



UNIVERSIDAD NACIONAL DE COLOMBIA

---

# **Análisis Mutacional de Microsatélites Humanos**

**Implicaciones Evolutivas, Poblacionales y Forenses**

**Manuel Paredes López**

Universidad Nacional de Colombia  
Facultad de Ciencias, Departamento de Química  
Bogotá, Colombia

2014

# **Análisis Mutacional de Microsatélites Humanos**

**Implicaciones Evolutivas, Poblacionales y Forenses**

**Manuel Paredes López**

Trabajo de Grado presentado como requisito para optar por el título de Doctor en  
Ciencias - Química

Director

Profesor: Orlando Acosta PhD

Línea de Investigación: Bioquímica

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Química

Bogotá, Colombia

2014

## **Dedicatoria**

A Lilian.....

... por su paciencia, su apoyo .... y su amor

# Tabla de contenido

<b>Resumen</b>	<b>1</b>
<b>Introducción</b>	<b>3</b>
<b>1. Los Microsatélites</b>	<b>5</b>
1.1 Estructura y distribución en el genoma	5
1.2 Clases de secuencias microsatélites	6
1.3 Papel funcional de los microsatélites	7
1.4 Expansión de trinucleótidos.	8
1.5 Variación de STRs en el genoma humano	10
1.6 Homoplasia	11
1.7 Mecanismos mutacionales de los microsatélites	13
1.8 Slippage de la ADN polimerasa	14
1.9 Reparación de la mutación por slippage.	16
1.10 Direccionalidad y dinámica de la mutación por slippage.	18
1.11 El equilibrio entre slippage y mutación puntual.	19
1.12 Efecto <i>allele-span</i> .	19
1.13 Modelos teóricos de mutación	22
1.13.1 Modelo de Alelos infinitos: (IA)	22
1.13.2 Modelo por pasos: <i>Stepwisemutationmodel</i> (SMM)	22
1.13.3 Modelo de dos fases: <i>Twophasemodel</i> (TP)	23
1.13.4 Modelo de K-alelos: (KA)	23
1.14 Edad, sexo y mutación)	23
1.15 Mutación somática vs germinal	26
1.16 Herramientas de análisis para el estudio de microsatélites	28
1.17 Microsatélites e identificación humana	31
1.17.1 Mutación y casos forenses	33
1.18 Caracterización de los loci microsatélites analizados	36
1.18.1 Kit AmpFISTR Identifier	36
1.18.2 Kit PowerPlex 16 HS	38
1.18.3 Alelos STR amplificados vía PCR	42

<b>2. Objetivos</b>	44
2.1 Objetivo General	44
2.2 Objetivos Específicos	44
<b>3. Metodología</b>	45
3.1 Universo:	45
3.2 Muestra para el cálculo de tasa mutacionales:	45
3.3 Conteo de meiosis	46
3.4 Criterios de exclusión	48
3.5 Muestras biológicas	48
3.6 Muestra poblacional analizada	49
3.7 Análisis de laboratorio	50
3.7.1 Métodos de extracción de ADN	50
3.7.2 Amplificación vía PCR-multiplex	51
3.7.3 Detección y asignación de alelos STR	51
3.7.4 Control de calidad	52
3.7.5 Criterios de análisis	55
3.7.5.1 Definición del tipo de mutación	55
3.7.5.2 Determinación del origen parental de la mutación	56
3.7.5.3 Determinación de la dinámica mutacional	56
3.7.6 Frecuencia de mutación	56
3.7.7 Tasa de mutación	57
3.7.8 Análisis estadísticos	57
3.7.9 Consentimiento informado	57
<b>4. Resultados y Análisis</b>	58
4.1 Distribución geográfica de mutaciones en Colombia	58
4.2 Origen parental de los eventos mutacionales	63
4.3 Edad y Mutación	64
4.4 Mecanismo mutacional: <i>Slippage vs null</i>	70
4.5 Dinámica mutacional	74
4.5.1 Relación ganancia vs pérdida de repeticiones	74
4.5.2 Direccionalidad	76

4.6	Tasas mutacionales	82
4.6.1	Tasas mutacionales para 15 STRs en Colombia	82
4.6.2	Tasas mutacionales según la secuencia de repetición	84
4.6.3	Tasas mutacionales según la longitud de la secuencia	86
4.6.4	Tasas mutacionales según el origen parental	87
4.6.5	Efecto de la tasa de mutación en el cálculo del IP	91
4.7	Frecuencias de mutación	93
<b>5.</b>	<b>Conclusiones</b>	97
<b>6.</b>	<b>Referencias</b>	101
	<b>Agradecimientos</b>	107
	<b>ANEXO 1: Secuencias de los microsatélites analizados</b>	108

## INDICE DE TABLAS

Tabla 1.	Localización cromosómica de los 16 loci del kit Identifiler	39
Tabla 2.	Localización cromosómica de los 15 loci del kit Powerplex16HS	40
Tabla 3.	Diferencia en la longitud de alelos intermedios en CSF	43
Tabla 4.	Número de mutaciones detectadas por Departamento	59
Tabla 5.	Distribución de 803 eventos mutacionales por región	61
Tabla 6.	Edad de la madre al primer nacimiento. Colombia	66
Tabla 7.	Eventos mutacionales detectados en el locus D19S433	73
Tabla 8.	Estado de pérdida o ganancia vs número de pasos	75
Tabla 9.	Relación de casos analizados y conteo de meiosis	83
Tabla 10.	Tasas mutacionales por locus	84
Tabla 11.	Agrupación de los 15 STR según su tasa mutacional	84
Tabla 12.	Tasas mutacionales según la secuencia y media de longitud	87
Tabla 13.	Tasas mutacionales para 15 loci STRs según origen	88
Tabla 14.	Tasas mutacionales de 15 loci STR en 5 poblaciones	91
Tabla 15.	Frecuencias de mutación en 14 STR en Colombia	97

## INDICE DE FIGURAS

Figura.1. Estructura básica de un microsatélite simple	5
Figura 2. <i>Crossing-over</i> desigual	14
Figura 3. <i>Frameshift mutations</i>	15
Figura 4. Replicación en slippage	17
Figura 5. Magnitud y dirección de los eventos mutacionales	20
Figura 6. Divisiones celulares en la ovogénesis y la espermatogénesis.	25
Figura 7. Estudio de paternidad	29
Figura. 8a. Plásmido FxC	30
Figura. 8b. Fragmentos obtenidos al amplificar proteína SSB bacteriana	30
Figura 9. Sobrelapamiento de loci kit Identifiler®.	37
Figura 10. Electroferograma línea celular 9947 <sup>a</sup> kit Identifiler®.	38
Figura 11. Electroferograma Kit PowerPlex16HS.	41
Figura 12. Sobrelapamiento de loci kit Powerplex16HS®.	41
Figura 13. Representación de la estructura general de un alelo STR	42
Figura 14 Trio simple	46
Figura 15 Varios hijos	47
Figura 16 Varios Hijos	47
Figura 17 Un progenitor	47
Figura 18. Esquema de la metodología utilizada	52

## INDICE DE GRAFICOS

Gráfico 1. Papel funcional de los STRs.	11
Gráfico 2. Distribución geográfica de 803 eventos	59
Gráfico 3. Distribución geográfica de 803 eventos en Colombia	62
Gráfico 4. Origen parental de los eventos mutacionales	63
Gráfico 5. Distribución de edad de los parentales	65
Gráfico 6. Tasas de mutación por rango de edad y origen parental	67
Gráfico 7. Distribución de eventos mutacionales según origen	70
Gráfico 8. Mecanismo mutacional tipo null	71
Gráfico 9. Eventos de Ganancia o Pérdida de repeticiones	74
Gráfico 10. Eventos mutacionales por pasos	76
Gráfico 11. Locus FGA Dinámica mutacional	78
Gráfico 12. Expansiones y contracciones locus: vWA31, D5S818, D7S820 y D13S317	78
Gráfico 13. Expansiones y contracciones locus: D8S1179, CSF1P0, D16S539	79
Gráfico 14. Expansiones y contracciones locus: D2S1338, D3S1358 y D18S51	79
Gráfico 15. Expansiones y contracciones locus: D19S433 y D21S11	80
Gráfico 16. Distribuciones unimodales locus: CSF1P0, vWA31 y D5S818	80
Gráfico 17. Distribuciones unimodales locus: D16S539, D13S317, D7S820 y D3S1358	81
Gráfico 18. Distribuciones bimodales locus: D19S433 y D21S11, FGA y D8S1179	81
Gráfico 19. Distribución multimodal locus: D18S51 y D2S1338	82
Gráfico 20. Tasa mutacional para 15 STRs	85
Gráfico 21. Tasa mutacional según la secuencia de repetición del STR	86
Gráfico 22. Tasa mutacional de 15 STRs en 5 poblaciones	92



## RESUMEN

Los laboratorios de genética forense, analizan actualmente polimorfismos de ADN del tipo Microsatélites (STR), para identificación humana. Si bien estos marcadores constituyen hoy las herramientas universales para los fines forenses, por su alta reproducibilidad, estabilidad y capacidad de automatización, presentan también una de las más altas tasas de mutación. Lo anterior genera la paradoja de contar por una parte, con sistemas de alta variabilidad poblacional y por lo tanto alto poder de identificación, pero al mismo tiempo, con la posibilidad de generar cambios inesperados entre los padres verdaderos y sus hijos. Hemos analizado 50.827 casos de investigación de la paternidad o la maternidad estudiados entre 2007 y 2012 en el Laboratorio de Genética Forense del Instituto Nacional de Medicina Legal de Colombia, caracterizando 811 eventos de mutación, una de las muestras más grandes hasta ahora estudiadas. Se presentan las tasas mutacionales de 15 loci STR de amplio uso forense y se hacen inferencias sobre la dinámica mutacional a partir de la información directa obtenida de transmisiones alélicas padre-hijo o madre-hijo.

**Palabras clave:**

Análisis Mutacional, Microsatélites, Dinámica Mutacional

## **ABSTRACT**

### **Mutational analysis of human microsatellites; Evolutionary, Populational and Forensic implications**

Forensic genetics laboratories currently analyze DNA polymorphisms like Microsatellite (STR) for human identification. For their high reproducibility and stability these markers are universally recognized as the most useful tool on the analysis of forensic evidences. However, the STRs also have one of the highest mutation rates. This fact, creates the paradox of having on the one hand, genetic systems of high population variability and high power of identification, but on the other hand, with the possibility of generate unexpected findings between biological parents and their children. In this work we analyzed 50,827 cases of paternity or maternity investigations, observed between 2007 and 2012 in the Laboratory of Forensic Genetics, in National Institute of Legal Medicine of Colombia, 811 mutational events were characterized, one of the largest samples studied so far. Mutational rates of 15 STR loci for forensic use are showed and some inferences on mutational dynamics are presented, based on information from direct allelic transmissions parent-child or mother-child.

#### **Key words**

Mutational analysis, Microsatellites, Mutation dynamics

## Introducción

En los últimos 25 años, los laboratorios de genética forense, han generado un gran cuerpo de datos genéticos a partir del análisis de polimorfismos de ADN utilizados en la identificación humana, a raíz de la información directamente recopilada de los casos forenses, y adicionalmente, por los datos provenientes de estudios de poblaciones de referencia forense requeridos para la valoración bioestadística de los resultados periciales. Este amplio volumen de información, permite detectar y hacer accesible al análisis, eventos de rara ocurrencia, como son las mutaciones genéticas y facilita la caracterización de la dinámica mutacional.

Entre los marcadores nucleares autosómicos, de herencia biparental, multi-alélicos o bi-alélicos, o los sistemas asociados a cromosomas sexuales bi o uni-parentales, hasta el ADN mitocondrial, son los Microsatélites autosómicos, los polimorfismos que han alcanzado la mayor difusión en el mundo forense, actualmente su estudio es universal, los mismos loci genéticos se analizan en todos los laboratorios forenses permitiendo la comparación de resultados y las contrapericias. Por la misma razón existe a partir de estos datos, varios estudios mutacionales sobre microsatélites, que han aportado evidencia sobre los mecanismos subyacentes al cambio evolutivo. Aun así, el estudio de la mutación es complejo porque los eventos mutacionales son de baja frecuencia en la naturaleza, lo Cuál limita su análisis. Para hacer inferencias biológicas validas, se requiere:

1. Contar con un gran número de transferencias meióticas fácilmente observables, idealmente a través de estudios directos entre padres e hijos.
2. Seleccionar un marcador genético de alta tasa mutacional que permita la observación de un número importante de eventos de los que puedan inferirse con mayor certeza patrones de conducta molecular.

Colombia, quizás por su compleja dinámica social, por su historia de conflictos no resueltos y no concluidos, por su confrontación y su actual polarización sociopolítica, genera una alta demanda de servicios forenses, donde las pruebas genéticas son

requeridas a diario por las autoridades judiciales; Irónicamente, esta coyuntura es también una importante fuente de conocimientos sobre el genoma humano, que debe ser administrado, clasificado y analizado, con el fin de aportar no solo al conocimiento fundamental de estas secuencias de ADN y su dinámica de cambio molecular, sino además, para resolver problemas concretos y cotidianos de la práctica pericial, que permitan perfeccionar el ejercicio forense en bien de las víctimas, en medio del conflicto o aun en la esperada era del postconflicto.

Los escenarios forenses donde se requiere valorar los eventos mutacionales son dos:

1. La mutación en línea germinal, detectable en la investigación del parentesco, a cuyo estudio se dedica este trabajo.
2. La mutación somática, de rara ocurrencia en la casuística forense

# 1. Los Microsatélites

## 1.1 Estructura y distribución en el genoma

Los microsatélites también conocidos como repeticiones cortas en tándem (STRs, *Short tandem repeats*)(1, 2, 3), o secuencias repetidas simples (SSR)(4), son secuencias de ADN que contienen bloques de repeticiones de 6 nucleótidos o menos, dispuestas sin interrupción un número limitado de veces (**Figura1**). Aunque se localizan a lo largo de todo el genoma, son particularmente abundantes en regiones centroméricas o en porciones heterocromáticas de los cromosomas y su densidad promedio por genoma es de un locus STR cada  $10^4$  o  $10^5$  bases; así, las secuencias microsatélites constituyen aproximadamente el 3% del genoma humano(5).

Se encuentran en todos los genomas conocidos, aunque son más frecuentes en eucariontes, la secuenciación de genomas completos en procariontes ha demostrado la presencia de microsatélites en un amplio número de especies bacterianas, esta información ha servido como herramienta de clasificación e identificación de cepas. Aunque no es claro el papel que cumplen en estos genomas, la expansión o contracción de microsatélites en bacterias se ha asociado a variaciones fenotípicas como el nivel de adherencia a la célula huésped y su invasividad entre otros (6). Finalmente, se han detectado microsatélites aun en genomas mitocondriales (7) y en cloroplastos (8).



**Figura.1.** Estructura básica de un microsatélite simple del tipo tetranucleótido

Su arreglo estructural y su frecuencia varían notablemente entre diferentes especies: Existen trectos homopoliméricos simples de purinas o pirimidinas, que son altamente

frecuentes en muchos organismos, poli (A) / poli(T) por ejemplo, son las secuencias repetidas en tándem más comunes en el genoma de primates (9). Aun así, la mayoría de los STRs encontrados en muchas especies son dinucleótidos (10), particularmente los arreglos tipo CA/TG. Tractos poli (CA) se encuentran en el 70% de los clones de librerías genómicas de ratón. En general, el motivo CA es el STR más común en animales, mientras que el poli (AT) es el más frecuentemente observado en plantas (11). Por su parte, en los STRs trinucleotídicos, las unidades de repetición CAG y AAT son las más frecuentes (12). Entre los tetranucleotídicos, los motivos ricos en A/T son los más abundantes (5).

## 1.2 Clases de secuencias microsatélites

Los microsatélites se clasifican según el tipo de arreglo de sus secuencias repetidas, generalmente se reconocen STRs de estructura simple porque se componen de una misma secuencia de repetición, independientemente del número de bases que la conformen (Ej. ATCATCATCATCATCATCATC). Otros, se consideran compuestos y están conformados por más de un tipo de repetición en tándem; en estos casos, generalmente las repeticiones son muy similares en tamaño y secuencia y varían solo en un nucleótido, lo cual hace pensar que tienen un origen común (Ej. CAATCAATCAATCAAGCAAGCAAG). Finalmente, un STR complejo presenta más de un tipo de repetición pero estas pueden ser de diferente tamaño y diferente secuencia. (Ej. CAATCAATCAATCAAGCAAGCAAGGCACACACAATCA) (13).

Otra forma de clasificación es aquella que considera secuencias perfectas, imperfectas, interrumpidas o compuestas: En un microsatélite perfecto la secuencia de repetición no está interrumpida por ninguna base que no pertenezca al motivo (Ej. ATATATATATATATA); siendo equivalente a las formas simples; en un microsatélite imperfecto, hay un par de bases entre los motivos repetidos que no concuerdan con la secuencia motivo (Ej. ATATATATATCATAT). En un microsatélite interrumpido hay una pequeña secuencia dentro de la secuencia repetitiva que no coincide con la secuencia motivo (Ej. ATATATATGCTCATATAT), mientras que en un compuesto, la

secuencia contiene dos secuencias de repeticiones distintas y adyacentes (Ej. ATATATATGTGTGTGT). (14)

Por otra parte, la longitud de alelos microsatélite parece ser dependiente de la localización genómica. En general, alelos ubicados en regiones no codificantes son más largos que los que se encuentran en regiones codificantes. Esto puede atribuirse a la selección negativa contra mutaciones que alteran el marco de lectura en regiones codificadoras. Un caso especial es el de los microsatélites trinucleotídicos, que en el humano, son muy frecuentemente incorporados a genes asociados a un grupo de enfermedades neurodegenerativas, donde la gravedad del cuadro clínico se relaciona con grandes expansiones en el número de los trinucleótidos. Al parecer, la selección limita la expansión de secuencias no – tripletes de microsatélites en secuencias codificantes. (16). Variantes de alelos microsatélites se han implicado en más de 40 enfermedades neurodegenerativas en el humano (17).

### **1.3 Papel funcional de los microsatélites**

En principio los STRs fueron considerados como ADN no codificante y por lo tanto, evolutivamente neutros, sin embargo, pueden ser intragénicos, e incluso intra-exónicos y se han asociado a procesos de regulación de la expresión, razón por la cual, no es muy exacto desligarlos de la función génica. De hecho, el 17% de los genes humanos poseen microsatélites en su marco abierto de lectura. Aun así, constituyen una gran fracción del ADN no codificante y son relativamente raros en regiones codificadoras de proteínas(10). De nuevo la genómica aporta datos clarificadores: Huang y cols., estudiaron 101 STRs en 54 especies de plantas y todos mapearon en regiones no expresivas. Situaciones similares se han observado en levaduras, anélidos, ratón, insectos y primates. (14).

Los STRs se han asociado a muchos niveles funcionales: Ensayos de hibridación, han demostrado la presencia de microsatélites ocupando las mismas posiciones cromosómicas entre especies diferentes, independientemente del motivo

microsatélite estudiado, lo Cuál sugiere un papel especial de los STRs en la organización del cromosoma; Se ha observado que generalmente, las regiones centroméricas de los cromosomas están compuestas de numerosas repeticiones en tándem lo Cuál afecta la organización del centrómero. Una estructura centromérica única, que contenía una familia divergente de repeticiones centrómero-específicas fue detectada en *Neurospora crassa* y es similar a la observada en *Drosophila* (18). El ADN repetitivo que flanquea el centrómero parece que interviene en la formación del cinetocoro y en la cohesión de las cromátides hermanas.

#### **1.4 Expansión de trinucleótidos.**

Los microsatélites más comúnmente observados en regiones codificadoras son trinucleótidos, esta predominancia puede explicarse sobre la base de la supresión de STR's no triméricos en regiones codificadoras, posiblemente debido a un cambio en el marco de lectura con el incremento o disminución en el número de repeticiones.

Más de una veintena de enfermedades neurodegenerativas se han asociado a la expansión de trinucleótidos (17). A diferencia de cómo se comportan otras secuencias STR, los trinucleótidos muestra una tendencia notable a la expansión hasta cientos de unidades repetitivas más que a la pérdida de unidades de repetición. Aun así, solo aquellos loci que forman estructuras secundarias (hairpins) estables se expanden. Las repeticiones (CAG/GTC)<sub>n</sub> y (CGG/GCC)<sub>n</sub> forman las estructuras secundarias más estables.

El modelo propuesto para la expansión de trinucleótidos es similar al slippage con la diferencia que no se forman loops de una sola repetición; en su lugar pueden formarse estructuras secundarias que contienen una gran cantidad de trinucleótidos. Por otra parte, las secuencias repetitivas generan estructuras secundarias inusuales, con patrones de asas plegadas complejas al parecer con efectos funcionales. Un ejemplo de esto son los "bucles" o *loops* de repeticiones (CCG)<sub>n</sub> observados en el síndrome de X-Frágil, que se forman en secuencias reguladoras del gen. Estas estructuras al parecer, tienen importantes efectos regulatorios de la expresión génica,



son favorables para la transcripción, ya que favorecen el reconocimiento de proteínas a motivos específicos en el ADN. (19). Por otra parte, expansiones de CAG dentro del marco de lectura, traducen a poli-glutaminas neurotóxicas, Se forman cuerpo de inclusión nucleares que contienen agregados de proteína con glutaminas expandidas, que generan enfermedades neurodegenerativas como la Enfermedad de Huntington (20).

En otro escenario, muchos STRs y Minisatélites se han propuesto como “sitios calientes” para recombinación (21). Los STRsdinucleotídicos son sitios preferenciales para recombinación, gracias a su alta afinidad con enzimas recombinasas. Al parecer, secuencias repetidas como GT, CA, CT, GA. GC o AT inducen conformaciones ADN-Z que facilitan el proceso de recombinación. También el número de unidades de repetición afecta la recombinación.

Los STRs pueden afectar también enzimas que controlan el ciclo celular. Por ejemplo, el gen humano CHK1 controla la progresión del ciclo celular, y su región codificadora contiene un tracto de (A)<sub>9</sub> que es un sitio potencial de mutaciones en tumores con inestabilidad de microsatélites. Alteraciones en el gen CHK1 en cáncer de colon y útero, fueron asociados a la presencia de alto grado de inestabilidad de trectos poli(A). También hay evidencia de interacciones entre reparación de ADN y puntos de verificación del ciclo celular: El sistema de reparación de malos apareamientos (*Mismatchrepair*, MMR) interactúa con el punto de verificación de células en G<sub>2</sub>, en respuesta a lesiones de ADN inducidas por (TG)<sub>6</sub> o NmetilNnitrosoguanidina. (22). Se han encontrado STRs (A)<sub>n</sub> en la región codificadora de genes MMR como MSH3, MSH6 y otros en eucariontes incluyendo al hombre (23).

Se han asociado microsatélites con mecanismos de regulación de la transcripción, por ejemplo motivos (CT)<sub>n</sub> en la región 5´UTR de algunos genes de *Arabidopsis* juegan un papel importante en la transcripción anti-sentido (14).

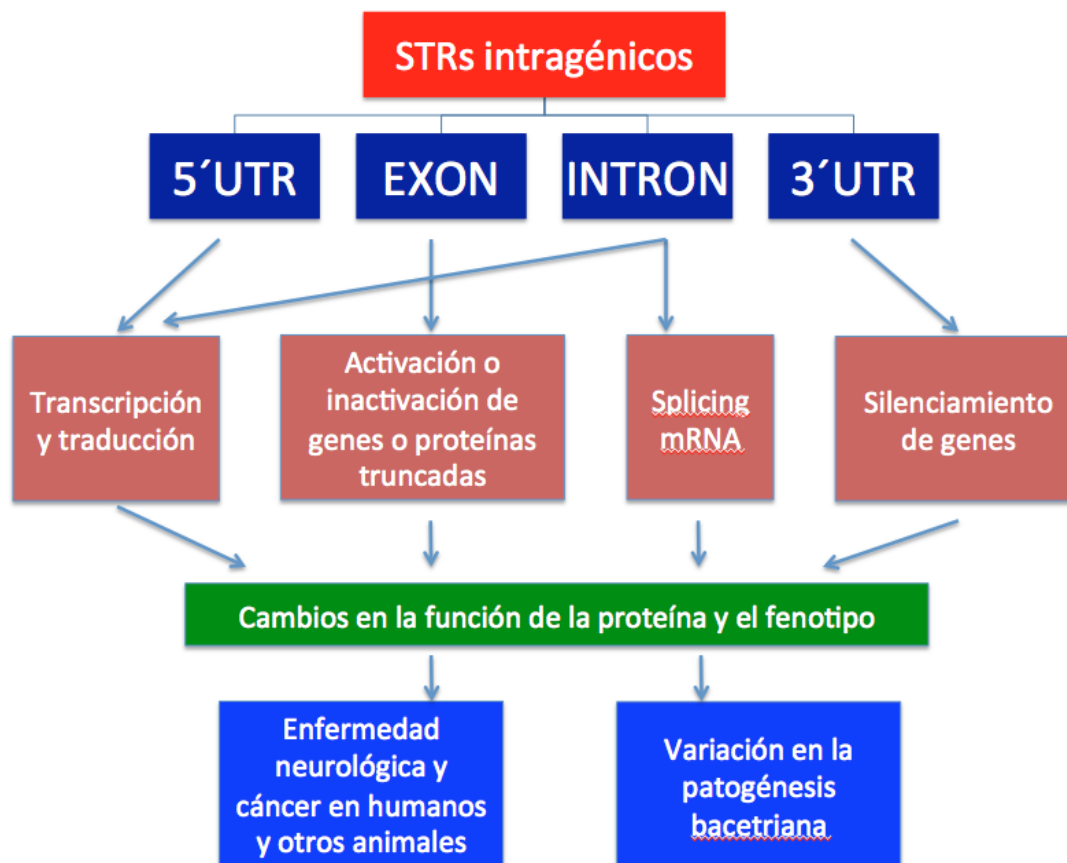
Hay evidencia de microsatélites asociados a patogenicidad y variabilidad genómica y a modulación de la expresión génica en microorganismos; motivos tetranucleotídicos están presentes en los ORFs de genes codificadores de proteínas que afectan lipopolisacáridos del *Haemophilus influenza*. La variación en el número de estos STRs afecta la cantidad de proteína expresada. Finalmente, secuencias repetitivas similares a STRs se han detectado en varios genes de virulencia de organismos patógenos(14).

En general, el papel funcional de los microsatélites en la expresión génica puede explicarse según la región del gen comprometida : La abundancia de STRs en la región 3'UTR y su expansión, comprometen la regulación de la expresión del gen. A nivel intrónico, afectando el mecanismo de *splicing* y por consiguiente la función del producto proteico, A nivel post-transcripcional, sobre los RNAs, que contienen expansiones como CAG, se ha planteado que pueden interferir con el procesamiento del transcrito primario, afectando la cantidad final de la proteína. Finalmente, la afectación directa de zonas expresivas por expansión de *repeats* dentro del marco de lectura, que generan poli-aminoácidos. Por otra parte, los defectos o deficiencias en los mecanismos reparadores tipo MMR hacen más propenso la secuencia repetida a la mutación como se observa en varios tumores. (24) (**Gráfico 1**)

### **1.5 Variación de STRs en el genoma humano**

Payser, B et al (25), compararon dos genomas humanos completos identificando 2.862.022 STRs perfectos con al menos 3 repeticiones. El 2,7% (78.429) fueron diferentes en las 2 secuencias. Variaron además en el número de repeticiones: STRs cortos fueron más comunes y mostraron niveles bajos de polimorfismo; los STRs largos fueron menos comunes pero mucho más polimórficos. Estimaron el efecto de la longitud del *repeat* y su secuencia, en la variación del STR; los tetra-nucleótidos fueron los más variables en el tamaño alélico (4,5% eran polimórficos). Los trinucleótidos fueron los menos variables; di y pentanucleótidos, tenían niveles de polimorfismo muy similar. Por otra parte los motivos de secuencia de mayor variación fueron: AC(8,5% polimórficas) AAC (2%), TCTA (23,3%), CTATT (26,9%) y TATCTA

(32,1%). Los tetra, penta y hexanucleótidos no mostraron variación por secuencia. Finalmente, los loci más variables tienden a tener repeticiones con más bajos contenidos de GC, principalmente en trinucleótidos.



**Gráfico 1.** Papel funcional de los STRs. (modificado de Li , 2004)

## 1.6 Homoplasia

La tecnología utilizada para documentar la variación es también determinante en la medida del polimorfismo, los estudios de secuenciación en diferentes niveles, permiten detectar la variación intra-alélica independientemente del tamaño del fragmento repetitivo, pero el análisis de STRs por fragmentos, así se realice en secuenciadores automatizados que permiten la asignación precisa del tamaño alélico, constituye en algunos casos un problema, si se tiene en cuenta que la

variación intra-alélica no puede ser detectada. Dos alelos iguales en tamaño pueden tener entonces secuencias variantes.

A nivel poblacional, se presenta un efecto similar, la homoplasia. Esta vez, por la imposibilidad de determinar el origen evolutivo de un alelo STR; Dos alelos diferentes en tamaño y origen, mutan hacia una misma forma alélica (idénticos por estado), para los evolucionistas, por convergencia, paralelismo o reversión. Esto genera un posible sesgo en los estudios poblacionales y obligaría a secuenciar los alelos sospechosos de ser homoplásicos, si se quiere inventariar la diversidad real de una población. (26). Un alelo puede convertirse en otro ya presente en la población, con lo cual, en teoría se podría complicar la interpretación histórica de la variación.

La homoplasia viene a ser más pronunciada cuando se comparan grupos o individuos genealógicamente y geográficamente distantes. La variación alélica detectada en el presente estudio se evaluó a partir del análisis de fragmentos STRs. Ocasionalmente se secuenciaron alelos, para definir la región del mismo que cambió entre un parental y su hijo, pero este tipo de estudio no permite evaluar la variación de secuencia entre alelos.

Aun así, Estoup y Cols. Evaluaron el efecto poblacional de la homoplasia de tamaño en microsatélites y concluyen que no representa un problema significativo en el análisis poblacional ya que la gran cantidad de variabilidad de los loci microsatélites, compensa el ruido que generaría la homoplasia en su evolución (42).

Los microsatélites son altamente mutables ya sea hacia la ganancia o la pérdida de repeticiones. La tasa de mutación se ha estimado en general en,  $10^{-3}$  a  $10^{-4}$  eventos por locus por generación, mucho mayor que las tasa de sustitución de nucleótidos simples, y varía entre loci, entre alelos y aun entre especies y depende parcialmente de su estructura intrínseca, como del número de unidades repetidas, la longitud en pares de bases y el motivo de repetición (27). Incluso la edad y el sexo del individuo parecen afectar el cambio molecular. (27, 28).

Weber y Wong (29), observaron que los STR dinucleotídicos tiene tasa de mutación de  $5,6 \times 10^{-4}$ , más bajas que las de los tetranucleotidos de  $2,1 \times 10^{-3}$ . Alelos microsatélites con longitudes más grandes, independientemente del número de repeticiones, usualmente poseen conformaciones más complejas y estables diferentes al ADN-B que pueden facilitar el *slippage* e incrementar la tasa de mutación. Más aún, como se dijo antes, las tasas de mutación de microsatélites iguales en longitud son inversamente proporcionales al tamaño de sus motivos de repetición(24).

