T. Prov. 23/128

T
16
56

UNIVERSIDAD DE GRANADA
Facultad de Ciencias
Fecha 11/04/00
ENTRADA NUM. 1526



UNIVERSIDAD DE GRANADA

DEPARTAMENTO DE MEDICINA LEGAL

"ANÁLISIS DEL POLIMORFISMO DE DIEZ LOCI TIPO MICROSATÉLITE EN LA POBLACIÓN ANDALUZA PARA SU APLICACIÓN EN IDENTIFICACIÓN GENÉTICA HUMANA"

BIBLIOTECA UNIVERSITARIA
GRANADA
Nº Documento D19682657
Nº Copia 121232556

Carmen Entrala Bernal

Granada, 2000

UNIVERSIDAD DE GRANADA
31 MAR. 2000
COMISION DE DOCTORADO

ANEXO IX

UNIVERSIDAD DE GRANADA

ESTUDIOS DE DOCTORADO

CONFORMIDAD DEL DEPARTAMENTO

DEPARTAMENTO: MEDICINA LEGAL

PROGRAMA DE DOCTORADO:

D/Dª ANTONIO F. HERNANDEZ JEREZ ..... **Secretario/a de este Departamento,**

CERTIFICO: Que este Departamento, en su reunión del día 30-marzo-2000 ....., **ha dado su conformidad para la presentación de la Tesis Doctoral de D/Dª** ..CARMEN ENTRALA BERNAL.....

Título de la Tesis ANALISIS DEL POLIMORFISMO DE DIEZ LOCI MICROSATELITE EN LA POBLACION ANDALUZA PARA SU APLICACION EN IDENTIFICACION GENETICA HUMANA

Director/es JOSE ANTONIO LORENTE ACOSTA, MIGUEL LORENTE ACOSTA

Tutor JOSE ANTONIO LORENTE ACOSTA .....

Y para que conste a los efectos pertinentes en la Comisión de Doctorado, expido la presente con el Vº. Bº. del Director de este Departamento en Granada a treinta..... **de** marzo..... **de mil** ~~novecientos noventa~~ **del año** 200

Vº. Bº.

EL DIRECTOR,

Fdo: ENRIQUE VILLANUEVA CAÑADAS

EL SECRETARIO,

Fdo: ANTONIO F. HERNANDEZ JEREZ

EXCMO. SR. PRESIDENTE DE LA COMISION DE DOCTORADO DE ESTA UNIVERSIDAD

UNIVERSIDAD DE GRANADA

ESTUDIOS DE DOCTORADO

AUTORIZACION PARA PRESENTACION DE TESIS

D. . .JOSE ANTONIO LORENTE ACOSTA ..... Director de la Tesis ANALISIS DEL POLIMORFISMO DE DIEZ LOCI MICROSATELITE EN LA POBLACION ANDALUZA PARA SU APLICACION EN IDENTIFICACION GENETICA HUMANA

de la que es autor D/Dª CARMEN ENTRALA BERNAL

Programa de Doctorado GENETICA

AUTORIZA la presentación de la referida Tesis para su defensa y mantenimiento de acuerdo con lo previsto en el Real Decreto 185/1985, de 23 enero, emitiendo el siguiente informe:

Y para que conste y surta sus efectos en el expediente correspondiente, expido la presente en Granada a 30 de marzo de mil novecientos noventa 2000

UNIVERSIDAD DE GRANADA ☆ FACULTAD DE MEDICINA ☆ Departamento de Medicina Legal

Fdo: JAL

RATIFICACION DEL TUTOR (cuando el Director no sea Profesor del Departamento)

D. ..... Tutor del Doctorando autor de la presente Tesis, ratifica la autorización del Director de la misma para su presentación por las siguientes razones:

Granada, ... de ..... de 19....

UNIVERSIDAD DE GRANADA

ESTUDIOS DE DOCTORADO

AUTORIZACION PARA PRESENTACION DE TESIS

D. MIGUEL LORENTE ACOSTA ..... Director de la Tesis ANALISIS DEL POLIMORFISMO DE DIEZ LOCI MICROSATELITE EN LA POBLACION ANDALUZA PARA SU APLICACION EN IDENTIFICACION GENETICA HUMANA .....

de la que es autor D/Dª CARMEN ENTRALA BERNAL .....

Programa de Doctorado .....GENETICA.....

.....

AUTORIZA la presentación de la referida Tesis para su defensa y mantenimiento de acuerdo con lo previsto en el Real Decreto 185/1985, de 23 enero, emitiendo el siguiente informe:

Y para que conste y surta sus efectos en el expediente correspondiente, expido la presente en Granada a ..30..... de marzo ..... de mil novecientos noventa ..2000.....

Fdo: Miguel Lorente.....

RATIFICACION DEL TUTOR (cuando el Director no sea Profesor del Departamento)

D. ..... Tutor del Doctorando autor de la presente Tesis, ratifica la autorización del Director de la misma para su presentación por las siguientes razones:

Granada, ... de ..... de 19....

INDICE

1. OBJETIVOS ----------1

1.1. Justificación del tema y objetivos ----------2

2. INTRODUCCION ----------5

2.1. Identificación genética en Medicina Legal ----------6

2.2. Estudio del ADN ----------8

2.2.1. ADN no codificante 10

2.2.1.1. Secuencias repetidas en tándem 13

2.2.1.2. Secuencias repetidas intercaladas 17

2.2.1.3. Significado biológico del ADN repetitivo 19

2.2.2. Polimorfismos del ADN.....20

2.2.2.1. ADN minisatélite.....21

2.2.2.2. ADN microsatélite 23

2.2.2.3. Origen genético de los polimorfismos en ADN repetitivo.....27

2.3. Técnicas de estudio de los polimorfismos del ADN en Genética Forense ----------30

2.3.1. Breve historia 30

2.3.2. Análisis por la reacción en cadena de la polimerasa 35

2.3.2.1. Componentes de la PCR.....38

2.3.2.2. Ventajas e inconvenientes del uso de la PCR en Genética Forense 42

2.3.3. Análisis de la variación interna de los loci minisatélite: MVR-PCR...44

2.3.4. Sistemas de análisis de los polimorfismos STR 48

2.3.4.1. Tipificación manual.....49

2.3.4.2. Marcaje fluorescente y detección semiautomática 50

2.3.4.3. Futuras tecnologías: Biochips.....55

2.4. Genética de Poblaciones ---- 62
2.4.1. Ley de Hardy-Weinberg ---- 63
2.4.2. Fuerzas evolutivas ---- 64
2.5. Problemas ético-legales ---- 67
2.4.1. Negativa del consentimiento ---- 69
2.4.2. Bancos de datos genéticos ---- 71

3. MATERIAL Y METODOS ---- 78
3.1. Muestras ---- 79
3.1.1. Muestras utilizadas para el estudio poblacional ---- 79
3.1.2. Muestras utilizadas para el estudio de validación ---- 79
3.2. Extracción de ADN ---- 84
3.2.1. Extracción orgánica mediante fenol/cloroformo-isoamilalcohol ---- 84
3.2.1.1. Extracción a partir de sangre total y manchas de sangre ---- 84
3.2.1.2. Extracción a partir de manchas de semen
e hisopos vaginales ---- 86
3.2.1.3. Extracción a partir de saliva ---- 88
3.2.1.4. Extracción a partir de hueso ---- 89
3.2.2. Extracción con Chelex ---- 90
3.3. Cuantificación del ADN ---- 92
3.4. Amplificación ---- 94
3.4.1. Kit SilverSTR™III ---- 95
3.4.2. Kit AmpFISTR Profiler Plus ---- 96
3.4.3. Comprobación de la amplificación ---- 98
3.5. Electroforesis ---- 99
3.5.1. Sistema de electroforesis en acrilamida
y detección con nitrato de plata ---- 99
3.5.2. Electroforesis capilar y detección por fluorescencia ---- 103

3.6. Análisis estadístico----------111

3.6.1. Frecuencias alélicas y genotípicas.....111

3.6.2. Tests para demostrar el equilibrio de Hardy-Weinberg 112

3.6.2.1. Test de bondad de ajuste 112

3.6.2.2. Test Exacto de Guo-Thompson.....115

3.6.3. Tests para demostrar la independencia de los alelos 116

3.6.4. Comparaciones poblacionales 118

3.6.5. Parámetros estadísticos de interés forense:
probabilidad de exclusión a priori, eficacia bioestadística,
probabilidad a posteriori, probabilidad de discriminación,
probabilidad de coincidencia,
heterocigosidad y diversidad genética 119

4. RESULTADOS ----------125

4.1. Resultados del estudio poblacional ----------126

4.1.1. Visualización de los resultados obtenidos
con el triplex SilverIII STR™ 126

4.1.2. Visualización de los resultados obtenidos
con el multiplex AmpFISTR Profiler Plus 128

4.1.3. Resultados obtenidos con el triplex SilverIII STR™.....131

4.1.4. Resultados obtenidos con el kit AmpFISTR Profiler Plus 140

4.1.4.1. Comparación poblacional 168

4.2. Resultados del estudio de validación----------173

4.2.2. Estudio de sensibilidad 173

4.2.3. Estudio de mezclas.....174

4.2.4. Estudio en indicios biológicos de interés forense.....179

5. DISCUSION ---------- 183

5.1. Estudio genético-poblacional ---------- 185

5.1.1. Triplex SilverIII STR™ 185

5.1.1.1. Discusión de los resultados analíticos 185

5.1.1.2. Discusión de los resultados estadísticos 189

5.1.2. Multiplex AmpFISTR Profiler Plus 194

5.1.2.1. Discusión de los resultados analíticos 194

5.1.2.2. Discusión de los resultados estadísticos 206

5.2. Estudio de validación ---------- 208

5.2.1. Estudio de sensibilidad 211

5.2.2. Estudio de mezclas 215

5.2.3. Estudio de muestras forenses 227

5.3. Discusión global ---------- 229

6. CONCLUSIONES ---------- 231

7. BIBLIOGRAFÍA ---------- 234

8. ANEXO ---------- 255

ABREVIATURAS

ADN	Acido desoxirribonucléico
AmpFLP	Amplificable Fragment Length Polymorphism
BSA	Bovine Serum Albumin (Albúmina Sérica Bovina)
CE	Capilar Electrophoresis (electroforesis capilar)
dNTPs	desoxiribonucleósido 5'-trifosfato
DTT	Ditiotreitol
EDTA	Acido EtilenDiaminoTetracético
esp	Esperados
frec	Frecuencia
G^2	Estadístico del test de razón de verosimilitud
gl	Grados de libertad
Het	Heterocigotos
Hom	Homocigotos
HLA	Human Leucocite Antigen (Antígeno Leucocitario Humano)
HWE	Hardy Weinberg Equilibrium (Equilibrio Hardy-Weinberg)
ISFH	International Society for Forensic Haemogenetics
K562	Linea celular humana, leucemia mieloide
Kb	Kilobase
LR	Likelihood ratio (razón de verosimilitud)
M	Molar
mA	Miliamperios
mg	Miligramos
min	Minutos
ml	Mililitros
MLP	Multilocus Probe (sonda multilocus)
mM	Milimolar

MVR	Minisatellite Variant Repeat
ng	Nanogramos
nm	Nanómetros
Obs	Observados
PAGE	Electroforesis en gel de poliacrilamida
pb	pares de bases
PCR	Reacción en cadena de la polimerasa
PD	Poder de discriminación
PE	Poder de exclusión
PDA	Piperazín diacrilamida
PIC	Contenido en información polimórfica
p.m.	Peso molecular
r.p.m.	Revoluciones por minuto
RFLP	Restriction Fragment Length Polymorphism (polimorfismos de longitud para fragmentos de restricción)
sd	Desviación estándar
SDS	Dodecil Sulfato Sódico
se	Error estándar
SLP	Single Locus Probes (sondas unilocus)
STR	Short Tandem Repeat (repeticiones cortas en tándem)
TBE	Tris-Bórico-EDTA
TEMED	N,N,N',N'-tetra methyl ethylene diamine
TWGDAM	Technical Work Group on DNA Methods
U.V.	Ultra Violeta
V	Voltio
VNTR	Variable Number of Tandem Repeats (repeticiones en tándem en número variable)

w	Vatios
w/v	Peso/Volúmen
χ^2	Estadístico del test de bondad de ajuste a distribuciones
µg	Microgramo
µl	Microlitro
µM	Micromolar

1. OBJETIVOS

1. JUSTIFICACIÓN DEL TEMA Y OBJETIVOS

La aplicación de las técnicas de análisis de ADN al campo de la Biología Forense se vislumbró en 1980 gracias al descubrimiento por Wyman y White del primer locus de ADN polimórfico. La enorme variabilidad que presentaban dichos loci entre los diferentes individuos hizo que rápidamente se pensara en su aplicabilidad a la identificación genética, ya fuese con fines de resolución de casos de criminalística o en el estudio biológico de la paternidad. De esta manera, la Medicina Legal, y en concreto la Biología Forense, sufrió una gran revolución al poseer las herramientas necesarias para la resolución de casos de criminalística que hasta entonces ni siquiera habían sido considerados. La denominada "técnica del ADN" (nombre con que se conoce al conjunto de técnicas utilizadas para el estudio de los polimorfismos del ADN con fines identificativos) fue aplicada por primera vez para resolver un caso judicial en 1985, por el genetista inglés Alec Jeffreys. Desde la resolución de este primer caso las técnicas de análisis de ADN han experimentado un avance considerable haciéndose cada vez más sensibles y alcanzando un nivel de fiabilidad cercano al 100%.

