Zerjal, T., Tyler-Smith, C., & Brinkmann, B. (2000). A short tandem repeat-based phylogeny for the human Y chromosome. *Am J Hum Genet*, 67(1), 182-196. doi: 10.1086/302953

Fuselli, S., Tarazona-Santos, E., Dupanloup, I., Soto, A., Luiselli, D., & Pettener, D. (2003). Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Molecular biology and evolution*, 20(10), 1682-1691.

Genealogy, I. S. o. G. (2014). Y-DNA Haplogroup Tree 2014. *Version: [10.26][Date: [10 April 2015]]*.

- Geppert, M., Ayub, Q., Xue, Y., Santos, S., Ribeiro-dos-Santos, A., Baeta, M., . . . Roewer, L. (2015). Identification of new SNPs in native South American populations by resequencing the Y chromosome. *Forensic Sci Int Genet*, *15*, 111-114. doi: 10.1016/j.fsigen.2014.09.014
- Gonçalves, R., Spinola, H., & Brehm, A. (2007). Y-chromosome lineages in São Tomé e Príncipe islands: Evidence of European influence. *American Journal of Human Biology*, *19*(3), 422-428.
- Gonzalez-Andrade, F., Roewer, L., Willuweit, S., Sanchez, D., & Martinez-Jarreta, B. (2009). Y-STR variation among ethnic groups from Ecuador: Mestizos, Kichwas, Afro-Ecuadorians and Waoranis. *Forensic Sci Int Genet*, *3*(3), e83-91. doi: 10.1016/j.fsigen.2008.08.003
- Gusmao, L., Brion, M., Gonzalez-Neira, A., Lareu, M., & Carracedo, A. (1999). Y chromosome specific polymorphisms in forensic analysis. *Leg Med (Tokyo)*, *1*(2), 55-60.
- Gusmao, L., Butler, J. M., Carracedo, A., Gill, P., Kayser, M., Mayr, W. R., . . . Genetics, D. N. A. C. o. t. I. S. o. F. (2006). DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci Int*, *157*(2-3), 187-197. doi: 10.1016/j.forsciint.2005.04.002
- Gusmao, L., Sanchez-Diz, P., Calafell, F., Martin, P., Alonso, C. A., Alvarez-Fernandez, F., . . . Amorim, A. (2005). Mutation rates at Y chromosome specific microsatellites. *Hum Mutat*, *26*(6), 520-528. doi: 10.1002/humu.20254
- Hammer, M. F., Karafet, T. M., Redd, A. J., Jarjanazi, H., Santachiara-Benerecetti, S., Soodyall, H., & Zegura, S. L. (2001). Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol*, *18*(7), 1189-1203.
- Hanson, E. K., & Ballantyne, J. (2004). A highly discriminating 21 locus Y-STR" megaplex" system designed to augment the minimal haplotype loci for forensic casework. *Journal of forensic sciences*, *49*(1), 40-51.
- IGAC, Instituto Geográfico Agustín Codazzi. (2010). Aspectos históricos y Socioculturales. *Geografía de Colombia*, 39-42.
- Jobling, M. A. (2001). Y-chromosomal SNP haplotype diversity in forensic analysis. *Forensic Sci Int*, *118*(2-3), 158-162.
- Jobling, M. A., Pandya, A., & Tyler-Smith, C. (1997). The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med*, *110*(3), 118-124.
- Jobling, M. A., & Tyler-Smith, C. (2000). New uses for new haplotypes the human Y chromosome, disease and selection. *Trends Genet*, *16*(8), 356-362.
- Jobling, M. A., & Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, *4*(8), 598-612.