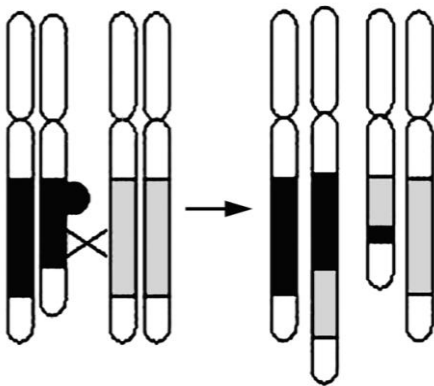
Una consecuencia de la alta tasa mutacional, es su alta variabilidad poblacional. La heterocigocidad promedio de un microsatélite de uso común en laboratorios forenses es generalmente mayor del 75%, para loci con 10 alelos. En Colombia se han estudiado con fines forenses muestras poblacionales de la principales regiones etno-geográficas del país y se ha estimado el nivel de heterocigocidad (Paredes 2003). Algunos loci hiper-variables poseen más de 30 alelos en población colombianay demuestran una heterocigocidad mayor al 95%, como el SE33 (30).

### **1.7 Mecanismos mutacionales de los microsatélites**

Varios mecanismos se han propuesto para explicar el comportamiento mutacional: errores durante la recombinación:el *crossing-over* desigual en la profase meiótica y el *slippageo* “resbalón” de la ADN polimerasa durante la replicación (31), es este último el que se ajusta más a la dinámica mutacional observada en estas secuencias. En realidad, los Microsatélites mutan por InDels de sus motivos de repetición a través de *slippage* de la polimerasa (32). En relación con la recombinación desigual como un mecanismo mutacional, es importante la observación que las tasas mutacionales de STRs del cromosoma Y, de uso común en pruebas forenses, que es en principio, un sistema genético no recombinante, son muy similares a las tasas observadas en STRs autosómicos, lo Cuál hace pensar que este mecanismo no es tan determinante como el *slippage* para generar mutaciones en los STRs. Más aún, Levinson y Gutman

(33), encontraron que cepas de *Escherichiacoli* con y sin sistema funcional de recombinación tenían una tasa similar de mutación.

Cuando ocurre un *crossing-over* desigual, pueden ocurrir cambios drásticos tales como la ganancia o pérdida de un gran número de repeticiones. Esto es debido a que ante una región repetitiva, se puede formar un *hairpin* durante la sinapsis, lo cual significa que solo partes, usualmente desiguales en longitud, de cada cromosoma serán intercambiadas y un cromosoma recibirá un fragmento más grande por el número mayor de repeticiones intercambiadas, y el cromosoma homólogo recibirá un número pequeño de repeticiones (14).**(Figura 2)**



**Figura 2.** Crossing-over desigual, los cromosomas resultantes son de mayor y menor longitud que los originales. Tomado de Oliveira, 2006.

## 1.8 Slippage de la ADN polimerasa

George Streisinger en 1966, generó un modelo original de mutaciones por desviación del marco de lectura (*frameshift*), donde proponía la formación de loops mal apareados, que podían generar ganancias o pérdidas de tamaño en la secuencia de ADN expresiva y su consecuente efecto sobre la traducción en la proteína. **(Figura 3)**

Advirtió además que las secuencias repetidas serían más probablemente objeto de este tipo de mutación (34).

*“...the insertion would be most likely to occur in a region of repeating bases or base*

***doublets through the pairing of a set of bases in one chain of the DNA molecule with the wrong, but complementary, set in the other chain....”***

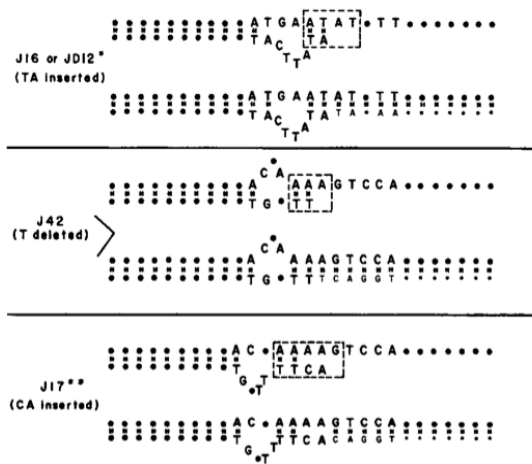


FIGURE 3. Possible mode of origin of some frameshift mutations. In the case of each mutation, a DNA molecule is shown with a gap in one of the two chains and a mispaired set of bases near the end of this chain. The critical area for each mutation is outlined by a dashed square. The appearance of the molecule after synthesis (and mutation) is shown in the second line, for each case.

In each case, the mutation could have occurred in the strand complementary to the one shown here by the same mechanism that is illustrated.

**Figura 3.** *Frameshift mutations*, del artículo original de Streisinger G, C.S.H.Symposia, 1966

Posteriormente Levinson y Gutman, en 1987 (33), aportaron las bases para explicar la dinámica mutacional de los microsatélites a través del mecanismo de la replicación en *slippage*, o *slippage* de la polimerasa, basado en el modelo de Streisinger. Según el modelo, existe un “mal apareamiento por deslizamiento de hebras” o *slipped-strand mispairing*, que afecta a la ADN polimerasa, la cuál, durante el proceso de replicación de una secuencia microsatélite, sufre una **detención del copiado**, que al parecer es ocasionada por varios factores que actúan de manera aditiva: la misma secuencia repetitiva, que posiblemente esté generando una estructura no usual de la molécula de ADN, motiva la detención de la actividad enzima y el complejo de replicación, durante el copiado y por otra parte, las características de la polimerasa que afectan su procesividad sobre la hebra molde.

A continuación se genera una **disociación del complejo de replicación**, que desplaza la enzima y aleja el extremo primado terminal de la hebra naciente, del sitio

activo de la polimerasa, lo cual también depende de su nivel de procesividad. Acto seguido, la enzima puede volver a su sitio original, continuando la copia desde la misma unidad de repetición que había dejado en el momento de la interrupción. Si este es el caso, el copiado es normal y no se genera mutación alguna. De otra manera, la enzima puede “caer” en una repetición previa a la última copiada o posterior a ella.

Estas situaciones dependen a su vez de un evento accesorio que es la formación de un **intermediario mal alineado transitorio** (IMAT) o un *loop* que generalmente contiene una unidad de repetición, la Cuál puede generarse sobre alguna de las cadenas sencillas naciente o molde. La estructura *slipped*, puede ser estabilizada por arreglos de las hebras de ADN, como *hairpins*, estructuras cruciformes, triplex o cuádruplex. El resultado es la generación de una inserción o una ganancia de una unidad de repetición o de otro modo, una delección o una pérdida de una unidad respectivamente. (**Figura 4**). La estabilidad de un *loop* dependerá de la rigidez de las hebras del ADN, de modo que la formación de un *loop* de dos nucleótidos es menos probable que la de un *loop* de 4 nucleótidos por razones estructurales (27). Algunos autores informan que en estudios *in vivo* sobre pedigríes, se observa un sesgo a favor de la ganancia de repeticiones más que a la pérdida durante el proceso (29).

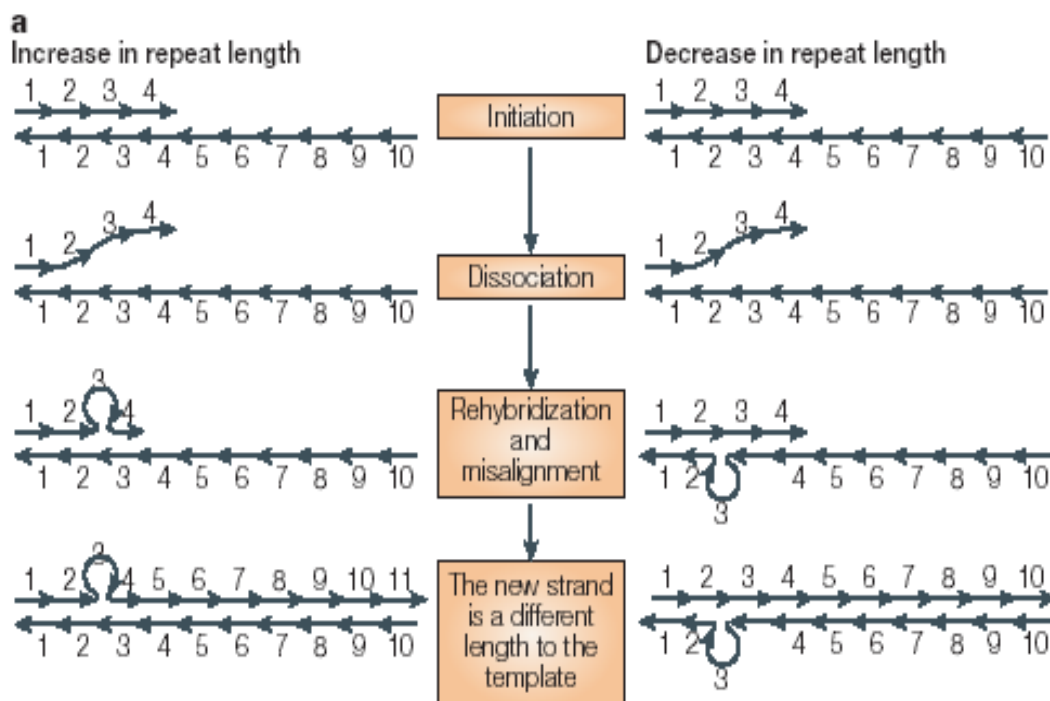
### **1.9 Reparación de la mutación por slippage.**

La mutación se hace real solo si tal *loop* no es corregido y el alelo nuevo, más corto o más largo, se convierte en un nuevo molde de replicación en un nuevo ciclo celular. La mutación depende ahora entonces, de que los mecanismos reparadores logren corregir eventos de *slippage*.

Se ha visto que MMR (*Mismatch repair*), tiene un impacto mucho más significativo que la corrección por prueba de lectura (*Proofreading*) (36). Esto demuestra que la inestabilidad de los microsatélites en humanos está asociada con genes que codifican factores del proceso reparativo MMR como los MSH. Cuando estos mutan,



incrementan la inestabilidad. Si el sistema es defectuoso, las secuencias codificadoras que contienen trectos de ADN repetitivo serán blancos preferidos para mutaciones en tumores humanos.



**Figura 4.** Replicación en slippage (Tomado de Ellegren, 2004)

La efectividad de la reparación tipo *mismatch* puede ser influenciada por la localización genómica del *mismatch*, por el DNA que rodea al *mismatch*, la presencia de señales de reconocimiento de cadena, incluso por el estado de metilación. Más sin embargo, es paradójico que el sistema MMR, el Cuál limita la mutación de secuencias STRs, podría ser particularmente vulnerable a mutación en virtud de tener los STRs en sí mismos, regiones codificadoras (23).

Ya que las regiones repetidas en el ADN son sitios más propensos a la mutación por slippage durante la replicación del ADN, la estabilidad de los microsatélites se regula usualmente en múltiples pasos in vivo a través del sistema MMR.

Las proteínas de MMR remueven las mutaciones por slippage con una alta eficiencia. En células de levaduras que carecen del sistema funcional MMR, se ha visto que tienen hasta 6mil veces incrementada la tasa mutacional. En general, la efectividad del MMR, está influenciada por la localización genómica del *mismatch* y el ADN alrededor del mismo, la presencia de señales de reconocimiento de hebras y aun por el estado de metilación. Sin embargo, el sistema MMR que limita la mutación de secuencias STR, es más vulnerable a mutaciones por la presencia de STRs en su propia región codificadora (24).

### **1.10 Direccionalidad y dinámica de la mutación por slippage.**

**El techo límite para la expansión (*lengthceiling*):** Suponiendo que la replicación en *slippage* es el mecanismo principal de mutación de los STRs y basándose en evidencia experimental (29, 37, 38) que demuestran una aparente direccionalidad de los eventos mutacionales hacia el incremento de longitud, Dermitzakis y Cols.(39) y Feldman (40) propusieron un modelo hipotético en el que se asume que repeticiones más largas, sufren o permiten más eventos de replicación tipo slippage que las más cortas.

Cuando un mecanismo de restricción a la longitud (*lengthconstrain*) interviene, elimina repeticiones aleatoriamente a partir de arreglos cortos o largos con igual probabilidad. Así, los arreglos largos pierden repeticiones a la misma tasa que los cortos, pero ganan repeticiones más rápido que los cortos. El autor no descarta que la selección natural sea la responsable de la restricción al crecimiento de los alelos STRs.

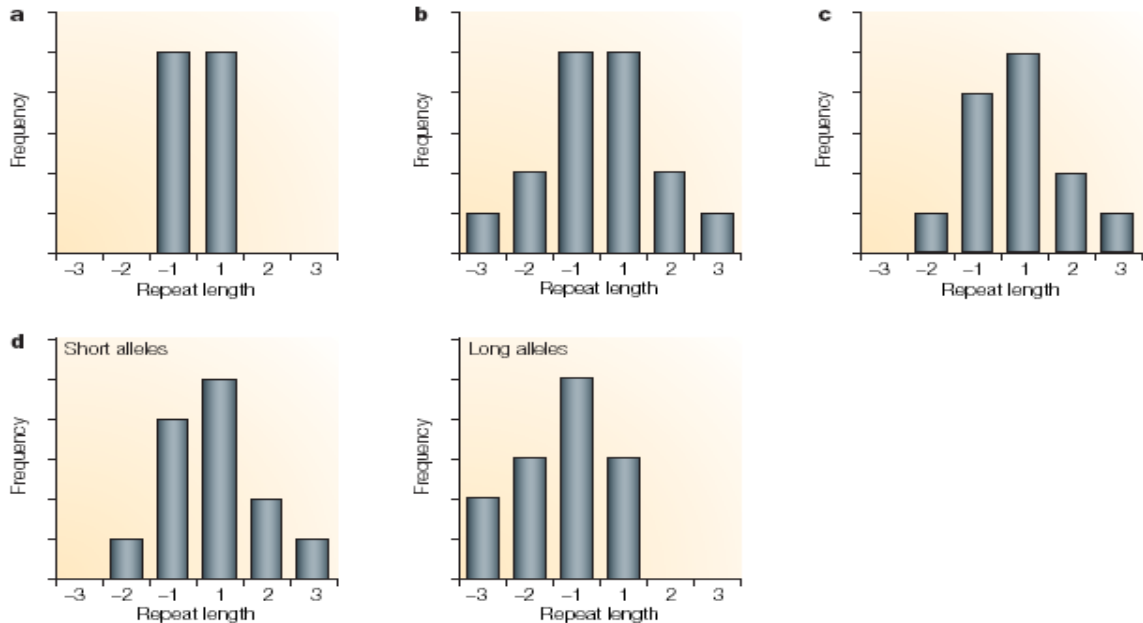
Al respecto, Li et al, (20) creen que la selección natural es una muy importante restricción que actúa sobre el tamaño de alelos STRs. En teoría, la selección natural puede actuar sobre alelos más grandes introduciendo una forma de “techo de longitud” (*lengthceiling*). Demostraron además que en *Arabidopsis thaliana*, la presión selectiva actúa de forma muy diferente en los extremos 5' y 3' del gen, agregando un efecto de polaridad en el proceso mutacional.

### **1.11 El equilibrio entre slippage y mutación puntual.**

Bell y Jurka, inicialmente en 1997 (43) y luego Kruglyak, (41) y Ellegren, (27), sugirieron que dentro de los genomas en equilibrio, la distribución de longitud de STRs es un delicado balance entre procesos mutacionales sesgados y mutaciones puntuales, actuando hacia la muerte del tracto repetitivo. Estas dos fuerza opuestas actúan sobre las secuencias microsatélites; en el equilibrio, habrá una distribución *steady-state* en las longitudes de las repeticiones dirigida por la tasa de slippage y la tasa de mutación puntual. La replicación en slippage favorece el crecimiento mientras que la mutación puntual interrumpe una región repetitiva larga en dos cortas. Cuando una mutación puntual ocurre dentro de una región de repeticiones perfectas, ella suprime las tasa de mutación por slippage por que altera la homogeneidad de la secuencia local repetitiva, que es necesaria para el malalineamiento del precursor *slipped-stranden* la generación de una mutación en longitud del STR (24).

### **1.12 Efecto *allele-span*.**

Rubinsztein, 1995 y Amos, 1996. Mecanismos que ligan la evolución de microsatélites al comportamiento de estas secuencias en una población. Una alternativa es que las tasa altas de mutación que demuestran los alelos más grandes puede que no se deban necesariamente a la longitud alélica por si misma. Los heretocigotos, pueden experimentar una relativa estabilidad en las secuencias microsatélites si la tasa de mutación no está afectada solo por el tamaño alélico si no también por el tamaño del otro alelo del locus (efecto *allele-span*). El modelo se basa en la posibilidad de intercambio desigual entre alelos del heterocigoto durante la meiosis, de modo que un alelo más largo pueda inducir la mutación sobre el más pequeño (37, 38).



**Figura 5.** Magnitud y dirección de los eventos mutacionales en microsatélites (Tomado de Ellegren, (4)

H. Ellegren (4), explica los efectos del modelo mutacional de microsatélites por pasos simples de una o varias repeticiones: Si se tiene en cuenta un modelo en el Cuál los alelos sufren *slippage* de modo que ganan o pierden una única unidad de repetición con la misma probabilidad (modelo de pasos simples), se generaría una distribución totalmente homogénea (Figura. 5 a). Otra alternativa consiste en que se generen cambios igualmente probables de ganancia o pérdida, pero esta vez involucren cambios de uno o más pasos (modelo de multipasos), lo cual generaría una distribución normal donde los alelos intermedios son los más favorecidos (Figura 5b) . Si existe direccionalidad a favor de la expansión de repeticiones se generarían distribuciones sesgadas (Figuras 5 c y d.)

Un umbral de mutación parece depender tanto del número de repeticiones involucradas como de la longitud en pares de bases del segmento. Algunos autores proponen (35), a través de datos experimentales y computacionales, que los dinucleótidos mutan por encima de 5 repeticiones y los mononucleótidos mutan desde 10 pb o repeticiones. Tres variables parecen estar involucradas en el mismo efecto: la longitud del fragmento, la longitud de la unidad de repetición y la secuencia del

mismo. Se menciona frecuentemente que se requiere de un mínimo número de unidades repetidas para que la polimerasa resbale y extienda la cadena original que copia.

Es interesante la observación de Ellegren(27), donde hablaba de una posible relación negativa entre la magnitud de la mutación y el tamaño alélico estandarizado. Él sugería que los alelos más largos son más propensos a disminuir en tamaño que los más cortos, basado en escasa evidencia obtenida de análisis mutacionales sobre STRs de peces, aves e incluso levaduras.

Recientemente, se han caracterizado los eventos mutacionales de forma másiva a través de estudios de secuenciación genómica total. Los hallazgos actuales confirman algunas hipótesis planteadas hace más de 12 años y descarta otras interpretaciones de la época:

Sun y cols. (44), reportaron recientemente el estudio mutacional más grande realizado hasta la fecha en 85.289 individuos de Islandia, que comprendió 2.058 cambios germinales en 2.477 microsatélites del tipo di y tetranucleotídico. La tasa de mutación por locus por generación, para estos últimos fue de  $10.01 \times 10^{-4}$ , 3,7 veces mayor que para los dinucleótidos ( $2,73 \times 10^{-4}$ ).

En relación con los dos tipos de STRs analizados, Sun encuentra que el 32% de las mutaciones en loci dinucleótidos fueron multipasos, comparado con el 1% en los tetranucleotídicos. Por otra parte, la tasa de mutación incrementó con la longitud del alelo, se cuadruplicó entre 30 y 70pb para di y entre 40 y 120pb para tetranucleótidos.

Luego, los loci con estructuras repetitivas uniformes o simples (Ej. CACACACA) mostraron una tasa 40% más alta que las estructuras compuestas como en el STR (CACATCACA), consistente con menos slippage para loci con repeticiones interrumpidas. Notaron que los alelos más cortos tienden a mutar hacia más largos y

alelos más largos tienden a disminuir su longitud. Este patrón contrasta con el observado en enfermedades por expansión de trinucleótidos donde los alelos más largos tienden a mutar a alelos aún más largos.

La tasa de mutación se correlacionó con la longitud del motivo de repetición, el número de repeticiones, el tamaño alélico, la distancia de exones, el género y la edad pero no con la tasa de recombinación, la distancia de los telómeros, la divergencia humano-chimpancé y la heterocigocidad parental.

### **1.13 Modelos teóricos de mutación**

El modelo mutacional que se adopte en un estudio, definirá la validez de las inferencias genético poblacionales que se hagan a partir de los datos genéticos obtenidos de los microsatélites. En general se usan cuatro modelos mutacionales:

#### **1.13.1 Modelo de Alelos infinitos: (IA)**

Cada mutación crea aleatoriamente un nuevo alelo. En microsatélites, las mutaciones alteran el número de repeticiones, por ejemplo un alelo con 12 repeticiones se considera genéticamente tan cercano a uno con 15 como a uno con 16 repeticiones. Es decir, la proximidad en términos del número de repeticiones no indica una mayor relación filogenética. Es el modelo clásico de Sewall Wright (45) para el cual propuso el uso de los estadísticos  $F_{ST}$ .

#### **1.13.2 Modelo de mutación por pasos: *Stepwisemutationmodel* (SMM)**

Cuando un microsatélite muta, gana o pierde una repetición, esto implica que dos alelos que difieren solo en una repetición, es decir, comparten un ancestro común más reciente, están más relacionados, que aquellos que difieren en varias unidades. Montgomery Slatkin 1995 (46), propuso una medida de la diferenciación genética:  $R_{ST}$ , similar a los  $F_{ST}$  de Wright (1951) y los  $G_{ST}$  de Nei de 1973 (47), pero basados en el modelo SMM.

El modelo es útil para estimar relaciones entre individuos y estructura poblacional, excepto en presencia de homoplasia, es decir cuando dos alelos son idénticos por estado pero no por descendencia. La homoplasia puede afectar los estudios poblacionales involucrando. El modelo descrito por Slatkin está basado en rasgos de distribuciones continuas, número de pares de bases o número de repeticiones y agrupa individuos de acuerdo al número de repeticiones.

### **1.13.3 Modelo de dos fases: *Two phase model* (TP)**

Di Rienzo en 1994 (48), introdujo este modelo como un complemento del modelo SMM para estudios sobre microsatélites. En él se establece que la mayoría de eventos mutacionales resultan en un incremento o disminución de una unidad de repetición, aunque también pueden ocurrir, en menor proporción, alteraciones que involucren grandes números de repeticiones.

### **1.13.4 Modelo de K-alelos: (KA)**

Crow y Kimura en 1970 (49) propusieron un modelo que asume, que si existen exactamente en un locus dado  $k$  posibles alelos, entonces la probabilidad de que un alelo dado mute en cualquier otro es  $\mu/k-1$ , donde  $\mu$  es la tasa de mutación.

El conocimiento de la dinámica mutacional tiene aplicación directa en los estudios evolutivos; la inferencia filogenética, requiere de una medida adecuada de las distancias génicas. Estas se basan generalmente en la presunción que los microsatélites siguen el modelo de mutación por pasos (SMM). Como se mencionó anteriormente, el SMM fue complementado hace dos décadas por Di Rienzo, quien propuso el modelo de mutación en dos fases (TPM); posteriormente, se incluyó el concepto de “techo” o límite superior de la expansión alélica.

## **1.14 Edad, sexo y mutación**

Desde 1912, Wilhelm Weinberg (50), notó que niños con Acondroplasia eran producto de gestaciones de padres normales que concebían en edades avanzadas. Desde entonces, se ha aceptado que, dado el notable efecto de la edad paterna

sobre la aparición de mutaciones, sugería que las tasas de mutación variaban entre los sexos. Una vez las técnicas moleculares estuvieron disponibles fue posible determinar el origen parental de una mutación en una enfermedad genética.

La evidencia molecular actual ha permitido relacionar la edad paterna avanzada al momento de la fecundación con la probabilidad de desarrollar autismo y esquizofrenia (51).

La causa de tal hallazgo es aparentemente muy obvia: el gran número de divisiones celulares en línea germinal pre-meióticas que se observan en el varón más que en la mujer, y la diferencia es por supuesto mayor en hombres más viejos. Según Vogel y Motulsky(52), hay 24 divisiones celulares desde el cigoto hasta un óvulo (23 replications cromosómicas ya que hay solo una en las dos divisiones meióticas).

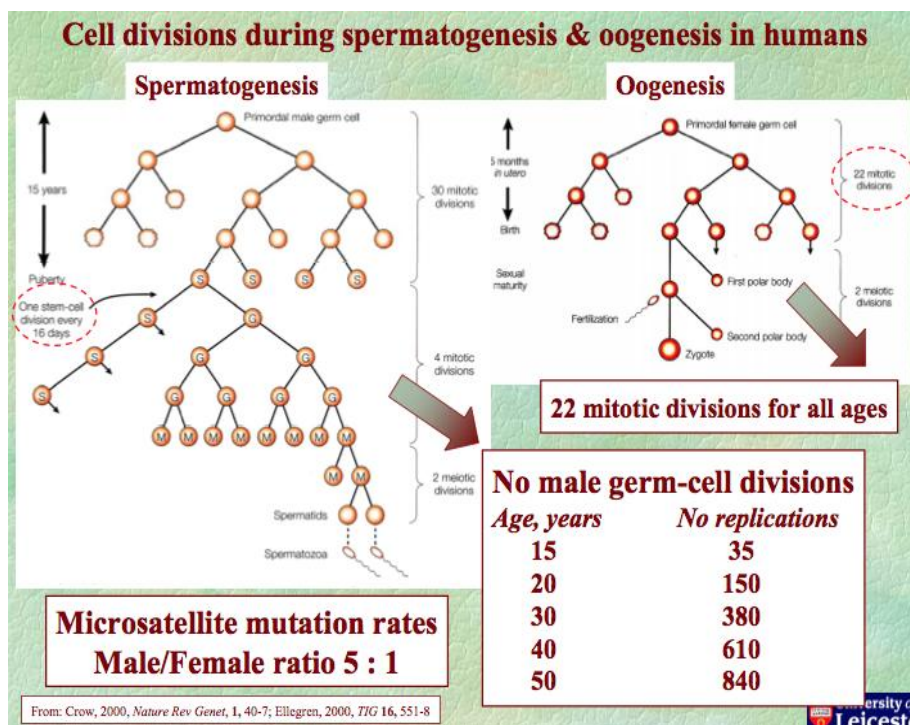
En el hombre, por el contrario, hay 30 divisiones previas a la pubertad (15 años), seguidas por divisiones de células madres (espermatogonias) que ocurren 23 veces por año, entonces 4 divisiones goniales y dos meióticas (una replicación). Por lo tanto, el número de replications cromosómicas desde el cigoto al espermatozoide producido por un hombre de edad T sería:

$$NT = 30 + 23 (T-15) + 5$$

A la edad de 20, sería 150; a los 30 años sería de 380; a los 40, de 610 y a los 50 de 840. A la edad de 30, el radio Hombre/mujer es de  $380/23 = 16,5$ . (Figura 6).

En cada división celular la replicación del genoma introduce la posibilidad de errores de copia. En humanos se ha reportado que el esperma de hombres adultos mayores tiene significativamente más mutaciones (52). La actividad reparadora declina también con la edad paterna avanzada y se compromete la misma replicación.





**Figura 6.** Divisiones celulares en la ovogénesis y la espermatogénesis.

Un estudio reciente parece corroborar esta observación histórica: Un hombre de 36 años generará dos veces más mutaciones a su hijo que un hombre de 20 años, y uno de 70 años, 8 veces más, según estima el grupo de KariStefansson, del laboratorio *deCODEGenetics* en Reikjavik, en un estudio recientemente publicado en *Nature*, realizado sobre 78 tríos padre-madre-hijo, en una muestra de población Islandesa. (53). No obstante la afirmación anterior, el estudio no prueba que los padres más viejos pasen genes deletéreos o enfermedades genéticas más probablemente que los jóvenes. Los datos corresponden exclusivamente a mutaciones tipo SNPs (*Single nucleotide polymorphisms*) y se obtuvieron por secuenciación genómica total. El promedio de edad para los padres fue de 29,7 años. Se calculó un promedio de 2 mutaciones nuevas por año.

### **1.15 Mutación somática vs germinal**

Las mutaciones de los microsatélites se generan en la línea germinal paterna o materna, aunque no se descartan las mutaciones somáticas, cigóticas o post-cigóticas ocurridas en el hijo.

La mutación somática generaría mosaicismo donde coexisten en proporciones diferentes alelos de diferente tamaño; se presentarían en forma de heterocigotos con señales electroforéticas muy desiguales, lo que indica que originalmente el individuo era homocigoto para ese locus y se conserva la doble dosis alélica, pero acompañada de un nuevo alelo que puede ser minoritario. La proporción de los alelos diferentes variaría según la población de células portadoras del alelo mutante y esto por supuesto, dependerá del momento en que se originó la mutación.

Por otra parte, si el individuo era heterocigoto, con un genotipo constituido por alelos seguidos en número, puede pasar que el alelo mutado se convierta en un alelo del mismo tamaño que el homólogo y de nuevo, se observen heterocigotos con dosis alélicas muy diferentes. Finalmente, si la mutación genera un alelo diferente a los dos del heterocigoto, habrá entonces un genotipo trialélico. Aun así, esta variación puede estar representando un evento ancestral de duplicación de un locus microsatélite con posterior divergencia, pero padres e hijos descendientes del individuo tendrían la misma constitución.

Los casos forenses podrían ser una fuente importante de evidencias de procesos mutacionales somáticos, pero en la práctica es muy difícil obtener un consentimiento para tomar muestras de fluidos o tejidos diferentes de un mismo individuo, las formastrialélicas llaman la atención con facilidad en muestras sanguíneas de referencia, pero sobre evidencias recuperadas de la escena del crimen o sobre víctimas, la presencia de un tercer alelo puede confundirse con una mezcla de fluidos proveniente de más de un aportante; (Ej. Violadores múltiples) aun así, el hecho que solo se observe en un locus pone en alerta al perito genetista y de coincidir con el

sospechoso del crimen, puede ser este un elemento somático altamente discriminante para identificar el origen de la evidencia.

Si el evento mutacional sucede en mitosis pre-meióticas, de la espermatogénesis, en divisiones previas a la generación del espermatozoides primario, particularmente durante la fase S del ciclo celular de la espermatogonia, el número de células portadoras de la mutación estará en función del número de divisiones celulares que resultan de ese linaje; suponiendo que se trate de un evento temprano sobre una célula embrionaria. Habría muchas células en el padre que porten la mutación y todas entrarían a meiosis, la consecuencia de esto es que tendríamos un número muy alto de espermatozoides que porten la mutación generando varias alternativas de herencia.

Volviendo al problema de la posibilidad de detección de estos eventos, se esperaría que en muestras de semen que se analizan de rutina en la investigación del delito sexual, se detecten las dos poblaciones moleculares.

Este mosaico gamético no se observa en la casuística forense. Se trata de un evento de muy baja frecuencia, en promedio se ven 20 eventos en 800 a mil casos de paternidad; en la investigación del delito sexual, se analizan 20 a 30 casos por mes, ya sean muestras postcoitales o manchas de semen recuperadas de la escena del crimen. A esta tasa, los peritos genetistas forenses esperarían observar un evento cada dos años aproximadamente en Colombia.

De nuevo como se esperaría en un evento somático, el mosaísmo germinal se presentaría en el análisis del semen, como un heterocigoto desbalanceado o un trialelo en el locus afectado, dependiendo de la población de células espermáticas involucradas con la mutación. Una población no determinada de mutantes pasará desapercibida para el analista y otros podrían confundirse en muestras mezcladas ya que no se podría saber si se trata de un alelo mutado o un componente menor de la mezcla aportado por uno de los contribuyentes en menor cuantía.