Un requisito fundamental para poder utilizar nuevos polimorfismos de ADN con fines identificativos y obtener resultados altamente fiables es el estudio y caracterización de dichos polimorfismos en la población que posteriormente se utilizará como referencia. De ahí la importancia que tienen en Genética Forense los estudios genético-poblacionales.

Por este motivo uno de los principales objetivos del presente trabajo de tesis doctoral ha sido la caracterización en nuestra población de referencia (Andalucía) de diez marcadores de ADN polimórficos tipo microsatélite (**D13S317, D7S820, D16S539, D3S1358, HUMVWA, HUMFGA, D8S1179, D21S11, D18S51, D5S818**). Este estudio implica la determinación de frecuencias alélicas y genotípicas de cada uno de los marcadores en una muestra de la población estudiada, la comprobación de que los marcadores estudiados se encuentran en equilibrio de Hardy-Weinberg mediante la aplicación de una serie de tests estadísticos, el estudio de independencia de los loci estudiados y el análisis de una serie de parámetros de interés forense. Además se hará un estudio comparativo de los resultados obtenidos para nuestra población con los obtenidos para otras poblaciones estudiadas y de esta manera veremos si existen variaciones significativas de unas poblaciones a otras.

El otro objetivo fundamental de este trabajo ha sido la realización de un estudio de validación de tres de los diez marcadores analizados (**D13S317, D7S820, D16S539**) para comprobar la efectividad de los mismos en los casos en los que tanto la cantidad como la calidad del ADN no son las idóneas. La necesidad de la realización de este tipo de estudio deriva del hecho de que frecuentemente los indicios biológicos encontrados en el lugar o escena de un delito presentan problemas a la hora de su análisis debido a la escasez de muestra en algunas ocasiones, o al pésimo estado en que se encuentran al haber estado sometidas a determinadas condiciones ambientales desfavorables para la conservación del ADN (humedad, calor, etc.). Por este motivo consideramos de gran importancia demostrar que los marcadores utilizados van a ser también de utilidad en aquellos casos en los que el ADN en estudio presente alguno de los problemas comentados.

El estudio de validación realizado constó de tres partes:

- *Determinación de la sensibilidad*, es decir, comprobación de la mínima cantidad de ADN amplificable para cada uno de los marcadores.

- *Estudio de mezclas de ADN* en diferentes proporciones.

- Estudio de aplicabilidad de los marcadores estudiados en *muestras de interés forense* como manchas de sangre, semen, saliva, caspa, etc.

2. INTRODUCCIÓN

2.1. IDENTIFICACION GENETICA EN MEDICINA LEGAL

El concepto de identificación de un individuo puede definirse como la determinación de una serie de rasgos o cualidades que hacen que se distinga del resto de los individuos de su clase. Por lo tanto, podría decirse que identificar es establecer la individualidad o identidad de una persona. No obstante, existe un pequeño matiz que diferencia los términos de identificación e individualización ya que, mientras que el primer concepto se centra en la persona y la considera como diferente de las de su grupo, el segundo la define como elemento del grupo al que pertenece (Lorente y Lorente, 1995).

En Medicina Legal todas las cuestiones relacionadas con ambos conceptos tienen enorme importancia, aunque el problema que se plantea en uno y otro caso es completamente diferente, originándose dos ramas diferentes de la Medicina Legal:

- La **Antropología** forense, que parte del sujeto vivo, del cadáver o de sus restos y que generalmente pretende establecer su identidad y algunas características de los hechos de los que forma parte como víctima.

- La **Criminalística** que estudia los indicios dejados en el lugar de los hechos, gracias a los cuales podemos contestar a las principales preguntas que surgen ante la comisión de un delito: ¿quién es el autor? (a veces también: ¿quién es la víctima?) y ¿cuáles fueron las circunstancias que rodearon al hecho?. En el proceso judicial, la criminalística trata de convertir los indicios en pruebas admisibles por los tribunales.

La criminalística, dentro del campo penal, exige la existencia de un hecho delictivo, mientras que en el caso de la antropología puede ser que éste no exista y que el problema sea exclusivamente el de la identificación (Lorente y Lorente, 1995).

Para la investigación médico-legal se necesitan elementos que estén directamente relacionados con los hechos y que aporten la información necesaria para llegar a la identificación del individuo. Estos elementos son los indicios orgánicos que al poseer material biológico posibilitan la aplicación de técnicas analíticas.

En general, la identificación humana en Medicina Legal es un proceso de comparación entre los resultados obtenidos en el análisis de los indicios y los obtenidos a partir de las muestras de origen conocido (muestras indubitadas). La forma por la que se llega a la identificación es a través de la individualización. Conforme se analizan características de la persona las posibilidades de que exista otra persona que comparta dichas características disminuyen.

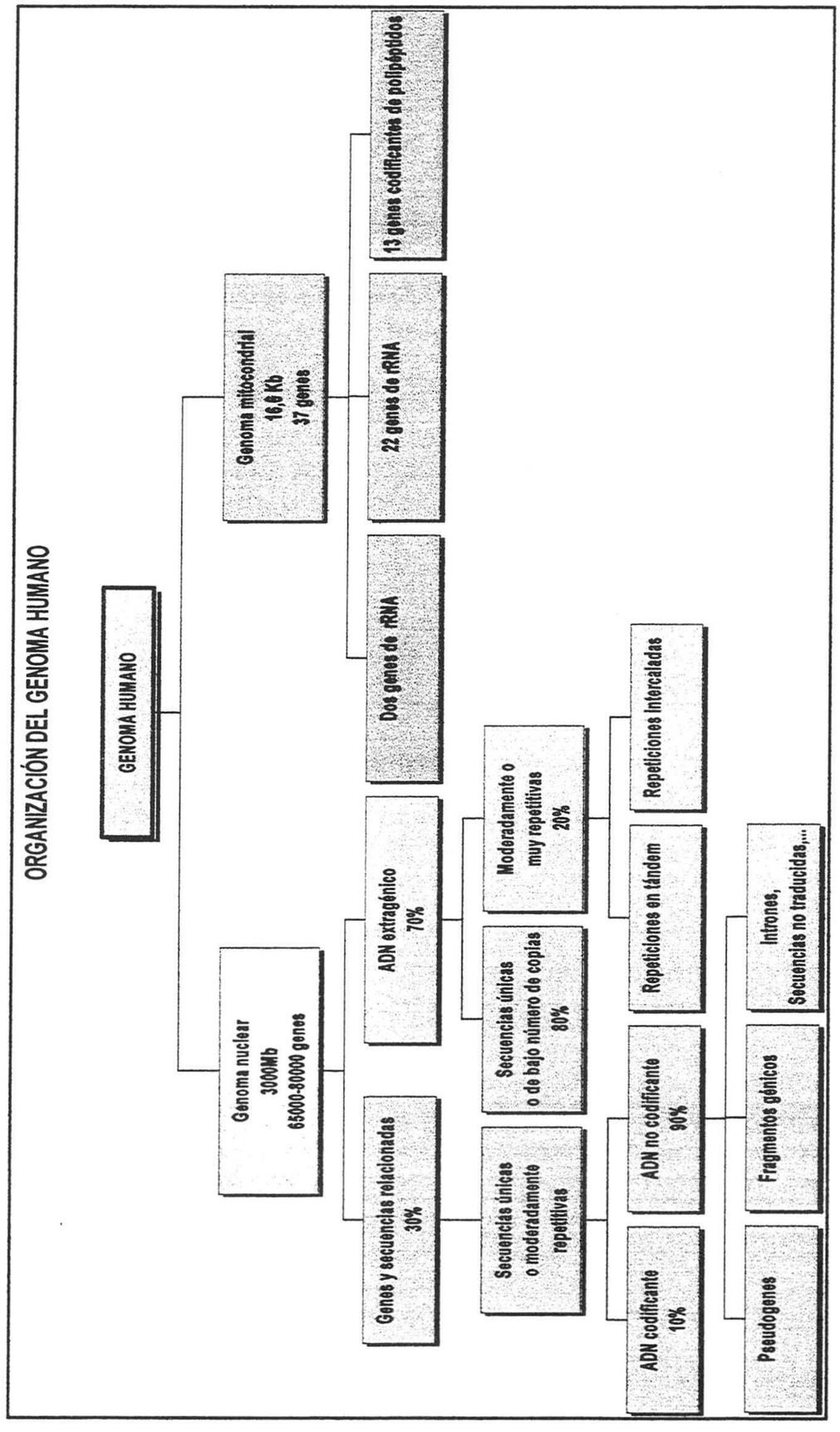
Las características analizadas con el estudio del ADN permiten, estadísticamente, identificar a un individuo con una probabilidad muy próxima al 100%. Aparte de identificar al individuo, la investigación medico-forense lo relaciona con una serie de hechos que forman parte de la investigación criminal, lo cual confiere a la Medicina Legal un componente social. La falta de una regulación específica en materia de la prueba pericial, remontándose al marco general de la ley de Enjuiciamiento Criminal, sin indicaciones específicas a la situación actual con la incorporación, sobre todo de la denominada "*tecnología del ADN*", puede hacer que las enormes posibilidades de la técnica queden mermadas a la hora de adoptar determinadas resoluciones en la investigación durante la fase de instrucción (Lorente y Lorente, 1995).

2.2. ESTUDIO DEL ADN

El término genoma humano se utiliza para describir la totalidad del ADN o material genético contenida en las células humanas. Este concepto engloba en realidad a dos genomas: un genoma complejo, el *nuclear*, y un genoma sencillo, el *mitocondrial*. El genoma nuclear proporciona la mayoría de la información genética esencial, que en su mayor parte es descodificada para especificar la síntesis de polipéptidos en los ribosomas citoplasmáticos. Las mitocondrias poseen sus propios ribosomas, y los escasos genes codificantes que contiene el genoma mitocondrial producen ARN mensajero que es traducido por ribosomas mitocondriales. No obstante, casi todos los polipéptidos mitocondriales vienen codificados en genes nucleares y son sintetizados en ribosomas citoplasmáticos antes de ser importados por la mitocondria.

El número de genes del genoma humano ha sido objeto de mucha especulación. Se sabe que el genoma mitocondrial contiene 37 genes pero se desconoce cual es exactamente el número de genes contenidos en el genoma nuclear. Cálculos teóricos basados en la carga mutacional que un genoma puede tolerar y en las tasas de mutación promedio observadas para el genoma humano estiman el límite superior en 100.000 genes. Los genes del genoma nuclear representan el 99,98% del número total de genes celulares. Si se considera que el tamaño medio de un gen humano es de 10 a 15 Kb y que no existe solapamiento, se obtiene que entre un 25% y un 35% del ADN nuclear está ocupado por genes. Dado que la mayor parte de los genes nucleares codifican polipéptidos y que la secuencia codificante para un polipéptido humano de tamaño medio puede considerarse cercana a los 500 codones (1,5 Kb), se puede llegar a la conclusión de que sólo el 3% del genoma nuclear tendría función codificante (ver figura 1.1) (Strachan et al, 1998).

Figura 1.1. Clasificación de las diferentes secuencias de ADN que componen el genoma humano.



2.2.1. ADN no-codificante

Entre un 80-90% del ADN contenido en el núcleo celular no cumple la definición clásica de gen y se ha denominado ADN no codificante. Debido a que por ahora no se le asocia función alguna también se le denomina ADN basura o ADN no esencial (Albert et al, 1989; Strachan et al, 1996). No obstante este ADN es una herramienta esencial para la investigación forense por toda una serie de características y peculiaridades que hacen de el un elemento ideal para la individualización.

En general el ADN no-codificante puede ser de dos tipos:

- ADN de copia sencilla o **no repetitivo** que actúa como ADN espaciador entre las regiones codificantes del genoma;
- ADN de copia múltiple, denominándose entonces ADN **repetitivo** (Singer, 1982a; Kao, 1985).

Entre las secuencias **no repetitivas** se encuentran los intrones, los cuales no contienen información útil para la síntesis de la cadena polipeptídica y son eliminados durante el proceso de "*splicing*". En las regiones que flanquean a los genes y en las zonas espaciadoras se han encontrado también secuencias extensas de ADN cuya función se desconoce. Las secuencias repetitivas se han encontrado en zonas como los orígenes de replicación los cuales son necesarios para el inicio de la replicación al entrar en la fase "S" de la mitosis (Huberman, 1995). Este tipo de secuencias fue descubierto gracias a diversos estudios.

Experimentos basados en técnicas de hibridación pusieron de manifiesto que tras romper el genoma en cortos fragmentos de ADN de doble hélice de aproximadamente 1000 pares de bases (pb), desnaturalizándolos después y comprobando la velocidad con la que vuelven a unirse para formar la doble cadena, en condiciones de alta astringencia que implicaban un apareamiento perfecto, aproximadamente el 70 % de las hebras de ADN se emparejaban tan lentamente como se podría esperar para una gran colección de una única secuencia de ADN no repetida, pero el 30 % restante se unía mucho más rápidamente, sugiriendo que contenían secuencias que estaban repetidas muchas veces en el genoma, y por lo tanto colisionaban con una hebra complementaria de forma relativamente rápida (Waring, 1966; Britten, 1968).