- Karafet, T., Xu, L., Du, R., Wang, W., Feng, S., Wells, R. S., . . . Hammer, M. F. (2001). Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet*, 69(3), 615-628. doi: 10.1086/323299
- Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., & Hammer, M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res*, 18(5), 830-838. doi: 10.1101/gr.7172008
- Kidd, J. R., Friedlaender, F., Pakstis, A. J., Furtado, M., Fang, R., Wang, X., . . . Kidd, K. K. (2011). Single nucleotide polymorphisms and haplotypes in Native American populations. *Am J Phys Anthropol*, 146(4), 495-502. doi: 10.1002/ajpa.21560
- Koblinsky, L., Liotti, T. F., & Oeser-Sweat, J. (2005). *DNA: Forensic and legal applications*: New Jersey: Wiley.
- Kovach, W. (1998). MVSP – A Multivariate Statistical Package for Windows, ver. 3.0. *Kovach Computing Services*(Pentraeth,Wales, U. K.).
- Lovo-Gomez, J., Blanco-Verea, A., Lareu, M. V., Brion, M., & Carracedo, A. (2007). The genetic male legacy from El Salvador. *Forensic Sci Int*, 171(2-3), 198-203. doi: 10.1016/j.forsciint.2006.07.005
- Mao, X., Bigham, A. W., Mei, R., Gutierrez, G., Weiss, K. M., Brutsaert, T. D., . . . McKeigue, P. M. (2007). A genomewide admixture mapping panel for Hispanic/Latino populations. *The American Journal of Human Genetics*, 80(6), 1171-1178.
- Martinez-Cortes, G., Salazar-Flores, J., Fernandez-Rodriguez, L. G., Rubi-Castellanos, R., Rodriguez-Loya, C., Velarde-Felix, J. S., . . . Rangel-Villalobos, H. (2012). Admixture and population structure in Mexican-Mestizos based on paternal lineages. *J Hum Genet*, 57(9), 568-574. doi: 10.1038/jhg.2012.67
- Mendizabal, I., Sandoval, K., Berniell-Lee, G., Calafell, F., Salas, A., Martínez-Fuentes, A., & Comas, D. (2008). Genetic origin, admixture, and asymmetry in maternal and paternal human lineages in Cuba. *BMC evolutionary biology*, 8(1), 213.
- Mesa, N. R., Mondragon, M. C., Soto, I. D., Parra, M. V., Duque, C., Ortiz-Barrientos, D., . . . Ruiz-Linares, A. (2000). Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. *Am J Hum Genet*, 67(5), 1277-1286. doi: 10.1016/S0002-9297(07)62955-3
- Muro, T., Iida, R., Fujihara, J., Yasuda, T., Watanabe, Y., Imamura, S., . . . Takeshita, H. (2011). Simultaneous determination of seven informative Y chromosome SNPs to differentiate East Asian, European, and African populations. *Leg Med (Tokyo)*, 13(3), 134-141. doi: 10.1016/j.legalmed.2011.01.001
- Muzzio, M., Muzzio, J. C., Bravi, C. M., & Bailliet, G. (2010). Technical note: A method for assignment of the weight of characters. *Am J Phys Anthropol*, 143(3), 488-492. doi: 10.1002/ajpa.21366

Nunez, C., Geppert, M., Baeta, M., Roewer, L., & Martinez-Jarreta, B. (2012). Y chromosome haplogroup diversity in a Mestizo population of Nicaragua. *Forensic Sci Int Genet*, 6(6), e192-195. doi: 10.1016/j.fsigen.2012.06.011

Nuñez, C., Baeta, M., Sosa, C., Casalod, Y., Ge, J., Budowle, B., & Martínez-Jarreta, B. (2010). Reconstructing the population history of Nicaragua by means of mtDNA, Y-chromosome STRs, and autosomal STR markers. *American journal of physical anthropology*, 143(4), 591-600.

Onofri, V., Alessandrini, F., Turchi, C., Pesaresi, M., Buscemi, L., & Tagliabracci, A. (2006). Development of multiplex PCRs for evolutionary and forensic applications of 37 human Y chromosome SNPs. *Forensic Sci Int*, 157(1), 23-35. doi: 10.1016/j.forsciint.2005.03.014

Palacio Castañeda, G. A. (2008). Urbanismo, naturaleza y territorio en la Bogotá Republicana. 1810-1910.