Habría que estudiar el semen del padre candidato de la mutación, para detectar si existen dos (o más) poblaciones celulares espermáticas diferentes para ese alelo, pero como se dijo antes, es muy improbable que el acusado de la paternidad acepte donar una muestra de semen para el estudio, unos días después de realizada la prueba, cuando se conozca el resultado de la misma, que de hecho no le favorecerá.

### **1.16 Herramientas de análisis para el estudio de microsatélites**

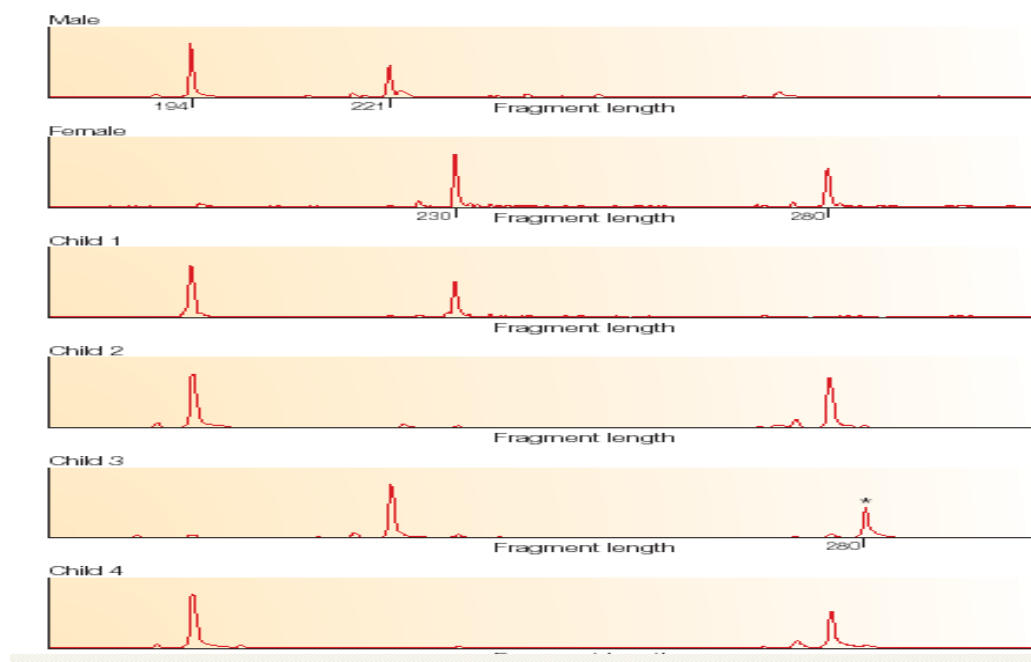
La medición de la variación en microsatélites tiene limitaciones: Entre los tipos de estudios que se realizan para medir la variación en STRs están: El análisis de *pedigrees* expandidos, que tiene el inconveniente de requerir de muchas observaciones en varias generaciones. La comparación de loci STR entre especies como humanos y otros primates como el chimpancé, ha sido objeto de crítica por la dificultad de comparar los mismos loci entre las especies estudiadas y por el escaso número de loci analizados.

Finalmente, el estudio de transferencias de alelos STR, en tríos padre/madre/hijo, en investigaciones de la paternidad o maternidad biológica, (55) permite reunir un número amplio de casos y estimar la tasa de mutación en relación con el número de transferencias alélicas o meiosis, de una forma directa y controlable. Afortunadamente, la alta tasa mutacional de los microsatélites hace posible la observación directa de eventos mutacionales. Es este el modelo de estudio mutacional que se utiliza en la presente investigación.

Como se mencionó antes, la mejor aproximación al evento mutacional en microsatélites es la observación directa de transmisiones alélicas de los padres a los hijos (4). Un alelo mutante se identifica por una transmisión mendeliana incompatible (**Figura7**).

Previamente, el caso ha sido analizado con un amplio número de microsatélites adicionales, en los Cuáles, no se ha observado un evento excluyente de la

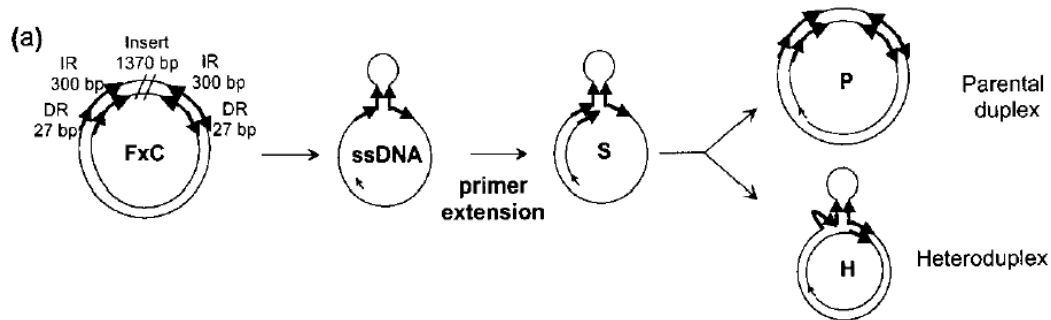
paternidad; así, el evento estudiado es generalmente único. Adicionalmente, el comportamiento de los alelos sugiere que se trata de un evento mutacional, ya que es posible con ayuda del secuenciador automático, asignar tamaños precisos en pares de bases y verificar que se ha ganado o perdido una sola repetición (evento de paso simple) o varias de ellas (evento multipasos).



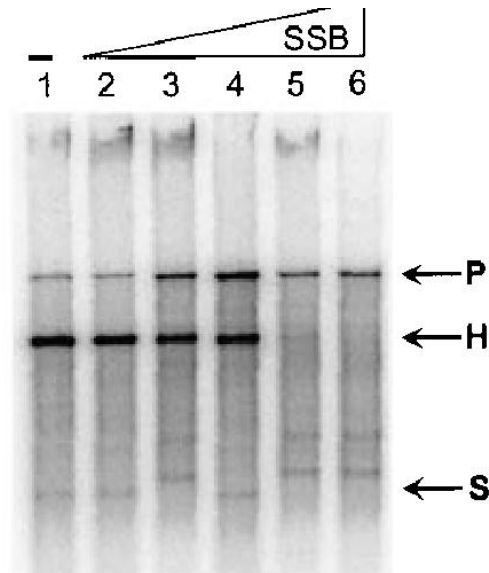
**Figura 7** Estudio de paternidad donde se observa una mutación entre el padre y el hijo tres\*  
(Tomado de Ellegren, 2004)

Una segunda alternativa de estudio es la reproducción *in vitro* del evento, utilizando una herramienta molecular construida, donde en forma controlada pueda observarse cambios discretos de peso molecular generados por *slippage*. Al respecto cito el trabajo de Viguera *et al*(56), en el Cuál se construyeron plásmidos con secuencias repetitivas cortas insertadas para inducir *slippage in vitro* y de forma controlada poder detectar los productos de amplificación mutantes y no mutantes. En la Figura 8a se observa el construido molecular y los fragmentos obtenidos luego de una separación en geles de agarosa. (Figura 8b). La forma H o heterodúplex se produce solo si hay *slippage* sobre la primera secuencia repetitiva de 27pb. Esta molécula genera una banda electroforética de fácil asignación. El modelo permite evaluar la actividad de

desplazamiento de hebra que poseen diferentes polimerasas, de modo que solo aquellas enzimas que logren ingresar dentro del *loop* generarán la molécula dúplex parental.



**Figura. 8a.** PlasmidoFxC: DR: repeticiones directas de 27pb, IR: Repeticiones inversas (Palindrómicas) de 300pb que permiten la formación de un loop, Insert: Inserto de 1,4Kb. Forma S: extensión detenida en la base del loop, Forma P: Duplex parental donde no se ha dado slippage y Forma H: heteroduplex resultante del evento de slippage. (Tomado de Viguera et al. 2001)



**Figura. 8b.** Fragmentos obtenidos al realizar un ciclo de amplificación, incluyendo la proteína SSB bacteriana que permite mantener el plásmido en cadena sencilla. (Tomado de Viguera et al. 2001)

## 1.17 Microsatélites e identificación humana

Debido a su alta variabilidad, herencia codominante, alta reproducibilidad, fácil detección, pequeño tamaño, entre otros, los microsatélites han sido explotados extensivamente como marcadores de ADN para mapeo genético, análisis de ligamiento, estudios genético poblacionales y evolutivos y se han constituido además como la herramienta de primera elección por los laboratorios forenses en identificación humana.

Desde el advenimiento de la Reacción en cadena de la polimerasa (PCR) en 1985, (57) el estudio de los microsatélites ha sido accesible al análisis másivo. Adicionalmente, la aparición de los Secuenciadores de ADN a principios de los años 90`s permitió el estudio automatizado de alelos STR`s. Una disciplina particularmente favorecida por estos avances tecnológicos ha sido la Genética forense. Inicialmente, cuando Sir AlecJeffreys propuso el análisis de secuencias minisatélites como una “huella genética”(DNA fingerprinting), que permitiría la identificación de personas, se abrió una puerta al conocimiento de la diversidad humana nunca antes pensada (58).

La diversidad de secuencias repetitivas en tándem anunciada por Jeffreys en los años 80`s, era de tal magnitud, que un patrón electroforético minisatélite de cerca de 30 bandas, logrado con la hibridización de sondas de ADN que reconocían múltiples *loci* en el genoma, podía encontrarse en 1 de cada  $6 \times 10^9$  personas en el planeta.

Si bien la idea de Jeffreys fue prontamente aplicada a investigaciones criminalísticas, rápidamente aparecieron las limitaciones metodológicas ya que el método requería de moldes de ADN mayores a 2 Kb, una situación difícil de lograr en la mayoría de las muestras biológicas presentes en la escena del crimen:manchas de semen, sangre, pelos, huesos antiguos etc.que se caracterizan precisamente por su alto grado de descomposición y por lo tanto de degradación del ADN que contienen.

Colombia fue uno de los primeros países en Latinoamérica en aplicar el estudio de polimorfismos de ADN con fines forenses. Desde el año 1994, se creó el primer laboratorio de genética forense en el Instituto Médico Legal de Bogotá y se

implementaron los protocolos metodológicos para el análisis de secuencias polimórficas, que se aplicaron a un número siempre creciente de investigaciones criminalísticas. Un año después, se estandarizaron los primeros sistemas de PCR multiplex para la amplificación de dos “tetraplex” sugeridos por el *Forensic Science Service*, en el laboratorio de Peter Gill en Birmingham – UK. Estos 8 STRs autosómicos, se analizaron en plataformas policromáticas sobre secuenciadores ABI, que detectaban 4 fluorocromos diferentes, asociados a los primers, lo que permitía la detección de alelos en los 8 loci, así estos se solapaban durante el corrido electroforético.

El modelo se utilizó en un muy amplio número de investigaciones criminales y de filiación. Adicionalmente, se realizaron estudios de frecuencias poblacionales de los alelos STR para los 8 loci. Luego, finalizando los años 90's, con la comercialización de kits PCR multiplex por la casa Applied Biosystems, como el AmpFISTR® Profiler® PCR Amplification Kit y el AmpFISTR® Cofiler® PCR Amplification Kit, se universalizó el uso de 5 STRs adicionales para un total de 13 loci, que el FBI acogió como el número mínimo de STRs, con poder de identificación humana suficiente para establecer bancos genéticos de perfiles de STR de uso criminalístico en la Unión americana, lo que dio origen al sistema informático: Combined DNA Index System (CODIS). [www.fbi.gov/about-us/lab/biometric-analysis/codis](http://www.fbi.gov/about-us/lab/biometric-analysis/codis) .

En el año 2003 se estudió la distribución de frecuencias de alelos microsatélite para los 13 loci autosómicos del CODIS en la población colombiana (59), con el fin de aportar significancia estadística a las coincidencias de perfiles de ADN detectadas en los casos forenses. Actualmente se evalúan las tasas mutacionales de los marcadores moleculares de uso rutinario en los laboratorios forenses.

La presencia de eventos mutacionales detectados sobre loci microsatélites en el estudio del parentesco ya sea para investigación de la paternidad o para identificación de individuos N.Ns, aunque es un evento poco frecuente, debe ser caracterizado en detalle y evaluado adecuadamente en el contexto de la



investigación de filiación para aportar pruebas de mayor contundencia científica a la administración de justicia.

La presencia de incompatibilidades con la herencia mendeliana, permiten demostrar que no existe relación de parentesco en muchos escenarios donde se investiga la filiación entre personas. La *International Society for Forensic Genetics*. [www.isfg.org](http://www.isfg.org) ha establecido criterios mundialmente aceptados para definir una “exclusión” de paternidad: Tres loci donde se demuestre una incompatibilidad entre el presunto padre y el hijo se asumen suficientes para generar un reporte de no paternidad. Porque tres? Porque es muy improbable que entre padre e hijo se observen tres eventos mutacionales en un número reducido de marcadores analizados (generalmente 15 loci STR); aunque poco común, se observan dos eventos de exclusión con cierta frecuencia, del mismo modo, un solo evento se observa con frecuencia en laboratorios que analizan número altos de casos de paternidad.

Aun así, estos umbrales son siempre arbitrarios y se basan en la probabilidad de ocurrencia de eventos de mutación. Cada mes, el Laboratorio de genética del Instituto Nacional de Medicina Legal y Ciencias Forense en Colombia reporta entre 15 y 25 eventos aislados de exclusión, analizando cerca de 700 a 1100 casos de paternidad en el mismo periodo. En los casos de exclusión de la paternidad el número de incompatibilidades más frecuente esta entre 8 y 10 de un total de 15 loci estudiados(61), véase más adelante.

#### **1.17.1 Mutación y casos forenses**

Cuál es el efecto que sobre la interpretación de resultados tiene un hallazgo de una posible mutación en un caso de paternidad o identificación: De detectarse un evento aislado de exclusión en la tipificación de 15 marcadores STRs entre dos individuos supuestamente emparentados como padre e hijo, el perito debe ampliar el número de loci al máximo posible del laboratorio, un número aceptable es el de 25 loci. Lo que normalmente se observa es que no aparece ninguna exclusión adicional y por el

contrario se eleva inmensamente la probabilidad de paternidad alcanzada con los loci compatibles.

El resultado probabilístico debe corregirse antes de informarlo a la autoridad, calculando una razón de verosimilitud o LR (*Likelihood ratio*) para el locus excluyente, asumiendo que se trata de un evento mutacional, lo Cuál se incorpora al valor total previamente calculado, ocasionando una disminución en el valor de probabilidad alcanzado, algunas veces en varios órdenes de magnitud. [http:// dna-view.com](http://dna-view.com).

Generalmente, este hecho guarda un patrón de ocurrencia fácil de detectar, la casi totalidad de los eventos son ganancias o pérdidas de una unidad de repetición, solo ocasionalmente se observan diferencias de dos o más unidades entre los alelos del padre y el hijo para el mismo locus. Un evento donde la diferencia es más de dos unidades debe hacer pensar en otro mecanismo mutacional como las mutaciones “*null*” o silenciosas; en ellas, generalmente hay una mutación puntual en la región de hibridación de uno de los *primers* usados para amplificar el locus, principalmente afectando el extremo 3´ de la zona de reconocimiento del iniciador.

Esto genera un falso homocigoto tanto en el padre como en el hijo que ha heredado el cromosoma con la mutación. Padre e hijo pueden presentarse como dos homocigotos diferentes (o iguales) ya que no se manifiesta el alelo no amplificado, que posee la mutación *null*. Así, no importa el número de repeticiones que diferencia al padre del hijo.

Ante la exclusión real de la paternidad, generalmente no existe duda; en la mayoría de los casos de exclusión, analizando el mismo número de *loci*, se encuentran en promedio 8-10 incompatibilidades (60), así que el hallazgo de una sola exclusión hace pensar inmediatamente en una mutación. Son muy pocos los casos donde, al ampliar los marcadores aparecen dos o tres resultados incompatibles adicionales. Este escenario hace pensar que el acusado es un pariente cercano del verdadero

padre biológico y por lo tanto no ha sido fácil excluirlo ya que comparte con el padre verdadero varios de sus alelos STR.

Afortunadamente, el estudio de casos de investigación de la paternidad, aunque genera el problema de investigación, también se constituye en una alternativa de estudio de la dinámica mutacional de los microsatélites. Es muy difícil reunir varias generaciones de individuos para seguir un evento mutacional; alternativamente, si se cuenta con muchos tríos de paternidad (8.000 a 11.000 casos nuevos /año en Colombia), es posible evaluar muchas meiosis (padre-hijo o madre-hijo), calcular tasas mutacionales, aislar y purificar los alelos comprometidos en el evento mutacional y secuenciar cada uno para luego con ayuda de herramientas bioinformáticas, caracterizar el evento molecular.

Se ha iniciado este trabajo en el Instituto Médico Legal, con el estudio sobre casos de paternidad utilizando como modelo microsatélites hipervariables de alta tasa mutacional, como el locus HUMACTBP2-SE33, un tetranucleótido complejo, en el que se detectaron 43 alelos distintos en población colombiana y 5 eventos mutacionales en cerca de 760 meiosis analizadas. (30), alcanzando una tasa de  $6.6 \times 10^{-3}$ , la más alta reportada para Cualquier microsatélite de uso forense(55). Recientemente publicamos las tasas mutacionales del locus FGA-FIBRA, uno de los marcadores de más alto polimorfismo en poblaciones colombianas (61).

Ahora, utilizando las bases de datos de perfiles genéticos del Instituto Médico Legal, es posible contar con un número importante de eventos de mutación en cerca de 16 loci microsatélite, que permiten observar la dinámica mutacional con más evidencias y por lo tanto hacer inferencias sobre la conducta molecular de estas particulares secuencias. Superando en parte la limitación del tamaño de la muestra, circunstancia que, en otros estudios, ha sido la talanquera para generar hipótesis más cercanas a la realidad y por otra parte, con la posibilidad de controlar el análisis en un mismo centro de investigación, bajo un mismo protocolo de estudio y con la alternativa de corroborar y confirmar los resultados.

## 1.18 Caracterización de los loci microsatélites analizados.

Los perfiles genéticos analizados en el trabajo actual son el resultado de la tipificación de muestras anónimas relacionadas con investigaciones de parentesco, en las Cuáles se amplificaron kits comerciales de identificación humana de las casas *AppliedBiosystems*, [www.lifetechnologies.com](http://www.lifetechnologies.com) y *Promega Co.*, [www.promega.com](http://www.promega.com) comúnmente usados en los laboratorios forenses. A continuación se listan los kits utilizados y los marcadores genéticos que contienen.

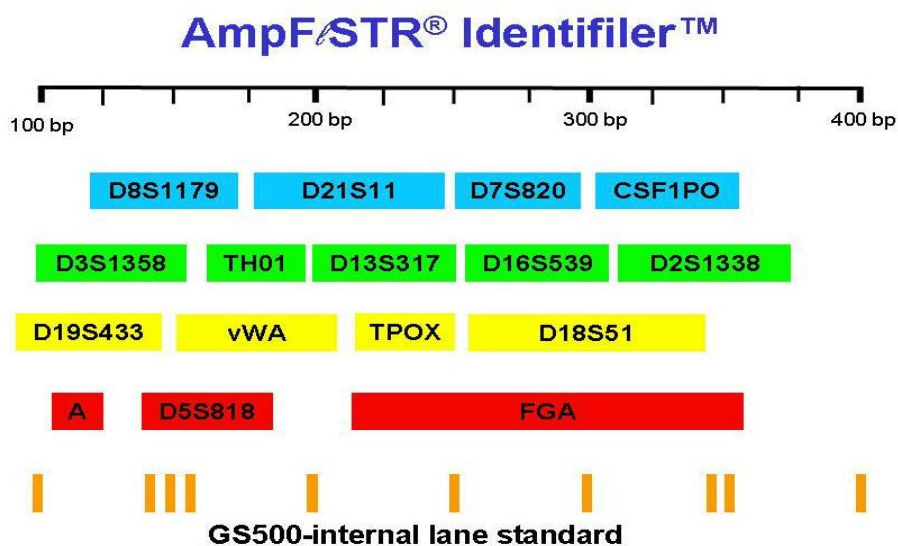
### 1.18.1 Kit AmpFLSTR® Identifiler® PCR Amplification Kit de la casa AppliedBiosystems.

Permite la amplificación simultánea en una reacción multiplex de 15 *loci* STR autosómicos y el gen de la Amelogenina, que mapea en cromosomas sexuales y es útil para controlar el sexo de la muestra. Los 15 marcadores autosomales mapean en 13 cromosomas; Dos de ellos se localizan en el cromosoma 2 p y q y dos más en el cromosoma 5 sobre el brazo largo. (**Tabla 1**)

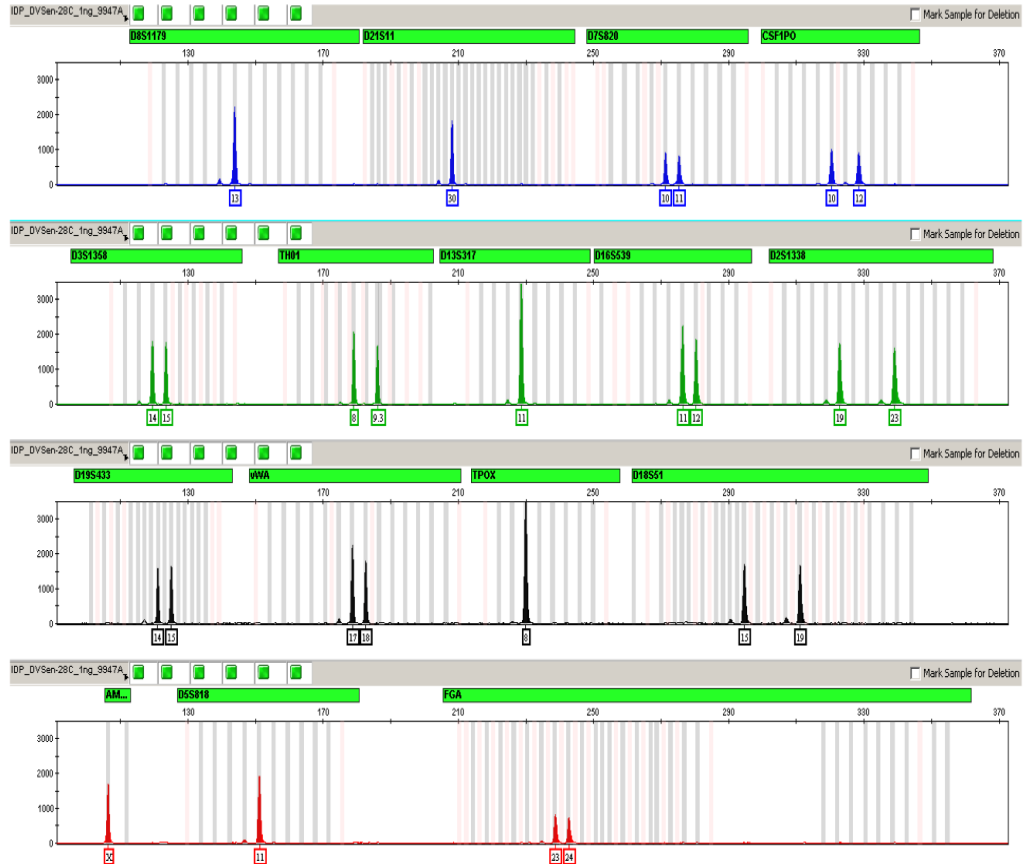
Utiliza *primers* marcados con 5 fluorocromos comercialmente registrados: 6-FAM, VIC, NED, PET y LIZ, esta tecnología permite diferenciar los alelos de un locus que se sobrelapan en tamaño con los alelos de otro locus durante el corrido electroforético (**Figura 9**). Están diseñados por lo tanto para ser detectados en plataformas de electroforesis capilar en analizadores genéticos automatizados, el sistema puede detectar las diferencias de emisión de cada fluoróforo y presentar un gráfico de picos (electroferograma) de colores diferentes para los alelos de cada locus así estos tengan el mismo tamaño en pares de bases y por la tanto la misma movilidad electroforética.

La **Figura10** muestra el perfil genético que se obtiene de la línea celular 9947<sup>a</sup>, incluida en el kit como control positivo de la reacción de PCR. Los alelos se detectan como señales de fluorescencia y son asignados a un tamaño específico utilizando el software *Genemapper ID-X*, en el Cuál se establecen parámetros de análisis que

permiten valorar la movilidad electroforética de un fragmento STR entre valores esperados, obtenidos por experimentos de repetitividad, que generan un rango de desviación aceptada, (*bins*) que aparecen en el electroferograma como barras grises detrás de cada alelo. La señal que migra fuera de estos parámetros es asignada como un “OL” *out of ladder*. La precisión del método (*local southern*) es cercana a 0,5 pb en cada asignación (62).



**Figura 9.** Sobrelapamiento de loci resuelto por el uso de 4 cromógenos distintos.  
kit Identifiler®.



**Figura 10.** Electroferograma que muestra el perfil genético para la línea celular 9947<sup>a</sup> con el kit Identifiler®.

### 1.18.2 Kit PowerPlex®16 HS System de Promega Co.

Permite la amplificación simultánea en una reacción multiplex de 15 *loci* STR autosómicos y el gen de la Amelogenina como el kit Identifiler. Los 15 marcadores autosomales mapean en 13 cromosomas; Dos de ellos se localizan en el cromosoma 21 q y dos más en el cromosoma 5 sobre el brazo largo. (**Tabla 2**)

Locus designation	Chromosome location	Alleles included in AmpF/STR® Identifiler® Allelic Ladder	Dye label	Control DNA 9947A
D8S1179	8	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	6-FAM™	13†
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		30‡
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 11
CSF1PO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 12
D3S1358	3p	12, 13, 14, 15, 16, 17, 18, 19	VIC®	14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		8, 9.3
D13S317	13q22-31	8, 9, 10, 11, 12, 13, 14, 15		11§
D16S539	16q24-qter	5, 8, 9, 10, 11, 12, 13, 14, 15		11, 12
D2S1338	2q35-37.1	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		19, 23
D19S433	19q12-13.1	9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2	NED™	14, 15
vWA	12p12-pter	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17, 18
TPOX	2p23-2per	6, 7, 8, 9, 10, 11, 12, 13		8††
D18S51	18q21.3	7, 9, 10, 10.2, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		15, 19
Amelogenin	X: p22.1-22.3 Y: p11.2	X, Y	PET®	X
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11‡‡
FGA	4q28	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 42.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		23, 24

**Tabla 1.** Localización cromosómica de los 16 loci del kit Identifiler

*Tomados de AmpFLSTR® Identifiler® PCR Amplification Kit User Manual*

STR Locus	Label	Chromosomal Location	GenBank® Locus and Locus Definition	Repeat Sequence <sup>1</sup> 5'→3'
Penta E	FL	15q	NA	AAAGA
D18S51	FL	18q21.3	HUMUT574	AGAA (22)
D21S11	FL	21q11-21q21	HUMD21LOC	TCTA Complex (22)
TH01	FL	11p15.5	HUMTH01, human tyrosine hydroxylase gene	AATG (22)
D3S1358	FL	3p	NA	TCTA Complex
FGA	TMR	4q28	HUMFIBRA, human fibrinogen alpha chain gene	TTTC Complex (22)
TPOX	TMR	2p24-2pter	HUMTPOX, human thyroid peroxidase gene	AATG
D8S1179	TMR	8q24.13	NA	TCTA Complex (22)
vWA	TMR	12p13.31	HUMVWFA31, human von Willebrand factor gene	TCTA Complex (22)
Amelogenin <sup>2</sup>	TMR	Xp22.1-22.3 and Y	HUMAMEL, human Y chromosomal gene for Amelogenin-like protein	NA
Penta D	JOE	21q	NA	AAAGA
CSF1PO	JOE	5q33.3-34	HUMCSF1PO, human c-fms proto-oncogene for CSF-1 receptor gene	AGAT
D16S539	JOE	16q24.1	NA	GATA
D7S820	JOE	7q11.21-22	NA	GATA
D13S317	JOE	13q22-q31	NA	TATC
D5S818	JOE	5q23.3-32	NA	AGAT

**Tabla 2.** Localización cromosómica de los 15 loci del kit PowerPlex16HS.

*Tomado de PowerPlex 16HS System Technical Manual.*

Utiliza *primers* marcados con 4 fluorocromos comercialmente registrados: FL, TMR, JOE y ROX , La **Figura 11** muestra el patrón de *ladders* alélicos de los 15 loci del PowerPlex, detectados con el software Genemapper IDX (63).La **Figura 12** muestra el solapamiento de los loci del kit PowerPlex 16HS.



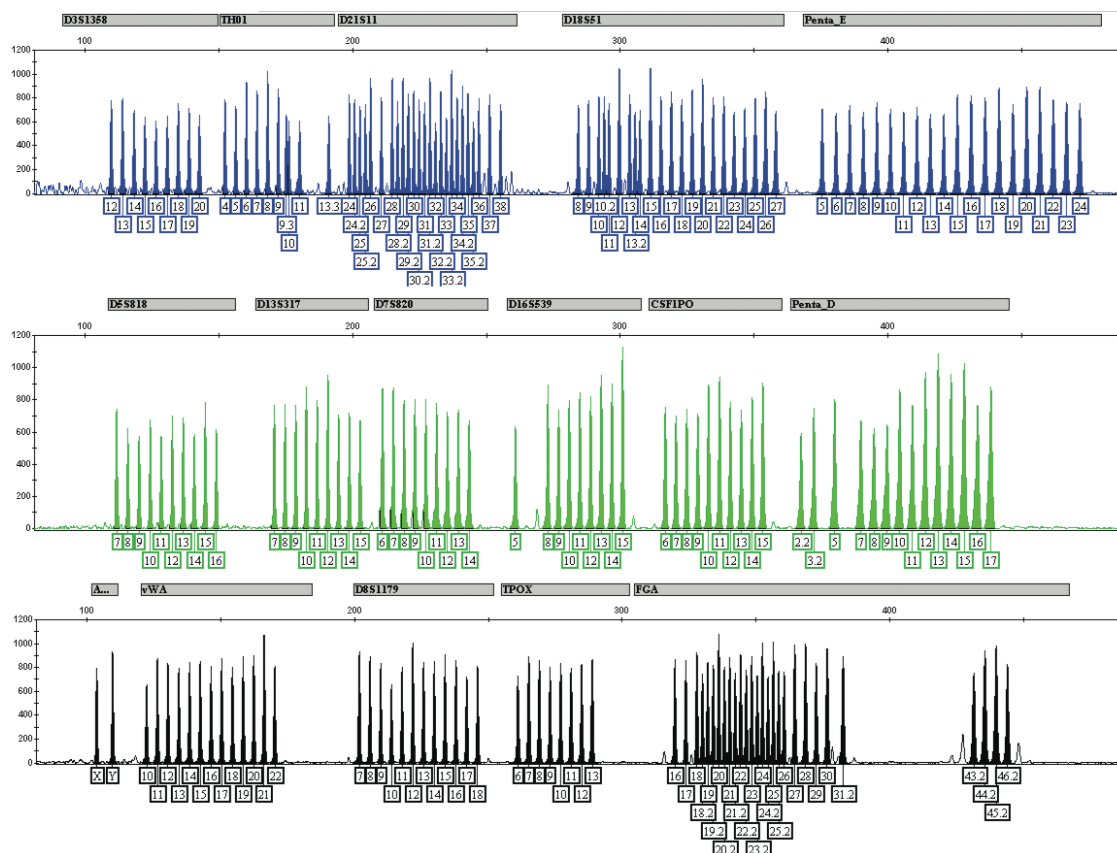


Figura 11. Electroferograma con los ladders alélicos de los 15 marcadores del Kit PowerPlex16HS.