Otros estudios evidenciaron que tras centrifugar el ADN genómico en gradientes de cloruro de cesio (CsCl) se separaban fracciones de ADN que flotaban a diferente nivel con relación al nivel principal de dicho ácido nucleico, encontrándose que se trataba de fragmentos con secuencias altamente repetitivas (Kit, 1961; Schildkraut, 1962).

Un tercer grupo de estudios consistió en la restricción enzimática de todo el ADN genómico seguida de una separación electroforética de los fragmentos en geles de agarosa y la posterior tinción con bromuro de etidio, demostrando la presencia de distintas bandas formadas por secuencias de repetición sobre un fondo difuminado de fragmentos de ADN (Southern, 1979).

Todas estas experiencias demostraron la existencia en el genoma de unas regiones formadas por variantes de cortas secuencias organizadas en repeticiones de una serie de nucleótidos; secuencias que pueden ser **altamente repetitivas** (aproximadamente más de 10^5 por genoma haploide) o **moderadamente repetitivas** (sobre 10^3 a 10^5 copias por genoma haploide).

Las repeticiones en tándem de una secuencia sencilla se denominaron "*ADN satélite*" debido a que las primeras secuencias de este tipo que fueron descubiertas tenían una proporción poco habitual de nucleótidos que hacía posible su separación del grueso del ADN de la célula como un componente minoritario o satélite, coincidiendo, por otra parte, con los niveles que permanecían separados del grueso del ADN cuando se centrifugaba en gradientes de CsCl, como si de un satélite se tratara.

El hecho de que estas secuencias de ADN carezcan de cualquier función conocida y la comprobación de que son capaces de propagarse y extenderse por ellas mismas a través del genoma sin un objetivo o beneficio aparente para él, hace que este tipo de ADN sea conocido también como "ADN egoísta". Además, la capacidad para moverse ha hecho que en el transcurso de la evolución las secuencias satélite se hayan modificado extraordinariamente, cambiando, incluso, sus posiciones en los cromosomas.

El ADN repetitivo ha sido clasificado a su vez en diversos tipos basándose en su organización estructural y frecuencia de reiteración de cada clase o tipo. Hay dos grandes clases, las **secuencias repetidas en tándem** y las **secuencias repetidas intercaladas,** que analizaremos a continuación con cierto detalle por su importancia en la identificación genética.

2.1.1.1. Secuencias repetidas en tándem

Suponen el 5 - 10 % del genoma de los mamíferos y, generalmente, se caracterizan por la presencia de una secuencia común repetida en tándem de manera continua en un fragmento de ADN (Singer, 1982a). Son características de los centrómeros y de los telómeros y parece que son importantes para determinar la estabilidad y la posición correcta de los cromosomas. Se encuentran distribuidas en el genoma de dos maneras:

-***Tipo I***: En grandes bloques caracterizados por una disposición repetitiva en tándem de diversas unidades.

-***Tipo II***: En pequeños bloques con la misma disposición en tándem formados por la misma unidad básica y dispersos por todo en genoma.

REPETICIONES EN TANDEM			
	Grado de repetición (por locus)	**Numero de loci**	**Longitud de la secuencia consenso**
SATELITES	10^3-10^7	1-2 por cromosoma	1-varios miles de pares de bases
MINISATELITES	10-10^3	Varios miles por cromosoma	9-100 pb
MICROSATELITES	10-10^2	>10^5 por genoma	1-6 pb

* **Secuencias repetidas en tándem de tipo I**. En función del tamaño de la secuencia consenso se distinguen tres grupos:

- ***Familias satélite***: El término "*ADN satélite*" se ha utilizado para describir grandes bloques de ADN altamente repetitivo organizadas en tándem. Hay cuatro grandes clases de satélites: I, II, III y IV (Singer, 1982a). Constituyen un 2-6% del genoma y debido a su diferente proporción de bases G/C con respecto al resto del ADN pueden ser aislados en un gradiente de Ag-Cs_2SO_4. Las secuencias consenso oscilan entre los 5 pb (satélite III) y los 170 pb (satélite I).

Por hibridación "in situ" de las distintas fracciones de ADN satélite frente a cromosomas humanos en metafase se observó que la localización de estas secuencias es básicamente la heterocromatina estructural de las *regiones centroméricas*.

- ***Satélite alfoide***: Tiene una densidad similar a la de los satélites II y III, sin embargo su secuencia consenso es de 340 pb. La secuencia muestra gran similitud con otras secuencias de primates aunque en estos la longitud es de 171 pb.

Se ha sugerido que el origen de esta secuencia estaría en la formación de un dímero de estas dos unidades de 171 pb que posteriormente darían lugar a la secuencia de 340 pb.

- ***Midisatélites***: son secuencias repetidas en tándem con una longitud total comprendida entre 250 y 500 Kpb. Son de copia única y altamente polimórficos.

* **Secuencias repetidas en tándem de tipo II: repeticiones cortas**. Se distinguen dos grupos:

- ***Minisatélites***: Son secuencias de ADN constituidas por una secuencia consenso de 9 a 100 pb repetida en tándem (Nakamura, 1987,a). Los minisatélites suelen encontrarse cerca de los telómeros, aunque también se han encontrado intercalados en otras posiciones cromosómicas. En los extremos de los telómeros existe una familia de minisatélites que se ha denominado "ADN telomérico" y que consta de un bloque de 10 a 15 Kb de repeticiones en tándem de una secuencia de 6 nucleótidos. Los hexanucleótidos son añadidos a los extremos de los telómeros por una enzima llamada *telomerasa* y parece que actúan como amortiguadores de la degradación de los extremos de los cromosomas.

Estas secuencias de ADN son altamente polimórficas o variables entre los distintos individuos, presentando índices de heterocigosidad mayores del 50% e incluso a veces, cercanos al 100% (Nakamura, 1987,a). A dichos polimorfismos se les denomina abreviadamente **VNTR** (Variable Number of Tandem Repeat).

- ***Microsatélites***: Contienen una secuencia consenso de 1 a 6 pb repetida en tándem. Al igual que los minisatélites son secuencias altamente polimórficas y en este caso a los polimorfismos de estas secuencias se les llama **STR** (Short Tandem Repeat). Entre las repeticiones de un solo nucleótido son frecuentes las de A y T (0.3% del genoma nuclear), mientras que las de C o G son bastante más raras. En cuanto a repeticiones de dinucleótidos son comunes las repeticiones de CA (0.5% del genoma nuclear) y suelen ser muy polimórficas. También son frecuentes las repeticiones CT/AG (0.2%). Las repeticiones de trinucleótidos o tetranucleótidos son comparativamente raras pero poseen un alto grado de polimorfismo, lo cual hace que sean muy utilizados como marcadores de identificación genética.

La mayoría de las repeticiones en tándem de trinucleótidos son útiles como marcadores polimórficos pero, además, ciertas repeticiones de CAG/CTG y CCG/GGC presentan un comportamiento anormal. En ambos casos, si el bloque de secuencias repetidas contiene menos unidades que un cierto tamaño umbral entonces se mantiene estable, pero si se supera la longitud correspondiente al valor umbral entonces se dispara la inestabilidad. Esta inestabilidad puede consistir tanto en la adición como en la pérdida de tripletes, aunque lo más frecuente es la adición. Además, el grado de inestabilidad es tal que estas secuencias casi nunca se transmiten inalteradas de padres a hijos.

Se ha visto que la expansión de repeticiones de tripletes intragénicos puede originar una variedad de enfermedades como el síndrome del X-frágil (Fu et al. 1991), la distrofia miotónica (Fu et al. 1992), la enfermedad de Huntington (Weber et al, 1992), etc. También se ha relacionado la inestabilidad de determinados tetra y pentanucleótidos con algunos tipos de cáncer como el colorectal (Goellner et al, 1997), gástrico (Silva et al, 1997), melanoma (Walker et al, 1994) y cáncer de mama (Plummer et al, 1997).

2.2.1.2. Secuencias repetidas intercaladas

El segundo tipo principal de ADN repetitivo, el cual se encuentra en el genoma humano en una proporción del 20 %, está formado por secuencias de repeticiones intercaladas individualmente en forma de unidades sencillas en diversos puntos del genoma. Estos elementos también han sido descritos en algunos mamíferos, fundamentalmente en los primates superiores y roedores.

Atendiendo a su tamaño han sido clasificadas en dos grandes grupos (Singer, 1982b):

- **SINE** o elementos nucleares cortos dispersos (**Short Interspersed Nuclear Elements**). Uno de los SINE más conocidos es la familia de repeticiones *Alu*. Esta repetición es específica de los primates y en el genoma humano ha alcanzado un elevado número de copias.

- **LINE** o elementos nucleares largos dispersos (**Long Interspersed Nuclear Elements**). Un ejemplo de LINE en humanos es el elemento LINE-1 o L-1, que también puede encontrarse en otros mamíferos como el ratón.

A diferencia de las repeticiones en tándem, que se encuentran todas ellas a continuación unas de otras a lo largo de todo el fragmento, entre los elementos intercalados existen fragmentos de ADN no-repetitivo, apareciendo los SINEs o LINEs cada cierto número de pares de bases más o menos largo.

El mecanismo por el que los SINEs y LINES están dispersos en forma de múltiples copias por todo el genoma no se conoce del todo. Sin embargo ambos aparecen como ejemplos de elementos genéticos móviles. Por su origen pueden ser transcritos a ARN y la secuencia copiada puede ser integrada como ADN en otro lugar del genoma, del mismo modo que los retrovirus se integran en el genoma

eucariótico (retrotransposición). No obstante, esta distinción estructural entre las secuencias repetidas en tándem y las intercaladas no debe entenderse como rígida, ya que se han descrito casos, como por ejemplo la familia de repeticiones *Sau 3a*, consistentes en 5 subunidades en tándem de aproximadamente 170 pb cada una que producen un elemento similar a un LINE de aproximadamente 850 pb de longitud que puede existir independientemente en un locus determinado del genoma sería entonces un VNTR) o encontrarse en varios loci como una serie de 850 pb cada una formando un LINE. Es significativo que una pequeña fracción de la unidad de 850 pb del genoma humano es extracromosómico, sugiriendo que tiene propiedades de un elemento genético móvil. Efectivamente, se sabe que la mayoría de los fragmentos repetidos intercalados derivan de unas cuantas secuencias de ADN transponibles que se han multiplicado hasta dar lugar a un cierto número de copias en nuestro genoma.

Existen muchas variedades de elementos transponibles, los cuales oscilan entre 2.000 y 10.000 pb de longitud, la mayoría de ellos están presentes entre 5 y 10 copias por célula diploide, constituyendo el 10 % del genoma de los eucariotas superiores.

2.2.1.3. **Significado biológico del ADN repetitivo**

A pesar de que las secuencias de ADN repetitivo no codifiquen para ninguna proteína, no puede decirse que carezcan de función. De hecho, se han propuesto varias funciones para estas regiones:

- Puntos calientes para la recombinación homóloga (Jeffreys et al, 1985a).

- Regulación de la transcripción (Gruss, 1981).

- Dirección del posicionamiento de las cromátidas homólogas durante la meiosis (Sen, 1988).

- Desarrollo de superestructuras locus específicas responsables del plegamiento de la cromatina (Vogt, 1990).

- Síntesis protéica (Ohno, 1984).

Aunque las funciones del ADN repetitivo mencionadas puedan ser de cierta importancia para la célula, la utilidad del mismo en el campo de la Biología Forense no deriva de las mismas, sino de una serie de características que hacen de el un elemento clave para la identificación. Entre estas características está el hecho de que, al ser secuencias no codificantes, no están sometidas al filtro que supone la *selección natural*, con lo que cualquier mutación que ocurra en estas secuencias permanecerá. Por esta razón existe un alto grado de variabilidad o polimorfismo para estos fragmentos entre los distintos individuos. En general, la utilidad de estos polimorfismos se extiende tanto al campo de la Medicina Forense como al Diagnóstico Clínico de enfermedades genéticas, a la construcción de mapas genéticos, etc.

2.2.2. Polimorfismos del ADN

Los primeros polimorfismos de ADN que se estudiaron 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. A estos polimorfismos se les llamó *polimorfismos de longitud para fragmentos de restricción* o, de manera abreviada, **RFLP** (Restriction Fragments Length Polymorphisms).

Posteriormente los estudios se centraron en otro tipo de polimorfismos basados en variaciones en el número de repeticiones de una secuencia central. En función de la longitud de la secuencia central se distinguen dos tipos de polimorfismos:

- **VNTR ("Variable Number of Tandem Repeats")**
- **STR ("Short Tandem Repeats")**

El descubrimiento de estos dos tipos de polimorfismos basados en técnicas de amplificación "in vitro" (PCR), hizo que los RFLPs fueran perdiendo importancia en la identificación forense hasta el punto de que en la actualidad están en franco desuso en los laboratorios de Genética Forense.

2.2.2.1. ADN Minisatélite

El primer locus polimórfico de esta clase fué detectado por Wyman y White en 1980 con una sonda de ADN arbitraria. Sin embargo la base estructural de estos polimorfismos fue descubierta al estudiar varios alelos de la región 5' del gen de la insulina del cromosoma 11 (Bell et al, 1982). Posteriormente, el análisis de otros loci hipervariables ha confirmado la existencia y estructura de este tipo de secuencias de ADN.

