



Universidad Nacional Autónoma de México
Facultad de Estudios Superiores Zaragoza



“Frecuencias alélicas de 16 marcadores polimórficos del sistema PowerPlex ESX 17 en la población del Valle de México para su aplicación en el campo de la identificación humana”

T E S I S

Para obtener el título de Químico Farmacéutico Biólogo

Presenta

Ricardo Jesús Ríos Rivas

Número de cuenta: 302293540

Directora de Tesis

M en C María Lourdes Vega Navarrete

Facultad de Estudios Superiores Zaragoza

Asesora de Tesis

Dra. Raquel Retana Ugalde

Facultad de Estudios Superiores Zaragoza



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Primero que nada quiero agradecer a mis padres el Sr. Eduardo Ríos Alemán y la Sra. Rosa María Rivas Huerta, por todo el amor, esfuerzo y sacrificio durante todos estos años. En verdad que los quiero mucho por ser los mejores padres del mundo.

A mi hermana Mariela Ríos Rivas a quien tanto quiero, por todo el apoyo que siempre me ha brindado y por siempre impulsarme a conseguir mis metas.

A la maestra Lourdes Vega Navarrete, a quien admiro y respeto. Muchas gracias por haberme permitido trabajar en su laboratorio para la realización de esta tesis, por haberme ayudado a abrir la puerta en el área de la Genética Forense, por enseñarme tanto, por darme la oportunidad de trabajar en uno de los lugares donde siempre había soñado y por la confianza que ha puesto en mí.

A la bióloga Dora Villegas Carmona por haberme brindado su amistad, por toda su ayuda y por guiarme en este camino de la Genética Forense.

A la química Erika Ramírez Flores por su amistad y apoyo que me brindo para realizar esta tesis.

Al Ing. Jorge A. Zogbi Velázquez por regalarme la oportunidad de estar en su laboratorio y proporcionarme todos los medios para realizar esta tesis. ¡Muchas Gracias!

Índice

Abreviaturas.....	5
Resumen.....	7
1.- Introducción	8
2.- Marco Teórico.....	11
2.1 Estructura y función del ADN.....	11
2.2 Herencia Mendeliana	13
2.3 Genoma humano	14
2.4 Polimorfismos	17
2.4.1 Polimorfismos de secuencia.....	17
2.4.2 Polimorfismos de longitud.....	17
2.4.3 Secuencias SINE.....	17
2.4.4 Secuencias LINE	18
2.4.5 Repeticiones tipo LTR	18
2.4.6 Transposones.....	18
2.4.7 ADN satélite.....	19
2.4.8 ADN minisatélite	19
2.4.9 ADN microsatélite.....	20
2.5 Análisis de polimorfismo por reacción en cadena de la polimerasa (PCR).....	21
2.5.1 Componentes de la PCR.....	23
2.5.1.1 ADN molde	23
2.5.1.2 Taq polimerasa	23
2.5.1.3 <i>Primers</i>	23
2.5.1.4 Cloruro de Magnesio (MgCl ₂)	24
2.5.1.5 Desoxinucleótidos trifosfatos (dNTPs).....	24

2.5.2 PCR Multiplex.....	24
2.6 Análisis de STRs por Electroforesis Capilar (EC).....	25
2.7 Aplicaciones de los polimorfismos del tipo STRs.....	27
2.7.1 Casos de paternidad	27
2.7.2 Investigación criminal	27
2.7.3 Personas desaparecidas	28
2.7.4 Seguimiento de trasplantes	30
2.8 Bases de Datos	33
2.8.1 Bases de datos alrededor del mundo	34
2.8.1.1 Base de datos de INTERPOL.....	34
2.8.1.2 Reino Unido.....	35
2.8.1.3 Canadá	35
2.8.1.4 Estados Unidos	36
2.8.1.5 México	37
2.8.2 Bases de datos poblacionales	38
2.9 Genética de poblaciones	39
a) Frecuencias alélicas y genotípicas	39
b) Ley de equilibrio de Hardy-Weinberg.....	39
c) Poder de discriminación (PD)	39
d) Probabilidad de identidad (P_1).....	40
e) Poder de exclusión (PE)	40
f) Heterocigosidad	40
g) Índice de paternidad	41
h) Índice de contenido polimórfico (PIC)	41
i) Número efectivo de alelos (NEA).....	41

2.10 Loci estudiados	42
3. Planteamiento del problema	44
4. Objetivos	45
5. Hipótesis	45
6. Metodología	46
6.1 Tipo de estudio	46
6.2 Población	46
6.3 Criterios de inclusión, exclusión y eliminación	46
6.4 Variables	47
6.5 Material y método	47
6.5.1 Material y Equipos	47
6.5.2 Método	48
7. Resultados	52
8. Discusión de resultados	62
9. Conclusiones	67
10. Recomendaciones	68
11. Referencias Bibliográficas.....	69
12. Anexo.....	78

Abreviaturas

A	Adenina
C	Citosina
CODIS	Sistema de índice combinado de ADN (Combined DNA Index System)
dNTPs	Desoxinucleótido trifosfato
EC	Electroforesis capilar
EDNAP	<i>European DNA Profiling</i>
ENFSI	<i>European Network of Forensic Science Institutes</i>
ESS	<i>European Standard Set</i>
FBI	<i>Federal Bureau of Investigation</i>
G	Guanina
IP	Índice de paternidad
LINE	<i>Elementos nucleares dispersos largos (Long Interspersed Nuclear Elements)</i>
LTRs	Repeticiones Terminales Largas <i>Long Terminal Repeats</i>)
Mg	Magnesio
MgCl ₂	Cloruro de magnesio
NEA	Número Efectivo de Alelos
ng	Nanogramos
pb	Par de bases
PC	Probabilidad de coincidencia
PCR	Reacción en cadena de la polimerasa
PD	Poder de discriminación
PE	Poder de exclusión
PIC	Contenido de información polimórfica
RFLPs	Polimorfismo de longitud de fragmentos de restricción
SINE	<i>Elementos nucleares dispersos cortos (Short Interspersed Nuclear Elements)</i>
STRs	Repeticiones cortas en tándem (Short tandem repeats)

T Timina
VNTRs Número Variable de Repeticiones en Tandem (Variable Number of Tandem Repeats)

Resumen

En el presente trabajo, se realizó un estudio genético-poblacional para los 16 marcadores del tipo STRs incluidos en el sistema **PowerPlex® ESX 17, de Promega Corp**, a partir de una población de 162 individuos que pertenecen a la población del Valle de México y que además no se encuentren relacionados biológicamente entre ellos.

Se calcularon de manera manual las frecuencias alélicas y parámetros estadísticos de interés forense (heterocigosidad observada, heterocigosidad esperada, homocigosidad, probabilidad de coincidencia, poder de discriminación, poder de exclusión, número efectivo de alelos, índice de paternidad, contenido de información polimórfica, el equilibrio de Hardy-Weinberg y el equilibrio de ligamiento).

Los resultados obtenidos demuestran que los marcadores D18S51 y D2S1338, no se encuentran en equilibrio de Hardy-Weinberg. También se puede observar que el marcador SE33 tiene un alto poder de discriminación (0.9913), alto poder de exclusión (0.6954) y una baja probabilidad de coincidencia al azar (0.0086) que lo hace ser un candidato para ser utilizado en seguimiento de trasplantes, pruebas de paternidad y en casos forenses dentro de la población de estudio.

1.- Introducción

Los avances científicos y tecnológicos dentro del campo de la genética han permitido que ésta pueda ser aplicada en la resolución de problemas de índole judicial, dando paso a lo que hoy en día se conoce como Genética Forense.

La Genética Forense es una especialidad que se basa en el estudio de la transmisión de caracteres hereditarios y el análisis de la variabilidad genética humana aplicada a la resolución de problemas judiciales y civiles. Precisa con un alto grado de certeza la relación entre los individuos, como en los casos de parentesco; o bien, entre la evidencia biológica recuperada en el lugar de los hechos y el o los probables responsables. ⁽¹⁾

Actualmente las regiones del ADN que más se utilizan en el área de la genética forense son los del tipo microsatélite (**Short Tandem Repeats**), del tipo tetranucleótido, debido a que tienen un alto grado de polimorfismo, alta heterocigosidad, herencia independiente (no ligados), se pueden amplificar varios STRs a la vez por reacciones multiplex-PCR, pueden ser analizados por electroforesis capilar y el tratamiento estadístico de los resultados es más sencillo que el de los polimorfismo de longitud de fragmentos de restricción (RFLPs).

Los RFLPs, fueron los primeros polimorfismos de ADN en estudiarse y eran analizados digiriendo el ADN con enzimas de restricción e hibridándolo posteriormente con sondas marcadas radioactivamente. De esta manera se podían estudiar los diferentes fragmentos que se originaban al cortar el ADN con una determinada enzima de restricción. La variabilidad entre individuos radicaba en el número de sitios de corte para esa enzima, de manera que para cada individuo se obtenía un patrón de bandas diferente.

Los microsatélites estudiados en este trabajo se transmiten por herencia mendeliana simple y constituyen un claro ejemplo de la diversidad genética entre individuos de la misma especie.

En general, las características que debe poseer un buen marcador genético desde el punto de vista forense, son las siguientes:

- ❖ Patrón de herencia bien establecido
- ❖ Elevado polimorfismo
- ❖ Alto grado de heterocigosidad
- ❖ Tasa de mutación baja
- ❖ Analizable mediante un método simple, rápido y reproducible
- ❖ Precisar de poco material para el análisis.⁽²⁾

Estas características se encuentran en las regiones microsatélites, las que se analizaron en el presente trabajo; sin embargo, también es importante conocer con qué frecuencia se encuentra cada variante alélica con el fin de establecer la rareza de un determinado perfil genético en la población. Información que es importante en estudios de identificación humana.

Por lo anterior es importante crear una base de datos del repertorio de frecuencias alélicas de una población para generar estimaciones precisas de qué tan frecuente es un perfil genético en la población.

En el presente trabajo se analizaron 16 marcadores genéticos de tipo STRs, los cuales se encuentran incluidos dentro del kit **PowerPlex® ESX 17**, de **Promega Corp.** que incluye el marcador **SE33**, el cual es uno de los marcadores genéticos más polimórficos e informativos que se utilizan en el área forense ⁽³⁾ y del que no se conoce la frecuencia de sus alelos en la población del Valle de México.

El estudio se realizó en una muestra representativa de 162 individuos que pertenecen a la población del Valle de México, los cuales fueron tomados al azar y no se encuentran relacionados biológicamente entre ellos. Una vez obtenidos los 162 perfiles genéticos, se procedió a realizar el análisis estadístico necesario para determinar si los marcadores genéticos incluidos dentro del sistema **PowerPlex® ESX 17**, podrían ser utilizados para realizar los diversos análisis de interés forense

dentro de la población referida y con ello decidir si dicho sistema puede ser una buena herramienta que ayude a resolver problemas de índole civil y judicial en la población de estudio.

2.- Marco Teórico

2.1 Estructura y función del ADN

El ADN contiene toda la información necesaria para constituir un individuo completo. Todos los seres vivos tienen como portador de la información genética al ADN y es el mismo en todas las células del individuo, es decir, el ADN de un hombre es el mismo en sangre, células de la piel, semen o en saliva. ⁽¹⁾

En 1953 James Watson y Francis Crick propusieron el modelo de la estructura tridimensional de la molécula del ADN. La molécula de ADN es un polinucleótido constituido por dos cadenas antiparalelas de unidades de desoxirribonucleótidos unidos covalentemente, dispuestos de una forma complementaria y adoptando una estructura enrollada de doble hélice. ^(1, 4, 5 y 6) (Fig. 1)

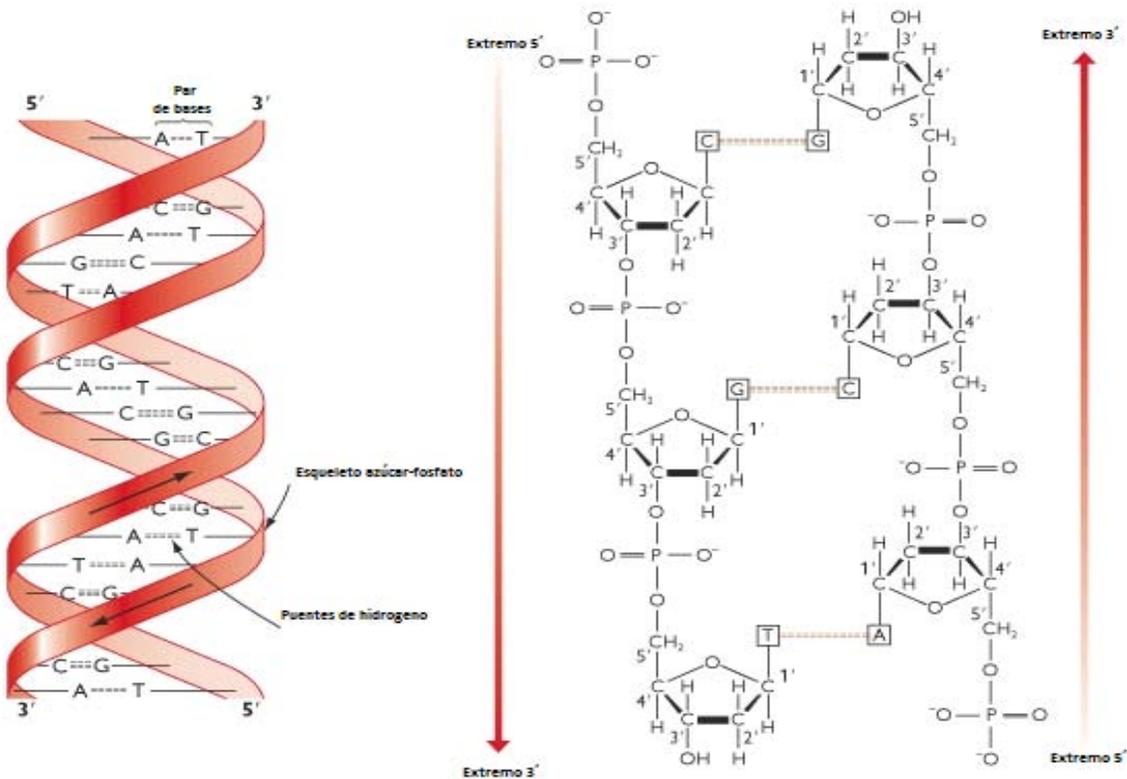


Figura 1. Estructura química de ADN. Modificado de Brown TA. ⁽⁷⁾

Los nucleótidos están unidos covalentemente por la formación de enlaces fosfodiéster donde el grupo fosfato en el carbono 5' de la desoxirribosa se une al carbono 3' de la siguiente desoxirribosa, formando un "esqueleto" de azúcar-fosfato-azúcar-fosfato dispuestos alternadamente. ^(1, 5,6 y 8)

Cada nucleótido está formado por tres componentes característicos: la desoxirribosa la cual es un azúcar de cinco carbonos que se encuentra unida a un grupo fosfato y a una base nitrogenada.

En la molécula de ADN podemos encontrar cuatro tipos de nucleótidos: adenina (A), citosina (C), guanina (G) y timina (T). La adenina y guanina son llamadas purinas y la citosina y timina pirimidinas. ^(5, 6 y 9) (Fig. 2)

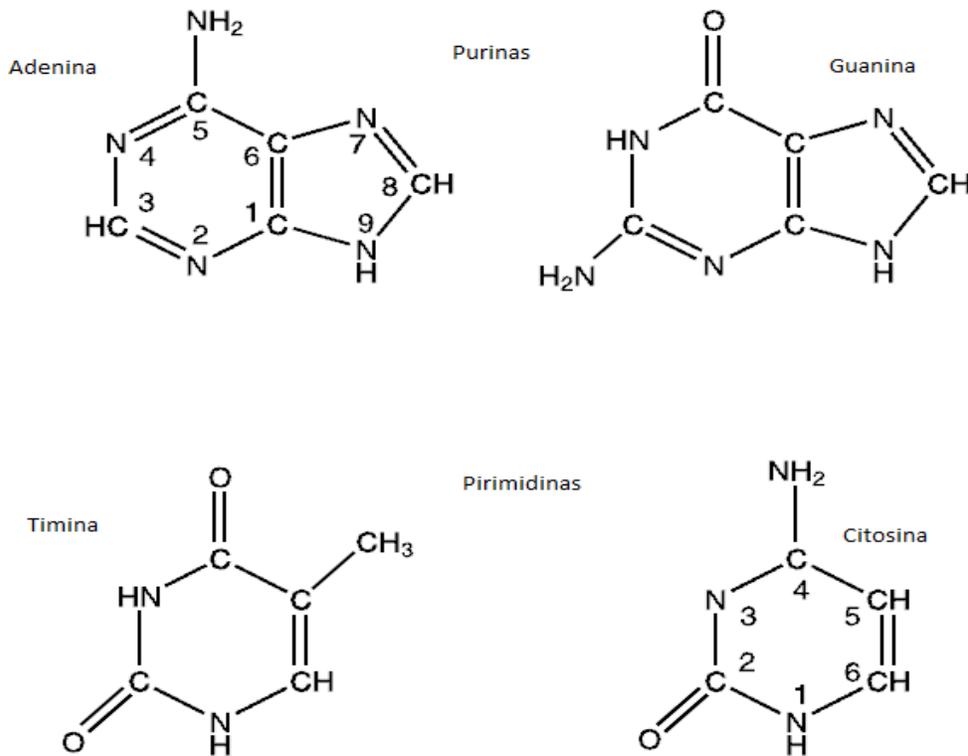


Figura 2.- Estructura química de las bases presentes en la molécula de ADN. Tomado de Kobilinsky L. ⁽¹⁰⁾

Las dos cadenas de polinucleótidos en la doble hélice de ADN se mantienen unidas por puentes de hidrógeno entre sus bases. Todas las bases están en consecuencia en el interior de la hélice y el esqueleto de azúcar-fosfato en el lado externo. Sin embargo, las bases no se aparean al azar: una adenina (A) siempre se aparea con una timina (T) y una guanina (G) con una citosina (C). Este apareamiento de las bases complementarias permite a los pares de bases colocarse en la disposición energética más favorable en el interior de la doble hélice. En esta disposición, cada par de bases presenta casi el mismo tamaño y mantiene de este modo el esqueleto azúcar-fosfato a una distancia semejante a lo largo de la molécula de ADN. Los dos esqueletos de azúcar-fosfato se enrollan entre sí para formar una doble hélice que contienen 10 bases por cada giro de hélice. Este enrollamiento también contribuye a la conformación energéticamente favorable de la doble hélice del ADN. ⁽⁵⁾

2.2 Herencia Mendeliana

Los gametos femenino y masculino (ovocito y espermatozoide), respectivamente contienen 23 cromosomas cada uno. Este es el número haploide para la especie humana.