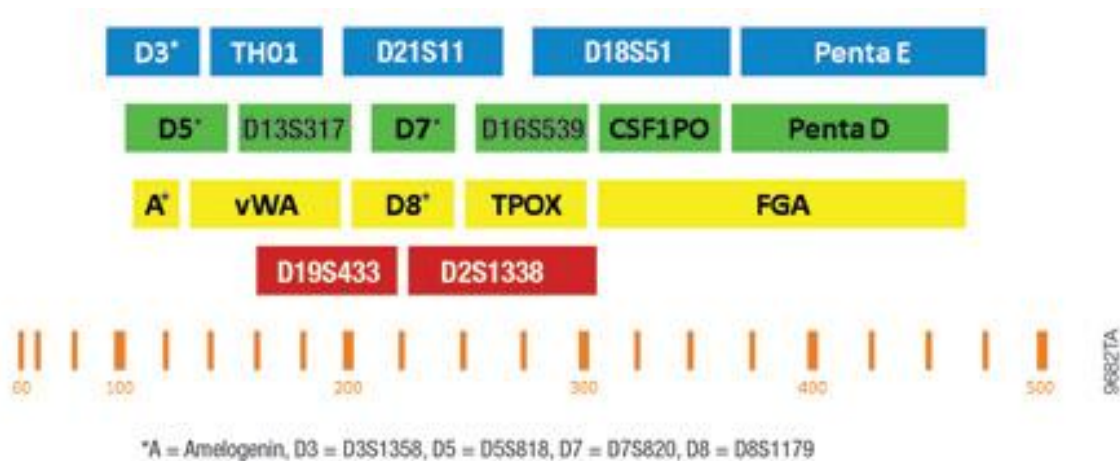


Figura 12. Sobrelapamiento de loci resuelto por el uso de 4 cromógenos distintos. kit Powerplex16HS®. Tomadas de PowerPlex 16HS System Technical Manual.

### 1.18.3 Alelos STR amplificados vía PCR.

Si bien un alelo, en un locus STR, corresponde a una serie de repeticiones en tándem, experimentalmente se detecta como un producto de amplificación, que posee una longitud, generalmente mucho mayor que la secuencia misma de repetición; así, la zona repetitiva está flanqueada en los extremos 3' y 5', por secuencia no repetidas que pueden alcanzar cientos de pares de bases.

Aun así, la diferencia entre uno y otro alelo del mismo locus STR, puede detectarse con facilidad porque las secuencias flanqueantes son constantes en longitud dentro del amplicón. Actualmente, la electroforesis capilar permite discriminar alelos STR hasta de 1pb de diferencia.

Cada fragmento amplificado contiene la zona repetitiva del STR, flanqueada por secuencias no repetitivas en cada extremo, que inician con la zona de hibridación de los *primers, forward y reverse*. (**Figura 13**)



**Figura 13.** Representación de la estructura general de un alelo STR amplificado

Así, el alelo 7 de un STR tetranucleotídico como el del locus D8S1179, puede estar conformado por 7 unidades ATCT, lo que suma 28 pb, pero el fragmento amplificado completo que se detecta en el analizador genético mide 420pb.

Alelos completos: es más común observar alelos que presentan diferencias por una o más unidades de repetición completas, por ejemplo en STRs tetranucleotídicos

simples, los alelos de un mismo locus pueden diferenciarse entre ellos solo por la ganancia o la pérdida de unidades de repetición completas.

En STRs tetranucleotídicos complejos, pueden observarse alelos incompletos donde la diferencia entre uno y otro, es solo una parte de una repetición, ya sea de 1, 2 o 3 pb. En este caso la nomenclatura acordada por la DNA Commision de la *International SocietyforForensicHaeomogenetics* (ISFH) (15) es: .1, .2 o .3 respectivamente. Ej. Para el locus CSF1P0, entre el alelo 10 y el 11 se han descrito 3 alelos intermedios incompletos: 10.1, 10.2 y 10.3. (**Tabla 3**).

ALELO	LONGITUD
10	307 bp
10.1	308 bp
10.2	309 bp
10.3	310 bp
11	311 bp

**Tabla 3.** Diferencia en la longitud de alelos intermedios en el locus STR, CSF1P0

## **2. Objetivos**

### **2.1 Objetivo General**

Realizar el análisis de 811 eventos de mutación detectados en 15 loci Microsatélites tetra-nucleótidos de uso común en identificación humana, caracterizando su comportamiento molecular, su impacto en situaciones de interés forense y su posible significado evolutivo y genético-poblacional.

### **2.2 Objetivos específicos**

1. Evaluar el efecto de variables demográficas como el sexo, la edad y el origen geográfico, sobre el comportamiento mutacional de los marcadores tipo STR estudiados en la población colombiana.
2. Calcular las tasas y frecuencias mutacionales de 15 loci STRs de amplio uso forense, para la población colombiana, a partir del análisis directo de transmisiones parentales, en una muestra de casos de paternidad y maternidad investigadas.
3. Realizar un estudio descriptivo de la dinámica mutacional de 15 loci STR de uso común en identificación humana y hacer inferencias poblacionales y evolutivas sobre el comportamiento de los eventos mutacionales detectados.

### 3. Metodología

#### 3.1 Universo:

Todos los casos de investigación de la paternidad o maternidad que fueron procesados desde el mes de enero del año 2007 al mes de diciembre del año 2012, en el laboratorio del Grupo de Genética Forense del Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, procedentes de todo el territorio nacional, con resultados no excluyentes de la paternidad o la maternidad. Esta es la misma muestra que se utilizó para el análisis de la dinámica mutacional de los 15 loci STR estudiados.

#### 3.2 Muestra para el cálculo de tasa mutacionales:

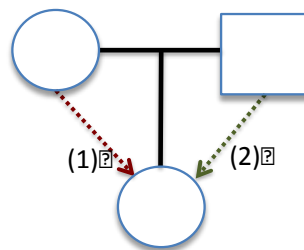
Todos los casos de investigación de paternidad o maternidad donde el resultado de la prueba genética fue de No-exclusión del parentesco investigado, en los Cuáles se encontró al menos una incompatibilidad aislada, es decir, donde en un panel de 15 marcadores tipo STR autosómicos, todos los loci estudiados, menos uno (o dos) de ellos, fueron compatibles con la paternidad o la maternidad biológica, según la investigación forense planteada. Todos estos casos fueron reprocesados desde la extracción del ADN y las exclusiones aisladas fueron verificadas. Adicionalmente, se amplió la tipificación con al menos 2 kits multiplex STRs comerciales de uso común en los laboratorios forenses. Finalmente, en todos los casos con mutaciones detectadas, se alcanzaron valores de certeza superiores a los estándares nacionales ( $W=99,9\%$ , Ley 721 de 2001), para la Probabilidad de Paternidad ( $W$ ), calculada con métodos bayesianos clásicos, utilizando un valor de probabilidad *a priori* de 0,5. El valor  $W$  promedio alcanzado fue de 99,9999%; igualmente el valor promedio observado para el Índice de paternidad (IP) o *likelihood ratio* fue alto.

### 3.3 Conteo de meiosis

Como población referente para el conteo de transferencias alélicas, se tomaron todos los casos de investigación de la paternidad o maternidad que fueron procesados desde el mes de enero del año 2009 al mes de diciembre del año 2012, en el laboratorio del Grupo de Genética Forense del Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, procedentes de todo el territorio nacional, con resultados no excluyentes de la paternidad o la maternidad. Esta muestra es inferior a la utilizada para el análisis de la dinámica mutacional, debido a que su validez está supeditada a la calidad de la documentación existente en bases de datos, sobre los tipos de casos que analizó el laboratorio, en los diferentes escenarios de la investigación del parentesco. Por otra parte, entre los años 2008 y 2009, se implementó el sistema informático (SIFMELCO – Sistema de información forense y de medicina legal de Colombia) que permite realizar conteos y consultas sobre las bases de datos de forma confiable.

Esta herramienta informática permitió el conteo detallado de meiosis maternas y paternas en cada modalidad de cotejo. Con el fin de estimar con mayor exactitud la tasa mutacional se incluyeron los siguientes escenarios:

1. Tríos simples: Un (1) Presunto padre, Una (1) madre y Un (1) hijo, donde se investigó la paternidad biológica: Cada progenitor aporta un alelo en cada locus, por lo tanto se cuenta una (1) meiosis para cada parental. Corresponde históricamente a la mayoría de los casos estudiados (>80%) (**Figura 14**).



**Figura 14**

2. Un (1) Padre y Una (1) madre con más de un hijo: Se cuenta una meiosis por cada hijo para cada parental. Ej. Padre, madre y tres hijos investigados (obviamente no se cuentan los hijos que resulten excluidos con el presunto padre); habría 6 meiosis analizadas en la prueba: 3 por el padre y 3 por la madre. Con dos hijos, se cuentan 4 meiosis (**Figuras 15 y 16**).

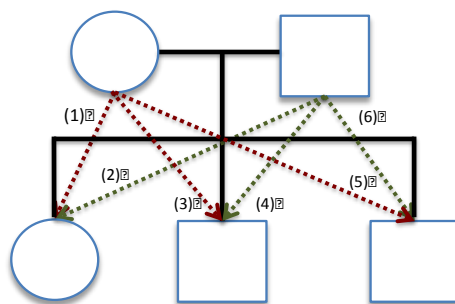


Figura 15

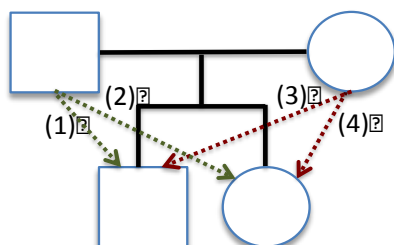


Figura 16

3. Paternidad con madre ausente (PMA). O Maternidad con padre ausente (MPA). En este caso solo se cuenta una meiosis por cada hijo analizado en cada caso (**Figura 17**).

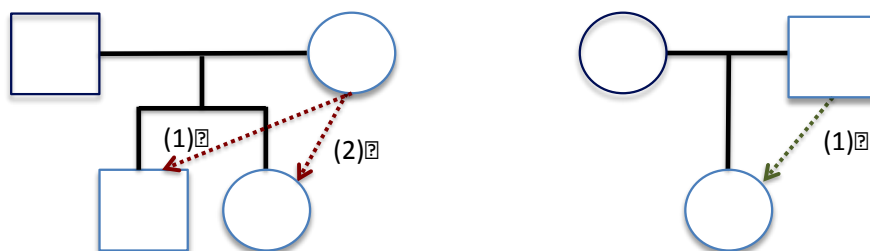


Figura 17

4. Finalmente, los conteos se multiplican, por cada hijo adicional que se analice en el estudio.
5. Presunto padre, presunta madre e hijo(s), cuando se investiga el parentesco en los dos posibles progenitores. En estos casos, una vez demostrada la no exclusión del parentesco padre o madre , se contaron una (1) meiosis por cada progenitor y este valor se multiplicó por el número de hijos presentes en el caso.

### **3.4 Criterios de exclusión:**

Todos los casos donde se encontraron 3 o más exclusiones de paternidad o maternidad fueron eliminados del estudio. Solo se incluyeron casos de tres incompatibilidades siempre que en una de ellas no se pudiera definir si el origen fue el padre o la madre.

Los individuos muestradantes, vinculados a investigaciones de la paternidad o maternidad, que no autorizaron el estudio en el consentimiento informado.

Todos los casos de investigación de la paternidad con presunto padre fallecido, a partir de los presuntos abuelos paternos del menor. Porque aunque es posible su detección, no puede saberse si la mutación se originó en el padre fallecido o en el(la) abuelo(a).

Todos los casos de reconstrucción de perfil genético del padre o del abuelo paterno fallecidos, por la imposibilidad de detectar eventos de mutación sobre perfiles deducidos parcialmente.

### **3.5 Muestras biológicas:**

Para las personas vivas, las muestras de sangre venosa se tomaron sobre tarjetas FTA<sup>®</sup>*Classiccard* de Whatmanclásicas, luego de que los usuarios firmaron el respectivo formato de consentimiento autorizando el uso anónimo de sus muestras.



Ocasionalmente, cuando el padre era fallecido, el estudio se realizó por orden judicial a partir de una muestra de sangre cadavérica de archivo o a partir de muestras óseas o dentales obtenidas por exhumación debidamente ordenada por el juez de familia respectivo.

### **3.6 Muestra poblacional analizada**

La población estudiada comprende 50.827 casos de investigación de la paternidad o la maternidad estudiados en un periodo de 6 años (2007-2012) en el Laboratorio de Genética Forense del Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia (INML), en el marco del Convenio interinstitucional con el Instituto Colombiano del Bienestar Familiar (ICBF).

Se caracterizaron 811 eventos de exclusión aislada de la paternidad en el periodo indicado. En todos ellos se obtuvo información directa del resultado del laboratorio y fue posible la verificación de cada evento sobre los registros electroferográficos que generaron los analizadores genéticos utilizados. Esta información permitió realizar el estudio de la dinámica mutacional de la totalidad de los 811 eventos detectados.

Por otra parte, el cálculo de las tasas mutacionales, se realizó a partir de 613 eventos (de los 811 totales), detectados en 32.216 casos analizados (de los 50.827 totales), entre enero de 2009 y diciembre de 2012. De los años 2007 y 2008 no se obtuvo información detallada que permitiera un conteo directo de las posibles meiosis paternas o maternas y en algunos casos no se reportaron a la base de datos los eventos de exclusión aislada, por lo cual no se tuvieron en cuenta para esta parte del estudio.

Los 32.216 casos generaron 64.772 cotejos padre/hijo o madre/hijo; 9.084 resultaron excluyentes de la paternidad o la maternidad (3 o más exclusiones), dejando un total de 55.688 meiosis, lo que generó un total 835.320 transferencias alélicas contemplando todos los 15 loci analizados. Los casos sin mutación fueron documentados uno a uno a partir de la información almacenada en el Sistema de

Información Forense de Medicina Legal Colombia SIFMELCO del Instituto Nacional de Medicina Legal y Ciencias Forenses. Dependiendo del número de hijos en cada caso se determinaron las meiosis maternas y paternas con las Cuáles se realizaron los cálculos de tasa de mutación.

### **3.7 Análisis de laboratorio**

#### **3.7.1. Métodos de extracción de ADN**

Sangre sobre tarjetas *FTA*<sup>®</sup> : se siguió el protocolo del fabricante, utilizando el reactivo *FTApurificationreagent*<sup>®</sup>, ambos de la casa *Wathman*, para el lavado inicial y TE (10mM Tris HCl, 0,1mM EDTA, pH 8.0) como buffer de lavado final o Agua deionizadaestéril.[www.whatman.com](http://www.whatman.com)

Las muestras de sangre cadavérica en mancha sobre tela, se procesaron por el protocolo de resinas quelantes tipo *Chelex*<sup>®</sup> 100 de *BioRad* de 100 mesh, forma sódica, con incubación inicial a 56°C y ebullición final por 8 min con agitación en cada paso. El ADN se tomó del sobrenadante luego de al menos 3 minutos de centrifugación a 13.000 rpm.

Las muestras de restos óseos se procesaron según el protocolo del INML basado en el del ForensicScienceService, Gill , P., et al 1994, Brevemente: Se realizó una limpieza mecánica y química con SDS al 10% de un fragmento de diáfisis de hueso largo, se pulverizaron 4 g de tejido óseo,previo pulimiento del periostio con motortool y fresa de acero. El hueso se fracturó y sepulverizó en un molinoReischt, en copas de acero inoxidable con revestimiento de Circonio.

El polvo de hueso se decalcificó en EDTA 0,7M y se incubó en un buffer de extracción que contiene ademásProteinasa K,250 uL(10mg/ml), Tween 20 (80uL) a 56°C o/n. Se adicionó posteriormente Fenol saturado con Tris Buffer y equilibrado con HCl a un pH: 8,6. La fase acuosa con ADN en solución se sometió luego a purificación con Cloroformo y Alcohol isoamílico y se procedió a su ultrafiltración en tubos *Centricon* o *Amicón 100*<sup>®</sup>de la casa *Whatman*con membranas de celulosa

regenerada (100.000 daltons). Los extractos fueron cuantificados por RT-PCR utilizando los kits: *QuantifilerDuo*<sup>®</sup> de la casa *AppliedBiosystems* o *Plexor*<sup>®</sup> de la casa *Promega Co* en equipos ABI 7500.

### **3.7.2. Amplificación vía PCR-multiplex**

Se siguieron los protocolos de los kits comerciales: *Identifiler*<sup>®</sup> y *SEfiler*<sup>®</sup> de la casa comercial *AppliedBiosystems Co* y los *Kits PowerPlex 16*<sup>®</sup> y *FFFL*<sup>®</sup> de la casa *Promega Co*. Los kits anteriores permiten la co-amplificación de un total de 22 loci STR diferentes,

En los últimos 2 años, los casos con exclusiones aisladas de paternidad o maternidad fueron también analizados con 6 marcadores STRs adicionales incluidos en kits multiplex de nueva generación como *NGMplus* de la casa *AppliedBiosystems* y *Powerplex 17 ESI* o *ESX*, para un total de 28 marcadores. Todos los productos de amplificación están marcados en uno de sus primers, con fluorocromos que posteriormente pueden ser detectados en secuenciadores o analizadores genéticos automatizados.

Se utilizaron termocicladores: 9700 y *Veritide* de la casa *AppliedBiosystems* y *PTC100* de *JM Research*, previamente verificados y calibrados. Toda reacción de PCR se montó con los controles positivos y negativos correspondientes.

### **3.7.3. Detección y asignación de alelos STR**

Las muestras amplificadas fueron corridas en secuenciadores o analizadores genéticos automatizados ABI 3100, ABI3130XL o ABI 3500XL de *AppliedBiosystems*, basados en electroforesis capilar, con *multiarrrays* de 16 y 24 capilares simultáneos. Las señales detectadas por la excitación de los fluorocromos a través de un láser de Argón, fueron colectadas por el Software *Data Collection* de la firma ABI y la asignación alélica fue realizada con el software *Genemapper ID-X v 1.2*. Cada kit contiene un *ladder* o escalerilla alélica que permite la asignación precisa

de los alelos, estableciendo “bins” o ventanas de aceptación con una precisión de 0,5pb.

#### **3.7.4. Control de calidad**

Todos los ensayos de Extracción y purificación de ADN, Amplificación vía PCR-multiplex de STRs, Cuantificación por RT-PCR y Detección electroforética, están validados y acreditados en el laboratorio por la norma ISO-IEC17025 v 2005 con entidades acreditadoras nacionales: SIC y ONAC.

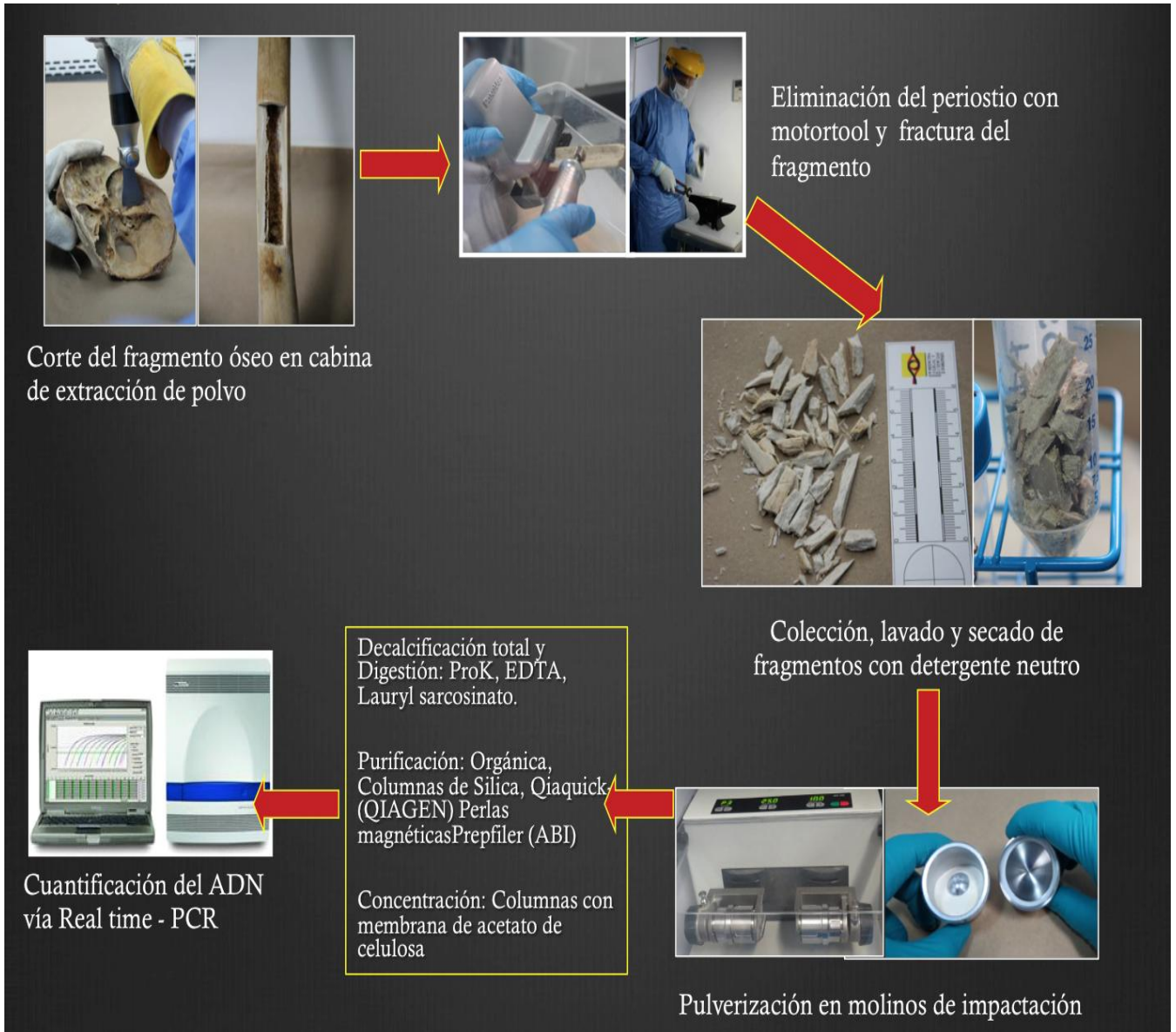
Todos los resultados excluyentes fueron repetidos desde las muestras originales del presunto padre e hijo, las exclusiones de segundo orden se reprocesaron con kits de otra casa comercial. En todos los casos donde se encontró una exclusión aislada, se analizaron al menos 22 marcadores en total. Todo informe pericial fue revisado por un segundo perito, basándose en el protocolo de revisión establecido en el sistema de calidad del laboratorio.

### **Figura 18. ESQUEMÁS DE LA METODOLOGIA UTILIZADA**



1. EXTRACCION MANUAL O AUTOMATIZADA DE ADN DE SANGRE EN TARJETAS FTA



## 2. EXTRACCION MANUAL O AUTOMATIZADA DE ADN DE TEJIDOS CALCIFICADOS



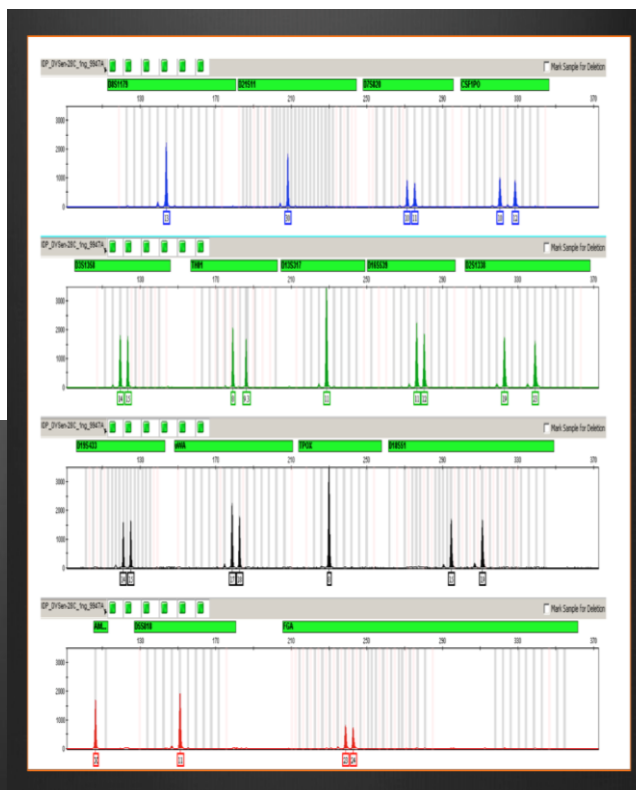
3. AMPLIFICACION VIA PCR- MULTIPLEX DE STRs Kit AmpFISTRidentifiler

Locus designation	Chromosome location	Alleles included in AmpFISTR® Allelic Ladder	Dye label	Control DNA 9947A
D8S1779	8	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	6-FAM™	13†
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		30‡
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 11
CSF1PO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 12
D3S1358	3p	12, 13, 14, 15, 16, 17, 18, 19	VIC®	14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		8, 9.3
D13S317	13q22-31	8, 9, 10, 11, 12, 13, 14, 15		11§
D16S539	16q24-qter	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15		11, 12
D2S1338	2q35-37.1	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		19, 23
D19S433	19q12-13.1	9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2	NED™	14, 15
vWA	12p12-pter	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17, 18
TPOX	2p23-2pter	6, 7, 8, 9, 10, 11, 12, 13		8††
D18S51	18q21.3	7, 8, 9, 10, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		15, 19
Amelogenin	X: p22.1-22.3 Y: p11.2	X, Y	PET®	X
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11‡‡
FGA	4q28	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		23, 24

† For CODIS purposes, profile reported as 13, 13.  
‡ For CODIS purposes, profile reported as 30, 30.  
§ For CODIS purposes, profile reported as 11, 11.  
†† For CODIS purposes, profile reported as 8, 8.  
‡‡ For CODIS purposes, profile reported as 11, 11.

#### 4. ELECTROFORESIS CAPILAR SEMI-AUTOMATIZADA



#### 3.7.5 Criterios de análisis

### 3.7.5.1 Definición del tipo de mutación

#### Mutaciones tipo “null”

Se asumieron como tipo “*null*” las exclusiones de segundo orden que tuvieran más de dos repeticiones de diferencia. Es decir, cuando el hijo y el presunto padre son homocigotos distintos y se presume que existe una mutación por substitución, en una base que hace parte de la zona de reconocimiento del primer, en su extremo 3'. Si con el segundo kit se recupera el alelo “silente” se asume que había una mutación tipo *null* y no se tuvo en cuenta en el inventario de mutaciones. Solo se incluyeron aquellas que no mostraron recuperación del heterocigoto con el uso de un kit alternativo.

Para el marcador D19S433, se asignaron como mutaciones tipo “null”, también aquellas que mostraban homocigotos tanto en padre o madre como en hijo, así fuesen de 2 o más pasos, ya que este marcador presentó un alto nivel de eventos de substitución como se verá en los resultados.

#### Mutaciones tipo “slipp”

Se asumieron como mutaciones tipo “*slipp*” toda incompatibilidad aislada de uno o de dos pasos de diferencia entre el alelo paterno o materno que se presume como el origen, y el alelo mutado del hijo. Se incluyeron también en esta categoría, casos donde se observaron cambios de más de dos pasos entre individuos heterocigotos diferentes.

Finalmente, exceptuando el marcador D19S433, se constituyó un tercer grupo llamado “indeterminado”, para incluir aquellos casos donde padre e hijo fueron homocigotos y la diferencia entre sus alelos era de uno o de dos pasos, ya que existe la posibilidad de que se trate de alelos nulos como de eventos tipo “*slipp*”, aunque estos últimos sean los más probables.

### 3.7.5.2 Determinación del origen parental de la mutación

Origen paterno



Se puede establecer sin duda que el padre genera la mutación.

Ej. Padre: 8, 11      Madre: 13,15      Hijo: 9, 13

Origen Materno

Se puede establecer sin duda que la madre genera la mutación

Ej. Padre: 8, 10      Madre: 13,15      Hijo: 10, 14

Origen indeterminado

No es posible asignar el origen con exactitud

Ej. Padre: 15, 16      Madre: 15, 16      Hijo: 14, 16

Ej. Padre: 16, 18      Madre: 16, 16      Hijo: 16, 17

Ej. Padre: 11, 11      Madre: 11,11      Hijo: 11, 12

### 3.7.5.3 Determinación de la dinámica de ganancia o pérdida de repeticiones

Ganancia:

Ej. Padre: 8, 12      Madre: 13,14      Hijo: 9, 13

Ganancia de 8 a 9 del padre

Pérdida

Ej. Padre: 15, 16      Madre: 15, 16      Hijo: 14, 16

Pérdida de 15 a 14 indeterminada en origen

Indeterminada

Ej. Padre: 8, 10      Madre: 13,15      Hijo: 10, 14

Ganancia 13 a 14? o Pérdida 15 a 14?

### 3.7.6 Frecuencia de mutación

Número de veces en que aparece una mutación en particular en la población estudiada.

### 3.7.7 Tasa de mutación

Definida como el número de eventos de exclusión aislada / número total de meiosis analizadas. Se calculó con intervalo de confianza del 95% para una distribución normal: <http://statpages.org/confint.html>.

### **3.7.8 Análisis estadísticos**

La significancia estadística de los hallazgos relacionados con la mutación y el origen parental, el estado de ganancia o pérdida de repeticiones, el tipo de mutación (slipp o null) se calculó con una prueba de Z asumiendo un alfa de 0,05.

### **3.7.9 Consentimiento informado**

Todos los datos provenientes de usuarios como la edad, sexo, origen geográfico del caso y perfiles genéticos, son anónimos en todos los documentos publicados. Los códigos originales de los casos se mantuvieron para realizar las verificaciones de los perfiles que presentaron eventos de mutación y para las consultas estadísticas a la base de datos. Para los archivos magnéticos que son soporte de la investigación, los códigos originales fueron reemplazados por una nueva numeración. Se utilizó un formato de consentimiento informado institucional, donde el usuario citado por la autoridad para la realización del estudio genético de la paternidad o maternidad, podría autorizar la toma voluntaria de su muestra sanguínea para la realización de la prueba judicial y adicionalmente autorizar o rechazar el uso de su perfil genético para estudios genético poblacionales, una vez informado del compromiso institucional de mantener los datos anónimos.