La búsqueda de loci VNTR puede hacerse mediante hibridación con fragmentos de ADN que contengan las repeticiones o bien usando sondas sintéticas complementarias a la unidad de repetición. El análisis se realiza mediante el uso de enzimas de restricción cuya diana no se encuentre dentro de las unidades de repetición. En este caso el tamaño de los fragmentos de restricción estará en función del número de repeticiones presentes en el locus analizado.

Desde el descubrimiento de la PCR el análisis de los VNTR se ha simplificado enormemente al ofrecer una serie de ventajas que no ofrecía el análisis por RFLP (Lorente y Lorente, 1995):

- Consume menos tiempo.

- La cantidad de ADN que se requiere es menor; de hecho, la PCR permite la obtención de resultados a partir de cantidades muy pequeñas de ADN con lo cual se posibilita el análisis en aquellos casos en los que los indicios biológicos encontrados son mínimos (un pelo, una mancha pequeña de sangre o semen, etc.).

- Al analizarse fragmentos pequeños, es posible amplificar material antiguo y parcialmente degradado.

- Permite estudiar la existencia de alelos nulos que no son detectados mediante Southern blot, bien porque escapan del gel durante la electroforesis al ser de pequeño tamaño o bien porque al estar formados por pocas unidades de repetición hibridan con un número bajo de moléculas de sonda. El hecho de no detectar este tipo de alelos podría estar originando un exceso de homocigotos al estar considerando como homocigotos a individuos heterocigotos con un alelo de este tipo en su genotipo (Chakraborty, 1994).

- Posibilidad de amplificar conjuntamente varios marcadores (reacciones tipo *multiplex*).

A los loci VNTR que, por poseer un rango de pesos moleculares comprendido entre 100-2000 pb, pueden ser amplificados mediante PCR se denominan **AmpFLP**. En la siguiente tabla se muestran algunos de los AmpFLPs usados en identificación forense.

Marcadores	Tamaño	NºAlelos	Referencias
D1S80	224-640	29	Kasai et al.1990, Budowle et al.1991.
ApoB	570-900	21	Boerwinkle 1989
YNZ22	170-870	12	Horn, 1989
COL2A1	650-890	5	Wu et al.1990, Rand et al.1992.

2.2.2.2. ADN Microsatélite

Los loci microsatélites son los de mayor utilidad hoy en día en el campo de la identificación genética. Según Richards y Sutherland (1992) cada 6-10 Kb del genoma humano puede encontrarse una secuencia de este tipo. En cuanto a los loci tri y tetraméricos (que son los de mayor interés forense en la actualidad) se estima que se distribuyen cada 15 Kb (Edwards et al, 1991).

Los STR se distribuyen ampliamente por todo el genoma encontrándose tanto en regiones génicas como extragénicas. Los STR situados en regiones extragénicas o no codificantes son los responsables del gran avance de la Genética Forense. Además de su aplicabilidad en dicho campo, estos loci son utilizados en el desarrollo de mapas de ligamiento y en el diagnóstico clínico de enfermedades de origen genético tales y como la distrofia muscular Duchenne/Becker, fibrosis quística, corea de Huntington y distrofia miotónica, (Mully et al, 1991; Morral y Estivill, 1992; Oudet et 1991; Weber et al, 1992).

El polimorfismo de los loci STR radica, al igual que en los VNTR, en las variaciones del número de repeticiones de la secuencia central. Pero además, los STR pueden variar en la longitud de la unidad de repetición y en el rigor con el que llegan a formar un modelo de repeticiones. Las unidades de repetición "simples" contienen unidades de idéntica longitud y secuencia, mientras que las llamadas unidades de repetición "complejas" pueden contener bloques de repetición de unidades de longitud variable y además una mayor o menor variación de las secuencias. En base a esto podemos decir que pueden encontrarse dos tipos de STR:

- Los que poseen un número bajo de alelos (<12) pero bien diferenciados,

- los que poseen un elevado número de alelos (>35) poco diferenciados.

A pesar de que los loci del segundo grupo son los más polimórficos, no son los más usados en identificación forense puesto que al poseer alelos que difieren en un número incompleto de repeticiones, y a veces intercaladas entre otras normales, requieren un análisis muy preciso y por lo tanto poco práctico. Los loci del primer grupo presentan alelos que difieren en una unidad de repetición aunque existen algunos loci como el HUMTH01 en el que alguno de sus alelos difieren en una repetición incompleta. Estos alelos son el 8.3, 9.3 y 10.3 y consisten en 8, 9 y 10 repeticiones AATG respectivamente más un trinucleótido ATG. De estas tres variantes alélicas la más frecuente es el 9.3 por lo que parece probable que las otras dos variantes pudieran surgir por deleción o inserción de una repetición completa del alelo 9.3 originando los alelos 8.3 y 10.3 respectivamente (Brinkmann, 1996).

Los STR más utilizados son los tri, tetra y pentanucleotídicos, ya que los dinucleotídicos presentan bandas sombra o tartamudas ("stutter bands") como consecuencia de la amplificación. Entre los tetranucleótidos más comunes destacan los siguientes STR: HUMTH01, HUMVWFA31, HUMFESFPS, HUMF13A01, HUMFXIIIB, HUMTPOX, HUMCSF1PO, etc. (Martinez Jarreta B et al, 1999)

A la hora de seleccionar un STR para uso forense se requiere que éste cumpla una serie de características como son (Budowle, 1997):

- una alta heterocigosidad

- una tasa de mutación baja

- que no origine artefactos durante la amplificación tales y como bandas sombra

- fácilmente amplificable

- de pequeño tamaño y,

- fácilmente coamplificable junto a otros STR.

Los STR analizados en el presente trabajo han sido analizados mediante PCR multiplex, tres de ellos con el kit Silver STR^{TM}III Triplex de Promega (**D16S539, D7S820 y D13S317**) y nueve con el kit AmpFlSTR Profiler Plus de Perkin Elmer Applied Biosystem (**D3S1358, HUMVWA, HUMFGA, D8S1179, D21S11, D18S51, D5S818, D13S317 y D7S820**). En el segundo kit además de los nueve loci STR viene incluido otro marcador que no es un STR y que sirve para la determinación del sexo en una muestra de ADN. Este locus llamado amelogenina es un fragmento del gen que codifica para la amelogenina humana y que se localiza en los cromosomas sexuales. El producto de amplificación resultante depende del diseño de los primers usados, originándose en el caso de muestras de ADN masculino dos fragmentos de distinto tamaño y en el caso de ADN femenino dos fragmentos idénticos.

Una vez que los STR han sido amplificados se procede al tipaje o asignación de alelos para lo cual existen dos métodos:

-*electroforesis en geles de poliacrilamida* y posterior tinción con nitrato de plata,

-*electroforesis en capilar* y asignación de alelos automatizada a través de un software.

Los loci STR analizados en este trabajo han sido tipados por ambos métodos.

A pesar de que la electroforesis capilar está sustituyendo a los métodos convencionales de electroforesis, en muchos laboratorios aún no se ha descartado el uso de estos últimos y es por ello por lo que se consideró conveniente la necesidad de incluir en este trabajo de tesis un estudio poblacional y validación de uno de estos marcadores.

Las principales ventajas que ofrece el análisis de los STR frente a los VNTR son:

- Al ser fragmentos de pequeño tamaño (menores de 400 pb) son menos susceptibles a la degradación de manera que es posible obtener resultados a partir de ADN sumamente degradado.

- Poseen alelos discretos, lo cual permite simplificar el análisis estadístico de las frecuencias alélicas en las poblaciones por un lado, y por otro facilitar la interpretación de los resultados al migrar cada alelo a una longitud determinada.

- El tiempo consumido se reduce considerablemente si comparamos con los VNTR y más aún si consideramos los últimos avances en electroforesis capilar.

2.2.2.3. Origen genético de los polimorfismos en ADN repetitivo

Según el tamaño de la unidad de repetición, los polimorfismos de las secuencias repetitivas pueden deberse a diferentes mecanismos genéticos. En loci **microsatélites** las tasas de mutación en la línea germinal presentan variaciones considerables que pueden alcanzar niveles de hasta 8×10^{-3}. En las repeticiones de tetranucleótidos se sabe que se forman alelos de nueva longitud sin que haya intercambio de los marcadores flanqueantes, lo cual implica que éstos no se originan por entrecruzamiento desigual. Parece que el mecanismo más probable para explicar la variación del tamaño es una forma de intercambio de información de secuencia que comienza con un apareamiento incorrecto por deslizamiento (Strachan, 1998).

Este fenómeno ocurre cuando en el apareamiento de las dos hebras complementarias las secuencias repetidas de ambas hebras se reconocen pero no se aparean con la que ocupa la posición correcta en la hebra complementaria. Esto provoca un "deslizamiento" de una hebra sobre la otra que da lugar a un apareamiento incorrecto. Aunque este mecanismo puede ocurrir tanto en ADN replicante como no replicante, parece más probable que ocurra en el replicante y en este caso a dicho mecanismo se denomina **deslizamiento replicativo**. La replicación con deslizamiento es responsable de deleciones y duplicaciones debido al apareamiento entre repeticiones no contiguas y se ha sugerido que constituye un mecanismo importante de evolución de secuencias de ADN y del genoma en general.

En loci **satélites y minisatélites** los polimorfismos de longitud pueden originarse como consecuencia de un **intercambio desigual entre cromátidas hermanas**. El mecanismo consiste en una forma de recombinación en la que el entrecruzamiento afecta a secuencias no alélicas situadas en cromátidas hermanas. Tiene lugar preferentemente en posiciones en las que los bloques de secuencias repetidas en tándem son de tamaño moderado o grande. Si los cromosomas se rompen y se reúnen mientras las cromátidas están unidas por apareamientos incorrectos de este tipo aparecerá una inserción o una deleción de una unidad de repetición completa. Una de las cromátidas quedará con una inserción y la otra con una deleción ya que el intercambio es recíproco.

Otro tipo de mecanismos genéticos que parecen implicados en la generación de polimorfismos en secuencias de ADN repetidas en tándem son las **reacciones de conversión génica.** La conversión génica consiste en una transferencia no recíproca de información entre un par de secuencias no alélicas de ADN (conversión génica interlocus) o entre secuencias alélicas (conversión génica interalélica). En este proceso una de las secuencias, denominada secuencia dadora, permanece inalterada mientras que la otra (secuencia aceptora) gana alguna secuencia procedente de la dadora. En una primera etapa se produciría la formación de un heterodúplex entre una hebra de ADN del gen dador y la hebra complementaria del gen aceptor. Seguidamente se produciría la conversión de un fragmento del gen aceptor mediante el mecanismo de reparación de apareamientos incorrectos, mecanismo enzimático basado en enzimas reparadoras que reconocen una zona mal apareada en un heterodúplex y la reparan base a base (Strachan, 1998).

Las mutaciones de la línea germinal pueden estudiarse detectando y caracterizando alelos minisatélite mutantes en gametos aislados. Para ello se hace una PCR de múltiples diluciones de alícuotas de ADN (cada una con unas 100 moléculas) extraído a partir de esperma humano. Los productos de PCR obtenidos a partir de cada alícuota permiten identificar cualquier nueva mutación de un tamaño lo suficientemente distinto del alelo progenitor como para poder definirla como un alelo nuevo (Monckton et al, 1994).

Parece que la mayoría de las mutaciones que ocurren en los loci implican ganancia de unos pocos elementos en un extremo del bloque de secuencias repetidas. Existe una desviación hacia la ganancia neta de elementos repetidos y las evidencias apuntan a la existencia de intercambios de secuencia no recíprocos entre alelos, lo cual sugiere una conversión génica interalélica.

2.3. TECNICAS DE ESTUDIO DE LOS POLIMORFISMOS DEL ADN

2.3.1. Historia de las técnicas de análisis genético utilizadas en Identificación Humana

La Hemogenética Forense nace a principios de siglo, cuando Karl Landsteiner describe el sistema ABO de los hematíes y Von Durgen y Hirschfeld descubren su transmisión hereditaria. Esta ciencia surgió como una rama de la Criminalística cuyo objetivo era la identificación genética tanto en casos de investigación criminal como en estudios biológicos de la paternidad. Inicialmente, las investigaciones se centraban en el estudio de antígenos eritrocitarios (sistema ABO, Rh, MN), proteínas séricas, enzimas eritrocitarias y sistema HLA. Con el estudio de dichos marcadores podía incluirse o excluirse una persona como posible sospechoso por poseer una combinación genética igual o diferente a la del vestigio biológico hallado en el lugar de los hechos.

Pero fue a mediados de siglo cuando gracias al descubrimiento del ADN y de su estructura y al posterior avance en las técnicas de análisis de dicha molécula la Hemogenética Forense evolucionó considerablemente hasta el punto de que hoy en día puede hablarse de una nueva subespecialidad dentro de la Medicina Forense: la Genética Forense. Dicha ciencia estudia básicamente unas regiones del ADN que presentan variabilidad entre los distintos individuos, es decir, estudia regiones polimórficas del ADN. Así, analizando un determinado número de regiones polimórficas, la probabilidad de que dos individuos sean genéticamente iguales es prácticamente nula (excepto en el caso de gemelos univitelinos).