Palha, T., Gusmao, L., Ribeiro-Rodrigues, E., Guerreiro, J. F., Ribeiro-Dos-Santos, A., & Santos, S. (2012). Disclosing the genetic structure of Brazil through analysis of male lineages with highly discriminating haplotypes. *PLoS One*, 7(7), e40007. doi: 10.1371/journal.pone.0040007

Price, A. L., Patterson, N., Yu, F., Cox, D. R., Waliszewska, A., McDonald, G. J., . . . Bedoya, G. (2007). A genomewide admixture map for Latino populations. *The American Journal of Human Genetics*, 80(6), 1024-1036.

Registro, Nacional, & Desaparecidos, d. (2015). Localización de Información Forense Estadística.

Rodas, C., Gelvez, N., & Keyeux, G. (2003). Mitochondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and Mestizo populations. *Human Biology*, 75(1), 13-30.

Roewer, L. (2009). Y chromosome STR typing in crime casework. *Forensic Sci Med Pathol*, 5(2), 77-84. doi: 10.1007/s12024-009-9089-5

Roewer, L., Kayser, M., Dieltjes, P., Nagy, M., Bakker, E., Krawczak, M., & de Knijff, P. (1996). Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet*, 5(7), 1029-1033.

Sanchez-Diz, P., Acosta, M. A., Fonseca, D., Fernandez, M., Gomez, Y., Jay, M., . . . Restrepo, C. M. (2009). Population data on 15 autosomal STRs in a sample from Colombia. *Forensic Sci Int Genet*, 3(3), e81-82. doi: 10.1016/j.fsigen.2008.08.002

Sanchez, J. J., Brion, M., Parson, W., Blanco-Verea, A. J., Borsting, C., Lareu, M., . . . Carracedo, A. (2004). Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications. *Forensic Sci Int*, 140(2-3), 241-250. doi: 10.1016/j.forsciint.2003.11.035

Sans, M. (2000). Admixture studies in Latin America: from the 20th to the 21st century. *Human Biology*, 155-177.

Seielstad, M. T., Minch, E., & Cavalli-Sforza, L. L. (1998). Genetic evidence for a higher female migration rate in humans. *Nat Genet*, 20(3), 278-280. doi: 10.1038/3088

Seldin, M. F., Tian, C., Shigeta, R., Scherbarth, H. R., Silva, G., Belmont, J. W., . . . Palatnik, S. A. (2007). Argentine population genetic structure: large variance in Amerindian contribution. *American journal of physical anthropology*, 132(3), 455-462.

Semino, O., Magri, C., Benuzzi, G., Lin, A. A., Al-Zahery, N., Battaglia, V., . . . Santachiara-Benerecetti, A. S. (2004). Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet*, 74(5), 1023-1034. doi: 10.1086/386295

Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., . . . Underhill, P. A. (2000). The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science*, 290(5494), 1155-1159.

Serdar, T., & Sema, D. (2009). Y-SNP haplogroups in the Antalya population in Turkish Republic. *Rom J Leg Med*, 17(1), 59-68.

Simms, T. M., Martinez, E., Herrera, K. J., Wright, M. R., Perez, O. A., Hernandez, M., . . . Herrera, R. J. (2011). Paternal lineages signal distinct genetic contributions from British Loyalists and continental Africans among different Bahamian islands. *American journal of physical anthropology*, 146(4), 594-608.

Simms, T. M., Wright, M. R., Hernandez, M., Perez, O. A., Ramirez, E. C., Martinez, E., & Herrera, R. J. (2012). Y-chromosomal diversity in Haiti and Jamaica: Contrasting levels of sex-biased gene flow. *American journal of physical anthropology*, 148(4), 618-631.

Sobrino, B., Brion, M., & Carracedo, A. (2005). SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci Int*, 154(2-3), 181-194. doi: 10.1016/j.forsciint.2004.10.020

Soler, N. L. (2008). De paso por la capital. *Revista Credencial Historia*. Edición No 224.

Soler, N. L. (2001). Algunos Datos Históricos. *Revista Credencial Historia*. Edición No 133.