## **4. Resultados y Análisis**

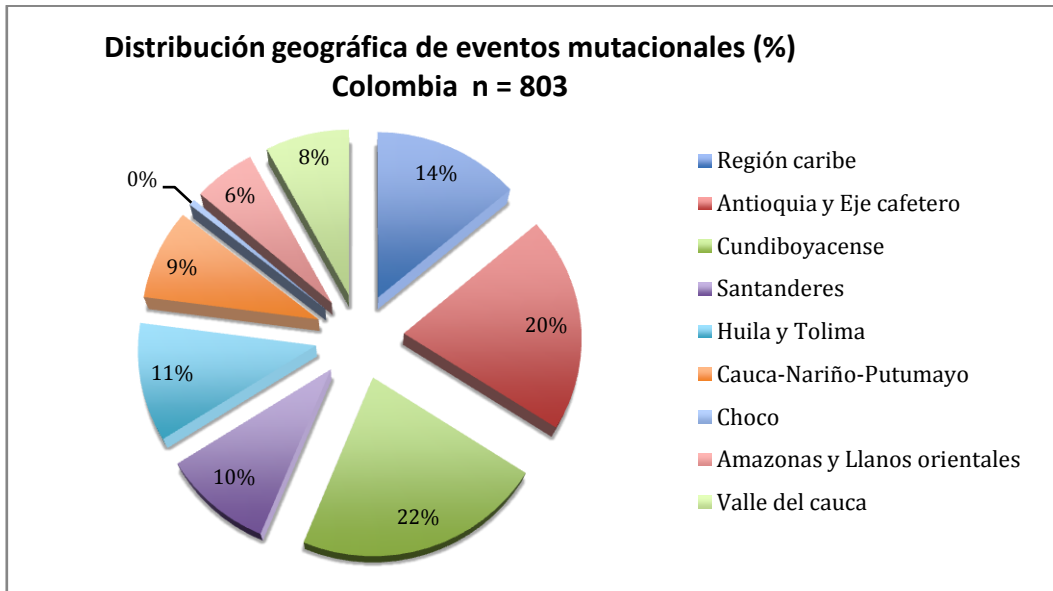
En 50.827 casos de investigación de la paternidad o la maternidad estudiados en un periodo de 6 años (2007-2012), se detectaron 811 eventos mutacionales en 15 STRs todos del tipo tetranucleótido, de uso común en los laboratorios forenses a nivel mundial.

#### **4.1 Distribución geográfica de los eventos mutacionales en Colombia**

Los eventos mutacionales detectados provienen de 30 de los 32 departamentos colombianos. En el Vichada y el Vaupés no se presentaron mutaciones; también, el número de casos de paternidad investigados en estos dos departamentos amazónicos fue mínimo. En la **Tabla 4**, se muestra el número de mutaciones detectadas por departamento, los datos se agrupan por regiones etno-geográficas y por porcentajes en la **Tabla 4** y el **Gráfico 2**. En 803 casos se contaba con la información de origen.

Por lo general, las mutaciones observadas siguen la misma distribución geográfica de las demandas por paternidad irresponsable en Colombia, lo que a su vez refleja la distribución de la población actual del país. (**Gráfico 3 y Tabla 5**). No se observa un patrón que sugiera la concentración de estos eventos en alguna población o subpoblación colombiana en particular. Considerando que se trata de polimorfismos de ADN no expresivos, no se espera que las mutaciones detectadas estén asociadas a factores geográficos o ambientales. Aun así, podrían realizarse estudios complementarios de ancestría para los individuos que originaron la mutación, una población de interés que sugiere el presente estudio es la de los Santanderes, donde el porcentaje de eventos mutacionales (78 eventos: 9,8% del total) supera notablemente el porcentaje de demandas de paternidad (6,2%). El estudio podría realizarse con los marcadores tipo InDels de ancestría que actualmente se analizan en poblaciones colombianas con el apoyo del ghep-ISFG. Finalmente, las mutaciones detectadas en las regiones amazónicas y de los llanos orientales del país, aunque presentan diferencias en porcentaje muy notables con respecto al

número de demandas de paternidad analizadas, la cantidad de casos es muy baja y esta diferencia podría no ser real.



**Gráfico 2.** Distribución geográfica de 803 eventos mutacionales en Colombia

Departamento	Nro. de mutaciones
BOGOTA	82
ANTIOQUIA	94
VALLE	63
CUNDINAMARCA	59
ATLANTICO	14
SANTANDER	58
BOLIVAR	21
NARIÑO	37
CORDOBA	21
TOLIMA	47
CAUCA	26

NORTE DE SANTANDER	20
BOYACA	35
MAGDALENA	15
HUILA	42
CESAR	20
CALDAS	40
RISARALDA	19
META	21
GUAJIRA	5
SUCRE	14
QUINDIO	8
CHOCO	5
CAQUETA	8
CASANARE	6
PUTUMAYO	7
ARAUCA	10
GUAVIARE	1
SAN ANDRES	2
AMAZONAS	1
VICHADA	0
VAUPES	0
GUAINIA	1
Suma	803
sin dato	8
Total	811

**Tabla 4.** Número de mutaciones detectadas por Departamento

Región geográfica analizada	% Población Censo - 2005	% de demandas (Infs.mensuales Convenio IML-ICBF)	% de mutaciones
Caribe e insular	21,4	12	13,8
Andina Nor-Occidental: Antioquia y Eje cafetero	17,8	19,4	20,3
Altiplano cundiboyacense	25,9	28,1	22,1
Nororiente - Santanderes	7,6	6,2	9,8
Andina Sur-oriental: Huila y Tolima	5,4	11,2	11,1
Andina Sur-Occidental: Cauca, Nariño y Putumayo	7,3	12,7	8,6
Pacífico: Choco	1	0,4	0,6
Amazonas y Llanos orientales	4,7	2,6	5,8
Valle del cauca	8,9	7,4	7,9
	100	100	100

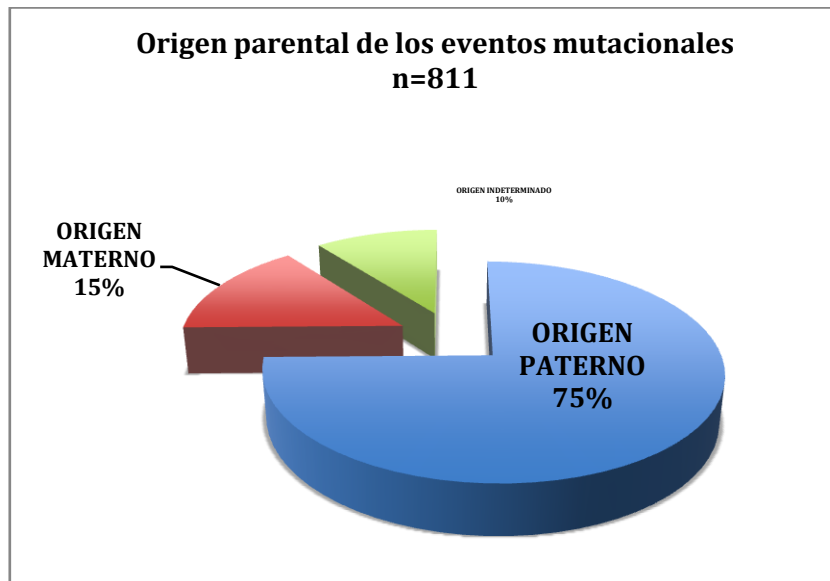
**Tabla 5.** Distribución de 803 eventos mutacionales por región etno-geográfica



**Gráfico 3.** Distribución geográfica de 803 eventos mutacionales en Colombia

## 4.2. Origen parental de los eventos mutacionales

En los 811 eventos estudiados, se determinó el origen parental de la mutación: en 608 casos (75%) las mutaciones se originaron en el padre, en 121 (14,9%) la madre fue el origen del evento y en 82 casos (10,1%), no fue posible determinar el parental (**Gráfico 4**).



**Gráfico 4.** Origen parental de los eventos mutacionales

La relación entre los orígenes, paterno: materno fue de 5:1. Si bien la diferencia es significativa ( $p > 0,05$ ), y siempre se ha encontrado un predominio de mutaciones provenientes de los padres, las proporciones varían con otros estudios mutacionales realizados sobre microsatélites:

Sun y Cols. en 2012 (44), analizaron 2477 loci STRs autosómicos, a partir de 85.289 islandeses y encontrando un radio de 3,3:1 entre hombre y mujer. Brinkmann(55), encontró un radio de 17:3 a favor de los padres. Estos dos estudios se ha hecho en condiciones distintas: el primero se presenta como el más amplio realizado a la fecha, sobre 24.832 tríos padre-madre-hijo pero sin evidencia de parentesco confirmado, ya que se trata de casos médicos donde se realizaron mapeos génicos de enfermedades basados en análisis de ligamiento y obtuvieron un error promedio



de genotipado de genotipos de  $1.8 \times 10^{-3}$ , que es casi la tasa mutacional esperada. Detectaron 466 eventos mutacionales en loci tetranucleotídicos y 1487 en dinucleótidos. El segundo estudio se realizó a partir de casos de paternidad en población alemana donde detectaron solo 23 eventos mutacionales.

Nuestros datos son más cercanos al estudio de Sun y Cols y al de Lu, realizado en población China en el 2012 (64) sobre 3.935 casos de paternidad en los que se demostró el parentesco, encontrando 195 mutaciones distribuidas en una relación 4.3:1 según su origen se paterno a materno respectivamente.

Por su parte, Huang y Cols.(10) analiza en 53 pedigrees multi-generacionales, 362 loci STR dinucleotídicos encontrando 97 mutaciones estimaron una proporción 1:1, pero la mayoría de los eventos no pudieron ser determinados en su origen.

Al parecer las proporciones encontradas en relación al progenitor origen de la mutación varían dependiendo del tamaño de la muestra analizada, en la Cuál, una muestra pequeña puede sesgar los datos poblacionales como posiblemente sucedió con el estudio de Brinkmann; por otra parte, la longitud de la repetición en el STR parece influir también en la diferencia de resultados. Llama la atención que en el estudio de Huang se obtengan proporciones iguales para los dos sexos en loci dinucleotídicos. Al parecer, para STRs de repeticiones de 2 bases, las mutaciones no solo están influenciadas por el número de eventos de replicación pre-meióticos que se observan en el varón. No se descarta la influencia epigenética como causa del mayor número de mutaciones STRs provenientes del padre (imprinting genómico).

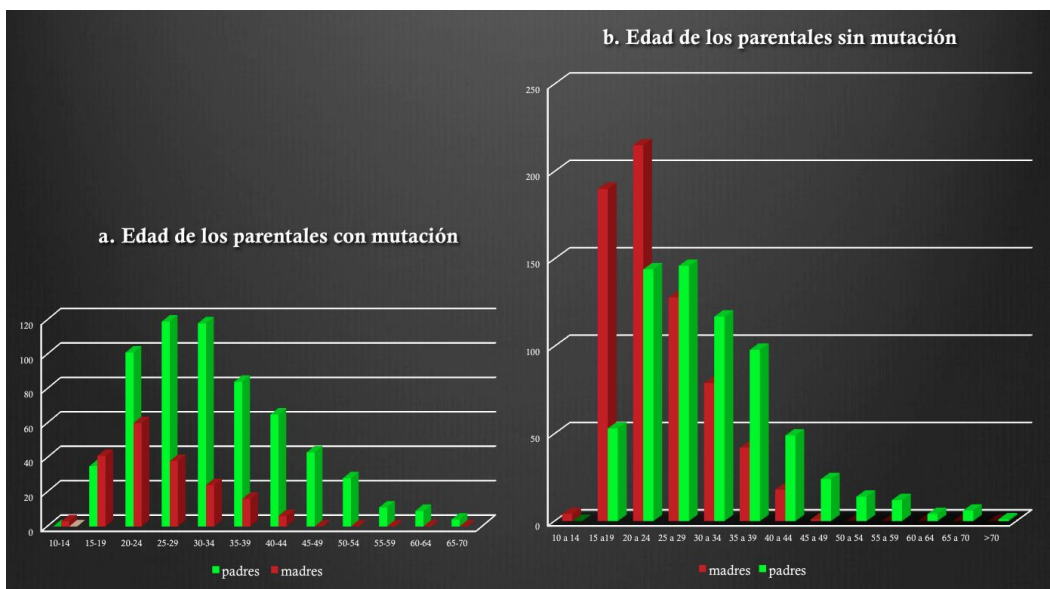
### **4.3. Edad y Mutación**

En 620 padres y 191 madres, de los 811 eventos estudiados, se logró información sobre la edad de los parentales que fueron el origen de la mutación. El valor registrado corresponde a la edad parental para la fecha de nacimiento del hijo.

Se detectaron eventos mutacionales en un muy amplio rango de edades, desde los padres adolescentes, hasta los padres adultos mayores. Por otra parte, la edad de los padres en casos de paternidad demostrada, sin mutación se utilizó para calcular el número de meiosis contenido en cada grupo de edad. **(Gráfico 5).**

La distribución de eventos mutacionales en los individuos estudiados esta sesgada por el número de casos que se presentan, según la edad de los parentales involucrados; se observan menos eventos mutacionales en edades mayores porque el número de demandas es también menor en estas edades.

Aun así, la diferencia en el sexo del origen parental se mantiene y es más notable en edades mayores (Gráfico 5a).



**Gráfico 5.** Distribución de edad de los parentales con mutación y sin mutación

El tema llama la atención si se tiene en cuenta que en muchos países europeos la edad de los hombres para tener hijos ha ido en aumento, la paternidad se ha ido retrasando. Para los padres primerizos en España ahora es de 34 años en promedio, comparada con 31 años de las madres españolas. En 2011 la edad media en que las mujeres de la península tuvieron su primer hijo fue de 31,06 años, frente a los 30,98 registrados en 2010 y los 30,87 de 2008. La cifra lleva subiendo varias décadas de forma sostenida. En 1975, primer año en el que el Instituto Nacional de Estadística

de España registró el dato, la media se situaba en los 28,7 años. [www.ine.es](http://www.ine.es) . España no es el único país donde se ha retrasado la paternidad: se trata de un fenómeno que comparte todo el mundo desarrollado y gran parte de los países emergentes o en vías de desarrollo.

En la Encuesta Nacional de Demografía y Salud realizada por Profamilia en 2010, [www.profamilia.org.co](http://www.profamilia.org.co) , la edad promedio de mujeres colombianas para su primer embarazo es de 21,6 años casi 10 años menos que las europeas, asumiendo un rango de fecundidad de 15 a 49 años. (**Tabla 6**). En nuestro estudio, la edad promedio de las mujeres que mutan es de 24,9; aunque no conocemos si se trata de primigestantes, la diferencia es notable en relación con la edad promedio de la madre colombianas, para su primer embarazo. La diferencia con respecto a las mujeres no mutantes analizadas en el presente estudio es menor: 24.3 años.

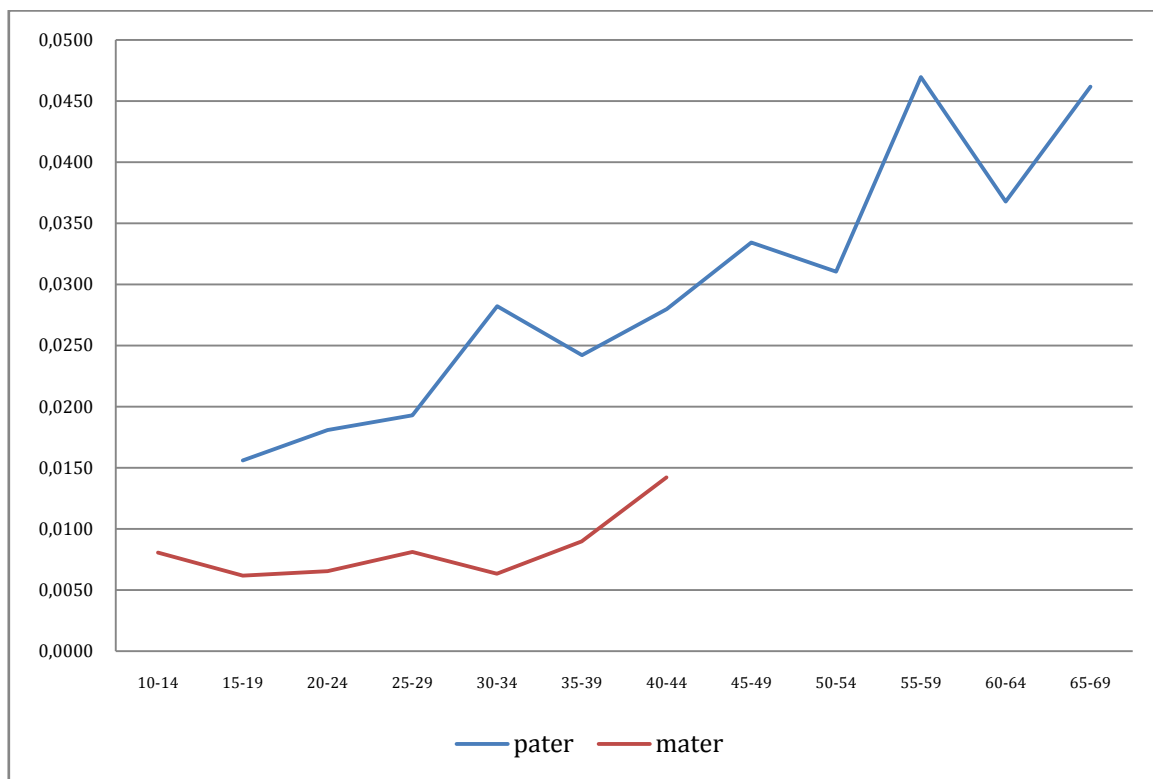
Cuadro 5.7 Edad al primer nacimiento								
Porcentaje de mujeres que ha tenido hijos antes de ciertas edades exactas, porcentaje que no ha tenido hijos y edad mediana al primer nacimiento, según área de residencia y edad actual, Colombia 2010								
Zona y edad actual	Porcentaje de mujeres que dio a luz antes de cumplir:					Porcentaje que nunca ha tenido hijos	Número de mujeres	Edad mediana al primer nacimiento
	15 años	18 años	20 años	22 años	25 años			
<b>ZONA URBANA</b>								
15-19	1.5	10.7	13.8	na	na	86.2	6,944	a
20-24	1.7	17.3	32.4	42.8	47.3	52.7	6,270	a
25-29	1.9	19.4	36.0	50.0	64.2	28.2	5,813	22.0
30-34	2.0	19.7	37.4	52.8	67.4	14.1	5,309	21.6
35-39	2.1	16.7	35.2	51.7	68.6	8.6	4,892	21.8
40-44	1.1	14.3	30.5	47.2	66.9	7.3	5,119	22.4
45-49	0.9	13.1	29.3	44.9	62.9	8.4	4,916	22.7
<b>ZONA RURAL</b>								
15-19	2.5	17.7	22.2	na	na	77.8	2,156	a
20-24	4.7	29.7	53.5	66.3	69.7	30.3	1,490	19.7
25-29	3.2	29.2	52.6	69.9	83.5	11.7	1,514	19.8
30-34	4.1	28.5	51.5	68.3	81.6	8.5	1,478	19.9
35-39	3.8	23.3	48.3	66.7	81.8	4.6	1,398	20.2
40-44	2.6	20.3	41.9	60.3	76.0	2.9	1,364	20.9
45-49	2.6	21.1	43.6	60.8	77.7	3.6	1,155	20.7
<b>TOTAL</b>								
15-19	1.8	12.4	15.8	na	na	84.2	9,100	a
20-24	2.3	19.7	36.4	47.3	51.6	48.4	7,760	23.0
25-29	2.2	21.4	39.4	54.1	68.2	24.8	7,327	21.4
30-34	2.4	21.7	40.4	56.1	70.5	12.9	6,787	21.1
35-39	2.4	18.2	38.1	55.0	71.5	7.7	6,290	21.4
40-44	1.4	15.5	32.9	50.0	68.8	6.4	6,483	22.0
45-49	1.3	14.6	32.0	47.9	65.7	7.5	6,071	22.3

na = No aplicable  
a = Se omite pues menos del 50 por ciento de las mujeres han tenido un nacimiento antes del comienzo del grupo de edad

**Tabla 6.** Edad de la madre al primer nacimiento. Colombia Tomado de: Encuesta Nacional de Demografía y Salud. Profamilia 2010.

Sobre la población total de parentales incluidos en el presente estudio, con y sin mutación, se contaron las transferencias meióticas en cada rango de edad y se obtuvo tasas mutacionales por edad.

Como se observa en el **Gráfico 6**, las tasas de mutación se incrementan con la edad tanto en el hombre como en la mujer, aunque la tendencia es mucho más evidente en el sexo masculino, donde la tasa se duplica en promedio cada 15 años entre las edades de 15 a 50 años; en edades superiores, la tasa se duplica en individuos con solo 5 años de diferencia. En la mujer la tasa de mutación permanece estable hasta los 40 años, cuando asciende notablemente.



**Gráfico 6.** Tasas de mutación por rango de edad y origen parental en Colombia

La edad promedio de los padres que mutan en el estudio de BerndBrinkman de 1998 (55) es 32 años y es estadísticamente significativa si se compara con la edad de los padres no mutados 28 años. En nuestro estudio, los padres con mutación tienen una

edad promedio de 33,4 años, aun mayor que la muestra alemana. La edad promedio para los padres no mutados en Colombia es de 31,1 años. La diferencia de nuevo es significativa. A pesar de la evidencia que relaciona la mayor edad paterna con la actividad mutacional, este efecto no parece ser generalizable a todo tipo de mutación.

En el estudio de Sun y Col. (44), los datos corresponden exclusivamente a mutaciones tipo SNPs (*single nucleotide polymorphisms*). El promedio de edad para los padres fue de 29,7 años y se determinó una relación clara entre mayoría de edad y mutabilidad en el varón; incluso, se llega a pronosticar una tasa de 2 mutaciones nuevas por año. En oposición, el estudio realizado por James Crow en 2006 (54), encuentra que si bien, para las mutaciones por sustitución de bases, la frecuencia incrementa con la edad paterna, las deleciones pequeñas no muestran ningún efecto importante de la edad y ocurren por igual en ambos sexos.

Ahora bien, para las secuencias microsatélites se han aportado también algunas evidencias que sugieren un proceso similar, aun así, son muchos ya los hallazgos que sugieren un proceso de mayor complejidad. Si se acepta que el origen de la variación de los microsatélites está en el proceso replicativo del ADN, puede esperarse que la tasa de mutación esté en relación con el número de divisiones celulares que se suceden en la línea germinal. Además se esperaría que las mutaciones fuesen más frecuentes en hombres que en mujeres e incluso, como para las secuencias expresivas ya mencionadas, el efecto es más notorio en hombres más viejos que en más jóvenes (4). Varios estudios de mutaciones en humanos encuentran 3 o 4 veces más mutaciones puntuales de origen paterno.

Al respecto, como se mencionó antes, B. Brinkmann en 1998 (55), también mencionaba que las tasas de mutación en microsatélites están determinadas por la edad y el sexo del parental, aunque en su estudio, se analizaron solo 23 eventos detectados en 6 loci STR. La edad promedio de los padres que mutaron fue mayor si se compara con la edad de los padres no mutados.

Por su parte, Huang y cols (10), no encontraron ningún sesgo del sexo asociado a las tasas de mutación. En estudios de microsatélites realizados en aves, algunos loci presentaron sesgos a favor del origen paterno y otros a favor del origen materno. Ellegren (4), concluye que la investigación que busca relacionar las tasas de mutación con la edad paterna no ha podido demostrar tal efecto en las secuencias repetitivas.

También sobre los microsatélites, Brohede y cols 2004 (65), abordaron el problema del efecto de la edad del varón sobre la dinámica mutacional. Analizaron la mutación del locus STR hipervariable (D21S1245) a través de la genotipificación del espermatozoides de individuos con mutación.

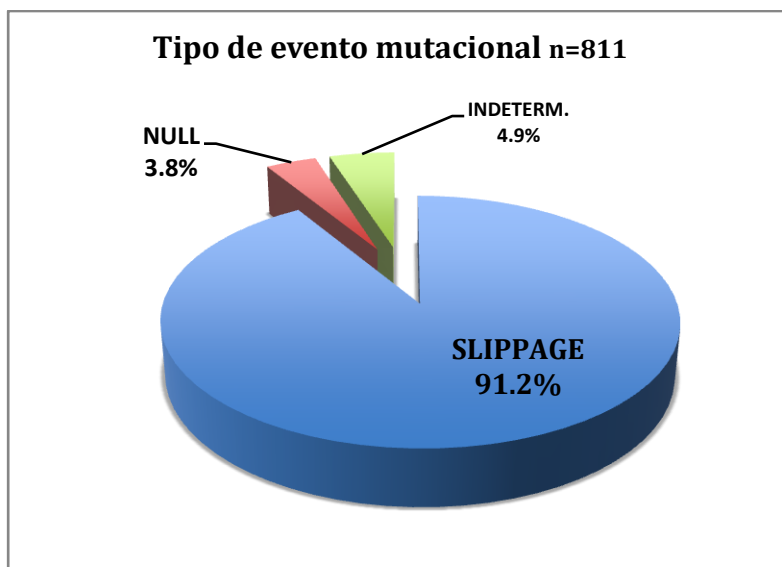
Reconocen que algunos loci pueden variar no solo en la tasa de mutación sino también en el patrón de mutación, además que la tasa de mutación puede variar con el sexo, la edad y posiblemente el *background* genético. Esta de acuerdo en que se requieren estudios de loci individuales y en individuos aislados, lo Cuál requiere básicamente el análisis de un gran número de transmisiones en línea germinal de un marcador particular en un parental dado. Analizaron 8 individuos de diferentes edades entre 18 y 23 años, representando los jóvenes, y un grupo entre 45 y 56 años, a los viejos.

En un estimado de 8623 espermatozoides de los 8 individuos, se detectaron 203 mutaciones en el locus estudiado, una tasa de 0,024. El locus D21S1245 es un STR complejo e interrumpido con un motivo (GAAA)<sub>n</sub>. Todas las mutaciones difieren del progenitor por múltiplos de 4. La tasa de mutación promedio de jóvenes fue más alta que la de adultos. Brohede, intenta explicar los resultados argumentando que los jóvenes presentaron alelos de diferente "linaje filogenético" que los de los adultos; el linaje, explica, son agrupaciones de los arreglos de la secuencia, que presentan diferentes tasas de mutación; sugieren que elementos que actúan en *cis*, es decir, dependientes de la secuencia de ADN del alelos STR pueden ejercer efectos sobre la tasa mutacional.

Finalmente, nuestro estudio no involucra otras secuencias de ADN diferentes a los STRs, ni diferencia las tasas de mutación entre loci de estructura diferente. Aun así, aporta evidencia contundente a favor del incremento de la mutabilidad a medida que avanza la edad en ambos sexos, principalmente en los varones, en secuencias microsatélites tetra-nucleotídicas.

#### 4.4. Mecanismo mutacional: *Slippage vs null*

Como hemos mencionado antes, en los 50.827 casos de investigación de la paternidad o la maternidad analizados, se detectaron 811 mutaciones, 31 de ellas (3.8%), presentaron un patrón de exclusión de segundo orden (padre e hijo son homocigotos diferentes), compatible con mutación tipo *null*. Las restantes, son 740 mutaciones tipo *slipp* (**Grafico 7**).

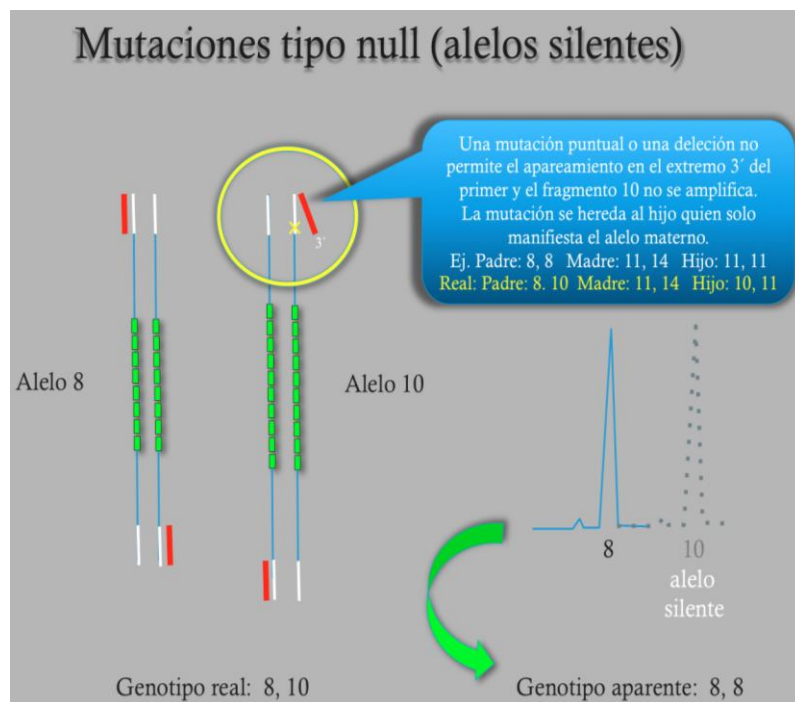


**Gráfico 7** . Distribución de eventos mutacionales según mecanismo de origen (n=811)

En 40 casos, no fue posible establecer el mecanismo mutacional ya que cumplían por igual con condiciones para las dos alternativas de cambio molecular. Por ejemplo: Si el Padre es 7,7, la Madre 8,9, y el Hijo 8, 8. En el Cuál, puede argumentarse la mutación por *slipp* del alelo 7 del padre, ganando una repetición en el hijo, al igual que una pérdida alélica (*allelic dropout*) producida por la falta de

hibridación de uno de los *primers* y la consecuente pérdida de la amplificación del alelo (alelo *null*). **(Gráfico 8)**

Cinco loci concentraron los 31 eventos tipo *null*: D2S1338, D16S539, D18S51, D19S433 y vWA31; llama la atención que el 58% (18 eventos) de las mutaciones nulas ocurrieron en el locus D19S433.



**Gráfico 8** . Mecanismo mutacional tipo null

El locus D19S433 es un microsatélite tetranucleótido, cuyo core es AAGG, que mapea en el cromosoma 19 humano, región 19q2. Sus alelos varían según el número de repeticiones de una secuencia simple:  $(AAGG)_n$ , precedida por cuatro tetranucleótidos así:  $(AAGG)(AAAG)(AAGG)(TAGG)$ . En la base de datos del Instituto Nacional de Estándares Técnicos de los EEUU: [nisthttp://www.cstl.nist.gov](http://www.cstl.nist.gov), están descritos los alelos completos: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 y 20, y



los alelos intermedios : 11.1, 12.1, 12.2, 13.1, 13.2, 13.3, 14.1, 14.2, 14.3, 15.2, 16.2, 17.2 , 18.2 y 19.2.

D19S433, es un marcador propuesto por la comunidad forense europea y se ha incorporado a varios kits comerciales, inicialmente de la casa AppliedBiosystems en los kits Identifiler y SGMplus y recientemente a algunos kits de la firma Promega. (PowerPlex ESI17 y ESX17).

A continuación se muestran todos los 18 eventos de mutación tipo null encontrados en el locus D19S433 (**Tabla 7**). En la tipificación inicial, se habían detectado varias exclusiones aisladas de segundo orden, que al parecer involucraban repeticiones incompletas. Ante este hallazgo se reprocesaron las muestras respectivas utilizando al menos 2 kits de casas comerciales diferentes. (*Identifiler y PowerPlex ESX-17*) Como resultado del análisis se recuperaron heterocigotos en 15 de los eventos previamente asignados como *slipp* de pasos incompletos. En los 3 restantes no se recuperó el alelo *dropout*, pero el número de pasos incompletos sugiere que se trata realmente de un alelo *null*.

Resaltamos el evento 5, en el Cuál, el padre presenta mutaciones tipo null para dos kits, en sus dos cromosomas. Un kit detectó un homocigoto 13.2 y el otro kit un homocigoto 14, este alelo fue también perdido en el hijo. En el caso 7, se observa un evento *null* proveniente de la madre y un evento *slipp* proveniente del padre. En el caso 8, los dos hijos reciben el alelo 15 mutado del padre y aparece como dos homocigotos diferentes.