En la meiosis el número de cromosomas se reduce de modo que se forman células que sólo contienen un miembro de cada par de cromosomas homólogos, es decir, el número diploide de 46 cromosomas es reducido a un número haploide de 23.

Después de la fertilización del ovocito por un espermatozoide se restablece el número diploide. En las células somáticas, los cromosomas se presentan como 23 pares homólogos para formar el número diploide de 46 cromosomas. Hay 22 pares de cromosomas apareados, los autosomas, y un par de cromosomas sexuales. Si el par de cromosomas sexuales es XX, el individuo es genéticamente femenino; si el par es XY, el individuo es genéticamente masculino. A cada miembro de cada par de cromosomas se le denomina homólogo. El par homólogo de cromosomas comparte los mismos *loci*.

Los genes localizados en cromosomas homólogos son distribuidos de tal manera que durante la meiosis son transferidos a los gametos de manera independiente. Dicho de otra manera, un gameto sólo recibe un cromosoma de cada par de cromosomas homólogos. A esto se le llama segregación independiente o ley de la segregación independiente.

La mitosis es el proceso por el cual una célula se divide para dar origen a dos células hijas que son genéticamente idénticas a la célula madre. Cada célula hija recibe el complemento total de 46 cromosomas. Debido a este proceso es que todas las células nucleadas del cuerpo de un individuo poseen la misma información genética. ^(11, 12 y 13)

2.3 Genoma humano

El genoma humano es el término usado para describir el total de la información genética (contenida en el ADN) en las células humanas. El genoma humano comprende dos genomas: un complejo genoma nuclear que representa el 99.9995% del total de la información genética y un genoma mitocondrial que representa el 0.0005% restante. ⁽¹⁴⁾

El genoma humano se compone aproximadamente de 3 200 000 000 pb de información que es organizada en 23 cromosomas los cuales son densos paquetes de ADN y proteínas (en su mayoría histonas) – los seres humanos contienen dos juegos (*sets*) de cromosomas – una versión de cada cromosoma heredada de cada uno de sus padres teniendo un total de 46 cromosomas, de los cuales 22 pares de cromosomas son autosómicos y 2 cromosomas sexuales que determinan el sexo (Fig. 3). Con base en los cromosomas sexuales los varones se designan como XY porque tienen una sola copia del cromosoma X y una sola copia del cromosoma Y, mientras que las mujeres tienen dos copias del cromosoma X y se designan como XX. ^(8 y 15)

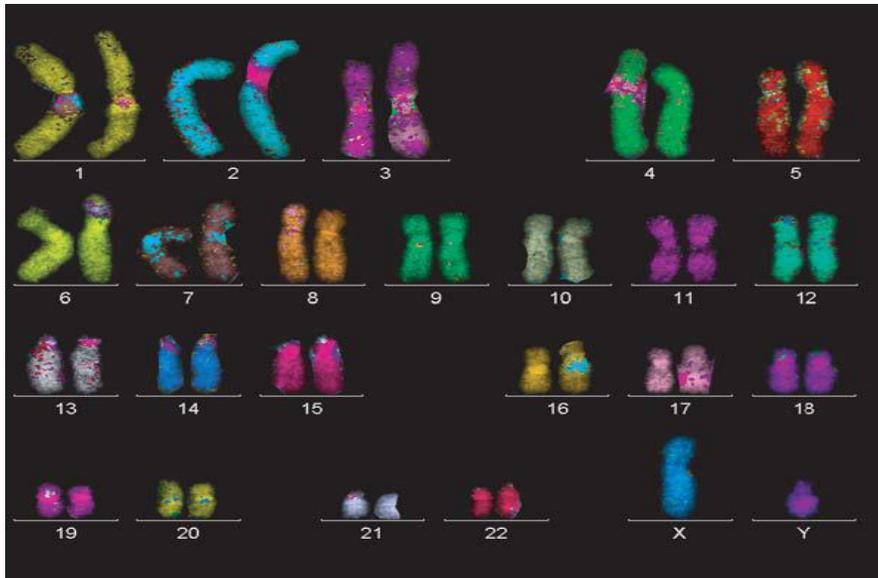


Figura 3.- Cariotipo masculino en donde se muestra los 22 pares de cromosomas autosómicos así como los cromosomas sexuales X y Y. Goodwin L. ⁽¹⁵⁾

Cada cromosoma contiene una cadena continua de ADN - el cromosoma 1 tiene un tamaño aproximadamente de 250 000 000 pb mientras que el cromosoma 22 tiene un tamaño aproximado de 50 000 000 pb. En términos físicos el intervalo de longitud de un cromosoma va de los 73mm a los 14mm. ⁽¹⁵⁾

En este trabajo que trata sobre la identificación humana, seguiremos un criterio de estudio del ADN y el genoma humano basado en su funcionalidad para identificar personas, de manera muy particular en el ADN nuclear. Por lo tanto el ADN nuclear desde el punto de vista que nos interesa se puede dividir en dos grandes grupos: ADN codificante o expresivo y ADN no codificante (Fig. 4).

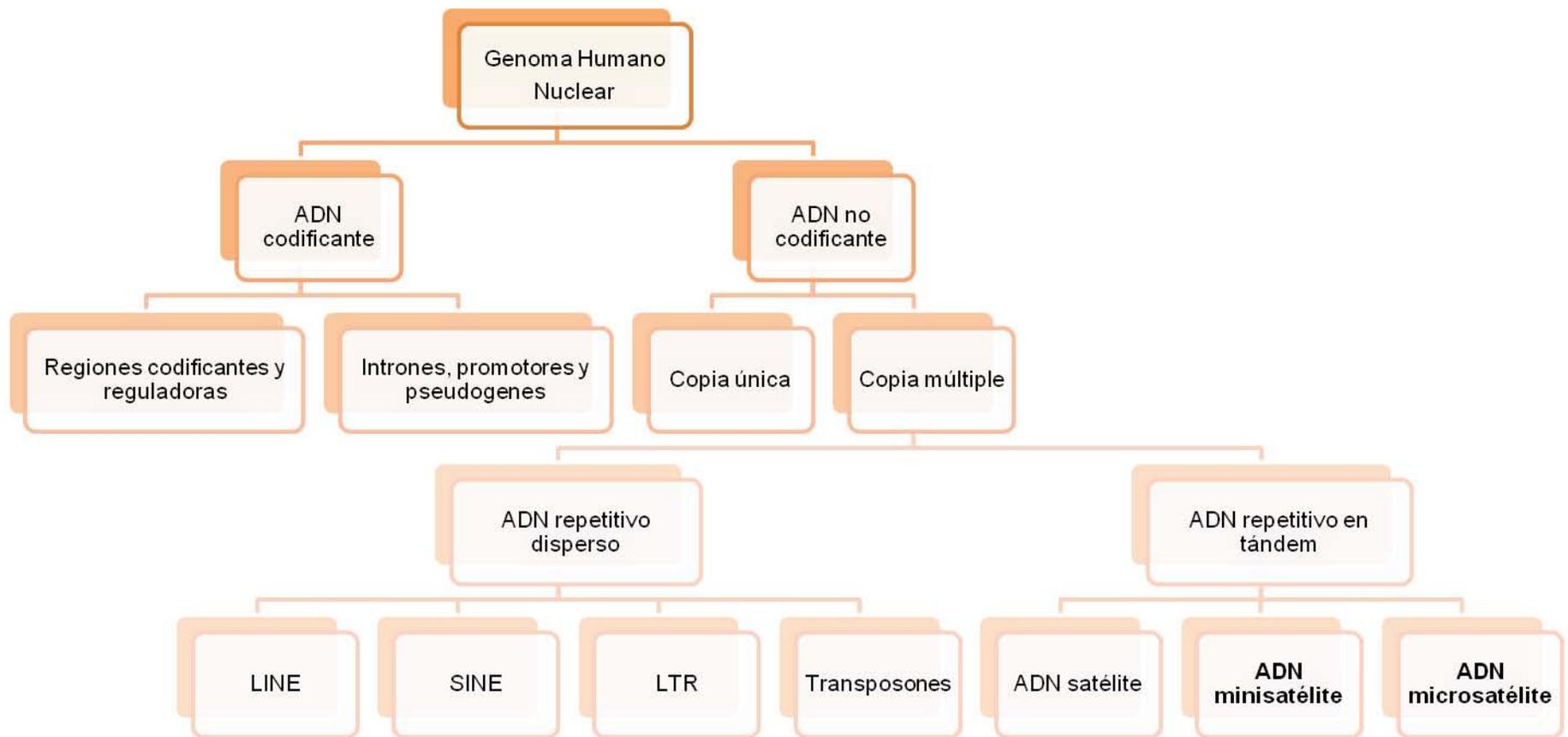


Figura 4.- Organización del Genoma Humano, según función y organización. Modificado de Goodwin L y Brown TA. ^(7 y 15)

2.4 Polimorfismos

Un polimorfismo expresa la variabilidad que existe dentro de un fragmento de ADN. Como regla general, cuantos más alelos haya, mayor polimorfismo y por ende, mayor poder de identificación. ⁽⁹⁾

Los polimorfismos se pueden clasificar como:

2.4.1 Polimorfismos de secuencia.

Se producen por el cambio de uno o más nucleótidos en una secuencia de ADN. Suelen ser poco polimórficos y son típicos de ADN expresivo.

2.4.2 Polimorfismos de longitud.

Se producen por la inserción o deleción de uno o más nucleótidos. Este tipo de polimorfismo es el más abundante en el ADN repetitivo, sobre todo en las regiones mini y micro satélite. ⁽¹⁶⁾

El ADN Repetitivo Disperso que constituye del 15-20% del genoma y consiste de secuencias de repeticiones intercaladas individualmente en forma de unidades sencillas que se distribuyen por diversos puntos del genoma y se divide en:

2.4.3 Secuencias SINE (*Short Interspersed Nuclear Elements*).

Consiste de repeticiones menores de 500 pb y también se le denomina *Alu*, por la presencia de un sitio de reconocimiento para la enzima de restricción *Alu* I. Estas secuencias son las más abundantes en el genoma humano y tienen una secuencia muy conservada de 300 pb rica en GC localizada en las regiones cromosómicas más activas a nivel transcripcional en las bandas R (de replicación precoz) de los cromosomas (elementos intercalados cortos, formados por menos de 500 pb). ^(1 y 16)

2.4.4 Secuencias LINE (*Long Interspersed Nuclear Elements*, o elementos nucleares dispersos largos).

Constituyen un 20% del genoma humano. Son secuencias con un tamaño de varias kilobases, agrupados en distintas familias.

El principal LINE es el llamado LINE-1 ó **L1**, formado por una secuencia de unas 6 kb repetida unas 800,000 veces en el genoma (aunque muchos de estos elementos no están completos, sino truncados y les falta la parte 5'), llegando a constituir alrededor de un 15% del genoma. Estos elementos, al contrario que los SINE, no son ricos en guaninas+citosinas (tienen un **42% de citosinas+guaninas**, que es cercano al contenido promedio del genoma humano) y se localizan predominantemente en las bandas G de los cromosomas. ^(1, 15 y 16)

2.4.5 Repeticiones tipo LTR

Los **HERV** (retrovirus endógenos humanos), representan copias de los retrovirus humanos que se han ido integrando en el genoma humano en el curso de la evolución y con frecuencia son el origen de proto-oncogenes celulares. Habitualmente representan copias truncadas del genoma de estos virus, y constituyen alrededor de un 8% del genoma (hay unas 450.000 copias). Como generalmente conservan alguna de las repeticiones terminales largas de estos genomas, se denominan también repeticiones tipo **LTR** (*Long Terminal Repeat*). ⁽⁷⁾

2.4.6 Transposones

Los transposones son secuencias móviles de ADN que pueden migrar a diferentes regiones del genoma. Estos elementos móviles del ADN han acompañado a los seres vivos durante su evolución contribuyendo decisivamente a los cambios genéticos. En general, se estima que una gran porción del ADN basura (no codificante) está formado por o procede de transposones. ⁽⁷⁾

2.4.7 ADN satélite

El genoma humano contiene en total unas 250 Mb de **ADN satélite** (llamado así porque al separar el ADN genómico en gradientes de densidad aparece como 3 bandas "satélites" de la banda principal). El ADN satélite está formado por la repetición de una secuencia de ADN miles de veces en tandem, es decir unas copias pegadas a otras. Esto da lugar a regiones repetidas con tamaños que van desde 100 kb hasta varias megabases. ⁽¹⁶⁾

2.4.8 ADN minisatélite

Estas secuencias de ADN fueron descritas por primera vez en 1985 por Jeffreys, quien reportó la presencia de secuencias de 0.1 a 40 kb con una unidad de repetición de 10-100 pb. Las cuales poseían un alto grado de variabilidad tanto en el número de repeticiones en tándem como en la secuencia de la unidad de repetición. ^(8, 12 y 17)

Nakamura y sus colaboradores denominaron a estas secuencias VNTR's (*Número Variable de Repeticiones en Tandem*), y las describieron como regiones altamente polimórficas y que presentaban índices de heterocigosidad mayores de 50% e incluso a veces, cercanos al 100%. ^(1, 15, 17 y 18)

La técnica usada por Jeffreys para analizar los VNTRs fue llamada RFLPs (Polimorfismos de longitud para fragmentos de restricción), debido a que involucra el uso de enzimas de restricción para cortar fragmentos de ADN, situadas dentro de las regiones del tipo VNTR y su posterior hibridación con sondas. ^(8 y 19)

Inicialmente se usaron *sondas multilocus* (MLPs) que detectan múltiples *loci* minisatélites bajo condiciones poco rigurosas de hibridación generando patrones complejos de bandas que corresponden a distintos *loci* con secuencias relacionadas entre sí y que serían específicos de cada individuo, a los que Jeffreys denominó "Huellas genéticas" ("*DNA fingerprints*")

Posteriormente Wong y Nakamura en 1987 introdujeron el uso de *sondas unilocus* (SLPs o *Single Locus Probes*) que proporcionan una o dos bandas según el carácter homocigoto o heterocigoto del individuo para ese *locus*, obteniendo un perfil unilocus de ADN (*DNA profiling*).^(1, 2, 16 y 18)

Sin embargo el uso de los minisatélites fue limitado en la práctica forense, debido a que para su análisis se requería de una gran cantidad de ADN (50–500 ng); por lo que trabajar con ADN degradado no funcionaba; la comparación entre laboratorios era complicada; y el análisis consumía mucho tiempo. Por lo que el uso de este tipo de marcadores ha sido reemplazado por el análisis de microsatélites.^(8 y 12)

2.4.9 ADN microsatélite

Los microsatélites se encuentran distribuidos por todo el genoma y fueron llamados así por su tamaño pequeño que va de 100 a 450 pb^(20, 21 y 22) lo que los hace especialmente idóneos para ser analizados mediante la técnica de PCR. También se conocen como STRs (*Short Tandem Repeats*) pues la unidad de repetición oscila entre 2 y 7 pb.^(23, 24 y 25)

Los STRs son nombrados de acuerdo a la longitud de su unidad de repetición, siendo los tetranucleótidos los más utilizados en la identificación humana.⁽⁸⁾ Sin embargo no sólo la secuencia de los STRs varía en la longitud de su unidad de repetición sino que también puede variar en cuanto a su patrón de repetición y de acuerdo a esta variación Urquhart categorizó a los STRs en repeticiones simples, repeticiones compuestas, repeticiones complejas y repeticiones hipervariables complejas.⁽²⁶⁾

El análisis de los STRs actualmente representa la herramienta de mayor uso en el campo de la genética forense debido a su alta heterocigosidad, gran poder de discriminación sobre todo cuando se utiliza un gran número de estos marcadores de manera simultánea (*PCR-multiplex*), se pueden generar resultados a partir de una pequeña cantidad de ADN, además de que su análisis es rápido y barato.^(22, 27, 28 y 29)

En la mayoría de los casos forenses donde las muestras biológicas son expuestas a diversas condiciones ambientales o al fuego por un largo periodo de tiempo, es muy probable que exista la presencia de inhibidores de PCR o el material genético se degrade lo que ocasionaría una pobre amplificación de los alelos/loci de gran tamaño (300-500 pb).^(22, 30 y 31)

Cuando existen inhibidores de la PCR en la muestra o bien cuando el ADN se encuentra degradado es muy probable obtener perfiles genéticos parciales con *allele dropout* o *locus dropout*. Este inconveniente se ha resuelto con el uso de *primers* que hibriden lo más cerca posible de la región blanco y con ello reducir el tamaño de los *amplicones* (miniSRTs).^(21, 22 y 31)

Actualmente los miniSTRs (<130 pb) representan una herramienta poderosa para analizar muestras con ADN degradado ya que suplementan la información de los STRs actualmente utilizados en el ámbito forense.^(21, 27, 32 y 33)

2.5 Análisis de polimorfismo por reacción en cadena de la polimerasa (PCR)

La reacción en cadena de la polimerasa (PCR) fue desarrollada por Kary B. Mullis en 1983. Esta técnica ha tenido un profundo impacto en el diagnóstico médico, análisis de muestras arqueológicas, estudios antropológicos, análisis de marcadores genéticos para aplicaciones forenses, pruebas de paternidad, estudios de expresión de genes, etc.^(34, 35 y 36)

La PCR es un proceso enzimático *in vitro* en el cual una región específica de ADN es replicada una y otra vez para así producir muchas copias de una secuencia en particular. Este proceso de “fotocopiado molecular” involucra el calentamiento y enfriamiento de las muestras en un preciso patrón de ciclos.^(8, 12 y 37)

La PCR consta por lo regular de 30 ciclos repetitivos, conformados cada uno de tres pasos o fases:

- ❖ **Desnaturalización:** En este paso se lleva a cabo la ruptura de los puentes de hidrógeno del ADN para desnaturalizarlo, para lo cual se realiza una incubación a 95 °C por un minuto. Este paso expone las bases nitrogenadas del ADN blanco.
- ❖ **Hibridación o alineamiento:** Ocurre la hibridación de las cadenas desnaturalizadas del ADN blanco con los *primers* (ADN sintético de hebra sencilla) a una temperatura que facilita el apareamiento de las bases nitrogenadas complementarias de ambas clases de ADN. Generalmente este paso se lleva a cabo a una temperatura que oscila entre 50-60°C.
- ❖ **Extensión:** Se lleva a cabo a 72°C que es la temperatura a la cual la **Taq polimerasa** extiende la longitud de los *primers*, añadiendo los diferentes nucleótidos libres en el orden que le va dictando la secuencia de nucleótidos de la cadena que actúa como molde (Fig. 5).