A pesar de que la Ciencia poseía las herramientas necesarias para el estudio del ADN su aplicación en la resolución de casos judiciales no se produjo hasta 1985 cuando el Ministerio del Interior Británico solicitó la ayuda de Alec J. Jeffreys, profesor de Genética de la Universidad de Leicester. Los primeros casos de Criminalística fueron resueltos gracias a la técnica de los RFLPs (fragmentos de restricción de longitud polimórfica). Jeffreys descubrió la existencia de unas regiones minisatélites hipervariables dispersas por el genoma humano que al ser tratadas con enzimas de restricción generaban fragmentos de longitud variable. Estudios posteriores realizados el mismo Jeffreys demostraron que las diferencias en el tamaño de estos fragmentos se debían a que estas regiones consistían en un determinado número de repeticiones en tándem de una secuencia central, el cual variaba de unos individuos a otros. El origen de dichas variaciones puede deberse a dos hechos:

- En primer lugar, la mutación de un nucleótido por otro en una región que actúe como secuencia diana para la enzima de restricción originará un fragmento de mayor tamaño al no reconocer la enzima dicha secuencia. O bien si corta en otro lugar podrá originar un fragmento de menor tamaño. Variaciones debidas a dichos cambios se denominan **polimorfismos mutacionales**.

- O puede que las variaciones se deban a la inserción o deleción de una secuencia dentro del fragmento que es cortado por la enzima de restricción en cuyo caso se denomina **polimorfismo de longitud o de inserción-deleción.**

El primer locus de ADN polimórfico fué descubierto usando una sonda de ADN arbitraria (Wyman y White, 1980). De esta manera observaron fragmentos de más de 15 longitudes diferentes en una pequeña muestra de individuos. Posteriormente se encontraron otros loci hipervariables como en la secuencia del gen de la insulina humana (Bell et al, 1982), en el oncogen "ras" (Capon et al, 1983), en el pseudogen de la zeta-globina (Proudfoot et al, 1982) y en el gen de la mioglobina (Jeffreys et al, 1985b). Estos loci hipervariables constaban de repeticiones en tándem de una secuencia de oligonucleótidos (11 a 60 pb), de manera que las diferentes longitudes de los fragmentos originados dependían del número de dichas repeticiones y se les denominó **VNTR** ("Variable Number of Tandem Repeat").

Tras el descubrimiento de los primeros VNTRs se vio que éstos podían ser aplicados a la medicina forense y sustituir a los marcadores clásicos (Jeffreys et al 1985b).

En un principio la manera de estudiar dichos marcadores se hizo por medio de la técnica llamada ***hibridación con sondas*** o ***Southern blot***. Esta técnica consta básicamente de las siguientes etapas:

1. *Digestión* del ADN con enzimas de restricción tras conseguir extraer un ADN de alta molecularidad.

2. Separación de los fragmentos obtenidos por medio de una *electroforesis* en gel de agarosa.

3. *Desnaturalización* de los fragmentos separados y cortados.

4. *Transferencia* de las cadenas simples a una membrana de nitrocelulosa o nylon y *fijación* de las mismas por medio de calor (80ºC).

5. *Prehibridación* con sondas de ADN inespecífico para bloquear los lugares de unión inespecíficos que pudiera haber en la membrana.

6. *Marcaje de la sonda* con nucleótidos radioactivos (^{32}P normalmente).

7. *Hibridación* de la sonda marcada y desnaturalizada con los fragmentos de ADN fijados a la membrana, y lavado de la membrana para eliminar el exceso de sonda o aquellas que hayan hibridado mal.

8. *Revelado* en placa radiográfica e *interpretación* de los resultados.

El tipo de sondas utilizadas puede ser de dos tipos:

- **Sondas Mono-locus**: son específicas para una región de un determinado cromosoma. Se unen a secuencias largas de nucleótidos y presentan mayor variabilidad que las sondas multi-locus. Como resultado se observan una o dos bandas por individuo, según sea homocigoto o heterocigoto. El patrón de bandas obtenido con estas sondas se denomina perfil unilocus de ADN o "DNA profiling"

- **Sondas Multi-locus**: hibridan con secuencias minisatélites presentes en varios loci de diferentes cromosomas. Son sondas de 10 a 15 nucleótidos que se repiten múltiples veces y tras el revelado se observan de 10 a 20 bandas por persona. Este patrón de múltiples bandas se conoce como huella genética multilocus o "DNA fingerprint".

A pesar de que el análisis SLP ha sido y es bastante útil en estudios de paternidad no puede decirse lo mismo de su aplicación a la Criminalística ya que presenta una serie de inconvenientes como son:

- En el caso de sondas monolocus la cantidad de ADN que se necesita está entre 20 y 100 ng, cantidad difícil de conseguir en casos de criminalística en los que los indicios biológicos encontrados son mínimos.

- En cuanto a la calidad del ADN, en la práctica forense es muy difícil encontrar en estado no degradado toda la cantidad de ADN que se necesita para un análisis con sondas mono-locus.

- El tiempo requerido para este tipo de análisis es de dos o tres días.

- El hecho de que se requieran cantidades elevadas de ADN hacen que normalmente, con el primer análisis se consume la totalidad de la muestra, con lo que se dificultan contrapericias y una posterior revisión del caso.

Todas estas limitaciones fueron superadas gracias a la aplicación en Genética Forense de una técnica, la Reacción en Cadena de la Polimerasa ("PCR"), que supuso una revolución en muchos campos de la Biología y de la Medicina.

Las sondas multi y mono-locus presentan una serie de ventajas e inconvenientes con respecto a una serie de parámetros como son:

- **Información aportada**: las sondas multi-locus tienen una mayor capacidad discriminativa al aparecer múltiples bandas. No obstante, las mono-locus son más específicas ya que el fragmento de ADN con el que hibridan es de mayor tamaño.

- **Cantidad y calidad del ADN**: cuando se usan sondas multi-locus se requiere aproximadamente un microgramo de ADN sin degradar mientras que en el caso de las mono-locus se necesita menos de 100 ng y este ADN no necesariamente debe estar en perfecto estado, siempre y cuando el fragmento complementario a la sonda esté intacto.

- **Especificidad entre especies**: las sondas multi-locus permiten su uso sobre el ADN humano y de cientos de animales superiores, mientras que las mono-locus son exclusivas de ADN humano.

2.3.2. ANALISIS POR LA REACCION EN CADENA DE LA POLIMERASA

La reacción en cadena de la polimerasa (conocida como PCR por sus siglas en inglés, *Polymerase Chain Reaction*) permite amplificar más de un millón de veces un ADN obtenido a partir de una región seleccionada del genoma, siempre y cuando se conozca una parte de su secuencia de nucleótidos. Esta técnica fue ideada en 1989 por Kary B. Mullis que obtuvo el premio Nobel de Química en 1993 por dicho invento.

Para la PCR se utilizan dos oligonucleótidos sintéticos de unos 15-20 nucleótidos que son complementarios a las zonas flanqueantes de la región que se quiere amplificar. Estos oligonucleótidos (habitualmente conocidos por su nombre en inglés, "primers") actúan como cebadores para la síntesis in vitro de ADN la cual está habitualmente catalizada por una enzima llamada *Taq polimerasa*. Dicha enzima se aísla de una bacteria termófila, denominada *Thermus Aquáticus,* que es capaz de crecer a temperaturas elevadas (79 - 85 ° C). A esta temperatura dicha enzima es capaz de mantener una media de extensión de más de 60 nucleótidos por segundo en regiones ricas en uniones G-C. La temperatura optima a la que actúa la *Taq polimerasa* permite el uso de elevadas temperaturas para la unión de los primers y para la extensión, de esta manera se aumenta el nivel de exigencia de la reacción y se reduce la extensión de los primers unidos inespecíficamente al ADN.

La reacción se lleva a cabo en una serie de ciclos cada uno de los cuales incluye tres fases o pasos (Mullis K,1990; Vosberg HP et al, 1989):

1. ***DESNATURALIZACIÓN***: Para que comience la reacción es necesario que el ADN molde se encuentre en forma de cadena sencilla. Esto se consigue aplicando temperaturas de 90 a 95ºC que producen la rotura de los puentes de hidrógeno intercatenarios y por lo tanto la separación de ambas cadenas. Para conseguir la completa separación de las hebras de toda la muestra esta temperatura debe mantenerse unos minutos. Si el ADN solo se desnaturaliza parcialmente éste tenderá a renaturalizarse rápidamente, evitando así una eficiente hibridación de los primers y una posterior extensión.

2. ***HIBRIDACIÓN***: Esta fase se denomina también fase de "annealing" o de emparejamiento. Una vez que el ADN está desnaturalizado se disminuye la temperatura hasta un rango comprendido entre los 40 y los 60ºC para que se pueda producir la unión de los primers a las secuencias flanqueantes del fragmento que se va a amplificar. La temperatura de fusión o annealing (Tm, "melting temperature") depende de varios factores y es relativamente específica para cada primer. La longitud de los primers y la secuencia son críticas en la designación de los parámetros de una amplificación, una fórmula simple para calcular la Tm es la siguiente:

$$Tm = 4(G+C) + 2(A+T)$$

No obstante, cada primer exige una serie de estudios experimentales para determinar su temperatura de annealing específica ya que si la temperatura es muy baja la unión se hará de forma inespecífica y si es muy alta no se producirá una unión completa.

3. ***EXTENSIÓN***: Durante este paso la *Taq polimerasa* incorpora nucleótidos en el extremo 3' del primer utilizando como molde la cadena de ADN previamente desnaturalizada. La temperatura a la que se lleva a cabo este paso suele ser de 72ºC ya que es la temperatura a la que la *Taq polimerasa* alcanza su máxima actividad. Normalmente una extensión de 20 segundos es suficiente para fragmentos menores de 500 pb, y 40 segundos para fragmentos por encima de 1.2Kb.

Un factor importante en el transcurso de las diferentes fases es el **tiempo de rampa**. Este se define como el tiempo invertido en pasar de una temperatura a otra y depende del diseño y de las características del aparato donde se realiza automáticamente este proceso, el termociclador. En las nuevas generaciones de termocicladores este factor se ha ido optimizando para hacerlo mínimo. Nosotros hemos utilizado para la realización de esta tesis un termociclador Gene-Amp PCR System 9600 de Perkin Elmer en el que la tasa media de calentamiento y enfriamiento corresponden a 0.15ºC/segundo y 0.46ºC/segundo, respectivamente.

2.3.2.1. COMPONENTES DE LA PCR

* *BUFFER DE AMPLIFICACIÓN*

Los buffer de PCR que se utilizan normalmente contienen KCl, Tris y $MgCl_2$. El buffer empleado por nosotros es un buffer comercial concentrado diez veces (10X) y que contiene:

- Gelatina..... 0.01% (w/v), en un volumen de 1,4 ml
- KCl ..... 500 mM
- Tris-ClH..... 100 mM, pH 8,3
- $MgCl_2$ ..... 15 mM

El $MgCl_2$ es el componente que más influye en la especificidad y rendimiento de la reacción ya que los iones Mg^{2+} son necesarios para la actividad de la *Taq polimerasa*, es decir, actúan como cofactores de la polimerasa.

La concentración óptima de $MgCl_2$ está en torno a 1.5 mM si se emplean concentraciones de 200 mM de cada uno de los dNTPs. No obstante, a veces es necesario probar con diferentes cantidades de Mg ya que un exceso del mismo origina una acumulación de productos inespecíficos y una cantidad insuficiente hace que disminuya el rendimiento de la amplificación.

* *PRIMERS*

A la hora de elegir unos primers para amplificar un determinado fragmento de ADN hay una serie de reglas a seguir (Coyne et al, 1996):

- La *longitud de cada uno de los primers* debe estar comprendida entre 18 y 24 bases ya que se ha comprobado que primers de mayor longitud (30-35 bases) no aumentan el rendimiento y los primers cortos carecen de suficiente especificidad.

- Ambos primers deben tener una *Tm* similar (como mucho la diferencia entre ambas temperatura debe ser de 5ºC).

- La *relación bases púricas: bases pirimidínicas* debe ser 1:1 (o como mucho 40-60%).

- La *secuencia de los primers* debe comenzar y terminar con 1-2 bases púricas.

- Para evitar la formación de *dímeros de primers* es necesario comprobar que los primers no contengan secuencias complementarias entre sí.

Los *dímeros de primers* consisten en fragmentos de doble cadena cuya longitud es muy próxima a la de la suma de los primers y se producen cuando un primer es extendido a continuación del otro. El mecanismo exacto por el que se forman estos dímeros no está completamente determinado. La observación de que primers con los extremos 3' complementarios favorecen su formación sugiere que el paso inicial se debe a interacciones transitorias que aproximan los extremos complementarios. Algunas polimerasas, incluida la Taq, han mostrado una débil actividad

polimerizadora no dirigida por un ADN patrón, la cual puede unir nucleótidos adicionales al doble extremo apareado. Si esta actividad puede producirse sobre una hebra sencilla de oligonucleótidos, resultaría una buena oportunidad para que la extensión formara un corto solapamiento en el extremo 3' con el otro primer, suficiente para promover la formación del dímero.

* *DESOXINUCLEÓTIDOS TRIFOSFATOS*

Las concentraciones de dNTPs que suelen usarse están en torno a 200 µM para cada uno de ellos. En un volumen de reacción de 25 µl con esta concentración de dNTPs se sintetizarían entre 6-6.5 µg de ADN. La concentración de dNTPs y de $MgCl_2$ va relacionadas ya que el Mg se une a los dNTPs con lo que concentraciones elevadas de dNTPs inhibirían la reacción al no tener la *Taq polimerasa* suficiente Mg como para incorporar dNTPs. Para una concentración de 200 µM de cada dNTP se suele añadir $MgCl_2$ a una concentración de 1.5 mM.