Nro. evento	Padre		Madre		Hijo 1		Hijo 2	
1	13.2	14	13	15 (13)	14	15 (14)		
2	15	15.2	14	15 (14)	15	15		
3	13,2	15 (13.2)	12	13	13	15 (13)		
4	11	12	14	14 (null?)	11	14 (11)		
5	13.2 (14)	14 (13.2)	12	13	12	14 (12)		
6	EXC		13	14 (13)	14 (15.2)		15.2	
7	13	14.2	13	15 (13)	14	15 (14)		
8	14	15 (14)	14	15.2	15 (15.2)		15.2	14 15 (14)
9	EXC		5.2 (14)	14	5.2 (13)		13	
10	14	15	13	15 (13)	15	15		
11	13	14	13.2	15 (13.2)	13	15 (13)		
12	13	15	13,2	15 (13.2)	15	15		
13	16	16,2	13.2 (14.2)	14,2	13.2 (16.2)		16,2	
14	14	14 (null?)	13,2	15,2	13,2	13.2 (null?)		
15	13	13 (null?)	14,2	15	15	15 (null?)		
16	15	17,2 (15)	15	16	15 (16)		16	
17	11	11 (null?)	13	15	13	13 (null?)		
18	14	14.2	12	13	12	14.2 (12)		

**Tabla 7.** Eventos de mutación tiponull detectados en el locus D19S433

Dauber en 2008 (69) presentó el hallazgo de 2 mutaciones en D19S433 causadas por una delección de 4pb en el extremo 3' de la región repetitiva, que no permite la unión del primer del kit Identifiler. En el mismo año Kitayama (70), reportó otro evento de mutación null en el D19S433, esta vez asociado a una substitución (G>A), 32 nucleótidos más arriba del extremo 3' de la zona repetitiva AAGG. Hasta tanto no se cuente con la secuencia de los alelos involucrados, no se descarta la presencia de los dos tipos de eventos: delección y substitución en las muestras de este estudio ya que no todos los *dropout* pudieron ser recuperados con otro kit.

Considerando el alto número de eventos silentes detectados, puede sugerirse que que existe en las zonas flanqueantes del D19S433 al menos una mutación de alta

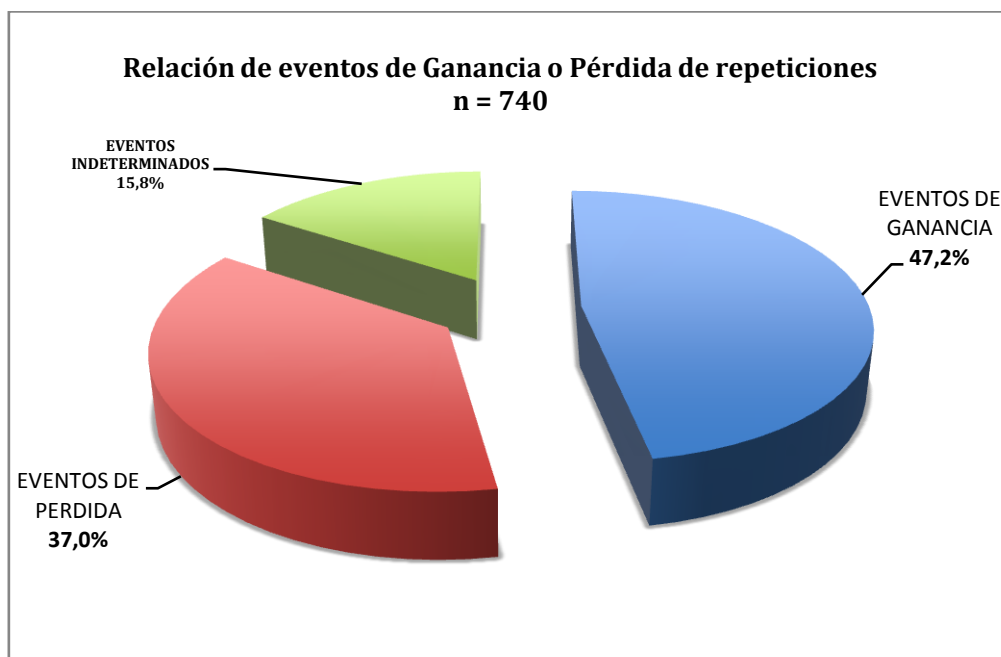
frecuencia en nuestra población. Una vez caracterizadas a nivel de su secuencia, podría determinarse la ancestría de estos eventos para determinar si se han originado en poblaciones específicas del país.

## 4.5 Dinámica mutacional

El análisis directo de un número amplio de mutaciones, permite observar comportamientos propios de la dinámica del proceso evolutivo de los microsatélites. Es evidente la gran variación existente entre loci y aun entre alelos del mismo locus a nivel de su actividad de cambio molecular.

### 4.5.1 Relación ganancia vs pérdida de repeticiones

De los 740 eventos tipo *slipp*, se definió el estado de ganancia o de pérdida de repeticiones en 623 casos así: 349 mutaciones se dieron por ganancia y 274 por pérdida (**Tabla 8**). En los 117 casos restantes (15,8%), no fue posible determinar este estado (**Gráfico 9**).



**Gráfico 9.** Eventos de Ganancia o Pérdida de repeticiones n = 811

Número de pasos	Ganancia	Pérdida	Indeterm.	Totales
4	0	1	0	1
3	0	1	0	1
2	4	9	1	14
1	345	263	116	724
Totales	349	274	117	740

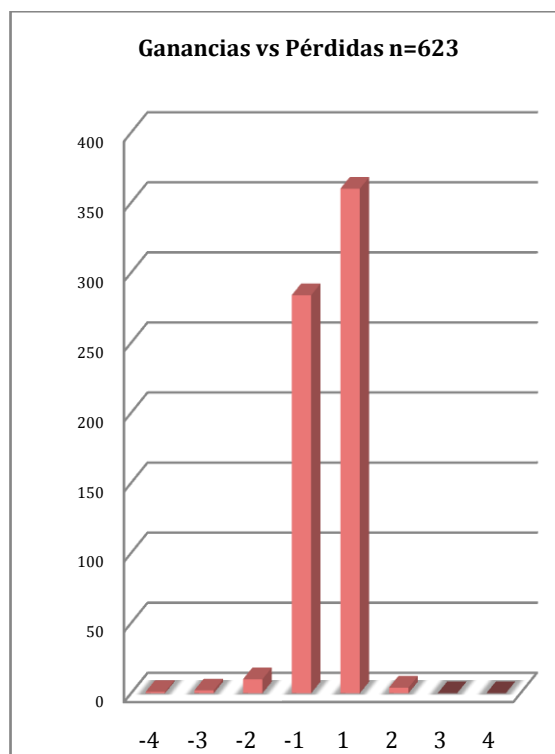
**Tabla 8.** Estado de pérdida o ganancia de repeticiones vs número de pasos

Aunque hay un moderado exceso de eventos de ganancia, la diferencia entre ganancias y pérdidas no es significativa ( $z = -1,54$ ) ( $P < 0,05$ ). Este hallazgo es compatible con la mayoría de los estudios mutacionales en microsatélites. No obstante, la visión particular del fenómeno a nivel de cada locus, permite detectar una alta heterogeneidad en la direccionalidad del proceso mutacional como se verá más adelante.

Por otra parte, los eventos multipasos solo representaron el 2,16% de todas las mutaciones tipo slipp, una proporción muy inferior a las reportadas en estudios similares; Ellegren (27) informó de eventos multipasos en 11% de los 102 eventos mutacionales detectados en su estudio. Huang (10), en un análisis de 97 mutaciones en humanos reportó 63% de cambios multipasos. La diferencia es dramática en relación con nuestros datos, consideramos que el tamaño de la muestra analizada en nuestro estudio permite hacer inferencias mucho más robustas sobre la dinámica del cambio evolutivo en microsatélites. Es importante aclarar que en el presente estudio, se eliminaron del conteo, todos los 32 eventos tipo *null*, porque fueron verificados con segundos juegos de *primers* y en todos se recuperaron los alelos en *dropout* y se rescató el heterocigoto. Estos eventos originalmente, sesgaban los resultados a favor de los eventos multipasos; es así como en los conteos iniciales, se observaron mutaciones de 2 y hasta de 9 pasos en la mayoría de estos casos, lo que en principio

parecían ser mutaciones tipo *slipp*. Consideramos que esta corrección permite ver la realidad de los cambios mutacionales. Los Microsatélites tipo tetranucleótidos en humanos mutan casi exclusivamente en pasos simples. (**Gráfico 10**)

Para Brohede y Cols. (71), las ganancias o pérdidas son independientes del tamaño alélico lo Cuál no indica una tendencia de los alelos más grandes a disminuir de tamaño. Para Brinkmann, existe una relación significativa entre las tasa mutacionales y el promedio geométrico del número de repeticiones en zonas no interrumpidas de los alelos STRs. Nuestros datos son parcialmente consistentes con esta última hipótesis y se ajustan más a lo propuesto por Bohede. TH01.

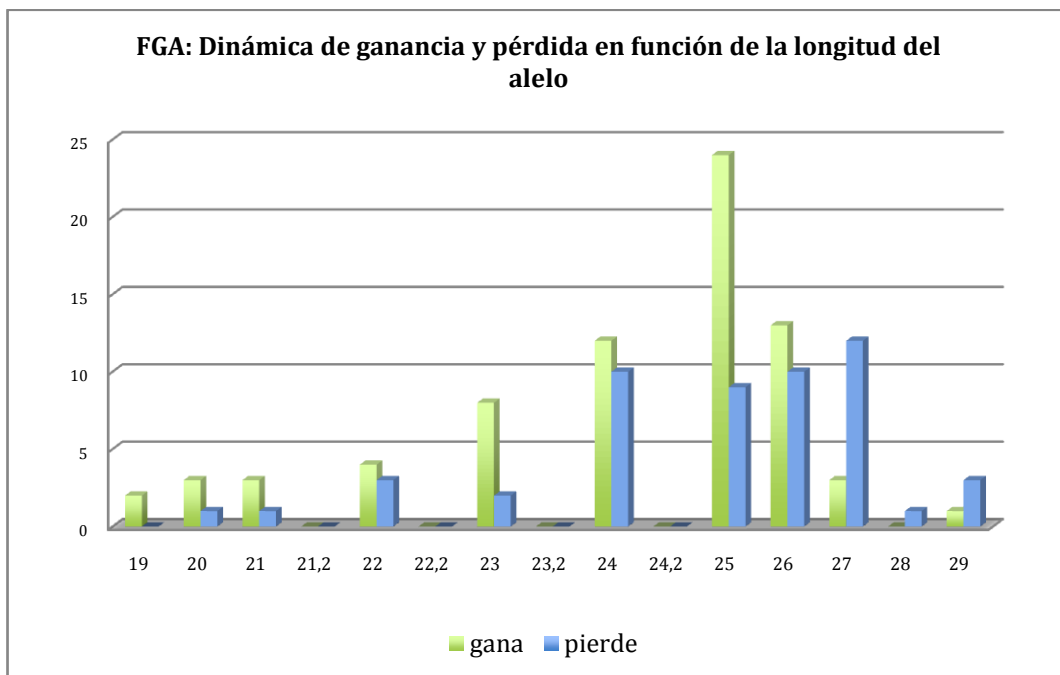


**Gráfico 10.** Eventos mutacionales por pasos

#### 4.5.2 Direccionalidad

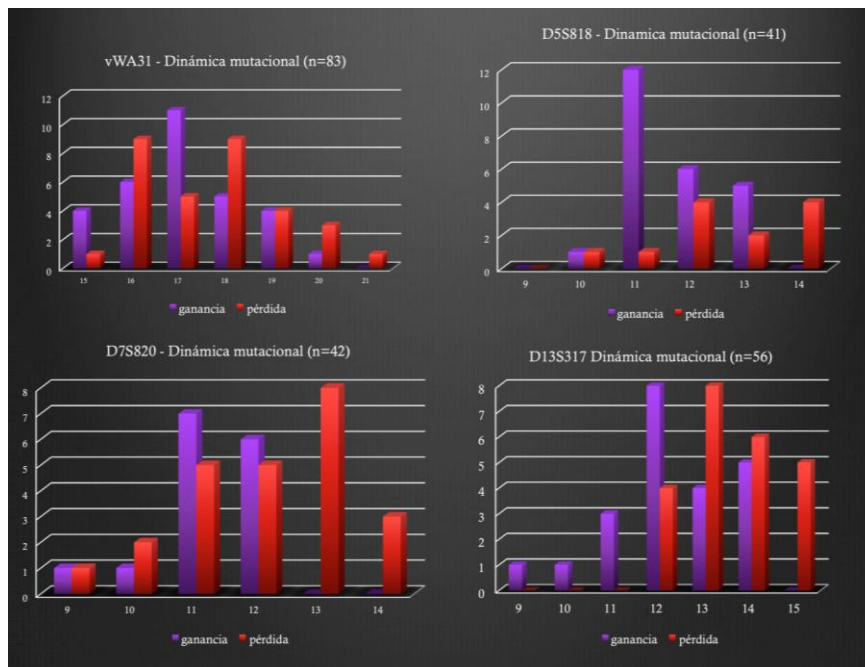
Puede aceptarse que en su mayoría, los STRs estudiados presentan un tendencia de mutación direccional. El patrón más común es aquel donde a medida que se incrementa la longitud del alelo, son más comunes los eventos de ganancia de repeticiones. Esto es cierto, si se acepta también que el proceso está limitado por un techo superior, donde, desde cierta longitud, la conducta cambia hacia la pérdida de unidades. El caso más documentado en esta muestra es el del locus FIBRA-FGA que demuestra la tasa mutacional más alta del grupo de polimorfismos analizados.

Las ganancias son predominantes y crecientes constantemente en número, a desde el alelo 19 hasta el 25. Este último parece ser el “techo mutacional”, desde aquí, disminuyen drásticamente los eventos de ganancia y se incrementan las pérdidas. En los alelos 24 y 26 se equilibran los eventos en las dos direcciones y acumulan entonces el mayor número de mutaciones. Este hallazgo no parece estar totalmente influenciado por las frecuencias alélicas. Para el locus FGA, los alelos más frecuentes en nuestras poblaciones son el 22, 23 y el 24. El alelo 25 tiene una frecuencia de 0,11 a 0,15 , pero el alelo 26 que parece tener también una muy alta tasa mutacional, se encuentra entre el 0,05 y 0,1 de frecuencia en poblaciones Colombianas (59). **(Gráfico 11)**

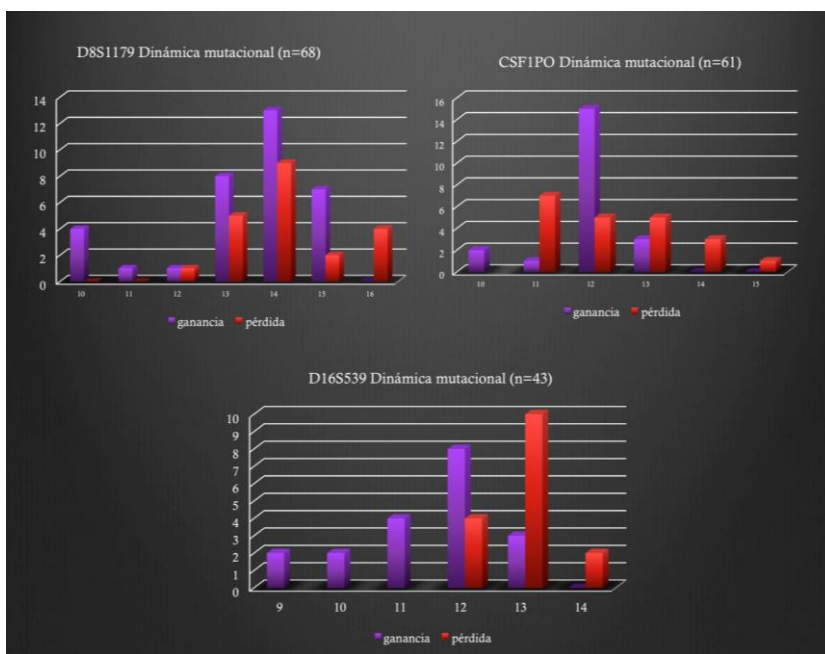


**Gráfico 11.** Locus FGA Dinámica mutacional

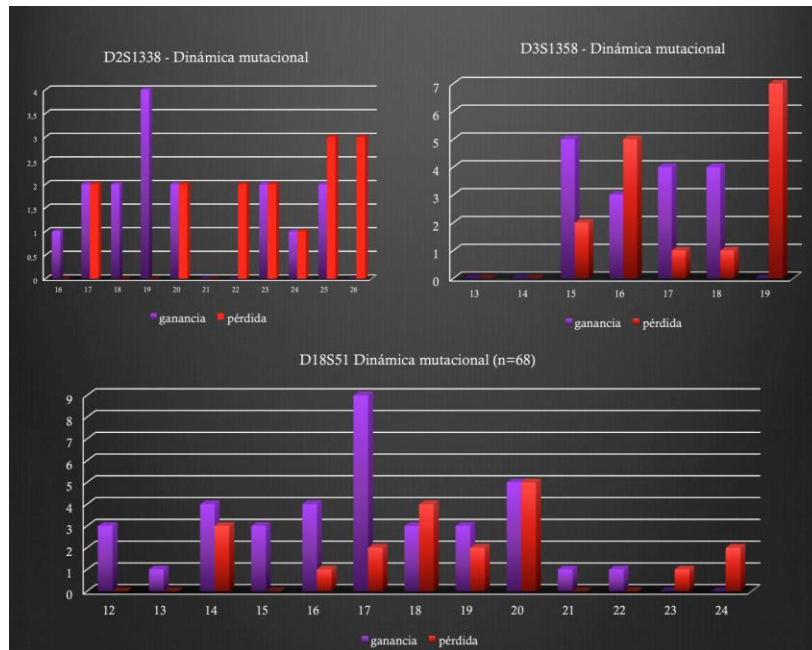
Otros loci analizados en el presente estudio siguen este comportamiento y puede observarse la tendencia unidireccional mencionada antes: (Gráficos 12, 13, 14 y 15 )



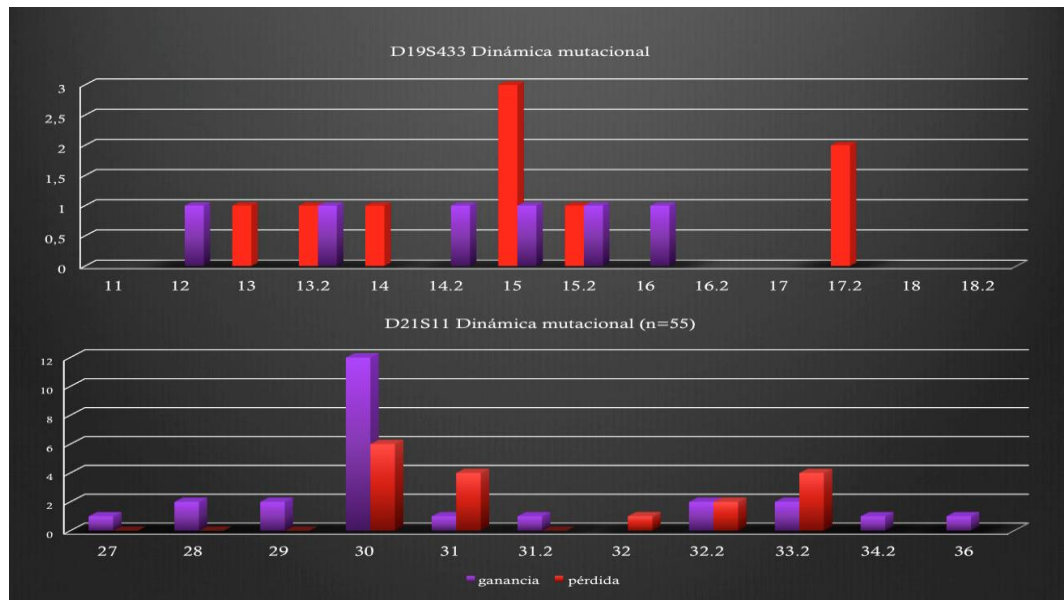
**Gráfico 12.** Distribución de expansiones y contracciones de alelos en los loci vWA31, D5S818, D7S820 y D13S317.



**Gráfico 13.** Distribución de expansiones y contracciones de alelos en los loci D8S1179, CSF1PO, y D16S539.



**Gráfico 14.** Distribución de expansiones y contracciones de alelos en los loci D2S1338, D3S1358, y D18S51.

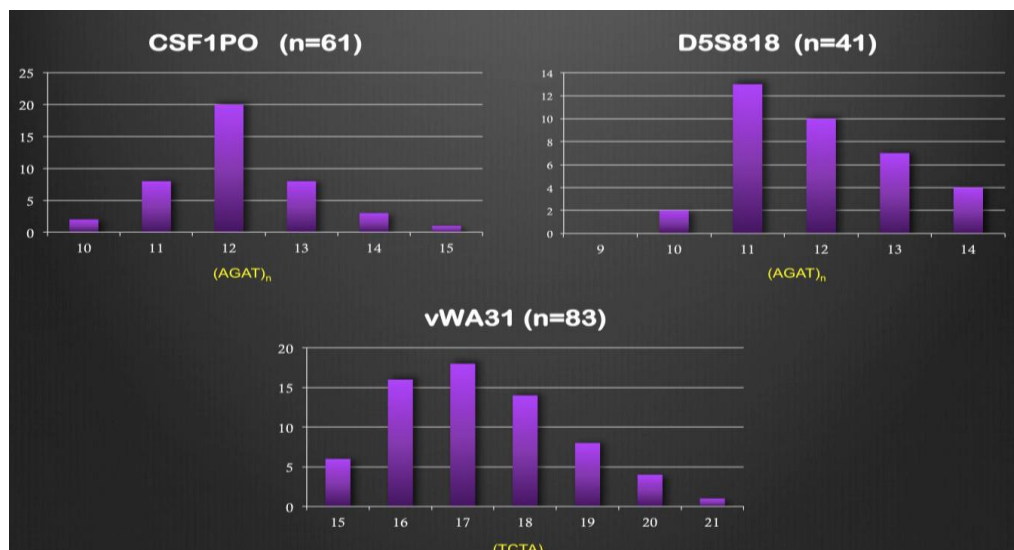


**Gráfico 15.** Distribución de expansiones y contracciones de alelos en los loci D19S433 y D21S11.

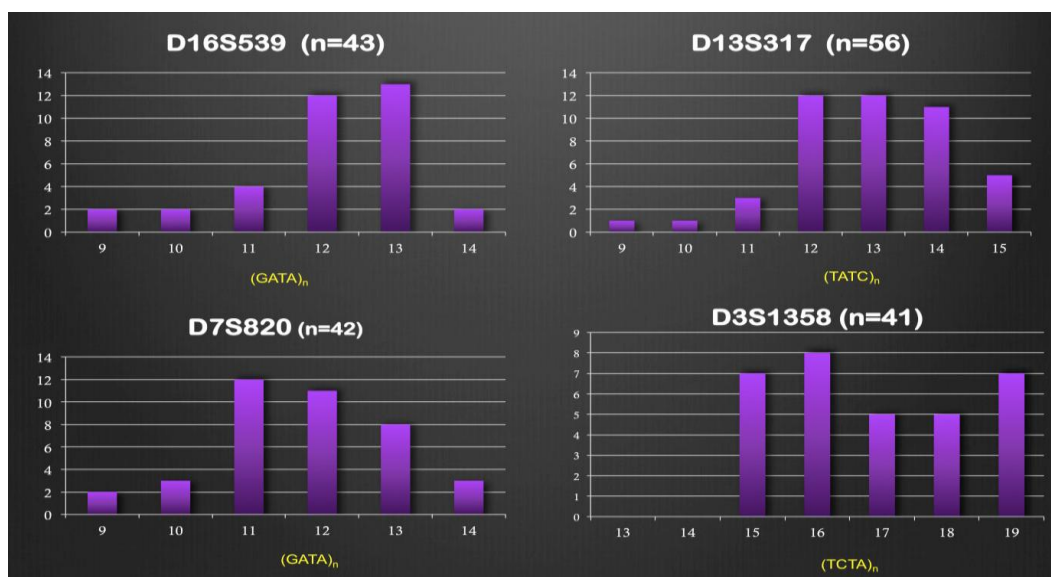
Es evidente que las ganancias aumentan con la longitud alélica hasta un techo definido, en la mayoría de los loci. Ahora bien, al observar la dinámica mutacional



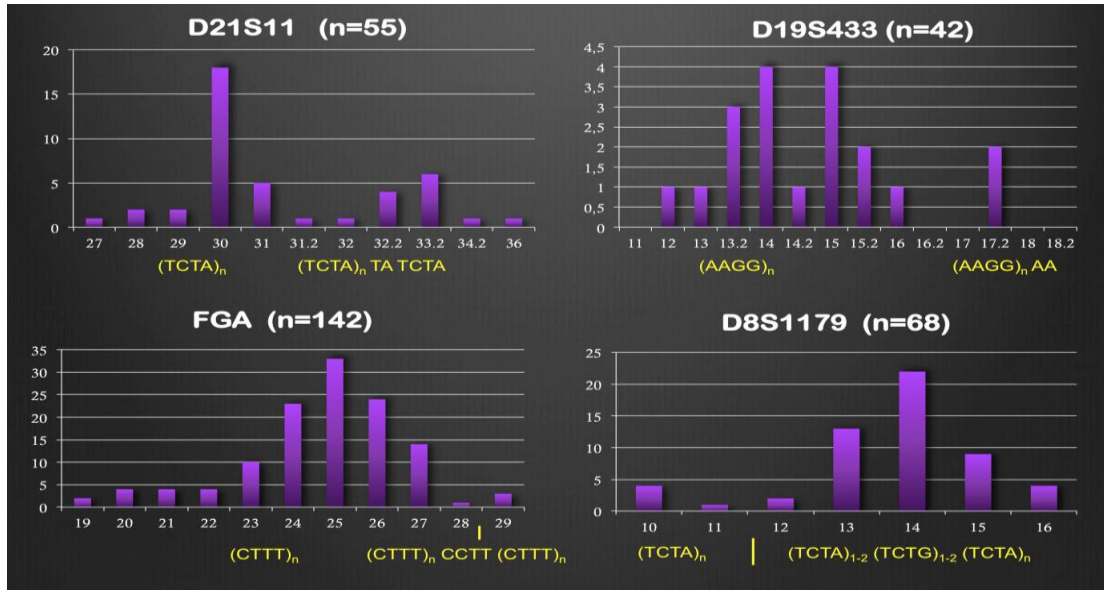
por alelo, integrando todos los eventos de expansión y contracción, las tasa mutacionales aumentan hasta cierto longitud alélica, luego de la Cuál hay un franco descenso de esta variable. La distribución de los alelos en función de su tasa mutacional demuestra un claro patrón unimodal en 10 loci. Los loci D2 D3 D18 D19 y D21 tienen una distribución bimodal o multimodal. (**Gráficos**16, 17, 18 y 19).



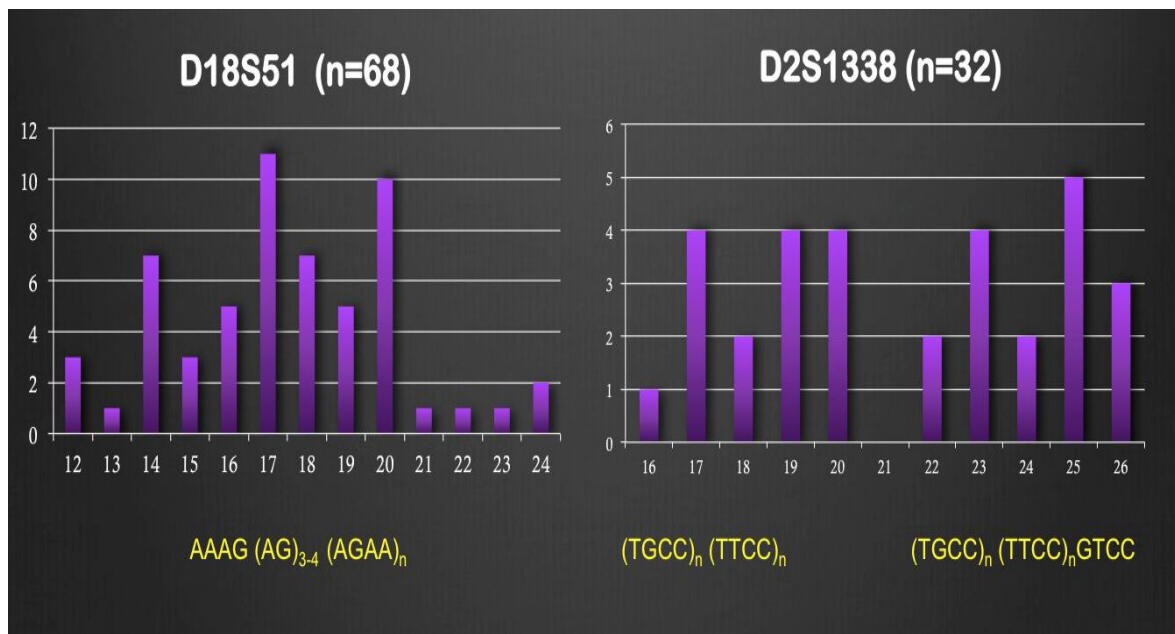
**Gráfico 16.** Distribución unimodal de eventos de mutación en los loci CSF1PO, D5S818 y vWA31.



**Gráfico 17.** Distribución unimodal de eventos de mutación en los loci D16S539, D13S317, D7S820 y D3S1358.



**Gráfico 18.** Distribución Bimodal de eventos de mutación en los loci D21S11, D19S433, FGA y D8S1179.



**Gráfico 19.** Distribución Multimodal de eventos de mutación en los loci D18S51 y D2S1338.

Los gráficos 16 a 19, muestran también las secuencias de repetición que caracterizan a cada locus, es interesante observar que las distribuciones unimodales se asocian a STRs de estructura simple; por su parte las distribuciones bimodales y

multimodales se asocian con niveles mayores de complejidad en la organización estructural del microsatélite.

#### 4.6. Tasas mutacionales

Teniendo en cuenta la variedad de escenarios en que puede presentarse un caso de investigación de la filiación, el número de cotejos presentes en esta población es bastante mayor al número de casos; se contaron 64772 cotejos, en los Cuáles se detectaron 9.084 incompatibilidades de la paternidad. La diferencia entre estos dos valores determina el número de casos donde se confirma la paternidad biológica y puede entonces estimarse el número de transferencias alélicas. Se contaron 23.932 meiosis paternas y 31.756 meiosis de origen materno. (**Tabla 9**).

<b>Año</b>	<b>Periodo</b>	<b>Nro de Casos</b>	<b>Nro de Cotejos</b>	<b>Nro de Exclusiones</b>	<b>Meiosis Paternas</b>	<b>Meiosis Maternas</b>	<b>Nro total de meiosis</b>
2009	Ene-Dic	8645	16885	2428	5942	8515	14457
2010	Ene-Dic	8746	17496	2037	6971	8488	15459
2011	Ene-Dic	7941	16205	2574	5766	7865	13631
2012	Ene-Dic	6884	14186	2045	5253	6888	12141
<b>Totales</b>		<b>32216</b>	<b>64772</b>	<b>9084</b>	<b>23932</b>	<b>31756</b>	<b>55688</b>

**Tabla9.** Relación de casos analizados y conteo de meiosis paternas y maternas

**4.6.1 Tasas mutacionales para 15 loci STR en población colombiana** Considerando las 613 mutaciones detectadas en los 4 años del estudio, se procedió al cálculo de tasas mutacionales en la población colombiana para los 15 loci STR mencionados, que pueden aplicarse a los estudios forenses de

investigación del parentesco ya sea en investigación de la paternidad o la maternidad o en la identificación de restos humanos y cadáveres en condición de no identificado (**Tabla 10**).