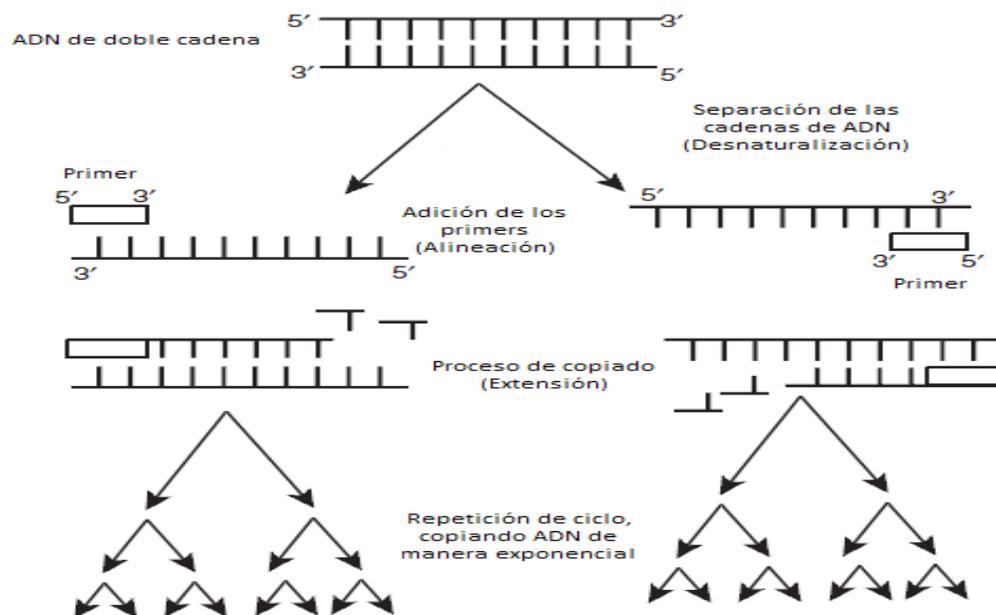


Figura 5.- Proceso de amplificación del ADN, mediante PCR. Modificado de Butler JM. ⁽⁸⁾

Esta técnica ha supuesto un avance metodológico de enorme repercusión en genética forense, ya que nos permite trabajar con pequeñas cantidades de ADN aun cuando este se encuentre degradado. Otras ventajas adicionales de la PCR son la rapidez, sencillez, facilidad de interpretación y menor costo de la determinación, además de proporcionar una buena sensibilidad y especificidad. ⁽²⁾

2.5.1 Componentes de la PCR

Para llevar a cabo una PCR es necesario contar con: la secuencia de ADN que se desea amplificar, *primers* o iniciadores, una ***Taq polimerasa*** termoestable, cloruro de magnesio, desoxinucleótidos trifosfatos (dNTPs) y un amortiguador. ⁽¹²⁾

2.5.1.1 ADN molde

La cantidad de ADN adicionado a una PCR dependerá de la sensibilidad de la reacción: para propósitos forenses la PCR es optimizada para trabajar con pequeñas cantidades de ADN. La mayoría de los kits comerciales de nueva generación requieren entre 0.1 a 1 ng de ADN genómico para obtener óptimos resultados. ^(8, 38 y 39)

2.5.1.2 Taq polimerasa

Esta es una enzima proveniente de la bacteria ***Thermus aquaticus*** la cual es resistente a altas temperaturas como las que se involucran en un proceso de PCR y trabajando de manera óptima entre 72-80°C.

Usando una enzima termoestable la PCR se convierte en una técnica muy sensible, aumentando su especificidad, sensibilidad y rendimiento. ^(37, 40 y 41)

2.5.1.3 Primers

Los *primers* o iniciadores usados en PCR definen la región del genoma que se ha de analizar. Los *primers* son pequeñas secuencias de ADN sintético de entre 18-30

nucleótidos, las cuales se alinean con la cadena molde de ADN muy cerca de la región blanco que se desea amplificar.

2.5.1.4 Cloruro de Magnesio (MgCl₂)

Es un componente crítico de la PCR ya que influye en la especificidad y el rendimiento de la reacción.

2.5.1.5 Desoxinucleótidos trifosfatos (dNTPs)

Los cuatro dNTPs (dATPs, dTTPs, dNCTPs y dGTPs) son los ladrillos con los que se construyen las nuevas cadenas de ADN. Una alta concentración de dNTPs disminuye la fidelidad con la que la polimerasa realiza su trabajo e incluso puede llegar a inhibir su actividad.

Los dNTPs pueden captar iones magnesio por lo que la concentración de ambos componentes debe guardar siempre la misma proporción aconsejándose que la concentración de Mg⁺² sea de 0.5 a 1 mM superior a la concentración de dNTPs. ^(12 y 34)

2.5.2 PCR Multiplex

Es una variante de la PCR en la cual dos o más *loci* son amplificados de manera simultánea por la inclusión de más de un par de *primers* en una misma reacción, la cual ha resultado ser una herramienta muy útil que se puede aplicar en el área forense ya que con la amplificación simultánea de hasta 24 *loci*, se puede obtener un alto poder de discriminación consumiendo una pequeña cantidad de la muestra. ^(30, 35, 38, 42 y 43)

Comercialmente se encuentran disponibles varios kits de “*STRs multiplex*” los cuales varían en los marcadores que incluyen, la combinación de colorantes fluorescentes, la cadena de ADN que es etiquetada, marcadores alélicos incluidas en el kit y lo más

importante, es que varían en la secuencia de los *primers* que se utilizarán para llevar a cabo la amplificación. ⁽⁴⁴⁾

Actualmente todos los laboratorios de genética forense trabajan con kits comerciales los cuales incluyen todos los reactivos necesarios para amplificar de manera conjunta una serie de marcadores. Con respecto a lo que sería amplificar uno a uno cada uno de los marcadores que se suelen utilizar, la *PCR multiplex* ofrece una serie de ventajas como lo son el ahorro en tiempo, menor costo, menor cantidad de muestra para el análisis y el analizar de manera simultánea varios *loci* aumenta el poder de discriminación.

2.6 Análisis de STRs por Electroforesis Capilar (EC)

La electroforesis es un método de separación de proteínas, ácidos nucleicos y otras sustancias químicas basado en la diferencia de la carga y masa de una molécula. En el caso del ADN, los grupos fosfato dentro de la molécula de ADN tienen una carga negativa.

Después de la amplificación, la longitud total de los productos de la PCR (*amplicones*) es medida para determinar el número de repeticiones presentes en cada alelo encontrado dentro de un perfil genético. Esta medición de la longitud se hace mediante una separación de los *amplicones* con base a su tamaño. ^(8 y 12)

En la EC los productos amplificados son transferidos dentro de un capilar de vidrio (50 μm de diámetro) mediante inyección electrocinética, donde se aplica de manera constante un voltaje positivo a la muestra por determinado tiempo con lo cual se permitirá a las moléculas de ADN migrar dentro del capilar, las moléculas viajarán a través de un polímero (POP-4) que llena el capilar y actúa como la matriz de tamizado (Fig. 6). La urea y la 2-pirrolidona en el POP-4, así como una temperatura de 60°C ayudan a prevenir la formación de alguna estructura secundaria durante la electroforesis.

Durante el proceso de electroforesis un rayo láser de argón incide en el capilar en una posición fija donde hay una ventana y los productos de PCR etiquetados con colorantes fluorescentes que pasan a través de esta ventana son excitados por el láser y emiten fluorescencia la cual es detectada por un dispositivo de carga acoplada (cámara CCD,) el cual convierte la fluorescencia en impulsos eléctricos que con ayuda de un software (**GeneMapper**), se representan gráficamente en un electroferograma. ^(8, 12 y 20)

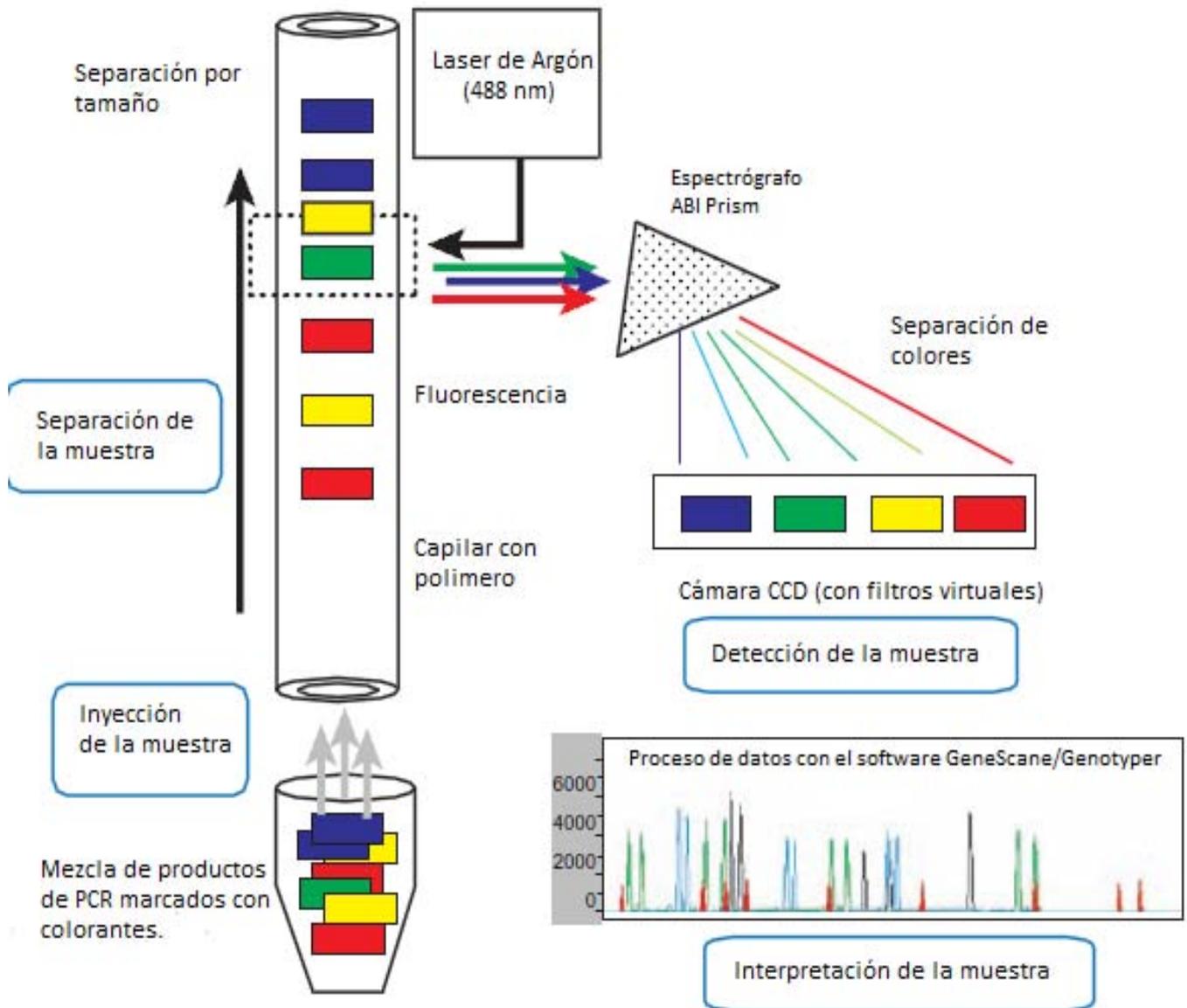


Figura 6.- Separación de detección de fragmentos STRs. Modificado de Butler JM. ⁽²⁰⁾

2.7 Aplicaciones de los polimorfismos del tipo STRs

2.7.1 Casos de paternidad

El objetivo de una prueba de paternidad, es determinar si el supuesto padre es el verdadero padre biológico, para ello los perfiles genéticos del niño, la madre y el supuesto padre son comparados. Partiendo del hecho de que el 50% de la información genética nuclear de un individuo se hereda de su padre y el 50% restante proviene de su madre, si el niño tiene un alelo que la madre no tenga, este será un alelo paterno obligado que tiene que venir del padre biológico. Si el hombre sometido a prueba (supuesto padre), no cuenta con los alelos paternos obligados, el resultado será una exclusión. Mientras que si el hombre sometido a prueba cuenta con los alelos paternos obligados, el resultado será una inclusión. En este último caso, se procederá a calcular un Índice de Paternidad (IP) lo cual ayudará a entender que tan fuerte es la asociación biológica entre el supuesto padre y el niño.

(11 y 45)

2.7.2 Investigación criminal

La utilización del ADN para realizar análisis en el ámbito de la administración de justicia ha adquirido, en pocos años, una importancia primordial sobre todo en aquellos procesos penales, en concreto en aquellos hechos delictivos que pueden dejar vestigios biológicos del autor sobre la víctima, el lugar o los instrumentos del delito, así como de la víctima sobre el autor o sus pertenencias (delitos violentos, como el homicidio, o contra la libertad sexual de las personas); pero también es un elemento particularmente útil para la identificación de cadáveres (desastres o accidentes).

En primera instancia se obtienen los perfiles genéticos de la víctima, del probable responsable y de los indicios biológicos relacionados a la investigación. Una vez

obtenidos cada uno de los perfiles genéticos, éstos serán comparados, pudiendo encontrar las situaciones siguientes:

Resultado	Descripción	Interpretación
Resultado nulo (Poco concluyente o no interpretable)	La comparación no es posible	No fue posible la obtención de un adecuado perfil genético en alguna de las muestras, debido a la insuficiente calidad o cantidad de ADN.
Resultado negativo (Exclusión)	Los perfiles son diferentes	Las muestras comparadas provienen de diferentes individuos.
Resultado positivo (Inclusión)	Los perfiles son los mismos	Las muestras comparadas provienen de un mismo individuo.

Cuadro 1. Posibles resultados en comparaciones de perfiles genéticos. Modificado de Gans J. and Urbas G.

El único resultado que necesita ser expresado alfa-numéricamente ante un juez, es la inclusión donde se determinará qué tan fuerte es la asociación entre la muestra de referencia y la evidencia y con ello determinar qué tan probable es, que ambas muestras provengan de un mismo individuo. ^(1, 11, 46 y 47)

2.7.3 Personas desaparecidas

Un desastre es un evento inesperado que causa serias heridas y muerte a un gran número de personas. Los desastres pueden ser desastres naturales o actos terroristas intencionales.

Generalmente un desastre requiere del manejo, análisis y comparación de un gran número de muestras biológicas y perfiles de ADN haciendo necesario el uso de sistemas electrónicos para el manejo de esta información.

Hay varios factores involucrados en un desastre que pueden cambiar el objetivo final de la identificación por ADN como lo son: número de víctimas, la fragmentación de los cuerpos, la degradación del ADN, la accesibilidad para la colección de muestras o el tipo de muestras de referencia disponibles.

La identificación en este tipo de eventos se puede realizar mediante la comparación de los perfiles genéticos obtenidos de las víctimas o de los fragmentos de las víctimas con respecto a los perfiles genéticos obtenidos de muestras de referencia a partir de efectos personales (coincidencia perfecta) o muestras de referencia de familiares (método de alelos compartidos). (Fig. 7)^(48 y 49)

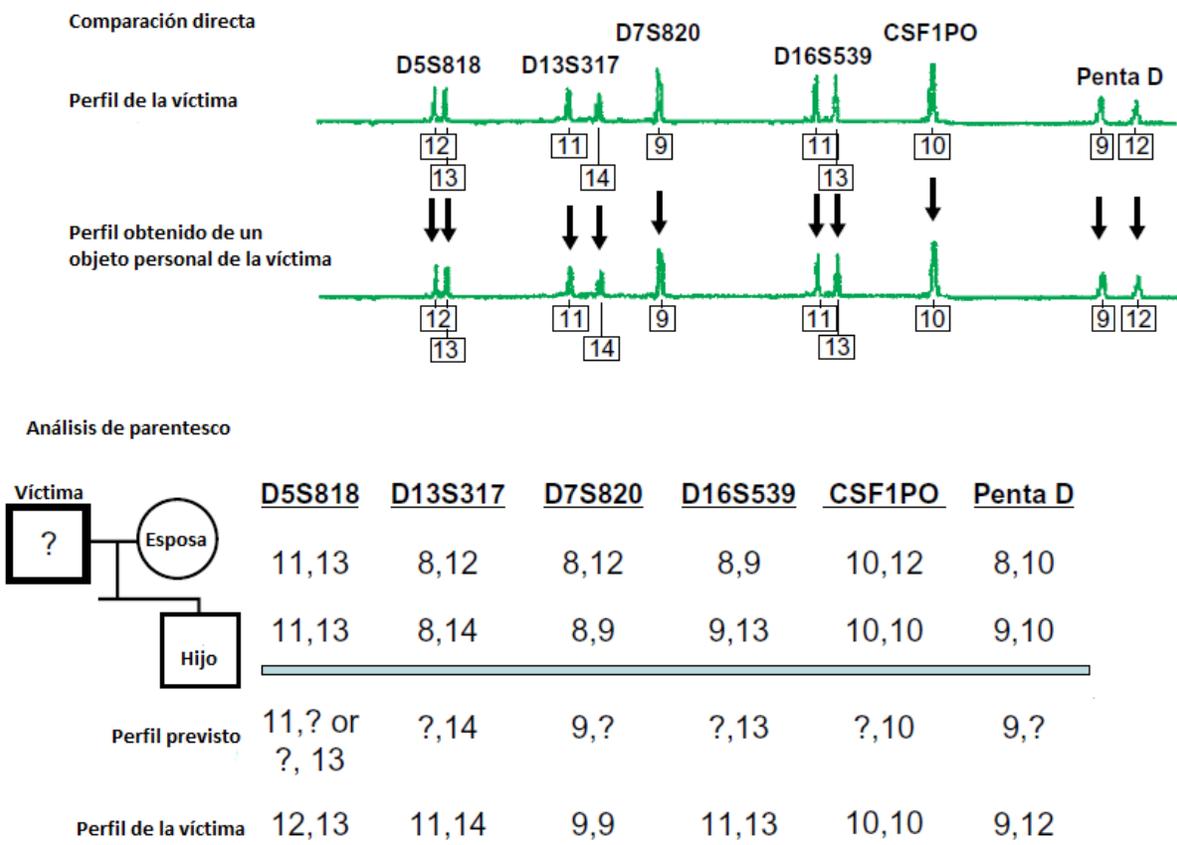


Figura 7.- Identificación de personas mediante comparación directa y mediante búsqueda por alelos compartidos. Modificado de Butler JM. ⁽¹¹⁾

En cuanto a la identificación por coincidencia perfecta, es necesario en primera instancia obtener el perfil genético de la víctima así como de los elementos biológicos identificados en algún objeto personal de la víctima. Posteriormente, se realiza un estudio de confronta entre dichos perfiles, esperando ver una correspondencia del 100% entre los perfiles genéticos obtenidos de la víctima y los elementos biológicos identificados en los objetos personales de la víctima, si éstos pertenecen a un mismo individuo. En caso contrario, los perfiles genéticos obtenidos de ambos tipos de muestras diferirán el uno del otro.