* *TAQ-POLIMERASA*

Las cantidades óptimas de *Taq polimerasa* necesarias para la síntesis de ADN están alrededor de 2 unidades en 25 µl de volumen final de reacción. La actividad de este enzima se ve influenciada por la concentración de dNTPs, de Mg^{2+} y de algunos iones monovalentes de manera que concentraciones elevadas de los mismos inhiben dicha actividad.

Por otro lado, pequeñas concentraciones de KCl estimulan la actividad sintética de la Taq en un 50-60% con un máximo aparente cuando su concentración es de 50 mM. Existen algunos datos relacionados con la influencia de ciertos reactivos que se emplean antes de la amplificación y que alteran la actividad de la Taq. Por ejemplo

concentraciones 1M de urea estimulan la actividad, el SDS a bajas concentraciones que la inhibe al igual que concentraciones mayores del 10% de etanol.

* *ADN MOLDE O "TEMPLATE"*

Es el ADN del cual queremos copiar un determinado fragmento, es, por tanto, el ADN que la Taq polimerasa utiliza como molde para la síntesis de nuevas cadenas polinucleotídicas. La cantidad de ADN necesaria para la PCR depende de varios factores:

- Del *marcador* que se va a amplificar: hay marcadores cuyos primers son más específicos o bien cuyas condiciones de amplificación están mejor optimizadas que las de otros. Por esta razón puede darse el caso de que cierta cantidad de ADN (sobre todo cuando jugamos con cantidades mínimas) amplifique para unos marcadores pero no para otros. Por ello, cuando en un laboratorio se va a utilizar un nuevo marcador es necesario hacer un estudio de validación en él que se incluye un estudio de sensibilidad. De dicho estudio de sensibilidad puede sacarse como conclusión cuál es la mínima cantidad de ADN que amplifica en condiciones estándar. De manera general, para casi todos los STR utilizados en Genética Forense la cantidad óptima de ADN que asegura un rendimiento adecuado está en torno a los 5 ng.

- *Calidad del ADN*: Cuando se trabaja con ADN cuya calidad es óptima no suele haber problemas en la amplificación y cantidades del mismo por encima e incluso por debajo de los 5 ng rinden buenos resultados. El problema aparece cuando la calidad del ADN obtenido no es la idónea, bien porque esté degradado o bien porque dicho ADN vaya ligado a una serie de contaminantes que pueden inhibir la actividad de la polimerasa. Si el ADN está degradado por la acción de enzimas de restricción el que obtengamos o no resultado en la amplificación va a depender de

que el fragmento a amplificar haya sido dañado o no. En el caso en el que tengamos ADN sin degradar pero unido a una serie de contaminantes habría que intentar diluir al máximo la muestra para disminuir dichos contaminantes, pero siempre dentro de un rango de ADN que no esté por debajo del límite de sensibilidad de la PCR. El problema de las cantidades mínimas de ADN y la presencia de contaminantes o inhibidores de la Taq es un hecho habitual en Criminalística y requiere un estudio pormenorizado de la muestra antes de la amplificación.

* ADYUVANTES DE LA PCR

Son elementos que mejoran el rendimiento y la especificidad de la PCR. Aunque algunos autores han recomendado el uso del DMSO y del glicerol, el adyuvante más extendido y utilizado es el BSA. A concentraciones por encima de 0.8 µg/µl el BSA incrementa la eficiencia de la PCR ya actúa como una proteína captadora de iones que pueden ser inhibidores de la Taq polimerasa.

2.3.2.2. VENTAJAS Y DESVENTAJAS DEL USO DE LA PCR EN ANÁLISIS DE INDICIOS BIOLOGICOS

Desde su aplicación práctica en 1989, la PCR ha supuesto una revolución en diversos campos de la Ciencia y de la Medicina, entre los que se encuentra la Genética Forense. El estudio de indicios biológicos por PCR ha permitido la resolución de un gran número de casos en Criminalística que hasta entonces eran desestimados por no poseer la suficiente cantidad de muestra para su análisis por RFLP. Con el uso de la PCR muestras tan mínimas como pueden ser un pelo con raíz, una minúscula mancha de sangre o semen e incluso caspa son suficientes en muchos casos para llevar a cabo un análisis de identificación genética.

La PCR ofrece una serie de ventajas, frente al uso de las técnicas de análisis genético utilizadas con anterioridad, como son (Devlin et al, 1990):

- Su capacidad para obtener resultados en casos en los que la cantidad de ADN es **mínima** o en casos en los que el ADN esté parcialmente **degradado**.

- Genera en un espacio corto de tiempo un **elevado número de copias** de la secuencia de ADN que es objeto de estudio, lo cual permite utilizar técnicas de visualización más sencillas y rápidas que el uso de sondas marcadas radioactivamente.

- Permite la determinación y agrupación alélica en **clases discretas**, lo que facilita la elaboración de bases de datos al ser la estandarización inmediata y posibilitar la aplicación de métodos bioestadísticos y programas elaborados.

El uso de marcadores microsatélites de pequeño peso molecular aumenta las probabilidades de obtener resultados positivos de amplificación cuando el ADN se encuentra degradado ya que puede ser que dichos fragmentos no hayan sido digeridos.

Esta ventaja es de gran importancia en Criminalística ya que normalmente los indicios biológicos encontrados han estado sometidos a diversos factores (calor y humedad) que favorecen el crecimiento bacteriano.

Sin embargo, una de las grandes ventajas de la PCR que es su elevada sensibilidad puede, en ocasiones, convertirse en un gran problema ya que se podría coamplificar un ADN extraño o ajeno al que nos interesa. No obstante, en los laboratorios de Genética Forense las medidas de precaución que se toman para evitar problemas de contaminación por manipulación son extremas.

2.3.3. Análisis de la variación interna de los loci microsatélites (MVR)

En 1984, Owerbach y Aagard llegaron a la conclusión mientras estudiaban la secuencia de un locus hipervariable cercano al gen de la insulina de que los loci minisatélites pueden estar constituidos por unidades de repetición que varían en su secuencia. Posteriormente se descubrieron otros loci minisatélites que presentaban dichas variaciones, por ejemplo un locus próximo al oncogén Ha-ras (Capon et al, 1983), el gen de la α-globina (Jarman et al, 1986) y todos los loci analizados por hibridación cruzada con sondas multilocus (Wong et al, 1987; Jeffreys et al, 1985a). Según algunos autores el origen de este tipo de variabilidad puede explicarse por mecanismos de conversión génica inducidos por las regiones de flanqueo (Smith, 1976; Dover, 1982).

La variabilidad en la secuencia de la unidad de repetición no se limita a los minisatélites ya que también se han encontrado en ADN satélite (Willard,1987), en bloques de repeticiones situadas en los telómeros (Allshire, 1989), en microsatélites (Weber, 1990) y en colas poli(dA) de retrotransposones (Epstein, 1990).

El primer locus minisatélite analizado mediante MVR fue el D1S8, cuya sonda se denominó MS32 (Wong et al 1987). Se localiza en el cromosoma 1 y está formado por una unidad de repetición de 29 pb, posee al menos 50 alelos distinguibles en longitud (entre 1,7 y 18 Kb) y una heterozigosidad estimada del 97,5%. La secuenciación de este minisatélite demostró que aproximadamente la mitad de las unidades de repetición compartían una transición A/G, la cual hace que se cree o destruya un punto de corte para la enzima de restricción *HaeIII.* Otra enzima utilizada en el análisis es la denominada HinfI, que corta una vez en todas las unidades de repetición.

La técnica de los MVR fue desarrollada en 1990 por Jeffreys y cols. y básicamente consistía en amplificar el ADN con unos primers específicos y marcarlo radioactivamente en uno de los extremos una vez amplificado. Tras el marcado se procedía a la digestión parcial con las enzimas HaeIII o HinfI, electroforesis y autorradiografía. La digestión con HinfI originaba un patrón de fragmentos en escalera continua; mientras que la digestión con HaeIII producía una escalera discontinua al no cortar en todas las repeticiones. El producto de este análisis puede codificarse como un código binario en el que un 1 representaría a las repeticiones cortadas por HaeIII y un 0 a las no cortadas por dicha enzima (Rodriguez-Calvo et al, 1996).

El problema que planteaba este análisis era que solo se podían analizar loci minisatélites que tuviesen unidades de repetición en los que se creasen o destruyesen sitios de restricción. Para evitar dicho problema Jeffreys y col. (1991) idearon un método basado en la utilización de unos cebadores específicos que se unían a las repeticiones *tipo-a* o *tipo-t*. La amplificación con el cebador específico para las repeticiones *tipo-a* junto con un cebador flanqueante del extremo 5' del minisatélite MS32 generaba un conjunto de productos de PCR para cada alelo que se extendían desde el extremo flanqueante hasta cada repetición de *tipo-a*. Lo mismo ocurría con las repeticiones de *tipo-t*. Al separar por electroforesis los productos de ambas amplificaciones se observaban dos escaleras complementarias de las que se deducía el código MVR.

Sin embargo, la dificultad que generó este método fue la amplificación preferencial de los alelos de menor tamaño ya que los cebadores específicos se unían a cualquier punto del minisatélite y a los productos más pequeños ya amplificados. Por este motivo esta técnica sufrió una serie de modificaciones que hicieron posible la amplificación de las unidades de repetición de mayor tamaño y, por lo tanto, la posibilidad de obtener el código binario completo. Además,

investigaciones posteriores revelaron otros tipos de repeticiones como las denominadas repeticiones tipo-N (Jeffreys et al, 1991; Tamaki et al, 1992), las tipo-J y las tipo-U.

Existen varias estrategias de detección de los productos amplificados, la más utilizada es la basada en la electroforesis en geles de agarosa seguida de hibridación con la correspondiente sonda marcada tanto de forma isotópica (Jeffreys et al, 1991; Tamaki et al, 1992) como no isotópica (Rodriguez-Calvo et al, 1996; Hopkins et al, 1994) y, finalmente, detección autorradiográfica de las bandas. Otros métodos empleados son la utilización de colorantes fluorescentes para la detección en secuenciadores automáticos y la electroforesis en minigeles de poliacrilamida seguida de tinción con nitrato de plata (Rodríguez-Calvo et al, 1996; Bellas et al, 1997).

No todos los loci son apropiados para un análisis MVR-PCR ya que se requieren una serie de características como son:

- La heterozigosidad basada en la longitud debe ser elevada (superior al 95%), para asegurarse de que la mayoría de los alelos son raros.

- El minisatélite debe tener variabilidad en la secuencia de las unidades de repetición; para que sea más informativo es mejor que solo tenga dos clases de repeticiones, y que éstas estén distribuidas regularmente a lo largo de todo el minisatélite.

- No debe poseer unidades de repetición con longitudes anormales puesto que en este caso las escaleras obtenidas tras la amplificación saldrían fuera del registro, haciendo el código diploide indescifrable (Jeffreys et al, 1993).

Aunque la mayoría de los minisatélites estudiados no cumplen alguno de estos tres requisitos, se han descrito algunos loci susceptibles al análisis por MVR-PCR, como son: MS31A (locus D7S21) (Neil et al, 1993), MS205 (locus D16S309) (Armour et al, 1993), G3 (locus D7S22) (Andreassen et al, 1997), etc.

Aplicaciones de los MVR

Es poco probable que la MVR-PCR pueda llegar a ser utilizada en pruebas de paternidad debido a la alta tasa de mutación que posee el minisatélite MS32 y por la presencia de unidades de repetición que no pueden amplificarse (repeticiones nulas o *tipo-0*). En cambio, su aplicación tanto en casos de criminalística como en estudios de identificación puede ser notable ya que se trata de una técnica sensible y rápida. Además, los códigos digitales poseen un elevado poder de discriminación y permiten la rápida elaboración de bases de datos.

Existen diversos trabajos en los que esta técnica ha sido aplicada con éxito en el estudio de diversos indicios biológicos como: manchas de semen y pelos (Yakamoto et al, 1994b), saliva de un sello de correos usado (Hopkins et al, 1994), manchas de sangre de diferente antigüedad y piezas dentarias sometidas a distintas temperaturas (Bellas et al, 1997). También ha sido utilizada para la identificación de cadáveres incinerados (Tamaki et al, 1995).

En conclusión, puede decirse que a pesar de que se necesite un estudio más amplio que evalúe el poder discriminativo de este método, la MVR-PCR es una técnica eficiente y reproducible para su aplicación en criminalística.

2.3.4. Sistemas de análisis de los polimorfismos STR

El análisis de los loci STR para su aplicación en identificación genética humana se ha simplificado enormemente gracias a la posibilidad de amplificar conjuntamente dos o más loci en una única reacción (multiplex-PCR) (Edwards, 1991). Los sistemas de amplificación multiplex ofrecen una serie de ventajas sobre los sistemas individuales:

- Disminuyen el tiempo de análisis
- Aumentan el poder de discriminación del sistema
- La cantidad de ADN que se requiere es menor
- Disminuyen la cantidad de reactivos con lo que se abarata el coste total del análisis.

Los productos de amplificación pueden ser separados en función de su tamaño mediante una técnica rutinaria en los laboratorios de Genética Forense, la electroforesis. Mediante esta técnica los fragmentos de ADN amplificado migran a través de un soporte, que suele ser de agarosa o de acrilamida, cuando son sometidos a la acción de un campo eléctrico. Los sistemas convencionales de electroforesis utilizan como soporte geles desnaturalizantes de poliacrilamida y la detección de los fragmentos separados se hace de manera manual mediante tinción de los geles con nitrato de plata.