<b>locus</b>	<b>Nro. de muts.</b>	<b>Tasa por locus</b>	<b>Intervalo de confianza 95%</b>
D2S1338	25	0,0004	0,0003-0,0007
D3S1358	31	0,0006	0,0004-0,0008
D5S818	32	0,0006	0,0004-0,0008
D7S820	31	0,0006	0,0004-0,0008
D8S1179	57	0,001	0,0008-0,0013
D13S317	45	0,0008	0,0006-0,0011
D16S539	32	0,0006	0,0004-0,0008
D18S51	55	0,001	0,0007-0,0013
D19S433	35	0,0006	0,0004-0,0009
D21S11	39	0,0007	0,0005-0,001
CSF1P0	41	0,0007	0,0005-0,001
FGA	112	0,002	0,0017-0,0024
VWA31	70	0,0013	0,001-0,0016
TPOX	6	0,0001	0-0,0002
TH01	2	0	0-0,0001
Total	613		

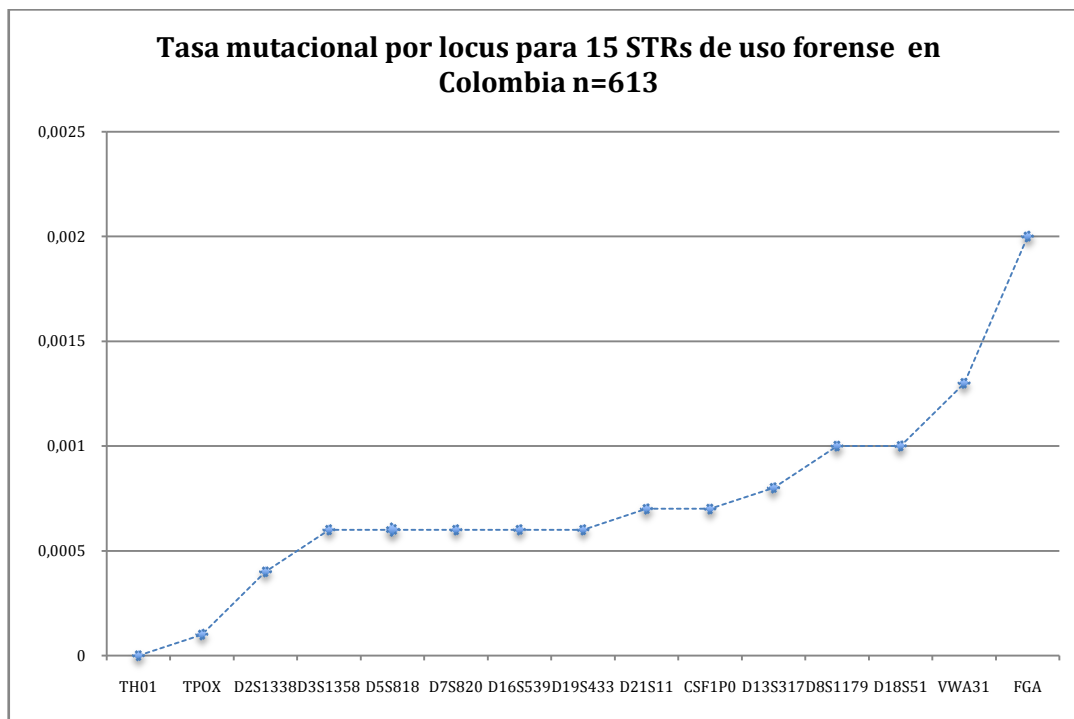
**Tabla 10.** Tasas mutacionales por locus (IC: 95%)

Se observan tres categorías en que pueden agruparse los loci analizados según su tasa mutacional (**Tabla 11**) (**Gráfico 20**):

#### Rangos de tasa mutacional

Baja tasa	Tasa media	Alta tasa
0 - 0,0001	0,0006 - 0,0008	0,001 - 0,002
TH01	D3S1358	D8S1179
TPOX	D5S818	D18S51
	D7S820	VWA31
	D16S539	FGA
	D19S433	
	D21S11	
	CSF1P0	
	D13S317	

**Tabla 11.** Agrupación de los 15 STRs según su tasa mutacional: Baja, Media y Alta



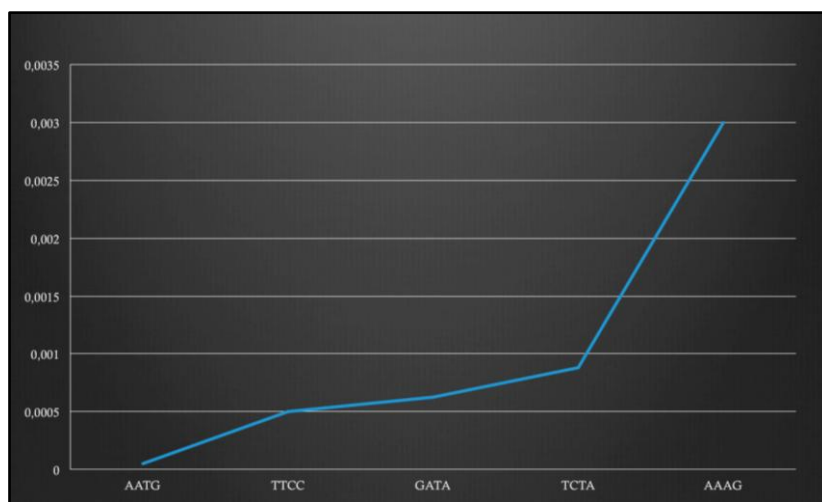
**Gráfico 20.** Tasas mutacionales para 15 STRs

#### 4.6.2 Tasas mutacionales según la secuencia de repetición del STR

Llama la atención que los tres grupos descritos aquí, pueden relacionarse también por su secuencia de repetición. Los loci TH01 y TPOX comparten el mismo *core* de repetición: (AATG)<sub>n</sub>, incluso en el mismo rango de extensión alélica: de 3 a 12 repeticiones, siendo independientes en su segregación; mapean en 11p y 2p respectivamente. Puede decirse entonces que la secuencia AATG presenta la más baja tasa de mutación en el grupo de marcadores estudiados.

En cuanto a los 4 STRs que muestran tasas de mutación mayores de  $1 \times 10^{-3}$  hasta  $2 \times 10^{-3}$ , los loci FGA y D18S51 también comparten la secuencia de repetición AAAG, la misma que caracteriza al locus SE33, que, aunque no hace parte del presente estudio, fue estudiado por nosotros anteriormente (30) y demuestra la más alta tasa de mutación en todas las series reportadas incluso en Colombia.

Los loci vWA31 y D8S1179 comparten el core TCTG y están también entre los que presentan mayor tasa mutacional. Por su parte los loci: D5, D7, D16 y D3 tienen un core común que es el AGTA/TCAT o GATA/CTAT todos presentan además una media geométrica de su longitud alélica de 10 repeticiones y tiene una misma tasa mutacional. (**Gráfico 21**).



**Gráfico 21.** Tasas mutacionales según la secuencia de repetición del STR

La secuencia del core y la media geométrica de la longitud alélica parecen ser las variables más relacionadas con la tasa mutacional (Tabla11).

#### 4.6.3 Tasas mutacionales según la longitud de la secuencia repetida continua del STR

Se calcularon promedios geométricos del número de repeticiones que se observan en los alelos de cada locus, en trayectos repetitivos no interrumpidos, de los Cuáles se ha propuesto, que son el principal sustrato del fenómeno del slippage. (Tabla 12). De esta manera, trayectos más largos con repeticiones homogéneas no interrumpidas, tendrían más alta tasa mutacional. Como se mencionó anteriormente, los 15 loci analizados se pueden agrupar en tres categorías según su tasa mutacional. Los loci TH01 y TPOX muestran solo trayectos de 7 repeticiones homogéneas y tienen la más baja tasa de mutación. Por su parte loci como el FGA, D18 y D8 tienen los trayectos homogéneos largos y presentan las tasas altas de mutación. No obstante, el efecto no se repite con el mismo patrón e todos los demás loci.

locus	Secuencias flanqueantes constantes	Secuencias de repetición cambiante	rango alélico	Media geométrica *	tasa mutacional	tasa %	Secuencia de repetición
D3S1358	[TCTA]2	[TCTG]1-4 [TCTA]7-16	12-19	15,33	0,0006	0,06	TCTA
D21S11	[TCTA]4 ... [TCTA]3 TA [TCTA]3 TCA [TCTA]2 TCCA TA ...	[TCTG]5-6 o 11-13 ... [TCTA]6-16	24-38	30,7	0,0007	0,07	
D8S1179		[TCTA]7-11 [TCTA]1-2 [TCTG]1-2 [TCTA]11-16	7-15	13	0,001	0,1	
D13S317		[TATC]5-17	8-15	11,2	0,0008	0,08	
VWA31	TCTA	TCTG TCTA [TCTG]4-16 [TCTA]3	11-24	17,2	0,0013	0,13	
D5S818		[AGAT]6-18	7-16	11,1	0,0006	0,06	GATA
D7S820	InDel T	[GATA]5-16	6-15	10	0,0006	0,06	
CSF1PO		[AGAT]5-16	6-15	10,09	0,0007	0,07	
D16S539		[GATA]4-16	5-8-15	10,3	0,0006	0,06	
D18S51	AAAG AGAGAG	[AGAA]8-27	7-27	16,3	0,001	0,1	AAAG/CTTT
FGA	[TTTC]3TTTT TTCT ...	[CTTT]7-17	17-33 y 42-51	22,7	0,002	0,2	
SE33		(AAAG)n			0,006	0,6	
D2S1338	[GTCC]1[TTCC]2	[TGCC]4-7 [TTCC]6-17	15-28	21,1	0,0004	0,04	TTCC/AAGG
D19S433	(AAGG)(AAAG)(AAGG)(TAGG)	(AAGG)7-18	9-17,2	12,7	0,0006	0,06	
TPOX		[AATG]3-12	6-13	9,2	0,0001	0,01	AATG
TH01		[AATG]3-12	4-11	7,8	0	0	

\* Media geométrica del número de repeticiones no interrumpidas

**Tabla 12.** Tasas mutacionales según la secuencia de repetición y la Media geométrica del número de repeticiones no interrumpidas del microsatélite.

#### 4.6.4 Tasas mutacionales según el origen parental

En la Tabla 13 se presentan las tasas mutacionales para los 15 marcadores STRs analizados, distribuidas según el origen parental de los eventos de mutación.

Las tasas mutacionales generales y por linaje, obtenidas para Colombia difieren en varios loci, de las tasas reportadas en otras poblaciones; (64, 66, 67, 68) y particularmente con los reportes nacionales que anualmente consolida la *American Association of Blood Banks (aabb)* en los Estados Unidos (66).

locus	Nro. de mutaciones Origen paterno	Tasa mutacional paterna	Intervalo de confianza 95%		Nro. de mutaciones Origen materno	Tasa mutacional materna	Intervalo de confianza 95%	
D2S1338	22	0,0009	0,0006	0,0014	3	0,0001	0	0,0003
D3S1358	24	0,001	0,0006	0,0015	2	0,0001	0	0,0002
D5S818	24	0,001	0,0006	0,0015	2	0,0001	0	0,0002
D7S820	21	0,0009	0,0005	0,0013	6	0,0002	0,0001	0,0004
D8S1179	44	0,0018	0,0013	0,0025	3	0,0001	0	0,0003
D13S317	33	0,0014	0,001	0,0019	5	0,0002	0,0001	0,0004
D16S539	22	0,0009	0,0006	0,0014	5	0,0002	0,0001	0,0004
D18S51	45	0,0019	0,0014	0,0025	8	0,0003	0,0001	0,0005
D19S433	23	0,001	0,0006	0,0014	11	0,0003	0,0002	0,0006
D21S11	17	0,0007	0,0004	0,0011	12	0,0004	0,0002	0,0007
CSF1P0	28	0,0012	0,0008	0,0017	6	0,0002	0,0001	0,0004
FGA	94	0,0039	0,0032	0,0048	6	0,0002	0,0001	0,0004
VWA31	54	0,0023	0,0017	0,0029	8	0,0003	0,0001	0,0005
TPOX	4	0,0002	0,00001	0,0004	2	0,0001	0	0,0002
TH01	2	0,0001	0,00001	0,0003	0	0	0	0
Total	457				79			



**Tabla 13** Tasas mutacionales para 15 loci STRs en Colombia según origen parental

Esta última información se ha convertido en referente mundial, para ajustar la valoración bioestadística que debe hacerse para presentar un resultado de paternidad cuando se detecta un evento mutacional, ya que este debe modificar el valor de la Razón de verosimilitud (RV) o *Likelihood ratio* (LR), que se estima en cada caso. El efecto del evento mutacional es muchas veces dramático sobre el valor del LR calculado con los loci no excluyente. Puede disminuirse en varios órdenes de magnitud, lo Cuál evita la sobrevaloración de un resultado que se presenta a la autoridad judicial, incluso abriendo la posibilidad de duda adicional sobre si el presunto padre es en realidad el padre biológico o un familiar cercano de este.

Considero conveniente realizar una crítica sobre los datoscolectados por la *aabb*, que, como se mencionó antes, son la base de datos de mayor uso en los laboratorios forenses a nivel mundial. Inicialmente, en el formato de envío de la información no se discriminan los tipos de casos donde se contaron las meiosis reportadas, solo se informa que muchos laboratorios trabajan principalmente casos de investigaciones de inmigración, paternidad, identificación, entre otros, pero se asume un control local de los datos antes de su envío. Por otra parte, no es posible verificar la exactitud de los conteos por parte de la *aabb*; la entidad acreditadora, menciona que debe seleccionar los datos que parecen más confiables. Esto hace dudar de la precisión de las tasas reportadas y genera la justificación de realizar este tipo de estudios localmente. Si el laboratorio no ve ninguna mutación, no aporta información sobre conteo de meiosis lo Cuál sesga notablemente la consolidación de los datos nacionales.

Adicionalmente, los datos se presentan agrupados por “raza”: Asiáticos, Caucásicos, Hispanos, Afrodescendientes. No existe un consenso para definir estas agrupaciones, y hay una alta probabilidad de error al agrupar las personas por su apariencia fenotípica o a veces por el aparente origen de su apellido. En hispanos, por ejemplo pueden ingresar a un individuo de apellido “González” que puede ser de

orígenes muy diversos como mexicano, cubano, uruguayo, español o filipino, etc. En la categoría de “Asiático” por ejemplo, se incluyen Turquía, India y China. De hecho, no existe un significado biológico de este tipo de agrupación.

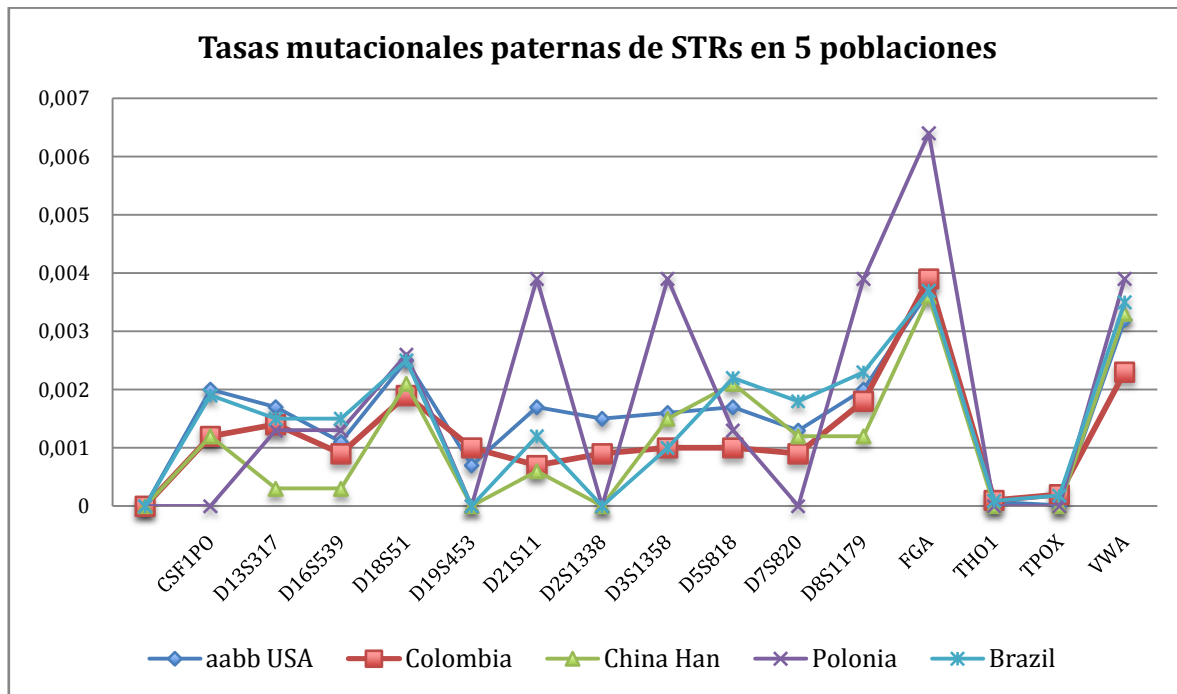
En nuestro país, es posible al crecer la base de datos, asociar los eventos mutacionales a regiones etnogeográficas, para estimar si existe alguna variación significativa, que amerite agrupar los casos con el criterio del origen étnico de las personas en Colombia. El estudio complementario de los eventos detectados en este estudio, con marcadores de ancestría tipo InDels, como se mencionó antes, podría aportar información de interés en este aspecto.

Las diferencias detectadas no son estadísticamente significativas en conjunto, no obstante, tienen un impacto importante como se dijo antes, en la valoración estadística de los resultados en pruebas de filiación e identificación de personas por pruebas de ADN. (**Tabla 14, Gráfico 22**) .

Locus	Colombia		aabb USA		China Han		Polonia		Brazil	
	Paterna	Materna	Paterna	Materna	Paterna	Materna	Paterna	Materna	Paterna	Materna
CSF1PO	0,0012	0,0002	0.002021	0.000283	0,0012	0	0	0	0,0019	0,00018
D13S317	0,0014	0,0002	0.001743	0.000436	0,0003	0	0,0013	0	0,0015	0,00018
D16S539	0,0009	0,0002	0.001127	0.000481	0,0003	0,0003	0,0013	0,0011	0,0015	0,00027
D18S51	0,0019	0,0003	0.002530	0.000748	0,0021	0,0003	0,0026	0	0,0025	0,00055
D19S453	0,001	0,0004	0.000745	0.000596			0	0		
D21S11	0,0007	0,0001	0.001709	0.001295	0,0006	0,0016	0,0039	0	0,0012	0,0016
D2S1338	0,0009	0,0001	0.001526	0.000245			0	0		
D3S1358	0,001	0,0001	0.001691	0.000211	0,0015	0	0,0039	0	0,001	0
D5S818	0,001	0,0001	0.001742	0.000300	0,0021	0	0,0013	0,0021	0,0022	0,00046
D7S820	0,0009	0,0002	0.001348	0.000073	0,0012	0,0003	0	0,0011	0,0018	0
D8S1179	0,0018	0,0001	0.002031	0.000333	0,0012	0	0,0039	0	0,0023	0,00055
FGA	0,0039	0,0002	0.003713	0.000522	0,0036	0,0006	0,0064	0,0021	0,0037	0,00027

THO	0,0001	0	0.000070	0.000043	0	0	0	0	0,00009	0
TPOX	0,0002	0,0001	0.000130	0.000081	0	0	0	0	0,00018	0
VWA	0,0023	0,0003	0.003258	0,000494	0,0033	0,0003	0,0039	0	0,0035	0,00055
n	457	79	?	?	138	32	22	6	258	51

**Tabla 14.** Tasas mutacionales de 15 STRs en 5 poblaciones



**Gráfico 22 .** Comparación de tasas mutacionales de 15 STR en 5 poblaciones.

**4.6.5 Efecto de la tasa mutacional en el cálculo del Índice de Paternidad (IP o LR) en investigaciones de la paternidad.**

Si se considera un caso tipo, donde se ha detectado una exclusión aislada en el locus D21S11 que, luego de la respectiva comprobación se asume como una mutación, el locus debe ser objeto de un cálculo del Índice de paternidad (IP), que es igual al RV o al LR ya mencionado, que tenga en cuenta la tasa de mutación específica para el marcador.

Veamos un ejemplo:

Padre: 31, 38      Madre: 28, 28      Hijo: 28, 32

El alelo 31 paterno es heredado al hijo ganando una repetición (+1), 31 a 32.

$\mu$  es la tasa de mutación

$\mu_1$  es la tasa de mutación paterna aabb para el locus D21S11: 0,0017

$\mu_2$  es la tasa de mutación paterna en Colombia para el locus D21S11: 0,0007

Utilizando el método de Charles Brenner: [www.dnaview.com](http://www.dnaview.com), se calculan los valores X y Y para estimar el índice de paternidad IP como  $X / Y$ .

### 1. Valor X

Probabilidad de los genotipos detectados, asumiendo la hipótesis que se trata de un trío “verdadero” donde presunto padre y madre son en realidad los padres biológicos del hijo.

$$X = P(\text{Genotipos} \mid \text{Trío verdadero})$$

Es proporcional a:

$$P(\text{Un hombre portador de los alelos 31, 38 transmita el alelo 32})$$

Esto es lo mismo que decir:

$P(\text{Transmitir el alelo 32}) \times P(\text{Mutación del locus}) \times P(\text{la mutación sea por incremento en la longitud del alelo original}) \times P(\text{el alelo mute en un número de pasos discreto})$ . Así tendríamos:

$$(1/2) \times \mu \times (1/2) \times (1/10)^{s-1}$$

donde s es el número de pasos de la mutación slipp, en este caso 1.

### 2. Valor Y

El denominador “Y” del índice indicaría la probabilidad del genotipo observado bajo la hipótesis alterna: el trío es falso, es decir, el padre no es el padre:

$$Y = P(\text{Genotipos} \mid \text{El Trío es falso})$$

Lo Cuál es proporcional a la probabilidad de encontrar en la población el alelo 32 asumiendo que no fue resultado de una mutación proveniente del presunto padre, sino heredado al hijo de otro padre desconocido sin mutación. Este valor es la misma frecuencia poblacional del alelo 32 del locus D21S11.

Cuando la mutación es de un solo paso, la ecuación de Brenner puede simplificarse así:

$$\mu / 4 \text{ Frecuencia del alelo 32}$$

en Colombia la frecuencia del alelo 32 en el locus D21S11 es de 0,004 (Paredes, 2003) (59).

El valor LR usando la  $\mu$  de la *aabb* sería:

$$0,0017 / 0,004 = 0,425$$

usando el  $\mu$  del presente estudio sería

$$0,0007 / 0,004 = 0,175$$

Ahora bien, si se asume un índice de paternidad IP de 10.000, acumulado con un grupo de marcadores no excluyentes, situación común en un caso de paternidad, lo que daría una Probabilidad de Paternidad (W) de 99,99% se reduciría en el primer caso así:

$$10.000 \times 0,425 = 4250 \text{ es decir una } W = 99,98\%$$

y en el segundo caso así:

$$10,000 \times 0,175 = 1750 \text{ es decir una } W \text{ de } 99,94\%$$

Si el umbral estándar que se exige al laboratorio es 99.99% como mínimo para presentar los resultados como paternidad No excluida, el laboratorio tendrá que analizar muchos más marcadores para alcanzar este valor si se usan las tasas mutacionales colombianas que las de la *aabb*. En otras palabras, el efecto de usar una tasa más alta de mutación que la real para una población determinada, es el de sobreestimar la probabilidad de paternidad calculada, lo Cuál va en contra de los intereses del acusado y a favor de la demandante.

#### 4.7 Frecuencias de mutación

Se presentan las frecuencias de mutación para los 15 loci STR analizados en población colombiana (Tabla 15). Teniendo en cuenta que la tasa de mutación por alelo requeriría para su cálculo de un número muy alto de eventos observados, se propone el uso de las frecuencias de mutación para modificar la ecuación de Brenner y ajustar en cada caso, la tasa mutacional de locus, en función de la frecuencia mutacional del alelo. Esto imprimiría un efecto más conservador en el cálculo de LR<sub>s</sub> en la investigación del parentesco.

Mutación	D2S1338	D3S1358	D5S818	D7S20	D8S1179	D13S317	D16S539	D18S51	D19S433	D21S11	CSF1PO	FGA	TPOX	VWA31	Mutación
8>9													0,1429		8>9
9>8				0,0256											9>8
9>10				0,0256		0,0217	0,0556								9>10
9>12							0,0278								9>12
10>7				0,0256											10>7
10>9			0,0278	0,0256											10>9
10>11			0,0278	0,0256	0,0727	0,0217	0,0556				0,0476				10>11
11>10			0,0278	0,1282							0,1667				11>10
11>12			0,3333	0,1795	0,0182	0,0652	0,1111				0,0238				11>12
12>10				0,0256											12>10
12>11			0,1111	0,1026	0,0182	0,0870	0,1111		0,0417		0,1190		0,4286		12>11
12>13			0,1667	0,1538	0,0182	0,1739	0,2222	0,0500			0,3571				12>13
12>17								0,0167							12>17
13>11						0,0217									13>11
13>12			0,0556	0,2051	0,0909	0,1739	0,2778				0,1190		0,4286		13>12
13>14			0,1389		0,1455	0,0870	0,0833	0,0167	0,0417		0,0714				13>14
13,2>12,2									0,0833						13,2>12,2
13,2>14,2									0,0417						13,2>14,2
14>13			0,1111	0,0769	0,1636	0,1304	0,0556	0,0500			0,0714				14>13
14>15					0,2364	0,1087		0,0667	0,1667						14>15
14,2>13,2									0,0417						14,2>13,2
15>13		0,0313													15>13
15>14		0,0313			0,0364	0,1087			0,0417		0,0238			0,0149	15>14
15>16		0,1250			0,1273			0,0500	0,1250					0,0597	15>16
15>17		0,0313												0,0149	15>17
15,2>14,2									0,0417						15,2>14,2
15,2>16,2									0,0417						15,2>16,2
16>14		0,0313													16>14
16>15		0,1250			0,0727			0,0167	0,0417					0,1343	16>15
16>17		0,0938						0,0667						0,0896	16>17

16>19					0,0149	16>19	
17>15					0,0299	17>15	
17>16	0,0625	0,0313		0,0333	0,0746	17>16	
17>18	0,0625	0,1250		0,1500	0,1642	17>18	
17,2>18,2				0,0833		17,2>18,2	
18>17		0,0313		0,0667	0,1343	18>17	
18>19	0,0625	0,1250		0,0500	0,0746	18>19	
19>12				0,0167		19>12	
19>18		0,2188		0,0333	0,0597	19>18	
19>20	0,0938			0,0500	0,0157	0,0597	19>20
19>23	0,0313					19>23	
20>17				0,0167		20>17	
20>18				0,0167		20>18	
20>19	0,0625			0,0833	0,0079	0,0448	20>19
20>21	0,0625			0,0833	0,0236	0,0149	20>21
21>20					0,0079	0,0149	21>20
21>22					0,0236		21>22
21>22				0,0167			21>22
22>20					0,0079		22>20
22>21	0,0625				0,0079		22>21
22>23				0,0167	0,0236		22>23
22>25					0,0079		22>25
23>19	0,0313						23>19
23>22	0,0313				0,0157		23>22
23>24	0,0625				0,0630		23>24
24>22	0,0313						24>22
24>23				0,0333	0,0787		24>23
24>25	0,0313				0,1024		24>25
25>23					0,0079		25>23
25>24	0,0938				0,0709		25>24
25>26	0,0625				0,1890		25>26
25>27					0,0079		25>27
26>27	0,0313						26>27
26>25	0,0625				0,0866		26>25
26>27					0,1024		26>27
27>25					0,0079		27>25
27>26					0,0866		27>26
27>28					0,0182	0,0236	27>28
28>27					0,0079		28>27
28>29					0,0409		28>29
28>30					0,0000		28>30



29>28		0,0157	29>28
29>30		0,0483	29>30
29>31		0,0182	29>31
30>28		0,0182	30>28
30>29		0,1227	30>29
30>31		0,2689	30>31
30>32		0,0182	30>32
31>30		0,0849	31>30
31>32		0,0363	31>32
31.2>29.2		0,0182	31.2>29.2
31.2>32.2		0,0218	31.2>32.2
32>31		0,0224	32>31
32.2>31.2		0,0363	32.2>31.2
32.2>33.2		0,0363	32.2>33.2
33>32		0,0000	33>32
33.2>32.2		0,0994	33.2>32.2
33.2>34.2		0,0545	33.2>34.2
34.2>35.2		0,0182	34.2>35.2
35.2>34.2		0,0000	35.2>34.2
36>37		0,0182	36>37
16 o 23>17 o 22	0,0313		16 o 23>17 o 22
20 o 22>21	0,0313		20 o 22>21
12 o 14>13		0,0833	12 o 14>13
13 o 15>14		0,0833	13 o 15>14
14 o 16>15		0,0417	14 o 16>15

**Tabla 15.** Frecuencias de mutación de 14 STRs en población colombiana

## 5. Conclusiones

1. La distribución geográfica de las mutaciones detectadas en 15 STRs no se aparta de la distribución actual de la población en Colombia. Una muestra más amplia debería ser estudiada para detectar si hay un patrón étnico o poblacional asociado a la distribución de las mutaciones analizadas, se requieren además estudios de ancestría con marcadores tipo InDels.
2. Más del 91% de los eventos estudiados son compatibles con un mecanismo tipo *slippage* de la polimerasa, solo el 3,8% tienen un patrón mutacional tipo null o alelo silente.
3. A medida que se avanza en edad, los parentales muestran mayores tasas de mutación, tanto en hombres como en mujeres. La mutación de STRs en humanos se incrementa con la edad principalmente en hombres donde la tasa se duplica cada 15 años de edad; en la mujer permanecesin variación hasta los 40 años cuando se duplica.
4. Es notable el predominio de los eventos de origen paterno en un radio de 5:1 con respecto a las mutaciones maternas. Este hallazgo es concordante con la literatura.

5. En el 10% de los casos no fue posible asignar el origen parental de la mutación, se requerirían análisis adicionales con marcadores de ligamiento para seguir los haplotipos y establecer con certeza el origen.
6. Se presentan las tasas mutacionales aplicables a la población colombiana para 15 STRs de uso común en la genética forense.
7. Aunque las tasas mutacionales no presentan diferencias altamente significativas, en relación con otras poblaciones, si hay un impacto importante de las diferencias encontradas en algunos loci como el D21S11, en la aplicación forense al momento de realizar los cálculos de índices de paternidad, donde se presente una mutación.
8. En un número importante de casos no es posible asignar el origen parental de la mutación, se requerirían análisis adicionales con marcadores de ligamiento para seguir los haplotipos y establecer con certeza el origen .
9. No hay diferencia significativa entre el número de ganancias vs el número de pérdidas. Aun así, hay muchos eventos indeterminados que no permiten conclusiones definitivas al respecto.
10. Se observa una direccionalidad en la dinámica mutacional hacia la ganancia de repeticiones en función de la longitud alélica hasta cierto nivel, luego del cuál se invierte el proceso y se incrementan los eventos de contracción del tamaño alélico.
11. El 45% de las mutaciones del locus D19S433 son tipo null. Se requiere de la secuenciación de estas mutaciones para caracterizar su naturaleza.
12. Predominan los eventos de mutación de un solo paso y no hay diferencia significativa entre la ganancia y la pérdida de repeticiones. Solo el 2,9% de los

eventos son multipasos a diferencia de los reportado en la literatura. Se requiere la secuenciación de estos eventos para definir su naturaleza.