Para el caso de identificación por alelos compartidos, lo que se busca es determinar si se comparten o no alelos entre las personas implicadas en el estudio (generalmente familiares directos; padre, madre e hijos), después del estudio de un determinado número de marcadores. Lo que está en investigación es la fuente de alelo materno obligado, así como del alelo paterno obligado en cada *locus* que se encuentra en estudio.

2.7.4 Seguimiento de trasplantes

Las quimeras se definen como individuos que muestran dos o más poblaciones de células genéticamente distintas.

El concepto de quimerismo en medicina se remonta a los orígenes de los trasplantes de órganos: un paciente trasplantado se convierte en una quimera porque posee órganos de dos seres vivos genéticamente diferentes. El paciente trasplantado vive con partes de por lo menos dos diferentes seres vivos: donador y receptor.

En los últimos años, las técnicas basadas en el análisis del ADN se han constituido en la principal herramienta para el estudio del quimerismo. Si bien el *Southern blot* era la técnica más usada para el estudio del quimerismo, éste ha sido sustituido por la reacción en cadena de la polimerasa (PCR) ⁽⁵⁰⁾. Además, también han ido cambiando los distintos tipos de polimorfismos de ADN utilizados para distinguir al donante del paciente.

En la actualidad los STRs son los polimorfismos de mayor utilidad para un correcto estudio del quimerismo ^(51 y 52). Los marcadores utilizados deben ser capaces de distinguir el ADN de dos individuos emparentados, esto es, ser informativos y la técnica empleada debe poder detectar material genético de ambos aun cuando la cantidad de uno de ellos esté muy poco representada en la muestra biológica, es decir ha de ser suficientemente sensible ^(51 y 53). El cumplimiento de estos dos conceptos es crucial para una correcta clasificación de los pacientes en cuanto a su “estado quimérico”. La sensibilidad obtenida va a depender de la técnica utilizada para el estudio. En general, la PCR es la técnica de mayor sensibilidad y, por tanto, la recomendada para el análisis del quimerismo.

La elección de los marcadores polimórficos que se van a analizar es igualmente importante para obtener un resultado correcto y evitar interpretaciones erróneas. Deben de cumplir dos requisitos: a) un alto índice de heterocigosidad, y b) un número elevado de alelos en la población. Lo ideal es que el donante y el paciente no tengan ningún alelo en común, y en el caso de que lo tuvieran, el donante y el paciente han de tener un alelo específico de cada uno no compartido. Los STRs presentan las características que ha de tener todo marcador genético susceptible de ser utilizado en el análisis del quimerismo; por otra parte, el pequeño tamaño de las secuencias hace que todas ellas se puedan detectar mediante PCR y, en muchos casos, incluso es posible el estudio simultáneo de varios loci STRs mediante la PCR multiplex. Si ésta se combina con una técnica automatizada de tinción multicolor de los cebadores y lectura en ABI Prism, podremos definir el tamaño de los diferentes fragmentos amplificados y cuantificar cada uno de ellos. El estado del quimerismo hematopoyético conviene analizarlo en poblaciones celulares puras de linfocitos T y de neutrófilos, que pueden fácilmente separarse mediante métodos físicos o inmunomagnéticos. Con estos métodos, en la fracción linfoide T, más del 95 % de las células son CD3+, detectadas mediante citometría de flujo. ⁽⁵¹⁾

Atendiendo a la presencia de células del donante en el receptor, el quimerismo puede clasificarse como:

- Quimerismo total o completo (QC): donde todas las células que se detectan proceden del donante.
- Quimerismo mixto (QM): en el que coexisten células del donante y el receptor en un compartimento celular dado. Se considera QM cuando el porcentaje de linfocitos T o de neutrófilos del receptor es superior al 5 %.
- Quimerismo dividido: cuando una o más líneas celulares proceden en su totalidad del donante y a su vez una o más proceden totalmente del receptor.
- Microquimerismo: en el que existe menos del 1 % de células del donante. Este grado de quimerismo sólo se puede detectar cuando se utilizan técnicas muy sensibles.

Una quimera establecida como completa en los primeros días después del trasplante, puede evolucionar a quimera mixta e incluso desaparecer, lo que se traduciría finalmente en un fallo del injerto. Por otra parte, el estado de quimera mixta puede evolucionar hacia el fallo del injerto, si predominan las células del receptor o hacia el estado de quimera completa con predominio casi absoluto (98 a 100 %) de la hematopoyesis del donante. Por lo tanto, dado que el estado de quimera puede variar en el tiempo, es importante monitorear su evolución. ⁽⁵⁰⁾

El estudio del quimerismo linfohematopoyético ha salido del marco de los laboratorios de investigación y se ha convertido en una importante herramienta clínica en la evaluación del éxito o fracaso de los trasplantes de células hematopoyéticas. Mediante estos estudios, podemos conocer si el sistema linfohematopoyético del donante ha sido capaz de implantarse en el receptor y si lo ha hecho desplazando al sistema linfohematopoyético del receptor o coexistiendo en equilibrio con éste. De esta manera, mediante determinaciones secuenciales, es posible conocer la evolución o comportamiento de la quimera con vistas a confirmar el fallo primario del injerto, o conocer, que puede haber un fallo secundario del mismo antes que otros indicadores se manifiesten. ⁽⁵⁰⁾

2.8 Bases de Datos

Gracias a los avances científicos y técnicos en las herramientas moleculares empleadas en genética forense se ha generado una cantidad de información extraordinaria que ha de almacenarse de modo racional y ordenado para su posterior uso. Es así que surge la necesidad de las denominadas *bases de datos*, lugares donde se almacena de modo ordenado y coherente cualquier tipo de información, que luego es rescatada de modo automático de acuerdo a parámetros establecidos previamente. ^(54 y 55)

El acceso a las *bases de datos* es más o menos restringido, dependiendo de la trascendencia de los datos almacenados, siendo los más protegidos aquellos que contienen información sobre la “identidad” de las personas.

El acceso a los datos debe de estar perfectamente controlado, y las conclusiones que se pueden obtener de los mismos van a depender de los programas informáticos que se autoricen por las diferentes legislaciones. Dependiendo del contenido de estas bases de datos se pueden diferenciar:

- ❖ **Bases de datos forenses criminales:** almacenan datos procedentes de personas que han sido procesadas o condenadas, así como de indicios biológicos encontrados en el lugar de los hechos. En algunos casos pueden considerarse perfiles de víctimas conocidas, con objeto de facilitar la resolución de delitos. La característica principal es que algunas de las muestras (evidencias) y datos considerados se obtienen sin el consentimiento de las personas implicadas. ⁽⁵⁴⁾

- ❖ **Bases de datos forenses civiles:** su único fin es la identificación de personas desaparecidas, lo cual se hace comparando el ADN de las personas no identificadas con el de los familiares. Por sus características humanitarias, se requiere que los familiares colaboren de modo voluntario y tras firmar un consentimiento informado, cosa que las diferencia de las bases de datos criminales siguiendo las consideraciones del protocolo de Helsinki. ^(54 y 56)

2.8.1 Bases de datos alrededor del mundo

2.8.1.1 Base de datos de INTERPOL

Las fuerzas policiales de los países miembros pueden enviar a la base de datos informatizada de INTERPOL sobre ADN los perfiles genéticos correspondientes a delincuentes, personas desaparecidas, cadáveres no identificados, o bien perfiles encontrados en el lugar del delito. Esta base de datos fue creada en 2002 con un único perfil de ADN, pero en 2012 contenía más de 136,000 aportados por 67 países miembros. Los países participantes utilizan activamente esta base de datos, que permite detectar periódicamente semejanzas entre los diversos perfiles enviados por los países miembros, como una herramienta para sus investigaciones policiales. En 2012 las consultas realizadas por los países miembros en la base de datos dieron lugar a 84 resultados positivos a escala internacional.

Los países miembros pueden acceder a la base de datos a través del sistema mundial de comunicación policial de INTERPOL, I-24/7.

INTERPOL actúa únicamente como conducto para el intercambio y el cotejo de información; no mantiene datos nominales que vinculen un perfil de ADN con una persona. Un perfil de ADN es simplemente una lista de números basada en la estructura de ADN de una persona, un código numérico que puede servir para distinguir distintos individuos; y no contiene información sobre las características físicas o psicológicas de la persona, ni sobre sus enfermedades o su predisposición a ellas. Los países miembros que utilizan esta base de datos de ADN siguen siendo los propietarios de los datos de los perfiles, y deciden sobre el acceso de otros países a ellos, así como sobre su envío y destrucción, de conformidad con sus respectivos ordenamientos jurídicos. ⁽⁵⁷⁾

2.8.1.2 Reino Unido

En 1995 se estableció en el Reino Unido la primera base nacional de datos de ADN (la *National DNA Database* o “NDNAD”), esta fue la primera base de datos en administrar información extraída de marcadores del tipo STRs. ^(58, 59, 60 y 61)

Inicialmente la NDNAD permitía sólo administrar la información genética de un conjunto de seis loci STRs (FGA, TH01, vWA, D8S1179, D18S51 y D21S11). En 1999 el conjunto de marcadores que podía administrarse en esta base de datos se incrementó a 10 (FGA, TH01, vWA, D8S1179, D18S51, D21S11, D3S1358, D16S539, D2S1338 y D19S433), ya para el año 2009 a petición de la Unión Europea y con el propósito de intercambiar información entre los países de Europa, se aumentó a 15 el número de marcadores que podían administrarse en esta base de datos. ⁽²⁰⁾

La NDNAD permite almacenar y administrar los perfiles genéticos obtenidos de muestras recolectadas del lugar de los hechos así como las muestras de referencia recolectadas a los individuos relacionados a un delito.

Es importante mencionar que hasta el 31 de marzo del 2013 en la NDNAD se habían almacenado **6, 737, 973** perfiles genéticos obtenidos de muestras de referencia y **428, 634** perfiles genéticos obtenidos de evidencias recolectadas en el lugar de los hechos. ⁽⁶¹⁾

2.8.1.3 Canadá

En 1998 fue promulgada la legislación donde se creó el estatuto para la creación y administración del *National DNA Data Bank*.

El *National DNA Data Bank* es administrado por la Real Policía Montada de Canadá en Ottawa y desde su inicio en el año 2000 ha ayudado a resolver un gran número de delitos por todo el país y alrededor del mundo. En el *National DNA Data Bank* es posible almacenar y administrar los perfiles genéticos obtenidos de muestras recolectadas del lugar de los hechos así como las muestras de referencia recolectadas a los individuos relacionados a un delito.

Hasta el 31 de enero del 2012 en el *National DNA Data Bank*, se habían almacenado 237,000 perfiles genéticos obtenidos de las muestras de referencia obtenidas de personas relacionadas a un delito y más de 71,000 perfiles genéticos de las muestras recolectadas en un lugar de hechos. Además se han encontrado 21,563 “*offender hits*” (coincidencia entre una muestra de referencia y un indicio) y 2,714 “*forensic hits*” (coincidencia entre dos indicios).⁽⁶²⁾

2.8.1.4 Estados Unidos

El **CODIS** (***Combined DNA Index System***) comienza a funcionar en 1990 como proyecto piloto cuando el FBI decidió crear una base de datos en la cual se administraran los perfiles genéticos obtenidos a partir de muestras relacionadas a un delito.

En 1994 en el Acta de Identificación a través del ADN se publica la ley 103 322, en la cual el congreso de los Estados Unidos autoriza al FBI establecer y administrar una base de datos nacional (***National DNA Index System, NDIS***), con el único fin de ayudar a la resolución de problemas judiciales. Del mismo modo esta ley autoriza al director del FBI, establecer los requisitos mínimos de calidad que deben cumplir los laboratorios que participan en el programa del CODIS.

Oficialmente el 13 de Octubre de 1998 el FBI comienza a operar la base de datos nacional de los Estados Unidos. El CODIS se divide en tres niveles: local, estatal y nacional. Siendo el nivel nacional el más alto dentro del CODIS y posibilita a los laboratorios participantes en el programa CODIS el comparar e intercambiar perfiles genéticos.

En el CODIS se pueden administrar los perfiles genéticos obtenidos de indicios biológicos relacionados a un delito, personas relacionadas a un delito, familiares de personas desaparecidas y de restos humanos no identificados. También es importante señalar que son 13 los marcadores (CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 y D21S11) del tipo STRs que se seleccionaron para ser usados dentro del CODIS, todos estos marcadores son altamente polimórficos y 11 de ellos se encuentran localizados en

diferentes cromosomas a excepción de CSF1PO y D5S818 que se encuentran en un mismo cromosoma. Todos los *loci* se encuentran localizados fuera de regiones codificantes y no tienen ninguna significancia funcional. (8, 12, 58, 59, 63, 64 y 65)

2.8.1.5 México

Actualmente México no cuenta con una base de datos nacional, sin embargo es importante señalar que en muchos Estados de la República Mexicana se han desarrollado sistemas computacionales que permiten administrar los perfiles genéticos obtenidos a partir de diversos tipos de muestras.

Un caso muy particular es el del Estado de México donde se ha hecho un esfuerzo muy importante para que el laboratorio de Genética del Instituto de Servicios Periciales cuente con un sistema confiable para la administración de los datos genéticos con los que se cuenta. Actualmente se dispone de una Plataforma de Identificación Genética (**GEIGEN FORENSE® V1.0**), que es un sistema digitalizado de extrema seguridad con arquitectura cimentada con estructuras de alta disponibilidad, con lo que se logra un eficiente y confiable manejo de los datos genéticos.

TECNOLOGÍA.

Sistema operativo UNIX. No existen virus informáticos, alta seguridad para el acceso de usuarios.

Base de datos ORACLE. Alta seguridad, posibilidad de reconstruir la información. Equipo de alta tecnología en una red interna para el manejo y el flujo de los datos.

Originalmente se proyectó una base de datos criminal, en la que fuera posible realizar una búsqueda y posibles *matches* o *hits*, con información genética idéntica, es decir, entre datos genéticos que provenían de una sola fuente (evidencia vs probable responsable); sin embargo las necesidades actuales y sobre todo la necesidad de poder identificar a personas que permanecen en calidad de desconocidas o buscar personas desaparecidas y más aún, poder intercambiar

información con otros Estados de la República, nos ha llevado a desarrollar nuevas aplicaciones en la plataforma, como lo es la posibilidad de establecer relaciones biológicas de parentesco directas entre los datos genéticos con los que cuenta nuestra base de datos, y estar en posibilidad de identificar individuos a través del análisis de sus familiares (padres o hijos).

Cabe mencionar que esta Plataforma es única en su tipo en todo el país y que hasta el día 22 de mayo del 2015 se han almacenado un poco más de 6, 500 perfiles genéticos.

Por último, es importante mencionar que los perfiles genéticos obtenidos de las muestras de referencia colectadas a familiares de personas desaparecidas y probables responsables de un delito de quienes se solicita realizar un estudio genético en el Laboratorio de Genética Forense del Instituto de Servicios Periciales del Estado de México se hace bajo la autorización por consentimiento informado donde se les hace saber que los datos genéticos obtenidos de sus muestras serán ingresados a una base de datos para realizar las investigaciones correspondientes.

2.8.2 Bases de datos poblacionales

El objetivo principal de generar una base de datos poblacional es poder encontrar los alelos más comunes y cuantas veces se repiten estos alelos en múltiples muestras con el fin de estimar los alelos presentes en la población de estudio y de esta manera establecer el peso estadístico de los perfiles genéticos y las confrontas de los mismo en la identificación de individuos.

Las bases de datos poblacionales son usadas para estimar la rareza de un perfil genético en una población con el fin de dar una indicación a un tribunal de la solidez de las pruebas de ADN. ⁽⁵⁹⁾

2.9 Genética de poblaciones

a) Frecuencias alélicas y genotípicas

La frecuencia de un alelo se define como el número de copias del alelo en la población, dividido por la suma de todos los alelos en esa población.

La frecuencia genotípica es la cantidad de veces que aparece el genotipo en una población. ⁽⁶⁶⁾

b) Ley de equilibrio de Hardy-Weinberg

La Ley de Hardy-Weinberg fue postulada en 1908 por George Hardy y Wilhelm Weinberg. Esta ley es una representación matemática de la relación entre las frecuencias alélicas y genotípicas de una población ideal y la cual dice que en toda población suficientemente grande, en la que los apareamientos se producen al azar y que no está sujeta a ninguna fuerza evolutiva como mutación, selección y migración, las frecuencia alélicas y genotípicas se mantienen constantes de una generación a otra. Cuando una población cumpla con lo que dice esta ley se dirá que dicha población se encuentra en Equilibrio de Hardy-Weinberg y cumpliendo esta condición se podrán determinar las frecuencias genotípicas a partir de las frecuencias alélicas, lo cual es muy útil en la genética forense para determinar la rareza de un perfil genético dentro de una población.