Su, B., Xiao, J., Underhill, P., Deka, R., Zhang, W., Akey, J., . . . Jin, L. (1999). Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. *Am J Hum Genet*, 65(6), 1718-1724. doi: 10.1086/302680

Suarez-Kurtz, G. (2005). Pharmacogenomics in admixed populations. *Trends in pharmacological sciences*, 26(4), 196-201.

- Swango, K. L., Timken, M. D., Chong, M. D., & Buoncristiani, M. R. (2006). A quantitative PCR assay for the assessment of DNA degradation in forensic samples. *Forensic science international*, 158(1), 14-26.
- Tarazona-Santos, E., Carvalho-Silva, D. R., Pettener, D., Luiselli, D., De Stefano, G. F., Labarga, C. M., . . . Santos, F. R. (2001). Genetic differentiation in South Amerindians is related to environmental and cultural diversity: evidence from the Y chromosome. *The American Journal of Human Genetics*, 68(6), 1485-1496.
- Torres, J. B., Doura, M. B., Keita, S. O., & Kittles, R. A. (2012). Y Chromosome lineages in men of West African descent. *PLoS One*, 7(1), e29687.
- Torres, J. B., Kittles, R. A., & Stone, A. (2007). Mitochondrial and Y Chromosome Diversity in the English-Speaking Caribbean. *Annals of human genetics*, 71(6), 782-790.
- Toscanini, U., Gusmao, L., Berardi, G., Amorim, A., Carracedo, A., Salas, A., & Raimondi, E. (2008). Y chromosome microsatellite genetic variation in two Native American populations from Argentina: population stratification and mutation data. *Forensic Sci Int Genet*, 2(4), 274-280. doi: 10.1016/j.fsigen.2008.03.001
- Underhill, P. A., & Kivisild, T. (2007). Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu. Rev. Genet.*, 41, 539-564.
- Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., . . . Oefner, P. J. (2000). Y chromosome sequence variation and the history of human populations. *Nat Genet*, 26(3), 358-361. doi: 10.1038/81685
- Usaquén, W. (2006). Caracterización Genética de 23 Poblaciones Humanas Colombianas A partir de 9 Sistemas Microsatélites. *Tesis para optar al título de Magister en Ciencias, Universidad Nacional de Colombia.*
- Valverde, L., Kohnemann, S., Cardoso, S., Pfeiffer, H., & de Pancorbo, M. M. (2013). Improving the analysis of Y-SNP haplogroups by a single highly informative 16 SNP multiplex PCR-minisequencing assay. *Electrophoresis*, 34(4), 605-612. doi: 10.1002/elps.201200433
- van Oven, M., Ralf, A., & Kayser, M. (2011). An efficient multiplex genotyping approach for detecting the major worldwide human Y-chromosome haplogroups. *International journal of legal medicine*, 125(6), 879-885.
- van Oven, M., Toscani, K., van den Tempel, N., Ralf, A., & Kayser, M. (2013). Multiplex genotyping assays for fine-resolution subtyping of the major human Y-chromosome haplogroups E, G, I, J, and R in anthropological, genealogical, and forensic investigations. *Electrophoresis*, 34(20-21), 3029-3038. doi: 10.1002/elps.201300210
- Wang, S., Ray, N., Rojas, W., Parra, M. V., Bedoya, G., Gallo, C., . . . Hurtado, A. M. (2008). Geographic patterns of genome admixture in Latin American Mestizos. *PLoS genetics*, 4(3), e1000037.

Wetton, J. H., Tsang, K. W., & Khan, H. (2005). Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs. *Forensic Sci Int*, 152(1), 45-53. doi: 10.1016/j.forsciint.2005.03.009

Wilgoren, J. (2002). Man Freed after DNA Clears Him of Murder. *New York Times*, 10(A), 10.

Wood, E. T., Stover, D. A., Ehret, C., Destro-Bisol, G., Spedini, G., McLeod, H., . . . Soodyall, H. (2005). Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *European Journal of Human Genetics*, 13(7), 867-876.

XTerminator™, B. (2007). BigDye® XTerminator™ Purification Kit Protocol Applied Biosystems