13. Los loci con repeticiones simples continuas presentan distribuciones unimodales de mutación. Loci con cambios del patrón de secuencia entre alelos, generan distribuciones bi o multimodales
14. Se presentan las tasas mutacionales aplicables a la población colombiana para 15 STRs de uso común en la genética forense.
15. Las tasas mutacionales varían, asociadas a motivos de repetición específico. Los loci que poseen el motivo AAAG/CTTT presenta la más alta tasa de mutación.
16. La variación en la tasas de mutación es multifactorial, destacándose, la secuencia del motivo, la longitud del motivo, la longitud del alelo (el número de repeticiones), la estructura alélica o complejidad del arreglo.
17. Se propone incluir en la ecuación de Brenner, la frecuencia mutacional de los alelos comprometidos en cada caso.
18. Se establecen las frecuencias de mutación para 15 STR de uso común en identificación humana para Colombia

## 6. Referencias

1. Litt M, Luty JA. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat with in the cardiac muscle actin gene. *Am J Hum Genet.* 1989; 44: 397–401.
2. Tautz, D. Hypervariability of simple sequences as a general source of polymorphic DNA markers. *NucleicAcidsResearch.* 1989; 17: 6463-6471.
3. Edwards, A, Civitello, A, Hammond, H. A., and Caskey, C. T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genetics.* 1991; 49:746-756.
4. Ellegren H. Microsatellites: Simple sequences with complex evolution. *NatureReviewsGenetics.* 2004; 5: 435-445.
5. Eckert K, Hile S. Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common Microsatellites present in the Human Genome. *Mol Carcinog.* 2009; 48(4): 379–388.
6. Bichara M, Wagner J, Lambert I.B. Mechanisms of tandem repeat instability in bacteria. *Mutation Research.* 2006; 598: 144–163.
7. Soranzo N, Provan J, Powell W. An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome.* 1999; 42:158–161.
8. Provan J, Powell W, Hollingsworth PM Chloroplast microsatellites: new tools

- for studies in plant ecology and evolution. *Trends Ecol Evol.* 2001; 16:142–147.
9. Toth, G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: Survey and Analysis. *Genome Res.* 2000; 10, 967–981.
  10. Huang Q, Xu F, Shen H, Deng H, Liu Y, Liu Y, Li J, Recker R, Deng H. Mutation patterns at dinucleotide microsatellite loci in humans. *Am. J. Hum. Genet.* 2002; 70: 625-634.
  11. Powell W, Machray GC, Provan J Polymorphism revealed by simple sequence repeats. *Trends in Plant Science.* 1996;1:215–222.
  12. Stallings R. L, Torney D.C, Hildebrand C.E, Longmire J.L, Deaven L.L, Jett J.H, Doggett N.A, Moyzis R.K. Physical mapping of human chromosomes by repetitive sequence fingerprinting. *Proc. Natl. Acad. Sci.* 1999; 87,6218- 6222
  13. Schlotterer C and Harr B. Microsatellite Instability. *Encyclopedia of Life Sciences*, Nature Publishing Group. 2001.
  14. Oliveira E. J, Padua J. G, Zucchi M. I, Vencovsky R. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology.* 2006; 29, 294–307.
  15. DNA Recommendations – 1994 report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) system. Technical Note. *Int J LegMed.* 107: 159-160.
  16. Metzgar D, Bytof J, Wills C. Selection against frame shift mutations limits microsatellite expansion in coding DNA. *Genome Res.* 2010; 10:72–80.
  17. Orr H and Zoghbi H. Trinucleotide repeats disorders. *Annual Review of Neuroscience* 2007; 30:575-621.
  18. Centola M and Carbon J. Cloning and characterization of centromeric DNA from *Neurospora crassa* *Mol Cell Biol.* 1994; 14(2): 1510-9.
  19. Catasti P, Gupta G, García AE, Ratliff R, Hong L, Yau P, Moyzis RK, Bradbury EM. Unusual structures of the tandem repetitive DNA sequences located at human centromeres. *Biochemistry.* 1994; 33(13):3819-30.
  20. Li Y, Korol A, Fahima T, Beiles A. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology.* 2002; 11: 2453–2465

21. Jeffreys AJ, Wilson V, Newumann R, and Keyte J. Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *NucleicAcids Res.* 1988;16:10953-10971.
22. Hawn MT, Umar A, Carethers JM, Marra G, Kunkel TA, Boland CR, Koi M. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res.* 1995; 55(17):3721-5.
23. Ou CY, Chang JG, Tseng HH, Wei HJ, Su TH, Hsu TY, Chang CP, Lee HH. Analysis of microsatellite instability in cervical cancer. *Int J Gynecol Cancer.* 1999; 9(1):67-71.
24. Bhargava A, Fuentes F. Mutational Dynamics of Microsatellites. *Mol Biotechnol.* 2010; 44:250–266
25. Payseur B, Jing P and R Haas. A Genomic portrait of human microsatellite variation. *Mol. Biol. Evol.* 2011; 28(1):303–312.
26. Selkoe KA and Toonen RJ, Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *EcolLett.* 2006; 9(5): 615-29.
27. Ellegren H. Heterogeneous mutation processes in human microsatellite DNA sequences. *NatureGenetics.* 2000; 24: 400-401.
28. Schlötterer C. Evolutionary dynamics of microsatellite DNA. *Chromosoma.* 2000; 109:365–371
29. Weber J and Wong C. Mutation of human short tandem repeats. *Human Molecular Genetics.* 1993; 2(8): 1123-1128.
30. Paredes M and Laverde L. Analysis of the SE33 hypervariable microsatellite locus (HumACTBP2) in Colombian populations. *International Association of Forensic Sciences. Book of abstracts.* 2008; 102-105.
31. Levinson G and Gutman G. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *MolBiolEvol.* 1987; 4 (3): 203-221.
32. Kelkar YD, Eckert KA, Chiaromonte F, Makova KD. A matter of life or death: How microsatellites emerge in and vanish from the human genome. *Genome Res* in press. 2011.
33. Levinson G and Gutman G. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12.

- NucleicAcids Res. 1987;15(13):5323-38.
34. Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, and Inouye M. Frameshift Mutations and the Genetic Code. Cold Spring Harbor Symposia on Quantitative Biology. 1966; 31:77-84.
  35. Ananda G, Walsh E, Jacob D, Krasilnikova M, Eckert K, Chiaromonte F, Makova K. Distinct mutational behaviors distinguish short tandem repeats from microsatellites in the human genome. *GenomeBiology and Evolution* Advance. 2012 ;13.
  36. Eisen J. Mechanistic basis for microsatellite instability. En: Goldstein D and Schlötterer Ch, editors. *Microsatellite Evolution and Applications*. 1<sup>st</sup>ed. New York: Oxford University Press; 2001. p.34-48.
  37. Rubinsztein D, Amos W, Leggo J, Goodburn S, Jain S, Li S, Margolis R, Ross C, Ferguson-Smith M. Microsatellite evolution — evidence for directionality and variation in rate between species. *Nature Genetics*. 1995; 10, 337 – 343.
  38. Amos W, Sawcer SJ, Feakes RW, Ribinsztein DC. Microsatellites show mutational bias and heterozygote instability. *Nature Genetics*. 1996; 13(4) 390-1.
  39. Dermitzakis E, Clark A, Batargias C, Magoulas A, Zouros E. Negative covariance suggest mutation bias in a two locus microsatellite system in the fish *Sparusaurata*. *Genetics*. 1998; 150:1567-1575.
  40. Feldman M, Bergman A, Pollock D, Goldstein D. Microsatellite genetic distances with range constraints: Analytic description and problems of estimation. *Genetics*. 1997; 145: 207-216.
  41. Kruglyak S, Durrett R, Schug M, Aquadro C, Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proc. Natl. Acad. Sci*. 1998; 95: 10774–10778.
  42. Estoup A, Jarne Ph, Cournet JM. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology*. 2002; 11:1591-1604.
  43. Bell GI and Jurka J. The length distribution of perfect dimer repetitive DNA is consistent with its evolution by an unbiased single-step mutation process. *J*



- MolEvol. 1997; 44(4): 414-21.
44. Sun J, Helgaso A, Masson G, Ebenesersdottir S, Li H, Mallik S, Gnerre P, Paterson N, Kong A, Reich D, Stefansson K. A direct characterization of human based on microsatellites. *Nature Genetics*. 2012; 44(10): 1161-1167.
  45. Wright S. Evolution in Mendelian Populations. *Genetics*. 1931; 16:97-159
  46. Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 1995; 139: 457-462.
  47. Nei M. Analysis of gene diversity in subdivided populations. *Proc. Nat. Acad. Sci USA*. 1973; 70(12): 3321-3323.
  48. Di Rienzo A, Peterson A, Garza J, Valdes A, Slatkin M, Freimewr B. Mutational processes of simple sequence repeat loci in human populations. *Proc. Nat. Acad. Sci USA* 1994; 91: 3166-3170.
  49. Crow J and Kimura M. An introduction to population genetics theory. Blackburn Press. 1970
  50. Weinberg W. Zurvererbung des zwergwuchses. *Arch. Rassenu.Gesell. Biol.*1912; 9:710.
  51. Frans E, Sandin S, Hultman C. Autism Risk Across Generations: A Population Based Study of Advancing Grandpaternal and Paternal Age. *JAMA Psychiatry*. 2013; 70(5): 516-521.
  52. Vogel and Motilsky Human Genetics. Problems and Approaches. 4th ed. Springer 2010.
  53. Kong A, Stefansson, K, et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature*. 2012; 488: 471-475.
  54. Crow J. Age and sex effects on human mutation rates: An old problem with new complexities. *J. Radiat.Res.* 2006; 47: Suppl.B75-B82.
  55. Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* 1998; 62:1408-1415.
  56. Viguera E, Canceill D, Ehrlich S. Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J.* 2001; 20: 2587-2595.
  57. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H, Specific Enzymatic

- Amplification of DNA in Vitro: The Polymerase Chain Reaction. Cold Spring Harbor Symposium in Quantitative Biology, 1986; 51:263-73.
58. Jeffreys A, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. *Nature*. 1985; 314(6006): 67-73.
59. Paredes et al. Analysis of the CODIS autosomal STR loci in four main Colombian regions. *Forensic Science International*. 2003; 137: 67-73
60. Instituto Nacional de Medicina Legal y Ciencias Forenses. Convenio ICBF-INML, Informe mensual.
61. Paredes M, Mutational dynamics of FGA-FIBRA locus. *Forensic Science international. Genetic Supplement Series 3*. 2011; e261-e262.
62. AmpFLSTR® Identifiler® PCR Amplification Kit User Manual
63. PowerPlex16HS System, Technical Manual.
64. Lu D, Liu Q, Wu W, Zhao H. Mutational analysis of 24 short tandem repeats in Chinese Han Population. *Int J Legal Med*. 2012; 126: 331-335.
65. Brohede J, Arnheim N, Ellegren H. Single-Molecule analysis of the hypermutable tetranucleotide repeat locus D21S1245 through sperm genotyping: A heterogeneous pattern of mutation but no clear male age effect. *Mol Biol Evol*. 2004; 21 (1): 58-64.
66. American Association of Blood Banks (aabb), Relationship testing annual report .2008. [www.aabb.org](http://www.aabb.org)
67. Mardini AC, Rodenbusch R, Schumacher S, et al. Mutation rate estimates for 13 STR loci in a large population from Rio Grande do Sul, Southern Brazil. *Int J Leg Med*. 2013; 127:45-47.
68. Wojtas M, Piniewska D, Polanska N, Stawowiak A, Sanak M. Mutations of microsatellite autosomal loci in paternity investigations of the Southern Poland population. *Forensic Science International*. 2013; 389-391.
69. Dauber E, Glock B, Mayr W. Two examples of null alleles at the D19S433 locus due to the same 4 bp deletion in the presumptive primer binding site of the AmpFISTR Identifiler kit. *For Sci Int: Genetics Supplement Series*. 2008; 1(1):107-108.
70. Mizuno N, Kitayama T, Nakahara H, et al. A D19S433 primer binding site

mutation and the frequency in Japanese of the silent allele it causes. J Forensic Sci. 2008; 53(5): 1068-1073.

71. Brohede J, Primmer C, Moller A, Ellegren H. Heterogeneity in the rate and pattern of germline mutation at individual microsatellite loci. Nucleic Acid Research. 2002; 30(9):1997-2003.

## **AGRADECIMIENTOS**

- Peritos del Convenio ICBF-IML Por la tipificación de los casos
- Dra. Lilian TorresMSc: Por su apoyo en la discusión y análisis de resultados.
- Dra. Erica Salguero CelisMSc: Por el análisis y verificación de resultados del locus D19S433.
- Dra. Lilia Judith LaverdeMSc: Por su estudio mutacional del locus SE33.
- Dra. Andrea PinzónMSc: Por su apoyo en el cálculo de edad del parental.
- Ing. Diego Ortiz: Por su labor en la programación del sistema SIFMELCO.
- Profesor Orlando Acosta PhD. U. Nacional: Por su asesoría científica.
- Instituto Nacional de Medicina Legal y Ciencias Forenses
  - Subdirección de Investigación Científica, por la financiación de proyecto.
  - Subdirección de Servicios Forenses, por garantizar la dedicación del tiempo requerido para el proyecto

Universidad Nacional de Colombia por su incalculable formación en la disciplina científica y su impacto en la sociedad colombiana.

## ANEXO 1. DESCRIPCION DE LOS 15 LOCI STR ANALIZADOS Y SU SECUENCIAS

### 1. locus **D8S1179** (GenBank AF216671)

Se trata de un STR simple cuya unidad de repetición es ATCT, mapea en el cromosoma 8q24.13, no se encuentra asociado a regiones expresivas. el fragmento amplificado en el kit mide entre 115 y 170pb

TCTGAAGTAAGTAAAACATTGATACAGGATCCTTGGGGTGTCGCTTTTTTGGCCA  
 GAAACCTCTGTAGCCAGTGGCGCCTTTGCCTGAGTTTTGCTCAGGCCCACTGGG  
 CTCTTTTTGCCACACGGCCGGGCAACTTATATGTATTTTTGTATTTTCATGTGTAC  
 ATTCGTATCTATCTGTCT **ATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT**  
**TA**TTCCCCACAGTGAAAATAATCTACAGGATAGGTAAATAAATTAAGGCATATTCA  
 CGCAATGGGATACGATACAGTGATGAAAATGAACTAATTATAGCTACGTGAAACT  
 ATACTCATGAACACAATTTGGTAAAAGAACTGGAAACAAGAATACATACGGTTT  
 TTGACAGCTGTACTATTTTACATTCCCAACAA

### >**D21S11** (GenBank AP000433; 29 repeats)

AACTGAAGGTTACTACTATCAGCTTCCGTTGTTCTAAGGGCTTCAGACTTGGACAG  
 CCACACTGCCAGCTTCCCTGATTCTTCAGCTTGTAGATGGTCTGTTATGGGACTT

TTCTCAGTCTCCATAAATATGTGAGTCAATTCCCCAAGTGAATTGCCTTCTATCTA  
 TCTATCTATCTGTCTGTCTGTCTGTCTGTCTGTCTATCTATCTATATCTATCTATCT  
 ATCATCTATCTATCCATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA  
 TCTATCGTCTATCTATCCAGTCTATCTACCTCCTATTAGTCTGTCTCTGGAGAACA  
 TTGACTAATACAACATCTTTAATATATCACAGTTTAATTTCAAGTTATATCATACCA  
 CTTCATAACATTATATAAAACCTTACAGTGTTTCTCCCTTCTCAGTGTTTATGGCTA  
 GTAATTTTTTACTGGGTGCCAGACACTAAT

>**D7S820** (GenBank AC004848-reverse complement; 13 GATA repeats)

ACTGCAACCTCCGCTTCTTGGGTCAAGTGGTTCTCCTGCCCCAGCCTCCTGAGT  
 AGCTGGGACTACAGGCATGTGCTACTGCATCCAGCTAATTTTTGTATTTTTTTTAG  
 AGACGGGGTTTCACCATGTTGGTCAGGCTGACTATGGAGTTATTTTAAGGTTAAT  
 ATATATAAAGGGTATGATAGAACAACCTTGTCATAGTTTAGAACGAACCTAACGATAG  
 ATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGACAGATT  
 GATAGTTTTTTTTAATCTCACTAAATAGTCTATAGTAAACATTTAATTACCAATATT  
 TGGTGCAATTCTGTCAATGAGGATAAATGTGGAATCGTTATAATTCTTAAGAATAT  
 ATATTCCCTCTGAGTTTTTGTACCTCAGATTTTAAGACCTCACAATTATCTCATA  
 AGGCTTAAAATCAATCATATTTTGTAGGATCAACTTATGGTATTTTTGCCTGTTTTTA  
 TTCCTTCTGGTGTGAAAACCTGATGCCTTCCATCGTGTA

0

**CSF1PO** (GenBank X14720; 12 AGAT repeats)

**Localización:** 5q33.1; human c-fmsproto-oncogene for CSF-1 receptor gene, 6th intron

**GenBank Accesion:** [X14720](#)

**Secuencias de primers:**

5'-AACCTGAGTCTGCCAAGGACTAGC-3' (AGAT strand)

5'-TTCCACACACCACTGGCCATCTTC-3' (TCTA strand)

CAGCTGGGATGTGGAGTGGTGTGAGGAGTGGCCACAGGGGAGCAGAGGAGGT  
 GGCAGAAGCCGGAGGTAAAGGTGTCTTAAAGTGAGAAAGAATAACTGCATCTTA  
 ACCTATTGGGAGGTCATTGTAAAGAGGAGAGTGATGGGGTCAGATTGTACAGAG  
 GAGGCACTTCGTGGTGGTCAGGAGCACACACTCCAGGGCAGTGTTCCAACCTG  
 AGTCTGCCAAGGACTAGCAGGTTGCTAACCACCCTGTGTCTCAGTTTTCTACCT  
 GTAAAATGAAGATATTAACAGTAACTGCCTTCATAGATAGAAGATAGATAGATTA  
 GATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGGAAG  
 TACTTAGAACAGGGTCTGACACAGGAAATGCTGTCCAAGTGTGCACCAGGAGAT  
 AGTATCTGAGAAGGCTCAGTCTGGCACCATGTGGGTTGGGTGGGAACCTGGAG  
 GCTGGAGAATGGGCTGAAGATGGCCAGTGGTGTGTGGAAGAGTCTGAGATGCA  
 GGGATGAGGAAGAGAAAGGAGATAAGGATGACCTCCAGGTCTCTGGCTATGGT  
 GATTGGGTGCA

### Reference Sequences for Commonly Used STR Markers

The following reference sequences have proven useful for primer design with miniSTRs.

>**FGA** (GenBank M64982; 21 repeats)

ACACCTTTAAAATTCCAAAGAAAGTTCTTCTTCTATATTTCTTTGGGATTACTAATT  
 GCTATTAGGACATCTTAACTGGCATTTCATGGAAGGCTGCAGGGCATAACATTATC  
 CAAAAGTCAAATGCCCCATAGGTTTTGAACTCACAGATTAAGTGTAAACCAAAT  
 AAAATTAGGCATATTTACAAGCTAGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT  
 TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT  
 TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT  
 CTTTCTTTTTTGCTGGCAATTACAGACAAATCACTCAGCAGCTACTTCAATAACCA  
 TATTTTCGATTTTCAGACCGTGATAATACCTACAACCGAGTGTGAGAGGATCTGAG  
 AAGCAGAATTGAAGTCCTGAAGCGCAAAGTCATAGAAAAGTACAGCATATCCA  
 GCTTCTGCAGAAAATGTTAGAGCTCAGTTGGTTGATATGAAACGACTGGAGGTA  
 AGTATGTGGCTGTGGTCCCGAGTGTCTTGTTTTTGA

>**TH01** (GenBank D00269; 9 TCAT repeats)

CCTACTCCCACCTGCCCTGCCTCCCTCTGCCCCAGCTGCCCTAGTCAGCACCC

CAACCAGCCTGCCTGCTTGGGGAGGCAGCCCCAAGGCCCTTCCCAGGCTCTAG  
 CAGCAGCTCATGGTGGGGGGTCCTGGGCAAATAGGGGGCAAATTCAAAGGGT  
 ATCTGGGCTCTGGGGTGATTCCCATTGGCCTGTTCCCTTATTTCCCTCATTC  
 ATTCATTCATTCATTCATTCATTCATTCATTCACCATGGAGTCTGTGTTCCCTGTG  
 ACCTGCACTCGGAAGCCCTGTGTACAGGGGACTGTGTGGGCCAGGCTGGATAA  
 TCGGGAGCTTTTCAGCCCACAGGAGGGGTCTTCGGTGCCTCCTTGGGCACTCA  
 GAACCTTGGGCTCCCTGGCACATTTAAAATGGGTTTTTTATTTA

>**TPOX** (GenBank M68651; 11 AATG repeats)

CCACTGCTGACTTCCCCAAGCTAACTGTGCCACAGAGTGGGGACCCCCTCCCA  
 GCTCTCACAACCCCCACCTTCCTCTGCTTCACTTTTCACCAACTGAAATATGGCC  
 AAAGGC AAAAACCCATGTTCCCCTGGCCTGTGGGTCCCCCATAGATCGTAAG  
 CCCAGGAGGAAGGGCTGTGTTTCAGGGCTGTGATCACTAGCACCCAGAACCGT  
 CGACTGGCACAGAACAGGCACTTAGGGAACCCTCACTGAATGAATGAATGAATG  
 AATGAATGAATGAATGAATGAATGAATGTTTGGGCAAATAAACGCTGACAAGGAC  
 AGAAGGGCCTAGCGGGAAGGGAACAGGAGTAAGACCAGCGCACAGCCCGACTT  
 GTGTT CAGAAGACCTGGGATTGGACCTGAGGAGTTCAATTTTGGATGAATCTCTT  
 AATTAACCTGTGTGGTTCCAGTTCCTCCCCTGAGCGCCCAGGACAGTAGAGTC  
 AACCTCACGTTTGAGCGTTGGGGACGCAAACACGAGAGTGCTTGGTGTGAGCA  
 C

>**VWA** (GenBank M25858; 18 repeats)

AGCAGTAGAGAGAACTAGAGGGATCATTTACTTCAAGCCCCTCATTTTATAGACA  
 TTAGTCTCCTACAATGTGCCGGGCACTTTGCCCTTATTATTTTGTGAACTCCT  
 CAGACTGATCCTATAAGGTAGAGTTCCCACCTTCCAGAAGAAGAAACAGGTCTA  
 GAGGATCCAAGTTGACTTGGCTGAGATGTGAAAGCCCTAGTGGATGATAAGAAT  
 AATCAGTATGTGACTTGGATTGATCTATCTGTCTGTCTGTCTGTCTATCTATCTAT  
 CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCCATCTATCCATCCATC  
 CTATGTATTTATCATCTGTCCTATCTCTATCTAACCTATGTATCTATTTATCATCTA  
 TCCTGTCTCTATCTATCCTTTGTATCTATCA

>**D3S1358** (NT\_005997 positions 754983-755121; 18 repeats)

CCATCTCTTAAAAAAAAAAATTAGCCGGACATGGTGGCTCATGCCTATAATCCCA  
GGTACTTGGGAGGCTGAGGCAGGAGGACTGCTTGAGCCCAGGAGTTTGAGGCT  
GTAGTGAGCTATGATTCCCCCACTGCAGTCCAATCTGGGTGACAGAGCAAGACC  
CTGTCTCATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG  
ATAGATAGATAGACAGACAGATAGATACATGCAAGCCTCTGTTGATTTTCATGAGT  
ATAAGAGATGCCCCCAAAGGCACAGGGAATACACACCACAGAAAAATAGATCCC  
TGGGCAGAAGTGGGCAAGTGAATATGGCCAGCATGCCATTCTGGAGCAGTGC  
CCTGGCAGCTGCAGTCCTCACCTGGGAATAGCTTTTCCACT

>**D5S818** (GenBank AC008512-reverse complement; 11 AGAT repeats)

CATTTGTAATTTTCTAATTCAAAGGAGTATATAATTATGTAATAATTTTAAAATTAAA  
TACTGAGACATGCATATGCTTTTAAAGCTTCTAATTAAAGTGGTGTCCCAGATAAT  
CTGTACTAATAAAAGTATATTTTAAATAGCAAGTATGTGACAAGGGTGATTTTCCTC  
TTTGGTATCCTTATGTAATATTTTGAAGATAGATAGATAGATAGATAGATAGATAG  
ATAGATAGATAGATAGAGGTATAAATAAGGATACAGATAAAGATACAAATGTTGT  
AAACTGTGGCTATGATTGGAATCACTTGGCTAAAAGCACTAAAGCATTTCCTCTG  
AGAGAGACAATTACTTTTTTGCTTAGGAACTACCTCAACAGCCTATTAGCATCT  
GAAATATGAGGTCCACTATCCAGATGGGAGAGGTTTAGAAAAAGAAGACTTATAT  
TACTCTGTATAATGAAATGATGGAGTATTTGGAG

>**D13S317** (GenBank AL353628-reverse complement; 11 TATC repeats)

GTATTGCAAGCACTTAGTTACATTTCTAGCATATAACACATGATCAATAAATATTTT  
GACATGAACAAATGGTAATTCTGCCTACAGCCAATGTGAATATTGGGATGGGTTG  
CTGGACATGGTATCACAGAAGTCTGGGATGTGGAGGAGAGTTCATTTCTTTAGT  
GGGCATCCGTGACTCTCTGGACTCTGACCCATCTAACGCCTATCTGTATTTACAA  
ATACATTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCAATCAATC  
ATCTATCTATCTTTCTGTCTGTCTTTTTGGGCTGCCTATGGCTCAACCCAAGTTGA  
AGGAGGAGATTTGACCAACAATTCAAGCTCTCTGAATATGTTTTGAAAATAATGTA  
TATTAATGAATGTACAAATTTCCCCACTTGTACTTTCAGACTGTTATCTGTGAGTT



AAAACCTCCTCCACTCTTTTTCTACCCAAATAA

>**D16S539** (GenBank AC024591-reverse complement; 11 GATA repeats)

GGAGGATGACTGTGTTCCCACTCTCAGTCCTGCCGAGGTGCCTGACAGCCCTG  
 CACCCAGGAGCTGGGGGGTCTAAGAGCTTGTA AAAAGTGTACAAGTGCCAGATG  
 CTCGTTGTGCACAAATCTAAATGCAGAAAAGCACTGAAAGAAGAATCCCGAAAAC  
 CACAGTTCCCATTTTTATATGGGAGCAAACAAAGCAGATCCCAAGCTCTTCCTCT  
 TCCCTAGATCAATACAGACAGACAGACAGGTGGATAGATAGATAGATAGATAGAT  
 AGATAGATAGATAGATAGATATCATTGAAAGACAAAACAGAGATGGATGATAGAT  
 ACATGCTTACAGATGCACACACAAACGCTAAATGGTATAAAAATGGAATCACTCT  
 GTAGGCTGTTTTACCACCTACTTTACTAAATTAATGAGTTATTGAGTATAATTTAAT  
 TTTATATACTAATTTGAAACTGTGTCATTAGGTTTTTAAGT

>**D18S51** (GenBank AP001534-reverse complement; 18 AGAA repeats)

TACAAAAAATACAAAAATTAGTTGGGCATGGTGGCACGTGCCTGTAGTCTCAGC  
 TACTTGCAGGGCTGAGGCAGGAGGAGTTCTTGAGCCCAGAAGGTTAAGGCTGC  
 AGTGAGCCATGTTTCATGCCACTGCACTTCACTCTGAGTGACAAATTGAGACCTTG  
 TCTCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG  
 AAAGAAAGAAAGAAAGAAAGAAAAAGAGAGAGGAAAGAAAGAGAAAAAGAAAAG  
 AAATAGTAGCAACTGTTATTGTAAGACATCTCCACACACCAGAGAAGTTAATTTTA  
 ATTTTAACATGTTAAGAACAGAGAGAAGCCAACATGTCCACCTTAGGCTGACGGT  
 TTGTTTATTTGTGTTGTTGCTGGTAGTCGGGTTTGTTATTTTTAAAGTAGCTTATC  
 CAATACTTCATTAACAATTTTCAGTAAGTTATTTTCATCTTTCAACATAAATACGCACA  
 AGGATTTCTTCTGGTCAAGACCAAATAATATTAGTCCATAGTAG

>**Penta D** (GenBank AP001752; 13 AAAGA repeats)

ATGGTGAGGCTGAAGTAGGATCACTTGAGCCTGGAAGGTCTGAAGCTGAAGTGA  
 GCCATGATCACACCACTACACTCCAGCCTAGGTGACAGAGCAAGACACCATCTC  
 AAGAAAGAAAAAAGAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG  
 AAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAACGAAGGGGAAAAGAGAATCA

TAAACATAAATGTAAAATTTCTCAAAAAAATCGTTATGACCATAGGTTAGGCAAAT  
 ATTTCTTAGATATCACAAAATCATGACCTATTAATAAATAATAAAGTAAGTTTC  
 ATCAAACTTAAAAGTTCTACTCTTCAAAAGATACCTTATAAAGAAAGTAAAAAGA  
 CACGCCACAGGCTAAGAGAAAGTACTTCTAATCACATATCTAAAAAGGACTTGT  
 GTCCAGATTAAGAATTCTTACACATCAATAAGACAACCCAATTAAAAATGGGCA  
 AAAGATTTGAAGAGATATTTAACCAAAGAAAACATATAAATGTG

>**Penta E** (GenBank AC027004-reverse complement; 5 AAAGA repeats)

ACTGGTCTACTTTGGGCTTAAAGTTGACGTCTCATTGCATTGAAAATTATTTGATA  
 AGAGAAAATAAAATACATTTTACCAACATGAAAGGGTACCAATAACAAGAAAATT  
 GTGGACAGGTGCGGTGATTCACGCCTGCAATCCTAGCACTTTGGGAGGCCGAT  
 GCAGGTGTATTACCTGAGCTCAGGAGATCAAGACCAGCCTGGGCAACATGGTGA  
 AACCCCGTCTCTACTAAAATACAAAAAATTAGCTGGGTGTGGTGGTAGGCACCT  
 GTAATCCCAGCTACTCTGGAGGCTGAAACAGGAGAATCACTTGAACCCAGGAGG  
 TGGAGATTGAAGTGAGCCGAGATCACGCCATTGCACTCCAGCCTGGGCGACTG  
 AGCAAGACTCAGTCTCAAAGAAAAGAAAAGAAAAGAAAAGAAAATTGTAAGGAGT  
 TTTCTCAATTAATAACCCAAATAAGAGAATTCTTTCCATGTATCAATCATGATACT

>**D2S1338** (GenBank AC010136; 20 repeats)

TAAAGACTTCATGGTCCTGACTACAGTGACAGCAAACCCTGAAAATGGCAATTC  
 CTAAGGCCCATAATCCAGCTGTGGGAGGGAGCCAGTGGATTTGGAACAGAAA  
 TGGCTTGGCCT-

TGCCTGCCTGCCTGCCTGCCTGCCTGCC-

TTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC-

GTCC-TTCCTTC-

CTCCTGCAATCCTTTAACTTACTGAATAACTCATGATTATGGGCCACCTGCAGGT  
 ACCATGCTAGGTACTAGGGATGTAGGCATGAACACTGACAAGGGCCTCTGGGAC  
 TGGCATTCTGGTAGGAAAAGGGGTGAGACAGGGAAGAAGCCAGCAAATGTATC  
 AACAGAAACAGTTCTAAGTGCTAGGAAGAAATGAACGTATTGATGTCACATAGA  
 TTTGAGGCTCAAAGCAATCGAAGGACTTTAATGGGGTGATCAGAGGAGGCC

gtgggaggaagccagtggtttgaaacagaaatggcttggccttgctgcctgcctgcc