La frecuencia para un genotipo homocigoto se calcula elevando al cuadrado de la frecuencia de un alelo en particular (p^2), mientras que la frecuencia de un genotipo heterocigoto puede ser calculada duplicando el producto de la frecuencia de los alelos en cuestión ($2pq$). ^(12, 15, 59 y 67)

c) Poder de discriminación (PD)

El poder de discriminación se caracteriza por la probabilidad de que dos individuos no relacionados y seleccionados al azar, puedan diferenciarse genéticamente

mediante el análisis de un marcador o un conjunto de marcadores. El PD depende del número de locus analizados y del polimorfismo de cada uno.

$$PD = 1 - P_I$$

$$PD = 1 - \sum_{i=1}^n p^2$$

Donde (P_I) es la probabilidad de identidad.

d) Probabilidad de identidad (P_I)

Capacidad de determinar que dos individuos dentro de una población compartan el mismo genotipo en un locus específico. La P_I para un solo locus puede ser calculado como la sumatoria de las frecuencias genotípicas al cuadrado. ^(8, 12 y 68)

$$P_I = \sum_{i=1}^n p^2$$

e) Poder de exclusión (PE)

La probabilidad de que un sistema genético específico proporcione evidencias que conduzcan a la exclusión de un individuo como probable responsable o bien como un probable padre. ^(8, 59 y 69)

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

Donde H es la heterocigosidad.

f) Heterocigosidad

Es la proporción de individuos heterocigotos en la población. La heterocigosidad se obtiene dividiendo el número de muestras que contienen alelos heterocigotos entre el total de las muestras. ⁽⁸⁾

g) Índice de paternidad

Refleja cuántas veces es más probable que un individuo analizado sea el padre biológico con respecto a que lo sea cualquier otro individuo de la población tomado al azar. ⁽⁸⁾

$$\mathbf{IP = 1 / 2h}$$

Donde h es la homocigosidad.

h) Índice de contenido polimórfico (PIC)

El PIC indica lo informativo que es un marcador midiendo su capacidad para identificar que alelo se ha heredado de cada uno de los parentales.

$$\mathbf{PIC = 1 - \sum_1^n P_i^2 - \left(\sum_1^n P_i^2 \right)^2 + \sum_1^n P_i^4}$$

i) Número efectivo de alelos (NEA)

El número efectivo de alelos, se define como la probabilidad de que dos alelos de un locus elegido al azar en la población sean idénticos por descendencia y esta probabilidad, es igual al inverso de la frecuencia de individuos homocigotos en la población.

$$\mathbf{NEA = 1 / h}$$

Donde h es la homocigosidad.

2.10 Loci estudiados

Los marcadores genéticos incluidos dentro del sistema **PowerPlex® ESX 17** son: D3S1358, TH01, D21S11, D18S51, D10S1248, D1S1656, D2S1338, D16S539, D22S1045, vWA, D8S1179, FGA, D2S441, D12S391, D19S433, SE33 y amelogenina (Fig. 8).⁽³⁹⁾

Hace un poco más de 15 años se propuso un conjunto de loci para todas las bases de datos de ADN de Europa con el único fin de poder intercambiar información entre los distintos países que conforman la Unión Europea, dicha propuesta fue realizada por el **European Network of Forensic Science Institutes (ENFSI)**. A dicho conjunto de loci se le nombró **European Standard Set (ESS)**, dentro del cual se incluían inicialmente los marcadores D3S1358, D8S1179, D18S51, D21S11, FGA, TH01 y vWA.⁽⁷⁰⁾

Debido al incremento de perfiles genéticos generados en Europa, en el año 2006 el **ENFSI** y el **European DNA Profiling (EDNAP)**, recomendaron que se incrementara el número de loci que conformaban el **ESS**. Los cinco marcadores recomendados por el **ENFSI** y **EDNAP** son: **D10S1248**, **D22S1045** y **D2S441** y son conocidos como mini STRs debido a que tienen un tamaño menor a los 125 pb; mientras que **D12S391** y **D1S1656** son conocidos como midi STRs y tienen un tamaño que va de los 125 a 185 pb. El propósito de la inclusión de estos cinco marcadores genéticos fue con varios propósitos: evitar falsas coincidencias en el caso de investigaciones criminales; minimizar al máximo las asociaciones fortuitas de parentesco en el caso de identificación de cadáveres desconocidos; y por último el obtener mejores resultados de muestras degradadas analizando marcadores del tipo mini y midi STRs.^(70, 71, 72, 73, 74 y 75)

Además también se incluye el locus **SE33**, el cual es un marcador tetranucleótido con un alto poder de discriminación y el cual es utilizado dentro de la base de datos nacional de Alemania.^(76, 77, 78 y 79)



Figura 8. Configuración del sistema PowerPlex® ESX 17. Tomado de Sprecher CJ. ⁽⁷⁰⁾

3. Planteamiento del problema

Un requisito fundamental para utilizar nuevos marcadores genéticos del tipo STRs en una población y con ello resolver asuntos en materia de criminalística, pruebas de paternidad, identificación de personas desaparecidas e identificación de restos humanos, es que se conozca la distribución de las frecuencias alélicas de dicho marcador y con ello estar en posibilidad de calcular diversos parámetros estadísticos de interés forense, los cuales permitirán determinar la frecuencia de un perfil genético en la población de estudio.

Por lo que en este estudio el planteamiento central es si las frecuencias alélicas de los marcadores incluidos en el sistema **PowerPlex® ESX 17**, están en equilibrio de Hardy-Weinberg y pueden ser utilizadas para determinar los distintos parámetros estadísticos de interés forense como: heterocigosidad, probabilidad de coincidencia, poder de discriminación, poder de exclusión, índice de paternidad. En particular las frecuencias del marcador SE33, el cual es uno de los marcadores genéticos más polimórficos e informativos utilizado en el área forense, el cual se sabe cuenta con un alto poder de discriminación en otras poblaciones estudiadas, además es el locus STRs más polimórfico estudiado a la fecha, dado que se han encontrado 178 alelos que van de 3 a 49 repeticiones, lo que hace que cada vez más laboratorios en Europa lo empiecen a adoptar. Siendo importante mencionar que hasta el momento no se encuentran reportadas las frecuencias alélicas del marcador SE33, para la población del Valle de México.

¿Los 16 loci de tipo STR incluidos en el sistema **PowerPlex® ESX 17** podrán ser utilizados para estudios de parentesco e identificación humana dentro de la población del Valle de México?

4. Objetivos

- ❖ Obtener las frecuencias alélicas de los 16 *loci* polimórficos del sistema **PowerPlex® ESX 17** en la población del Valle de México.
- ❖ Determinar si las frecuencias alélicas y genotípicas se encuentran dentro del equilibrio de Hardy-Weinberg.
- ❖ Determinar los parámetros estadísticos de interés forense: heterocigosidad observada (HO), heterocigosidad esperada (HE), homocigosidad (Hom), probabilidad de coincidencia (PM), poder de discriminación (PD), poder de exclusión (PE), número efectivo de alelos (NEA), índice de paternidad (IP) y contenido de información polimórfica (PIC).
- ❖ Determinar la heterocigosidad de cada uno de los *loci* que componen el sistema **PowerPlex® ESX 17**.
- ❖ Determinar si las frecuencias alélicas y genotípicas se encuentran en desequilibrio de ligamiento.

5. Hipótesis

- ❖ Se espera no haya diferencia significativa entre los genotipos observados y los esperados en cada uno de los *loci*, por lo que dicha población se encontrará en equilibrio de Hardy-Weinberg.
- ❖ La heterocigosidad observada en la muestra poblacional del Valle de México tomada para este estudio, será superior al 60%.
- ❖ Se espera no haya diferencia significativa entre los genotipos observados y los esperados en cada uno de los *loci*, por lo que dicha población no se encontrará en desequilibrio de ligamiento.

6. Metodología

6.1 Tipo de estudio

- ❖ Observacional
- ❖ Prospectivo
- ❖ Transversal
- ❖ Descriptivo

6.2 Población

En el presente estudio se analizaron 162 muestras de sangre fijada en papel FTA, provenientes de hombres y mujeres del Valle de México, los cuales fueron tomados al azar y no se encontraban relacionados biológicamente entre ellos. Dichas muestras fueron tomadas del banco de muestras del **Centro de Estudios e Investigaciones ANIGEN**, el cual se encuentra ubicado en Municipio Libre No.366 Col. Sta. Cruz Atoyac C.P. 03310 México, D.F. La autorización de toma de muestras fue hecha bajo consentimiento informado donde se hacía saber a los donadores que sus muestras serían utilizadas para realizar un estudio genético poblacional. (**Ver Anexo**)

6.3 Criterios de inclusión, exclusión y eliminación

- ❖ **Inclusión:** Muestras de sangre fijadas en papel FTA provenientes de individuos no emparentados entre si y que residan en el Valle de México.
- ❖ **Exclusión:** Individuos relacionados biológicamente, así como a individuos que no residan en el Valle de México.
- ❖ **Eliminación:** Todas aquellas muestras contaminadas o cuyo ADN se encuentre degradado.

6.4 Variables

Independiente: El genotipo de cada individuo de la muestra poblacional.

Dependiente:

- ❖ Frecuencias alélicas y genotípicas
- ❖ Heterocigosidad
- ❖ Equilibrio de enlace
- ❖ Parámetros estadísticos de interés forense

6.5 Material y método

6.5.1 Material y Equipos

- ❖ Analizador automático de biología molecular 3100-Avant de Applied Biosystem.
- ❖ Termociclador GeneAmp® PCR System 9700
- ❖ Gabinetes de Bioseguridad Air Clean 600 PCR Workstation
- ❖ Micropipeta Eppendorf de 100-1000 μ L
- ❖ Micropipeta Eppendorf de 20-200 μ L
- ❖ Micropipeta Eppendorf de 20 μ L
- ❖ Micropipeta Eppendorf de 10 μ L
- ❖ Puntas para pipeta de 10 μ L
- ❖ Puntas para pipeta de 100 μ L
- ❖ Puntas para pipeta de 1000 μ L
- ❖ Vortex THERMO MIXER
- ❖ Tubos Eppendorf de 1.5mL
- ❖ Harris Micro-Punch para papel FTA
- ❖ Placas de 96 pozos
- ❖ Septas para placa de 96 pozos
- ❖ Tapas para charola de amplificación
- ❖ Baño seco BOCKEL

- ❖ Papel FTA
- ❖ Encendedor
- ❖ Lancetas
- ❖ Guantes de nitrilo
- ❖ Torundas
- ❖ Alcohol etílico al 70%
- ❖ Cubrebocas
- ❖ Hi-Di™ formamide
- ❖ Pop-4
- ❖ PowerPlex® ESX 17 System
- ❖ Congelador Revco
- ❖ Refrigerador Revco
- ❖ Papel absorbente
- ❖ Palillos
- ❖ Tubos de 0.2mL para PCR

6.5.2 Método

Obtención de la muestra

Se solicitó a cada donante que rotulara una tarjeta FTA, posteriormente se limpió el área de punción para la obtención de sangre con una torunda con alcohol al 70%. Se realizó la punción con una lanceta y se depositaron aproximadamente 5 gotas de sangre sobre la tarjeta FTA, posteriormente dejar que la muestra de sangre seque para que pueda ser guardada para su posterior uso.

Lavado de los fragmentos de tarjeta FTA

Lo primero que se hizo fue irradiar los gabinetes de bioseguridad con luz UV durante 15 minutos. En el área de fraccionamiento de indicios se etiquetó los tubos Eppendorf de 1.5mL con la clave correspondiente a la muestra a trabajar, posteriormente con un Harris Micro-Punch se cortaron 10 discos de papel FTA

impregnado con muestra y se colocaron dentro de los tubos etiquetados. En el área de extracción a cada muestra se le adicionó 400µL del reactivo de FTA, dejando incubar a temperatura ambiente y agitación (1000 rpm) durante 5 minutos en el *thermomixer*. Una pasados los 5 minutos, se decantó el reactivo para luego colocar 400µL de agua inyectable libre de nucleasas a cada muestra incubando a temperatura ambiente y agitación (1000 rpm) durante 5 minutos en le *thermomixer* para después decantar con sumo cuidado el agua, procurando limpiar el remanente del reactivo en la boca del tubo con un papel absorbente limpio (repitiendo este último paso 2 veces más). Una vez terminados los lavados correspondientes se retiró toda el agua que se pudiera con ayuda de una micropipeta y punta de 1000µL para posteriormente colocar los tubos con la muestra a 60°C durante 30 minutos, para eliminar restos de agua. Finalmente la muestra está lista para amplificarse con el protocolo recomendado por el fabricante

Amplificación del ADN

El sistema utilizado fue el **PowerPlex® ESX 17**, el kit se mantuvo en congelación hasta el momento de ser usado.

Lo primero que se hizo fue descongelar el **PowerPlex ESX 5X Master Mix**, **PowerPlex ESX 10X Primer Pair Mix**, y el agua grado amplificación, completamente. Posteriormente se determinó el número de reacciones incluyendo el control positivo y negativo, debiendo considerar 1 o 2 reacciones más para la compensación en el pipeteo, se preparó la mezcla de reacción en un tubo Eppendorf estéril color ámbar de 1.5mL (agitando con vortex por 15 segundos a cada uno de los tubos de reactivo).

COMPONENTES DE LA MEZCLA DE AMPLIFICACIÓN	VOLUMEN POR REACCIÓN
Agua, Grado Amplificación	17.5 µL
PowerPlex ESX 5X Mater Mix	5 µL
PowerPlex ESX 17 10X Primer Pair Mix	2.5 µL
Volumen total de la reacción	25 µL

Cuadro 2. Mezcla de amplificación del sistema PowerPlex® ESX 17

La mezcla de reacción se adicionó a cada uno de los tubos de 0.2mL, de acuerdo al número de muestras a amplificar, posteriormente con la ayuda de un palillo irradiado con luz UV se agregó un disco de papel FTA con muestra. La amplificación se llevó a cabo en un termociclador GeneAmp® PCR System 9700, siguiendo el protocolo recomendado por el fabricante, el cual se muestra a continuación:

Protocolo de amplificación

96°C por 2 minutos

94°C por 30 segundos

59°C por dos minutos

72°C por 90 segundos

Por 24 ciclos, después:

60°C por 45 minutos

4°C (se conserva a 4°C hasta el momento de su análisis)

Detección de los fragmentos amplificados utilizando el *Analizador 3100-Avant* de *Applied Biosystem*.

El kit se mantuvo en congelación hasta el momento de ser usado. Se preparó una mezcla de **CC5 Internal Lane Standard 500** y **formamida Hi-Di** en un tubo Eppendorf, las cantidades de cada reactivo fueron de acuerdo al número de muestras, la proporción de cada reactivo fue calculada de la siguiente manera:

$$[(2.0\mu\text{L CC5 ILS 500}) \times (\#\text{muestras})] + [(20.0\mu\text{L Hi-Di}^{\text{TM}} \text{ formamide}) \times (\#\text{muestras})]$$

La mezcla se agitó durante 15 segundos y enseguida se fraccionaron 18 μ L en una placa de 96 pozos de acuerdo con el número de muestras a correr. Posteriormente se colocaron 2 μ L de cada producto de PCR en el pozo correspondiente y 0.5 μ L de **PowerPlex ESX 17 Allelic Ladder**, en cada uno de los pozos destinados para éste. La placa de 96 pozos fue sellada con una septa y se procedió a realizar la hoja de corrida en el **Data Collection**, para posteriormente comenzar con el corrimiento electroforético.

Los resultados obtenidos fueron analizados con ayuda del software **GeneMapper ID v.3.2**.

Análisis estadístico

En el presente estudio, además de calcular las frecuencias alélicas, también fueron calculados la heterocigosidad observada (HO), heterocigosidad esperada (HE), homocigosidad (Hom), probabilidad de coincidencia (PM), poder de discriminación (PD), poder de exclusión (PE), número efectivo de alelos (NEA), índice de paternidad (IP), contenido de información polimórfica (PIC), el equilibrio de Hardy-Weinberg y el equilibrio de ligamiento (LE), para todos los marcadores incluidos en el sistema **PowerPlex® ESX 17**.

7. Resultados

En el cuadro 3 se muestran las frecuencias obtenidas de los 16 marcadores STRs contenidos en el sistema **PowerPlex® ESX 17**. Dichas frecuencias son utilizadas para calcular diversos parámetros estadísticos de interés para la genética forense.

En el cuadro 4 podemos observar que 14 de los 16 marcadores estudiados se encuentran en equilibrio de Hardy-Weinberg, lo que nos indica que las frecuencias obtenidas en estos marcadores pueden ser utilizadas para calcular diversos parámetros estadísticos de interés forense. Los marcadores D18S51 y D2S1338 estuvieron en desequilibrio.

En el cuadro 5 se muestran los valores obtenidos de los diversos parámetros estadísticos de los 16 STRs analizados.

Analizando los resultados expuestos en la gráfica 1, se puede observar que el marcador que presenta el índice más alto de heterocigosidad observada es el D19S433, seguido del FGA y del SE33, mientras que el marcador con el índice de heterocigosidad observada más bajo es el D10S1248.

Lo que respecta a los resultados expuestos en la gráfica 2, se puede observar que el marcador que presenta el índice más alto de heterocigosidad esperada es el SE33, seguido del D1S1656 y del FGA, mientras que el marcador con el índice de heterocigosidad esperada más bajo es el D22S1045.

En cuanto a lo mostrado en la gráfica 3, se puede observar que el marcador con un valor más bajo de poder de coincidencia al azar es el marcador SE33, seguido del FGA y el D1S1656, mientras que el marcador D22S1045 tiene el índice de probabilidad de coincidencia al azar más alto. Lo que indica que la probabilidad de encontrar dos individuos con el mismo genotipo en el marcador SE33 es menor con respecto a la de los demás marcadores, también es importante mencionar que un valor bajo de probabilidad de coincidencia al azar reduce la frecuencia de asociaciones fortuitas de parentesco o *adventitious matches*.^(8, 50, 51, 79 y 80)

Con lo que respecta a la gráfica 4, se puede observar que todos los marcadores del sistema **PowerPlex® ESX 17** cuentan con un alto poder de discriminación. El marcador con el índice de poder de discriminación más alto es el SE33 seguido del

D1S1656 y el FGA, mientras que el marcador con el menor valor de poder de discriminación es el D22S1045.

En la gráfica 5, se presentan los valores del poder de exclusión para cada uno de los 16 STRs analizados donde se observa que el marcador con el índice más alto es el marcador D19S433, seguido del FGA y el SE33.

En la gráfica 6, se observan los valores del número efectivo de alelos para cada uno de los marcadores analizados, este valor nos habla del número de alelos con capacidad de pasar a la siguiente generación observando que el índice más alto lo tiene el marcador D19S433, seguido del FGA y el SE33.