2.3.4.1. Tipificación manual

La tipificación manual de STR mediante electroforesis desnaturalizante en geles de poliacrilamida y su posterior tinción con nitrato de plata presenta las ventajas de que es un método sencillo y que no requiere un equipamiento costoso. Es una técnica accesible a cualquier laboratorio básico de Genética Forense y que permite la resolución tanto de estudios de paternidad como de identificación genética. El uso conjunto de tres de estos sistemas multiplex: **CTT** (CSF1PO, TPOX, TH01), **FFv** (F13A01, FESFPS, vWA) y el **SilverSTR™ III** Triplex (D16S539, D7S820, D13S317) permite que se alcancen probabilidades de coincidencia de 1 en 1.03×10^9 en poblaciones de caucasianos-americanos e índices de paternidad de 521 (Promega, Technical Manual, 1998). Sin embargo, presentan otra serie de inconvenientes que dificultan la interpretación de los datos: uno de ellos es el desdoblamiento de bandas y el otro la aparición de bandas tartamudas o "sttuter bands".

El desdoblamiento de bandas se debe al hecho de que a veces las dos cadenas de ADN de un determinado fragmento amplificado pueden migrar a diferentes tasas debido a la diferencia en la secuencia de bases. Este hecho suele observarse en los fragmentos de mayor tamaño.

Las *sttuter bands* son bandas que poseen una unidad de repetición menos que el alelo verdadero. Son generadas por un deslizamiento de la polimerasa durante la amplificación y dificultan enormemente la interpretación de los resultados en los casos de mezclas de muestras.

2.3.4.2. Sistemas de electroforesis capilar y detección por fluorescencia

La electroforesis capilar (normalmente representada por su acrónimo en inglés, CE) es un método alternativo para la separación de fragmentos y obtención de secuencias de ADN que suple, en parte, los inconvenientes de los sistemas convencionales. En este caso el soporte o medio de separación es un polímero incluido en un capilar de silica de unos 50μm y de longitud variable lo cual hace que la cantidad de calor generado sea menor y que puedan aplicarse voltajes mayores. La gran ventaja en cuanto a rapidez de la técnica se debe a que la preparación del gel y la carga de muestras se hacen de manera automática. Por otro lado se consigue una mayor sensibilidad, al poderse obtener resultados interpretables en los casos en los que el número de copias obtenidas por PCR es bastante bajo, y aumenta la capacidad para diferenciar alelos que se distinguen en un par de bases. Además, los resultados obtenidos son analizados por un software evitándose así problemas de interpretación y permitiendo que éstos queden almacenados para posibles futuros análisis.

Para que pueda llevarse a cabo el análisis por electroforesis capilar es necesario que el ADN sea amplificado utilizando un par de primers o más (en el caso de los multiplex) marcados en el extremo 5' con unas moléculas llamadas *fluorocromos* los cuales emiten fluorescencia a diferentes longitudes de onda cuando son excitados por láser. En los kits de amplificación los fluorocromos van ligados a uno de los primers (con lo cual se evitan los problemas derivados de la diferente movilidad de las dos cadenas de ADN), mientras que en kits de secuenciación pueden ir marcados o bien los primers o bien los dideoxinucleótidos. En el kit utilizado en esta tesis (AmpFISTR Profiler Plus, Perkin Elmer) los primers vienen marcados con los fluorocromos: *5-FAM, JOE, NED y ROX*.

El software de recogida de datos define ciertas áreas en una *cámara CCD* para recoger exclusivamente las emisiones procedentes del grupo de fluorocromos que estemos utilizando. Estas áreas se conocen con el nombre de filtros virtuales y su funcionamiento es similar al de los filtros físicos que se utilizan para separar las diferentes longitudes de onda. A pesar de que existe cierto solapamiento entre los diferentes rangos, este problema es minimizado gracias a la creación en el ordenador de una *matriz* matemática. Dicha matriz consta de cuatro filas y cuatro columnas y los números indican el grado de solapamiento entre las longitudes de onda emitidas por los cuatro fluorocromos. Los valores de la matriz varían según el analizador, el filtro virtual y las condiciones de electroforesis por lo que ésta debe ser creada para cada analizador y para unas determinadas condiciones electroforéticas.

El analizador ABI Prism 310 utiliza seis tipos de filtros virtuales denominados *A, B, C, D, E y F*. Los filtros A, C, D y F son usados en aplicaciones de GeneScan que es un software para análisis de fragmentos, mientras que el A, B, y E se usan para aplicaciones de secuenciación. En nuestro caso, para el análisis de fragmentos amplificados con el kit Profiler Plus el filtro utilizado fue el F. Cada filtro combina cuatro o menos fluorocromos (incluyendo el reservado para el estándar de tamaño) asignando un color a cada uno de ellos.

El *capilar* en el que se lleva a cabo la electroforesis se encuentra cubierto por poliamida opaca excepto por una pequeña zona denominada "ventana del capilar" la cual es atravesada por el láser y permite que las muestras sean excitadas a su paso por la misma.

Finalmente los datos analizados son enviados al ordenador el cual los transforma en secuencias de ADN o en fragmentos analizados con sus correspondientes alelos asignados.

DESCRIPCIÓN DEL ANALIZADOR

El instrumento de electroforesis capilar utilizado en esta tesis ha sido el analizador genético ABI Prism 310 de Perkin Elmer Applied Biosystems. Fundamentalmente el analizador genético consta de dos partes: una en la que se lleva a cabo todo el proceso de inyección de la muestra, electroforesis y detección de la fluorescencia y otra que es el sistema informático acoplado al analizador.

Una vez que las muestras amplificadas han sido preparadas y desnaturalizadas se colocan en un soporte con capacidad para 48 o 96 tubos y que encaja en la parte del analizador destinada a la carga de muestras (conocida como distribuidor de muestras o *autosampler,* que es su denominación inglesa). El autosampler se encarga de mover el soporte y, por lo tanto, los tubos de manera que el capilar pueda ser introducido en los mismos. Además dispone de:

- 2 viales de 4 ml: uno en el que se añade el buffer y otro con agua destilada para lavar el electrodo negativo y el capilar entre unos análisis y otros.

- Un tubo eppendorf de 1.5 ml que es el vial de deshecho, al cual irá la carga de polímero utilizada para cada muestra una vez terminado el análisis.

Encima del autosampler y junto al capilar a una distancia de unos 2 mm se encuentra el electrodo negativo que está conectado a una fuente de alto voltaje (15000 V) y proporciona la corriente necesaria para la electroforesis.

Existe otra zona en el aparato que es la encargada de bombear el polímero hacia el capilar. Esta zona consta de una jeringa que se rellena de polímero y un bloque de metacrilato con conductos internos y donde se encuentra además el electrodo positivo o ánodo al cual va acoplado el reservorio de buffer. A través de los

conductos internos los electrodos inmersos en el buffer quedan conectados durante la electroforesis, generándose así la diferencia de potencial necesaria para la separación de los fragmentos de ADN.

Entre las válvulas del bloque de metacrilato se encuentran:

- Válvula del buffer del ánodo: se mantiene cerrada mientras se realiza el relleno del capilar y se abre para contactar el polímero con el buffer durante la electroforesis.

- Válvula de deshecho: permite la salida de exceso de polímero.

La última zona importante en el analizador es la dedicada a la electroforesis y detección de fluorescencia la cual consta de:

- Una placa o plato que proporciona al capilar la temperatura adecuada para la electroforesis (60ºC en análisis de fragmentos y 50ºC en secuenciación)

- El capilar de silica relleno de polímero. Es el soporte de la electroforesis.

- Un láser de argón.

- Una cámara CCD.

- Una ventana detectora en la cual se hace coincidir la ventana del capilar.

Una vez que las muestras han sido analizadas los datos generados son recogidos por el *ordenador.* El ordenador acoplado al secuenciador ABI Prism 310 es un modelo de Macintosh que permite la instalación de diversos programas informáticos con diferentes funciones. El ABI Prism 310 Data Collection Software es el programa que recoge lo que se denomina el dato bruto o raw data y que es la fluorescencia detectada en la cámara CCD cuando los fragmentos marcados, separados por electroforesis y excitados por el láser pasan por la ventana de detección. Existen dos programas que analizan el dato bruto y lo convierten en secuencia de ADN o en fragmentos de un determinado tamaño:

- **ABI Prism®DNA Sequencing analysis**: para secuenciación
- **GeneScan® Analysis**: para análisis de fragmentos.

En este trabajo, puesto que se han analizado fragmentos de ADN, los programas utilizados han sido el ABI Prism® 310 Data Collection y el GeneScan® Analysis. Una vez analizados los fragmentos y con sus correspondientes tamaños (en pb) se hizo una asignación automática de los alelos mediante otro programa informático llamado Genotyper®2.0.

2.3.4.3. Futuras tecnologías: Biochips

Las técnicas de análisis genético se encuentran hoy en día en continuo desarrollo y evolución. La necesidad de técnicas que permitan el aislamiento y análisis de los casi cien mil genes que componen el genoma humano justifica la existencia de líneas de investigación destinadas al descubrimiento de nuevos métodos que permitan monitorizar elevados volúmenes de información genética en paralelo y que reduzcan tanto el tiempo empleado como el coste por análisis.

Desde el análisis de los primeros polimorfismos de ADN con fines identificativos, la Genética Forense ha sufrido una gran evolución. Los expertos en la materia han sido testigos de cómo el descubrimiento de la PCR revolucionó las técnicas de identificación genética. Es probable que la próxima revolución la constituyan los llamados biochips o microarrays y, por este motivo, hemos creído interesante dedicar un pequeño capítulo de esta tesis a lo que cada día se confirma más como la tecnología del futuro en Genética Forense.

Los **biochips** surgen como consecuencia de una combinación entre técnicas microelectrónicas empleadas para la fabricación de microprocesadores informáticos y materiales biológicos. En general puede decirse que la principal característica de los chips es su capacidad para generar información en muy poco espacio, ya que posibilitan el procesamiento de multitud de ensayos simultáneamente. Esta característica es la que hace que los biochips sean probablemente la tecnología del futuro en el campo de las investigaciones biomédicas.

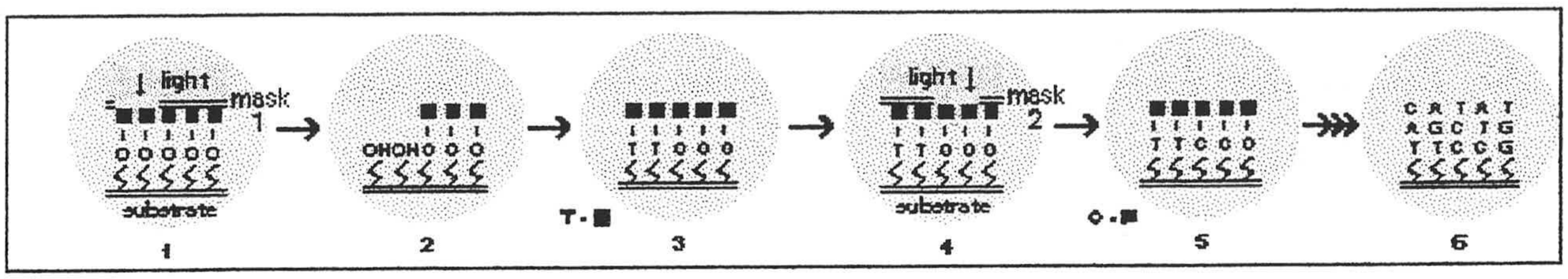
La fabricación de los biochips es similar a la de los chips informáticos: por medio de la técnica denominada **fotolitografía** se depositan circuitos microscópicos sobre láminas de silicio. En el caso concreto de los biochips, estas láminas son de vidrio y lo que se deposita en dichas láminas son cadenas de ADN. Las láminas de vidrio presentan una serie de ventajas como son (Cheung et al, 1999):

- Posibilidad de unir las cadenas de ADN a la superficie del cristal, convenientemente tratada, mediante enlaces covalente.
- Su capacidad para aguantar altas temperaturas y lavados de elevada fuerza iónica.
- Al ser un material no poroso el volumen de hibridación puede reducirse al mínimo.
- Su baja fluorescencia evita ruidos de fondo.

Hay compañías comerciales que han desarrollado otras estrategias para la fabricación de biochips. No obstante, los más usados actualmente y con mayor número de aplicaciones son los basados en *técnicas fotolitográficas*.