En la gráfica 7, se muestran los valores del índice de paternidad siendo el marcador D19S433 quien tiene el valor más alto, seguido del FGA y el SE33.

En cuanto a la gráfica 8, se muestran los valores del índice de contenido polimórfico donde se observa que el marcador D19S433 tiene el valor más alto, seguido del FGA y el SE33.

Por último, en el cuadro 6 se muestran a detalle los valores p de desequilibrio de ligamiento entre pares de *loci* en la población del Valle de México. En la mayoría de parejas de *loci* analizados se obtienen valores de p superiores a 0.05. Los valores significativos para una $p < 0.05$ se muestran en color rojo; los datos subrayados indican los valores significativos tras la corrección de Bonferroni.

Cuadro 3. Frecuencias alélicas obtenidas de los 16 loci STRs contenidos en el sistema PowerPlex® ESX 17

Alelo	D3S1358	TH01	D21S11	D18S51	D10S1248	D1S1656	D2S1338	D16S539	D22S1045	vWA	D8S1179	FGA	D2S441	D12S391	D19S433	SE33
6		0.3364														
7		0.3240														
8		0.0648						0.0061			0.0092					
8.3		0.0030														
9		0.0802						0.1018			0.0030					
9.3		0.1759														
10		0.0154			0.0030	0.0030		0.1944			0.1141		0.4907			
11				0.0030	0.0030	0.0370		0.2530	0.0493		0.0586		0.2716		0.0030	
11.2															0.0030	
11.3													0.0308			
12				0.0987	0.0308	0.0555		0.3086	0.0030		0.0987		0.0308		0.0432	
12.2															0.0154	
12.3													0.0030			
13	0.0123			0.1111	0.2314	0.0833		0.1234	0.0061		0.3179		0.0123		0.2067	0.0092
13.2															0.1296	
14	0.0555			0.1697	0.4320	0.1358		0.0092	0.0370	0.0462	0.2345		0.1358		0.2654	0.0061
14.2															0.0401	
14.3						0.0030										
15	0.4660			0.1759	0.2129	0.1481		0.0030	0.4567	0.0493	0.1327		0.0216	0.0092	0.1358	0.0648
15.2	0.0030														0.0864	0.0030
15.3						0.0123									0.0030	0.0030
16	0.2808			0.1172	0.0740	0.1759	0.0277		0.3919	0.3703	0.0246		0.0030	0.0216	0.0462	0.0432
16.2															0.0216	
16.3						0.0401										
17	0.1296			0.1975	0.0123	0.0679	0.1419		0.0493	0.2962	0.0061			0.0462	0.0030	0.1049
17.3						0.1790								0.0030		0.0061
18	0.0462			0.0555		0.0092	0.0586		0.0030	0.1728		0.0030		0.1388		0.0740
18.3						0.0462								0.0030		
19	0.0061			0.0308			0.2345		0.0030	0.0586		0.0586		0.3549		0.0895
19.3						0.0030								0.0030		
20				0.0216			0.1234			0.0061		0.0833		0.2191		0.0679
20.2																0.0092
21				0.0061			0.0432					0.0617		0.0679		0.0308
21.2																0.0061
22												0.1327		0.0586		0.0061
22.2												0.0061				0.0030
23							0.2006					0.1327		0.0493		0.0030
24							0.0462					0.1759		0.0123		
24.2																0.0123
25							0.0308					0.1975		0.0061		
25.2																0.0586
26								0.0061				0.1018		0.0061		
26.2				0.0030												0.0956
27				0.0123								0.0401				
27.2																0.1080
28				0.0864								0.0030				
28.2																0.0586
29				0.1666												
29.2				0.0030												0.0586
30				0.3364								0.0030				
30.2				0.0216												0.0524
31				0.0709												
31.2				0.1080												0.0154
32				0.0061												
32.2				0.1327												0.0061
33.2				0.0493												0.0030
34.2				0.0030												

Cuadro 4. Resultados de la prueba de X^2 para el sistema **PowerPlex® ESX 17**

Locus	X^2 calculada	gl	X^2 Tablas ($\alpha=0.05$)	Resultado
D3S1358	4,2446	3	7.8147	Equilibrio
TH01	4,4587	3	7.8147	Equilibrio
D21S11	4.0433	6	12.5916	Equilibrio
D18S51	19.2587	10	18.3070	Desequilibrio
D10S1248	7,1332	3	7.8147	Equilibrio
D1S1656	9.2996	6	12.5916	Equilibrio
D2S1338	14.5186	6	12.5916	Desequilibrio
D16S539	0,4818	3	7.8147	Equilibrio
D22S1045	0.0224	1	3.8415	Equilibrio
vWA	7.6312	3	7.8147	Equilibrio
D8S1179	3.0845	3	7.8147	Equilibrio
FGA	6.4943	3	7.8147	Equilibrio
D2S441	1.8110	3	7.8147	Equilibrio
DS12S391	1.1866	6	12.5916	Equilibrio
D19S433	2.7399	3	7.8147	Equilibrio
SE33	10.5470	10	18.3070	Equilibrio

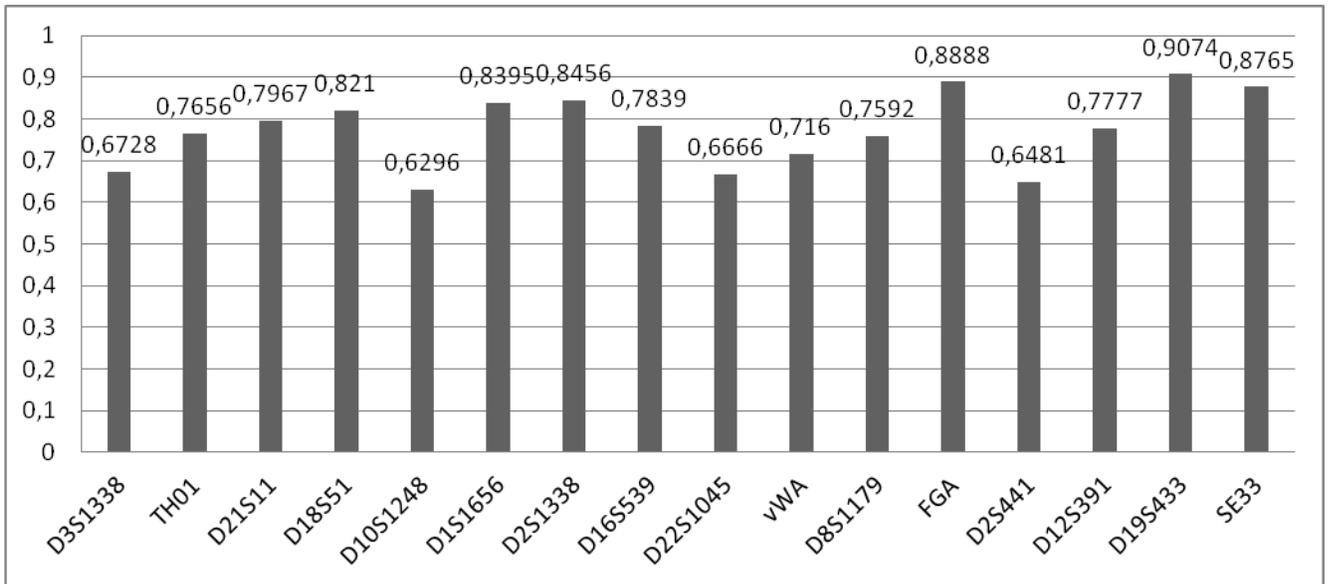
Cuadro 5. Parámetros estadísticos calculados.

LOCUS	D3S1358	TH01	D21S11	D18S51	D10S1248	D1S1656	D2S1338	D16S539	D22S1045	vWA	D8S1179	FGA	D2S441	D12S391	D19S433	SE33
HO	0,6728	0,7656	0,7967	0,821	0,6296	0,8395	0,8456	0,7839	0,6666	0,7160	0,7592	0,8888	0,6481	0,7777	0,9074	0,8765
HE	0,6816	0,7399	0,8142	0,8608	0,7077	0,8765	0,8531	0,7771	0,6313	0,7371	0,7993	0,8685	0,6644	0,7932	0,8377	0,9291
HOM	0,3271	0,2345	0,2037	0,1790	0,3703	0,1604	0,1543	0,2160	0,3333	0,2839	0,2407	0,1111	0,3518	0,2222	0,0925	0,1234
PM	0,1485	0,1104	0,0548	0,0346	0,1309	0,0271	0,0372	0,0843	0,2045	0,1107	0,0664	0,0308	0,1612	0,0667	0,0443	0,0086
PD	0,8514	0,8895	0,9452	0,9653	0,8690	0,9728	0,9627	0,9156	0,7954	0,8892	0,9335	0,9691	0,8387	0,9332	0,9556	0,9913
PE	0,3856	0,5056	0,5528	0,5927	0,3381	0,6249	0,6362	0,5329	0,3785	0,4380	0,4964	0,7205	0,3579	0,5235	0,7605	0,6954
NEA	3,0571	4,2643	4,9091	0,5865	2,7	6,2344	6,48	4,6287	3	3,5217	4,1545	9	2,842	4,5	10,8108	8,1
IP	1,5285	2,1321	2,4545	2,7932	1,35	3,1172	3,24	2,3143	1,5	1,7608	2,077	4,5	1,4212	2,25	5,4054	4,05
PIC	0,5773	0,7135	0,7565	0,7879	0,5112	0,8145	0,8224	0,7394	0,5679	0,6419	0,7047	0,8766	0,5396	0,7308	0,8990	0,8615

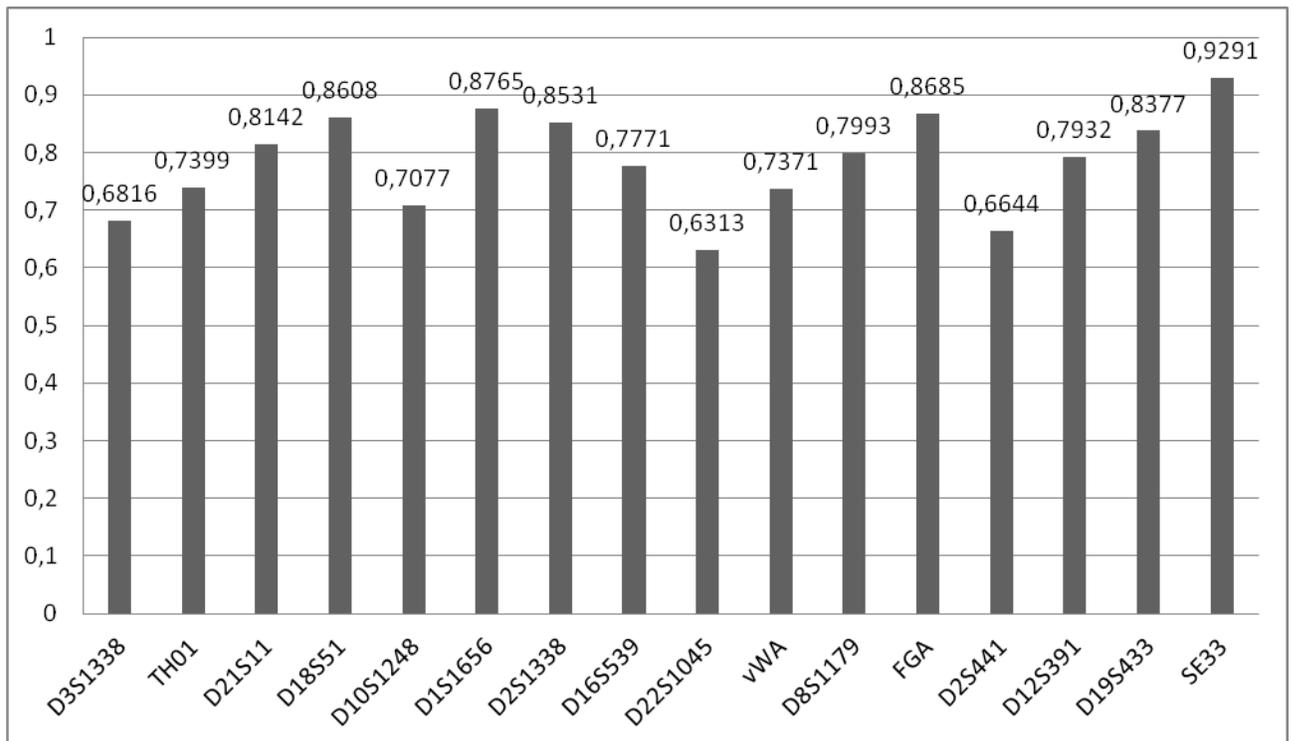
HO: Heterocigosidad observada, HE: Heterocigosidad esperada, HOM: Homocigosidad, PM: Probabilidad de coincidencia, PD: Poder de discriminación, PE: Poder de exclusión, NEA: Número efectivo de alelos, IP: Índice de paternidad, PIC: Índice de contenido polimórfico.

Cuadro 6. Valores p de desequilibrio de ligamiento entre pares de loci en la población del Valle de México.

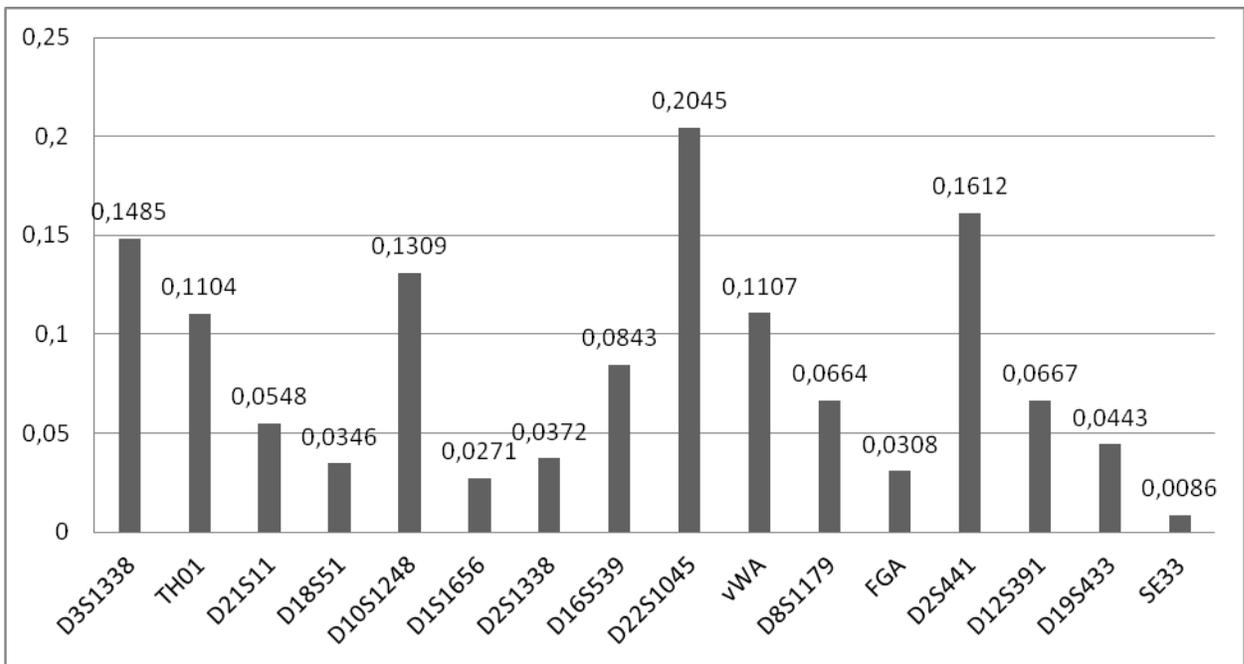
	D3S1358	TH01	D21S11	D18S51	D10S1248	D1S1656	D2S1338	D16S539	D22S1045	vWA	D8S1179	FGA	D2S441	D12S391	D19S433
D3S1358															
TH01	0.28723														
D21S11	0.00584	0.10020													
D18S51	0.6480	0.24772	0.02119												
D10S1248	0.04624	0.30475	0.16079	0.00277											
D1S1656	0.21347	0.60050	0.32980	0.56970	0.47376										
D2S1338	0.56554	0.48644	0.43604	0.37376	0.52416	0.28297									
D16S539	0.43931	0.12436	0.15386	0.10366	0.78178	0.22020	0.72703								
D22S1045	0.78723	0.60743	0.78178	0.08881	0.22832	0.28752	0.04663	0.38010							
vWA	0.03911	0.11990	0.20129	0.49980	0.49069	0.98881	0.49178	0.54822	0.42614						
D8S1179	0.27307	0.62218	0.14357	0.16901	0.04228	0.23396	0.59832	0.84990	0.41525	0.49752					
FGA	0.89218	0.74040	0.52624	0.72574	0.84465	0.08257	0.68277	0.90990	0.63297	0.57861	0.35782				
D2S441	0.64416	0.42218	0.07782	0.80574	0.23624	0.31802	0.12455	0.34485	0.85495	0.42059	0.25941	0.11149			
D12S391	0.96822	0.42158	0.81554	0.61010	0.81634	0.13218	0.41248	0.21634	0.01673	0.27673	0.27990	0.19881	0.39307		
D19S433	0.07149	0.20772	0.19505	0.08198	0.62178	0.94257	0.43188	0.10158	0.58059	0.02931	0.71010	0.95564	0.88594	0.48188	
SE33	0.99198	0.08436	0.98990	0.18218	0.87248	0.91950	0.21792	0.95089	0.30832	0.55871	0.80564	0.77010	0.49950	0.05713	0.67238



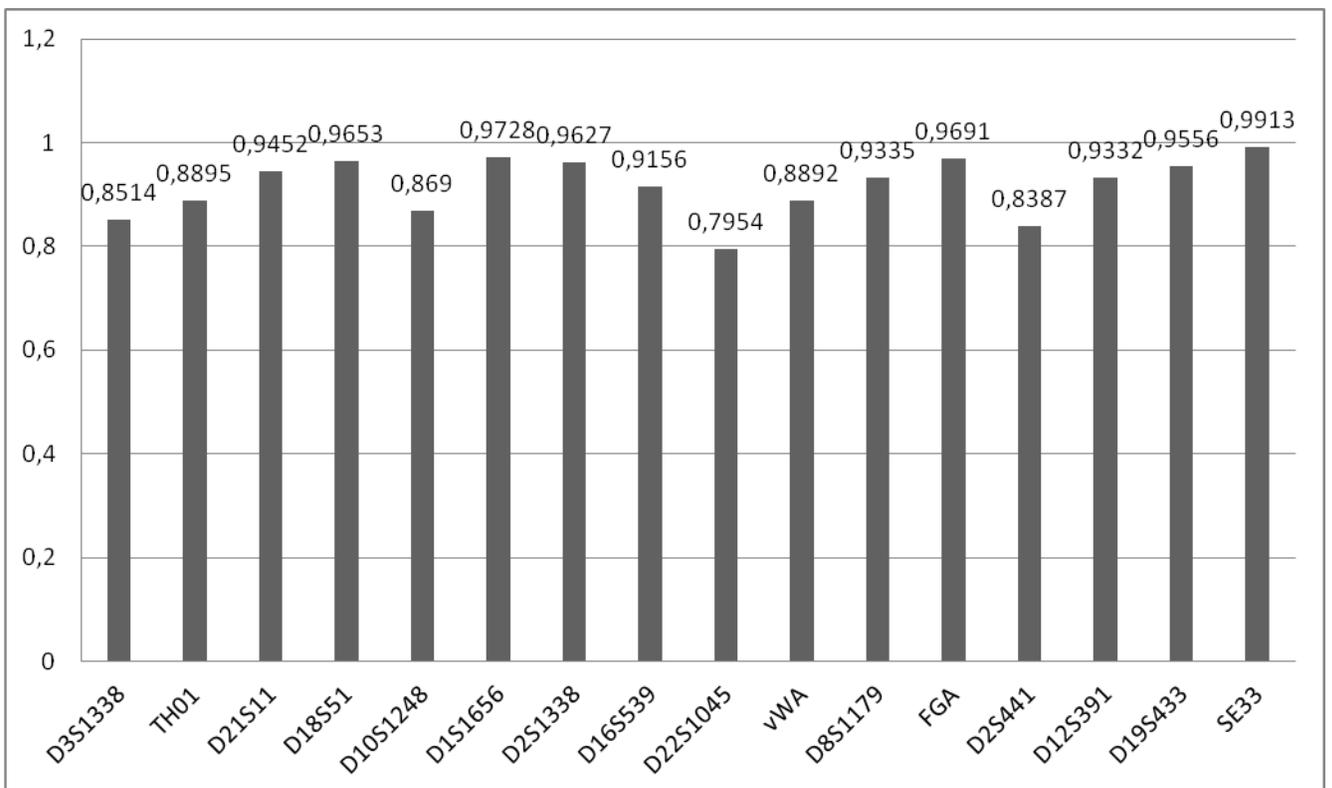
Gráfica 1. Heterocigosidad Observada



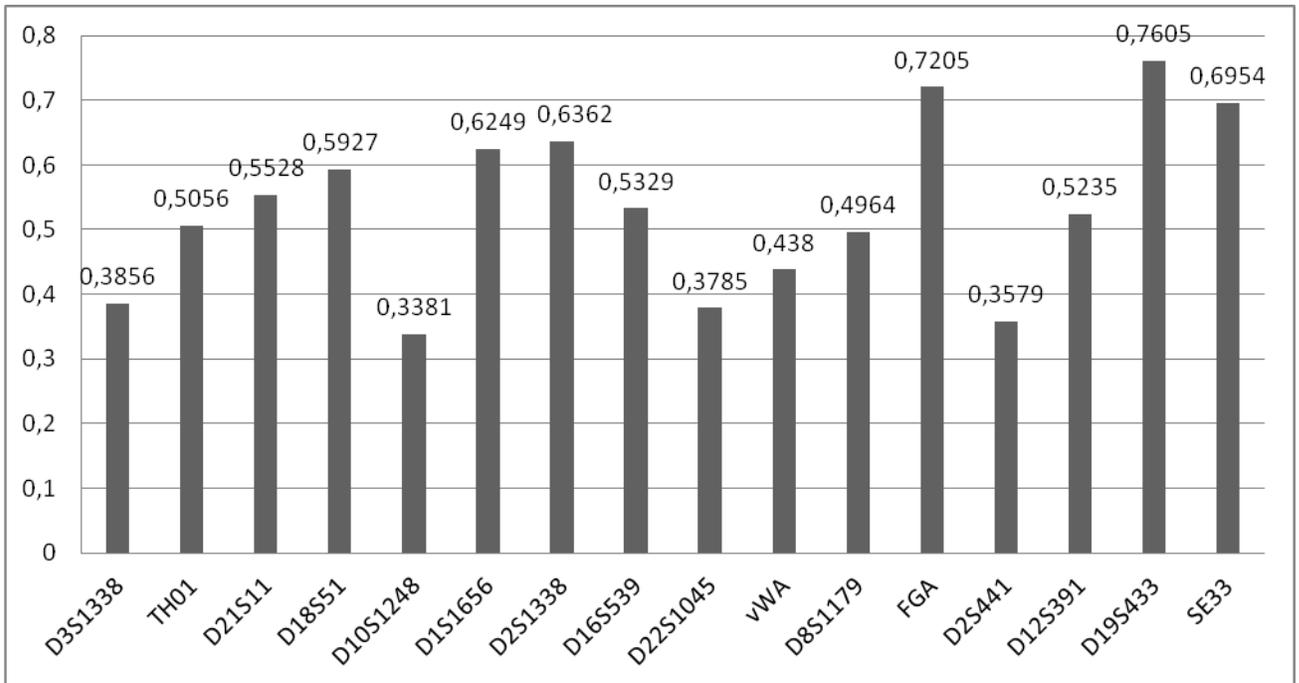
Gráfica 2. Heterocigosidad Esperada



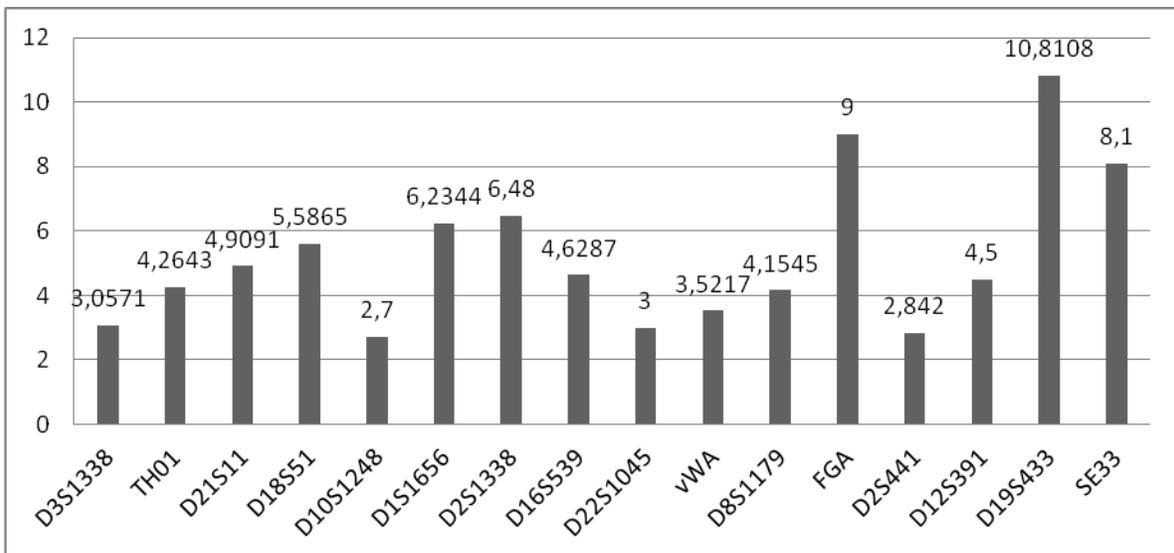
Gráfica 3. Probabilidad de coincidencia al azar



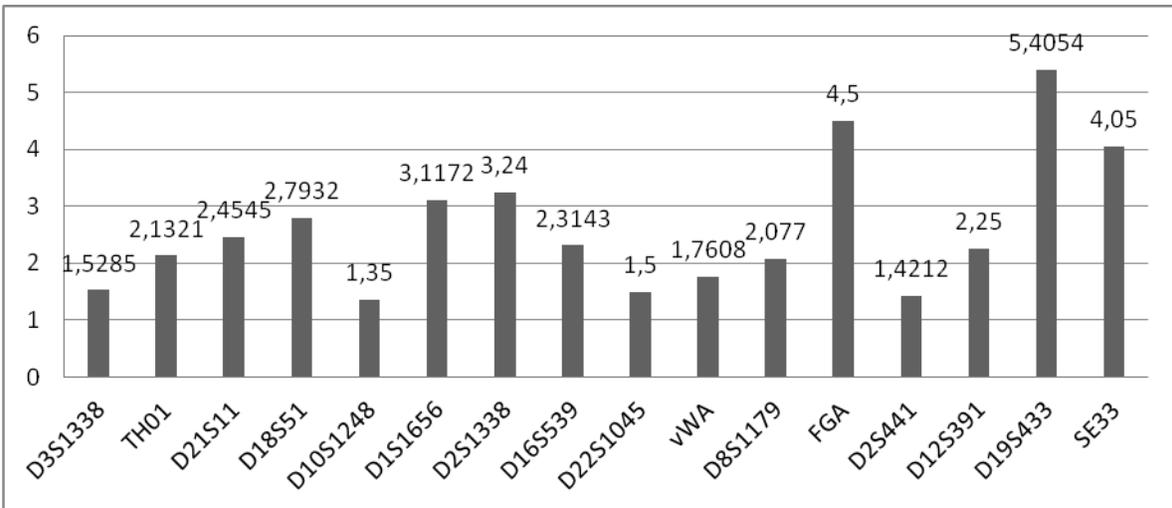
Gráfica 4. Poder de discriminación



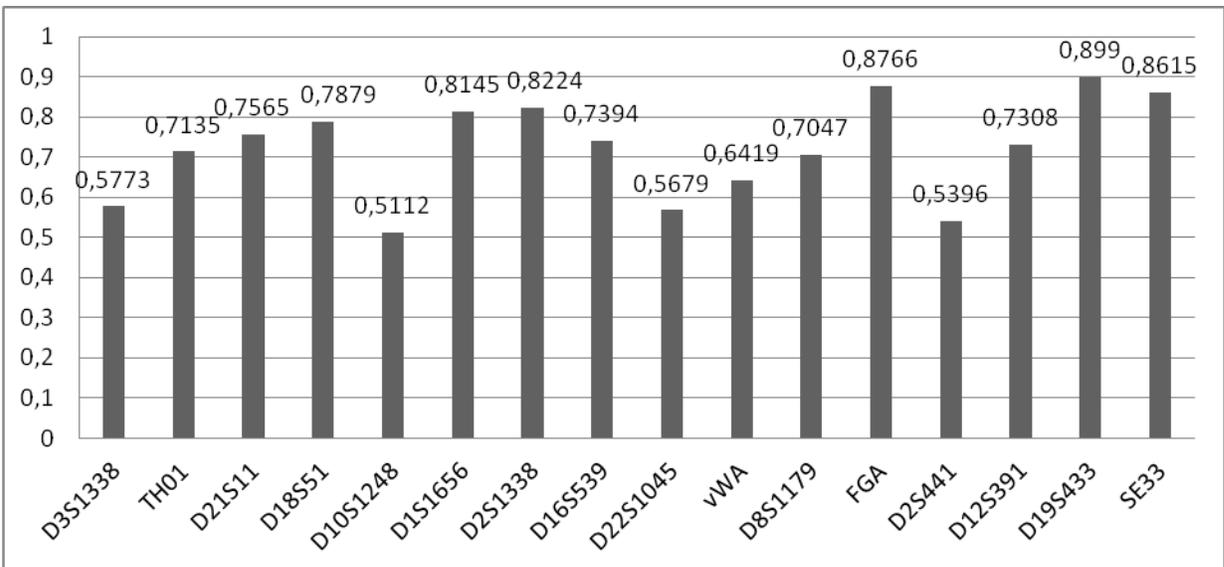
Gráfica 5. Poder de exclusión



Gráfica 6. Número efectivo de alelos



Gráfica 7. Índice de paternidad



Gráfica 8. Índice de contenido polimórfico

8. Discusión de resultados

Debido a que en los últimos años se ha incrementado la aplicación de la genética forense para realizar análisis relacionados a la resolución de problemas judiciales así como de procesos civiles (demandas de paternidad), ha surgido la necesidad de crear bases de datos poblacionales para estimar la rareza de un perfil genético en una población con el fin de dar una indicación a un tribunal de la solidez de las pruebas de ADN.⁽¹⁾

Para que las frecuencias alélicas obtenidas de la población de estudio puedan ser utilizadas en el análisis genético, es necesario que estén en equilibrio de Hardy-Weinberg. La Ley de Hardy-Weinberg dice que la frecuencia de los alelos en una población permanecerían iguales (en equilibrio) a través de las generaciones, siempre y cuando se cumpla con las siguientes condiciones: que la población tenga un tamaño bastante grande (población infinita), que el apareamiento ocurra al azar, así como la ausencia de fuerzas evolutivas (mutación, deriva genética y selección).^(8 y 66)

Para determinar si los marcadores genéticos incluidos dentro del sistema **PowerPlex® ESX 17**, se encontraban dentro del equilibrio de Hardy-Weinberg se procedió a realizar una prueba de chi-cuadrada con la cual se determinó si existía o no una diferencia estadísticamente significativa entre las frecuencias alélicas observadas y las frecuencias alélicas esperadas.^(8 y 81)

De los 16 *loci* analizados, 14 se encuentran dentro del equilibrio de Hardy-Weinberg, los *loci* D18S51 y D2S1338 son los que se encuentran en desequilibrio.

La heterocigosidad observada (HO) refleja la frecuencia de individuos heterocigotos encontrados en la población de estudio, mientras que la heterocigosidad esperada (HE), es una estimación más precisa de la

heterocigosidad en la población. La heterocigosidad es la probabilidad de que dos alelos del mismo *locus* tomados al azar de la población sean diferentes. Siendo los marcadores D19S433, FGA y SE33 los que tienen un índice más alto de heterocigosidad, mientras que los marcadores con el índice de heterocigosidad más bajo son el D10S1248, D2S441 y D22S1045.

El poder de discriminación se caracteriza por la probabilidad de que dos individuos no relacionados y seleccionados al azar, puedan diferenciarse genéticamente mediante el análisis de un marcador o un conjunto de marcadores. Siendo los marcadores SE33, D1S1656 y FGA quienes cuentan con los índices más altos de poder de discriminación, mientras que los marcadores con el índice de poder de discriminación más bajo son el D22S1045, D2S441 y D2S1338. ^(8 y 12)

En cuanto al poder de exclusión podemos decir que se define como la probabilidad de excluir a un individuo que esté involucrado en algún caso forense. En casos de paternidad, por ejemplo, sería la probabilidad de excluir un “falso padre”. Este parámetro nos permite establecer la proporción de individuos falsamente implicados en un peritaje. Siendo los marcadores D19S433, FGA y SE33 los que presentan el índice de poder de exclusión más alto, mientras que los marcadores con el índice de poder de exclusión más bajo son el D10S1248, D2S441 y el D22S1045. ⁽⁸⁾

La probabilidad de coincidencia al azar también es conocida como *match probability*, se define como la probabilidad de que dos individuos elegidos al azar tengan idénticos genotipos en uno o más sistemas genéticos, es decir, es la probabilidad de que dos individuos no relacionados compartan un mismo perfil de ADN. Esta probabilidad puede aumentar cuando tenemos un perfil de ADN parcial por degradación o cuando tenemos la reducción del número de *locus* informativos. Además, también puede aumentar si el probable responsable y el culpable real tienen parentesco o si se originan de la misma subpoblación. El único caso

excluíble a estos valores es el de los gemelos monocigóticos, ya que comparten el mismo material genético. ⁽¹²⁾

Los marcadores con el índice más bajo de probabilidad de coincidencia son el SE33, D1S1656 y FGA, mientras que los marcadores con el índice de probabilidad de coincidencia más alto son el D22S1045, D2S441 y D13S1338.

El PIC es un índice que evalúa la informatividad de un marcador en la población de acuerdo con las frecuencias de los alelos e indica la capacidad de un marcador genético para identificar que alelo se ha heredado de cada uno de los padres. Según la clasificación de Botstein, marcadores con valores de PIC superiores a 0,5 son considerados muy informativos, con valores entre 0,25 y 0,50 medianamente informativos y con valores inferiores a 0,25, poco informativos. Los *loci* con el PIC más alto son el D19S433, FGA y el SE33. ^(82 y 83)

El índice de paternidad (IP) es el valor que expresa cuántas veces es más probable que el probable padre sea el padre biológico a que lo sea cualquier otro sujeto en la población de estudio. El IP se modificará principalmente en función de la frecuencia con que se observen los alelos heredados del padre en la población general; a menor frecuencia alélica, mayor el IP y mayor la probabilidad de que sea el padre biológico. En este parámetro los marcadores con el índice de paternidad más alto son el D19S433, FGA y SE33, mientras que los marcadores con el índice de paternidad más bajo son el D10S1248, D2S441 y el D22S1048. ⁽⁸⁾ En cuanto al número efectivo de alelos se define como la probabilidad de que los alelos de un locus elegido al azar en la población sean idénticos por descendencia. Para este índice los marcadores con los valores más altos son el D19S433, FGA y el SE33. ⁽⁸⁴⁾

El desequilibrio de ligamiento es un término utilizado para describir el grado de asociación no aleatoria de alelos específicos en diferentes *loci* dentro de un mismo cromosoma. Dicho de otra manera se dice que existe desequilibrio de ligamiento

cuando dos *loci* que se encuentran físicamente cerca uno del otro en un mismo cromosoma y no se transmiten de manera independiente a la siguiente generación.

De acuerdo con los resultados de este estudio, los siguientes marcadores están ligados entre sí:

D3S1358 con D21S11, D10S1248 y vWA

D21S11 con D18S51

D18S51 con D10S1248

D10S1248 con D8S1179

D2S1338 con D22S1045

D22S1045 con D12S391

vWA con D19S433

Después de aplicar la corrección de Bonferroni ($0.05/16 = 0.003125$), todos los resultados dejan de ser significativos, a excepción de la pareja **D18S51-D10S1248**, lo que indica que solo este par de marcadores se encuentra en desequilibrio de ligamiento.

El desequilibrio de ligamiento puede ser debido a efectos como: subestructura poblacional, selección y la deriva genética. ⁽⁸⁵⁻⁸⁶⁾

Es de gran importancia mencionar que el incremento del uso de la genética forense con fines de identificación ha hecho que se genere una gran cantidad de información que actualmente es administrada en las denominadas *bases de datos* las cuales son capaces de realizar comparaciones en cuestión de segundos entre los distintos perfiles genéticos almacenados en ellas. De dichos comparativos se ha observado que se pueden encontrar falsas coincidencias o asociaciones fortuitas de parentesco sobre todo entre perfiles parciales o perfiles obtenidos con un bajo número de *loci*. Los eventos antes mencionados podrían disminuirse al máximo usando el mayor número de marcadores genéticos y más aún cuando

estos tienen un alto poder de discriminación, un alto poder de exclusión y una baja probabilidad de coincidencia al azar tal y como los tiene el marcador SE33.

9. Conclusiones

Las frecuencias alélicas de los marcadores D3S1358, TH01, D21S11, D10S1248, D1S1656, D16S539, D22S1045, vWA, D8S1179, FGA, D2S441, D12S391, D19S433 y SE33, se encuentran dentro del equilibrio de Hardy-Weinberg.

Se observó que la heterocigosidad de los 16 marcadores genéticos incluidos en el sistema PowerPlex® ESX 17 es mayor al 60%, lo que indica que en la población estudiada existe una gran variabilidad genética.

El marcador SE33 del cual no se tenía registro alguno hasta el momento de sus frecuencias alélicas para población mexicana, mostró ser un marcador genético con un alto poder de discriminación, alto poder de exclusión y una baja probabilidad de coincidencia al azar que lo hace ser un candidato para ser utilizado en seguimiento de trasplantes, pruebas de paternidad y en casos forenses.

Los marcadores **D18S51- D10S1248**, son el único par de marcadores de los 16 *loci* incluidos en el sistema PowerPlex® ESX 17, que se encuentran en desequilibrio de ligamiento.

10. Recomendaciones

El único marcador genético del cual no habían sido reportadas sus frecuencias para la población del Valle de México, es el SE33 que muestra ser un locus con una heterocigosidad observada de 0.8765, un alto poder de discriminación (0.9913), un alto poder de exclusión (0.6954) y una baja probabilidad de coincidencia (0.0086). Por lo anterior, se recomienda el uso del marcador SE33 ya que sería de gran utilidad en pruebas de parentesco así como en la resolución de problemas judiciales en los que se requiera un análisis de ADN.

11. Referencias Bibliográficas

1. Lorente JA, Vega ML, Rosas GO. Genética Forense: La Ciencia al Servicio de la Justicia. Mens Bioquím. 2007; 31.
2. Cal ML. Análisis de polimorfismos de ADN microsatélite de cromosoma Y. Estudio de la población de Galicia y aplicaciones forenses [Tesis de Doctorado]. [Santiago de Compostela (España)]: Universidad de Santiago de Compostela; 2001. 155 p.
3. Krause M, Heide KG, Krawczak M. SE33 allele and genotype frequencies in the population of Schleswig-Holstein (Northern Germany). International Congress Series. 1288; 2006: 400-402.
4. Watson JD, Crick FHC. Molecular Structure of Deoxyribose Nucleic Acids. Nature. 1953 April 25; 171: 738-740.
5. Alberts B. Molecular Biology of the cell. 5a ed. New York: Garland Science, 2007.
6. Nelson DL, Cox MM. Principios de Bioquímica. 5a ed. Barcelona: Ed Omega; 2007.
7. Brown TA. Genomes. 2ª ed. Oxford: Wiley-Liss; 2002.
8. Butler JM. Forensic DNA typing: biology, technology and genetics of STR markers. 2a ed. London: Elsevier; 2005.
9. Lorente JA. Un detective llamado ADN. Madrid: Ed Temas de Hoy; 2004.
10. Kobilinsky L. Forensic DNA Analysis. New York: Chelsea House; 2007.
11. Butler JM, Advanced Topics in Forensic DNA Typing: Interpretation. San Diego CA: Elsevier; 2015.
12. Kobilinsky L. DNA: Forensic and Legal Applications. New Jersey: Wiley-Interscience; 2005.
13. Sadler TW. Embriología médica con orientación clínica. 9a ed. Buenos Aires: Ed. Médica Panamericana; 2004.
14. Strachan T. Human Molecular Genetics. 4a ed. New York: Garland-Science; 2011.

15. Goodwin L, Linacre A, Hadi S. An introduction to forensic genetics. 2a ed. England: Wiley; 2011.
16. Fernández F. “Análisis de 15 loci tipo short tandem repeats (STR) en la población de Paraguay y su uso en la identificación forense”. [Tesis de Doctorado]. [Granada (España)]: Universidad de Granada; 2008. 246 p.
17. Nakamura Y, Leppert M, O’Connell P, Wolff R, Holm T. Variable number of tandem repeat (VNTR) markers for gene mapping. *Science*. 1987; 235: 1616-1622.
18. Gill P, Jeffreys AJ, Werrett. Forensic application of DNA “fingerprints”. *Nature*. 1985 Dec 12; 318 (12): 577-579.
19. Jeffreys AJ, Wilson V, Neumann R, Keyte J, Applications of human minisatellites by the polymerise chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Res*. 1988 Nov 4; 16 (23): 10953-10971.
20. Butler JM. *Advanced Topics in Forensic DNA Typing: Methodology*. San Diego: Elsevier; 2011.
21. Coble MD, Butler JM. Characterization of new miniSTR loci to aid analysis of degraded DNA. *J. Forensic Sci*. 2005; 50 (1): 43-53.
22. Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J Forensic Sci*. 2003; 48 (5): 1054-1064 p.
23. Edwards AI, Civitello A, Hammond HA, Caskey T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet*. 1991 June 10; 49: 746-756.
24. Kupferschmid TD, Calicchio T, Budowle B. Maine Caucasian population DNA database using twelve short tandem repeat loci. *J Forensic Sci*. 1999; 44(2):392–395.
25. Butler JM. Short tandem repeats typing technologies used in human identity testing. *BioTechniques*. 2007; 43 (4).

26. Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in Short Tandem Repeats sequences a survey of twelve microsatellites loci for use as forensic identification markers. *Int Leg Med.* 1994; 107: 13-20.
27. Nastainczyk M, Schulz S, Kleiber M, Immel UD. STR analysis of degraded DNA using a miniplex. *Forensic Sci Int Genet.* 2009 August 22: 53-54.
28. Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B. Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J Forensic Sci* 2001; 46 (3): 647–660.
29. Alford RL, Hammond HA, Coto I, Caskey CT. Rapid and efficient resolution of parentage by amplification of short tandem repeats. *Am J Hum Genet.* 1994 March 11; 55: 190-195.
30. Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W. A new “miniSTR-multiplex” displaying reduced amplicon lengths for the analysis of degraded DNA. *Int J Leg Med.* 2005 June 14; 120: 115-120.
31. Chung DT, Drábek J, Opel KL, Butler JM, McCord. A Study on the effects of degradation and template concentration on the amplification efficiency of the STR miniplex primer sets. *J Forensic Sci.* 2004 July; 49 (4): 733-740.
32. Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM, et al. Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise. *Forensic Sci Int.* 2006; 164: 33-44.
33. Herzog U, Augustin C, Püschel K. Allele frequencies of six miniSTR markers in a population sample from Northern German and its application on forensic stain analysis. *Forensic Sci Int.* 2008: 331-333.
34. Rodríguez IP, Barrera HA. La reacción en cadena de la polimerasa a dos décadas de su invención. *Ciencia UANL.* 2004; 7(3): 323-335.
35. Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH. Multiplex PCR: Critical Parameters and Step-by-Step Protocol. *BioTechniques.* 1997 Sept; 23 (3): 504-511.

36. Pinilla G, Cubillos K, Rodríguez M. Bodas de plata de la reacción en cadena de la polimerasa (PCR). *Ciencias Biomedicas*. 2008; 6 (9): 65-75.
37. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science*. 1988 Jan 29; 239 (4839): 487-491.
38. Promega. Technical Manual PowerPlex® Fusion System. [En línea]. Disponible en: <https://worldwide.promega.com/~media/files/resources/protocols/technical%20manuals/101/powerplex%20fusion%20system%20protocol.pdf> [Consulta: 15-04-2015]
39. Promega. Technical Manual PowerPlex® ESX 17 System. [En línea]. Disponible en: <https://worldwide.promega.com/~media/files/resources/protocols/technical%20manuals/101/powerplex%20esx%2017%20fast%20system%20protocol.pdf> [Consulta: 15-04-2015]
40. Eckert KA, Kunkel TA. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res*; 18 (13): 3739-3744.
41. Innis MA, Myambo KB, Gelfand DH, Brow MAD. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci*. 1988 Dec; 85: 9436-9440.
42. Markoulatos P, Siafakas N, Moncany M. Multiplex Polymerase Chain Reaction: A Practical Approach. *Journal of Clinical Laboratory Analysis*. 2002; 16: 47-51.
43. Schoske R, Vallone PM, Ruitberg CM, Butler JM. Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci. *Anal Bioanal Chem*. 2003; 375: 333-343.

44. Butler JM, Devaney JM, Marino MA, Vallone PM. Quality control of PCR primers used in multiplex STR amplification reactions. *Forensic Sci Int.* 2001; 119: 87-96.
45. Barbaro A. Progress in Forensic Genetics: New Markers Validation Studies and Population Data. [PhD Thesis]. [Santiago de Compostela (España)]: Universidad de Santiago de Caompostela; 2012. 177 p.
46. Gangs J, Urbas G. DNA identification in the Criminal Justice System. *Australian Institute of Criminology.* 2002; 226: 1-6.
47. Lee JW, Lee HS, Park M, Hwang JJ. Evaluation of DNA match probability in criminal case. *Forensic Sci Int.* 2001; 116: 139-148.
48. Alonso A, Martín P, Albarrán C, García P, Fernández de Simón L, et al. Challenges of DNA Profiling in Mass Disaster Investigations. *Croat Med J.* 2005; 46 (4): 540-548.
49. Budowle B, Bieber FR, Eisenberg AJ. Forensic aspects of mass disaster: Strategic considerations for DNA-based human identification. *Legal Med.* 2005; 7: 230-243.
50. Amor AM, Martínez G. Importancia del estudio del quimerismo en el trasplante alogénico de la médula ósea. *Rev Cubana Hematol Inmunol Hemoter.* 2003; 19: 2-3.
51. Urbano A, Fernández F, Briones J, Aymerich M. Estudio del quimerismo hematopoyético en el trasplante alogénico de médula ósea. *Hematologica.* 2004; 89: 65-71.
52. Scharf SJ, Smith AG, Hansen JA, McFarland C, Erlich HA. Quantitative determination of bone marrow transplant engraftment using fluorescent polymerase chain reaction primer for human identity markers. *Blood.* 1995 April 1; 85 (7): 1954-1963.
53. Thiede C, Florek M, Bornhäuser M, Ritter M, Mohr B, et al. Rapid quantification of mixed chimerism using multiplex amplification of short tandem repeat markers and fluorescence detection. *Bone Marrow Trasplantation.* 1999; 23: 1055-1060.

54. Álvarez MJ, Martínez LJ, Saiz M, Álvarez JC, Lorente JA. Nuevas aplicaciones en identificación genética. *Caud Med Forense*. 2010; 16 (1-2): 5-18.
55. Ge J, Eisenberg A, Budowle B. Developing criteria and data to determine best options for expanding the core CODIS loci. *Investigative Genetics*. 2012; 3 (1): 2-14.
56. Lorente JA, Entrala C, Alvarez JC, Arce B, Heinrich B, et al. Identification of Missing Persons: The Spanish "Phoenix" Program. *Croat Med J*. 2001; 42: 267-270.
57. INTERPOL (La función del AND en la resolución de asuntos penales). [En línea]. Disponible en: <http://www.interpol.int/es/Especialidades/Polic%C3%ADa-cient%C3%ADfica/ADN> [Consulta: 1-05-2015]
58. Carey L, Mitnik L. Trends in DNA forensic analysis. *Electrophor*. 2002; 23: 1386-1397.
59. Buckleton J, Triggs CM, Walsh SJ. *Forensic DNA Evidence Interpretation*. New Zeland: CRC Press; 2005.
60. Wallance H. The UK National DNA Database. *Science & Society*. 2006; 7: 26-30.
61. National DNA Database Strategy Board Annual Report 2012-13. [En línea]. Disponible en: [https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/252885/NDNAD Annual Report 2012-13.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/252885/NDNAD_Annual_Report_2012-13.pdf) [Consulta 1-05-2015]
62. National DNA Data Bank. [En línea]. Disponible en: <http://www.publicsafety.gc.ca/cnt/cntrng-crm/crrctns/protctn-gnst-hgh-rsk-ffndrs/htnl-dna-bnk-eng.aspx> [Consulta 1-05-2015]
63. Ge J, Sun H, Li H, Liu C, Yan J, Budowle B. Future directions of forensic DNA databases. *Croat Med J*. 2014; 55: 163-166.
64. Miller K, Brown BL, Budowle B. The Combined DNA Index System. *International Congress Series*. 2003; 1239: 617-620.

65. Song YS, Patil A, Murphy JD, Slatkin M. Average Probability that a “Cold Hit” in a DNA Database Search Results in an Erroneous Attribution. *J Forensic Sci.* 2009; 54 (1).
66. Hamilton MB. *Population Genetics*. UK: Wiley-Blackwell; 2009.
67. Guo SW, Thompson EA. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. 1992 June; 48: 361-372.
68. Huston KA. Statistical Analysis of STR data. *Gene Print* 1998; 1: 14-15.
69. Quintero A, Padilla JR, Hernández G, Valle Y, Valdez LL, et al. Datos poblacionales de cinco STRs de la serie INTERPOL en una población mestiza del Occidente de México. *Revista de Inves Clin.* 2009; 61(2): 104-109.
70. Sprecher CJ, McLaren RS, Rabbach DR, Ensenberger MG, Fulmer PM, et al. The PowerPlex® ESX and ESI Systems: Meeting the New European Standard. [En línea]. Disponible en: <http://worldwide.promega.com/resources/profiles-in-dna/2009/the-powerplex-esx-and-esi-systems-meeting-the-new-european-standard/> [Consulta 25-04-2015]
71. Sprecher CJ, McLaren RS, Rabbach D, Krenke B, Ensenberger MG, et al. PowerPlex1 ESX and ESI Systems: A suite of new STR systems designed to meet the changing needs of the DNA-typing community. *Forensic Sci Int.* 2009: 2-4.
72. Hill CR, Diewer DL, Kline MC, Sprecher CJ, McLaren RS, et al. Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex1 ESX 17 and ESI 17 Systems. *Forensic Sci Int.* 2010.
73. Poetsch M, Bayer K, Ergin Z, Milbrath M, Schwark T, et al. First experiences using the new Powerplex® ESX17 and ESI17 kits in casework analysis and allele frequencies for two different regions in Germany. *Int J Legal Med.* 2010 June 22.

74. Gill P, Fereday L, Morling N, Schneider PM. The evolution of DNA databases—Recommendations for new European STR loci. *Forensic Sci Int.* 2006; 156: 242-244.
75. Harder M, Rennenberg R, Meyer P, Kyora BK, von Wurmb N. STR-typing of ancient skeletal remains: which multiplex-PCR kits is the best?. *Croat Med J.* 2012; 53:416-422.
76. Lászik A, Sótonyi P, Rand S, Hohoff C. Frequency data for the STR locus ACTBP2 (SE33) in eight populations. *Int J Legal Med.* 2001; 115: 94-96.
77. Butler JM, Hill CR, Kline MC, Dueder DL, Sprecher CJ, et al. The single most polymorphic STR Locus: SE33 performance in U.S. populations. *Forensic Sci Int.* 2009: 23-24.
78. Heinrich M, Müller M, Rand S, Brinkmann B, Hohoff C. Allelic drop-out in the STR system ACTBP2 (SE33) as a result of mutations in the primer binding region. *Int J Legal Med.* 2004; 118: 361-363.
79. Davis C, Ge J, King J, Malik N, Weirich V, Eisenberg AJ, et al. Variants observed for STR locus SE33: A concordance study. *Forensic Sci Int.* 2012; 6: 494-497.
80. Leclair B, Fregeau CJ, Bowen KL, Fournay RM. Enhanced Kinship Analysis and STR-based DNA Typing for Human Identification in Mass Fatality Incidents: The Swissair Flight 111 Disaster. *J Forensic Sci.* 2004 September; 49(5):1-15.
81. Fung WK. *Statistical DNA Forensics: Theory, Methods and Computation.* England: Wiley; 2008.
82. Botstein D, White RL, Skolnick M, Davis RW. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am J Hum Genet.* 1980; 32: 314-315.
83. Guo X, Elston RC. Linkage Information Content of Polymorphic Genetic Markers. *Hum Hered.* 1999; 49: 112-118.
84. Kimura M, Crown JF. The Number of Alleles that can be Maintained in a Finite Population. *Genetics.* 1964; 49: 725-738.

85. Fujii K, Iwashima Y, Kitayama T, Nakahara H, Mizuno N, Sekiguchi K. Allele frequencies for 22 autosomal short tandem repeat loci obtained by PowerPlex Fusion in a sample of 1501 individuals from Japanese population. *Legal Med.* 2014; 16: 234-237.
86. Bright JA, et al. A guide to Forensic DNA Interpretation and Linkage. [En línea]. Disponible en : <http://www.promega.com/resources/profiles-in-dna/2014/a-guide-to-forensic-dna-interpretation-and-linkage/>

12. Anexo

Declaración de consentimiento informado

C. _____ de _____ años de edad y que es habitante de la Ciudad de México; manifiesta que ha sido informado/a sobre los beneficios que podría suponer la extracción de sangre en papel FTA para cubrir los objetivos del Proyecto de Investigación titulado “Análisis genético de 22 Loci STR en la Población de la Ciudad de México” con el fin de mejorar el análisis forense en el sistema de justicia mexicano.

He sido también informado/a de que mis datos personales serán protegidos y no se utilizarán con ningún otro fin que no sea el de la investigación, antes mencionada.

Tomando ello en consideración, OTORGO mi CONSENTIMIENTO a que esta extracción tenga lugar y sea utilizada para cubrir los objetivos específicos en el proyecto.

México DF, 2013

Donante