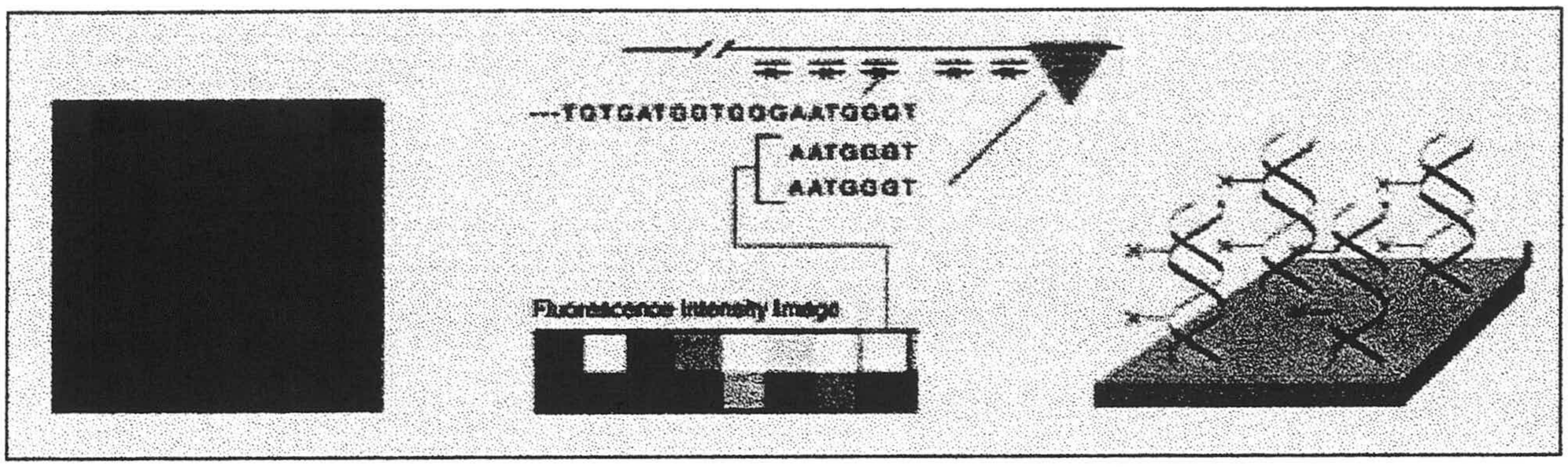
Estos chips, como se dijo anteriormente, consisten en una pequeña lámina de vidrio que posee unos grupos reactivos, a los que posteriormente se unirán los nucleótidos, protegidos mediante una película realizada con un agente químico fotodegradable. Mediante una máscara que se sitúan sobre el chip, se logra enfocar un haz de luz hacia unas posiciones y regiones determinadas, degradando al agente químico protector en dichas zonas y dejándolo intacto en las zonas protegidas por la máscara. A continuación se añade al chip un medio que contiene uno de los cuatro nucleótidos que se unirá, mediante enlace covalente, al cristal con los grupos reactivos que hayan quedado desprotegidos. Cada grupo añadido lleva una molécula receptora fotodegradable.

Todo este proceso se va repitiendo con los diferentes nucleótidos y máscaras hasta generar unos oligonucleótidos con las secuencias que nos interesen. El siguiente esquema muestra el todo el proceso explicado anteriormente.



Cada "casilla" del chip posee una cadena de un oligonucleótido de manera que solamente aquel fragmento de ADN que hibride perfectamente con ella permanecerá unido tras los diversos lavados.

Previamente a la hibridación, el ADN de la muestra a estudiar debe haber sido amplificado y marcado fluorescentemente en uno de sus extremos. Una vez marcado se incuba en el recipiente que contiene al chip tras lo cual se lava varias veces para eliminar los fragmentos que no hayan hibridado y se introduce el chip en un escáner en el que se detectan los patrones de hibridación. Esta detección se realiza en base a la fluorescencia emitida por los fluorocromos de la muestra cuando son excitados por luz y en aquellos pocillos en los que la unión haya sido completa la fluorescencia será mayor que los que contengan alguna base desapareada.



Un ordenador conectado al escáner es el que identifica las secuencias sonda por su posición el chip.

Tipos de Chips

En función de la técnica utilizada para la fabricación de los biochips, existen actualmente en el mercado cuatro tipos de chips de ADN que están siendo desarrollados por diferentes compañías comerciales:

- Genechips. Fueron los primeros en aparecer y están siendo desarrollados por la compañía Affymetrix. La técnica empleada es la fotolitografía (que ha sido explicada anteriormente). Por ahora, parece que éstos son los más utilizados y con mayores aplicaciones.

- Matrices electrónicamente activas. Están siendo desarrolladas por Nanogen y permiten que el usuario diseñe sus propios chips mediante un programa informático específico. También se denominan bioelectrochips debido a que el método utilizado

es electrónico. Este método posiciona el ADN y lo fuerza a hibridar, las hebras de ADN del chip se mantienen unidas aprovechando su carga eléctrica negativa. Al añadir la muestra se invierte la polaridad de la placa, de manera que sólo permanecerán unidas aquellas cadenas que hayan hibridado. Las posibles aplicaciones de estos chips son el diagnóstico de enfermedades infecciosas, cáncer, descubrimiento de fármacos y determinación de expresión génica.

- GEM arrays. Han sido producidos por Incyte-Synteni. Están destinados a la cuantificación de expresión génica. Para su formación se parte de una librería de ADNc del sistema que se quiera estudiar y se deposita sobre una superficie de cristal. El empleo de distintos fluorocromos posibilita por un lado, la diferenciación de expresión entre dos muestras de ARNm y por el otro, la cuantificación del nivel de expresión de cada gen en función de la intensidad de fluorescencia.(Duggan et al, 1999)

- Espectrochips. Se utilizan en secuenciación, empleando espectrometría de masas para obtener el peso molecular de los fragmentos. Han sido desarrollados por la compañía Sequenom.

Aplicaciones de los biochips

A pesar de ser una tecnología muy reciente y que, por lo tanto, está aún en vías de experimentación, actualmente los biochips están siendo aplicados en:

- **Monitorización de expresión génica**: permite determinar cual es el patrón de expresión génica y cuantificar el nivel de expresión de manera simultánea para un elevado número de genes. Esto permite realizar estudios comparativos de activación de determinados genes en tejidos sanos y enfermos y determinar así la función de los mismos (Duggan et al, 1999)

- **Detección de mutaciones y polimorfismos**: Permite el estudio de todos los posibles polimorfismos y la detección de mutaciones en genes complejos (Hacia, 1999).

- **Secuenciación**: Mientras que se han diseñando algunos biochips para secuenciación de fragmentos cortos de ADN (Pastinen et al. 1997), no existe aún en el mercado ningún biochip que permita secuenciar *de novo* secuencias largas de ADN.

- **Diagnóstico clínico y detección de microorganismos**: Posibilitan la identificación rápida empleando unos marcadores genéticos de los patógenos.

- **Screening y toxicología de fármacos**: el empleo de los biochips permite el analizar los cambios de expresión génica que se dan durante la administración de un fármaco de forma rápida, así como la localización de nuevas posibles dianas terapéuticas y los efectos toxicológicos asociados.

- **Seguimiento de terapia**: los biochips permiten valorar rasgos genéticos que pueden tener incidencia en la respuesta a una terapia.

- **Medicina preventiva**: El conocimiento y posible diagnóstico de ciertos caracteres genéticos asociados a determinadas patologías permite una prevención de las mismas antes de que aparezcan los síntomas.

UNIVERSIDAD DE GRANADA
3 1 MAR. 2000
COMISION DE DOCTORADO

2.4. GENETICA DE POBLACIONES

El objetivo fundamental de la Genética de Poblaciones es el análisis de las consecuencias de las leyes de la herencia mendeliana sobre la composición genética de las poblaciones y la influencia que ejercen procesos como la **mutación**, **selección**, **migración** y **deriva genética** sobre las frecuencias génicas de las mismas. El estudio se centra en el conjunto de individuos y no en el individuo aislado, ya que éste no puede ser considerado como unidad evolutiva al no sufrir cambios genéticos o evolutivos a lo largo de su existencia. En cambio, la constitución genética de una población puede cambiar y evolucionar de generación en generación.

La Genética de Poblaciones basa sus estudios en un modelo de población ideal que podría ser definida como un *grupo de individuos que se reproducen entre sí de forma aleatoria y que pueden presentar diferencias en las frecuencias génicas con respecto a otros grupos vecinos* (Moreno Suárez, 1999). No obstante, este concepto es un modelo teórico que rara vez se da en la Naturaleza ya que no todos los individuos de una población se casan con individuos del mismo grupo al ser los límites o fronteras entre poblaciones cada vez menos frecuentes. Estos límites son esenciales en el estudio de la Genética de Poblaciones puesto que definen unidades reproductoras de individuos. En general pueden dividirse en dos grupos: fronteras naturales y fronteras sociales.

Las *fronteras naturales* vienen determinadas por el propio ambiente físico, ya sean cadenas montañosas, océanos, islas, etc., e impiden la transmisión de genes de unas poblaciones a otras. Hoy en día, este tipo de fronteras apenas se dan debido a la facilidad y rapidez de las comunicaciones.

Las *fronteras sociales* condicionan la relación entre individuos debido a diferencias de cultura, clase social, religión, etc. pero al igual que las naturales no son impenetrables.

2.4.1. Ley de Hardy-Weinberg

Esta ley es el principio básico de la Genética de Poblaciones y fue postulada en 1908 de manera independiente por dos científicos: George Hardy, matemático inglés, y Wilhelm Weinberg, médico alemán. Según esta ley en toda población de gran tamaño en la que los apareamientos se dan al azar (población panmíctica) y en la que hay ausencia de fuerzas evolutivas como la mutación, selección y migración, las frecuencias génicas y genotípicas se mantienen constantes de generación en generación. Una población que cumpla dichas características se dice que está en equilibrio de Hardy-Weinberg y sus frecuencias genotípicas estarán determinadas por sus frecuencias génicas. En el caso de un locus con dos alelos codominantes cuyas frecuencias génicas son p y q respectivamente, la frecuencia de equilibrio para los tres genotipos se establece por la siguiente fórmula:

$$(p+q)^2 = p^2 + 2pq + q^2$$

La finalidad de los estudios poblacionales realizados por genetistas forenses o antropólogos es definir las poblaciones analizadas mediante las frecuencias génicas de determinados loci polimórficos y aplicar una serie de tests estadísticos destinados a la confirmación de equilibrio de Hardy-Weinberg para esa población y a la correlación de unos loci con otros. En este tipo de estudios es muy importante que el tamaño de la muestra sea grande (por encima de 100 individuos)para que los resultados sean fiables. Además, en el caso de estudios poblacionales con fines forenses es importante que la muestra sea tomada al azar y que entre los individuos analizados no exista ningún tipo de relación familiar ya que la consanguinidad puede alterar las frecuencias génicas.

2.4.2. Fuerzas evolutivas que pueden influir en el equilibrio Hardy-Weinberg

La asunción de equilibrio de Hardy-Weinberg en una población implica el hecho de que en esa población no existiría evolución, entendida ésta como un cambio de las frecuencias génicas. En esta estabilidad de las frecuencias génicas influye la ausencia de fuerzas evolutivas como son la mutación, selección, flujo y deriva genética (Moreno Suárez, 1999).

- **Mutación**. Una mutación puede definirse como un cambio en el material genético de un locus determinado. Cuando el locus que sufre la mutación codifica alguna proteína funcional del organismo el efecto de tal mutación se hace notable y es probable que tal mutación no pase a través del filtro que supone la selección natural (mutaciones deletéreas). En cambio, mutaciones en regiones o locus de ADN no codificante pasan inadvertidas de generación en generación (mutaciones neutras).

Los genes pueden mutar en cualquier tipo de célula, ya sea somática o sexual. En el primer caso el cambio producido no se transmitirá a la descendencia y, por lo tanto, desde el punto de vista evolutivo estas mutaciones no tienen importancia. En cambio, cuando la mutación se da en células sexuales (óvulos o espermatozoides) el gen alterado puede transmitirse a la descendencia y alterar el estado de equilibrio.

La frecuencia de aparición de una mutación se denomina frecuencia de mutación y se define como el número de nuevas mutaciones por locus y por generación. En la especie humana las tasas de mutación son muy bajas.

- **Selección Natural.** Una de las causas más importantes del cambio de las frecuencias génicas reside en la capacidad que tienen sus portadores para producir descendencia superviviente. La enorme variedad de mecanismos responsables de la modificación del éxito reproductivo de un genotipo se conoce como selección.

Estos mecanismos incluyen factores como cambios de temperatura, humedad, en la disponibilidad de alimento, diferencias en la atracción sexual, etc.

Cualquier variante alélica que aumente la probabilidad de que un individuo sobreviva y se reproduzca verá aumentada su frecuencia en las siguientes generaciones ya que dejarán más descendientes que los individuos que carezcan de ella. La fuerza que actúa sobre cada genotipo para reducir su valor adaptativo se denomina coeficiente de selección, **s**.

$$s = 1-\omega$$

Siendo ω la eficacia biológica, un parámetro que mide el mayor o menor éxito reproductor de un genotipo. Cuando en un mismo ambiente un genotipo puede dar lugar a más descendencia que otro, se dice que tiene una eficacia mayor. Por lo tanto, el coeficiente de selección mide la reducción de la eficacia.

- **Flujo Genético**. El flujo o migración de genes de una población a otra es otro de los factores que pueden contribuir en el cambio de las frecuencias génicas de una población. Este flujo de genes entre dos poblaciones puede darse en una sola dirección o en dos, y, en este segundo caso, dará lugar a poblaciones genéticamente idénticas. Si dos poblaciones genéticamente diferentes están aisladas geográficamente pero interconectadas por una serie de poblaciones intermedias es probable que se origine un gradiente geográfico de frecuencias génicas. Estos gradientes suelen conocerse con el nombre de geoclinas y tienden a desaparecer

cuando el intercambio es muy fuerte y, por lo tanto, las poblaciones intermedias son genéticamente idénticas.

El flujo genético ha tenido una gran influencia en el proceso de evolución humana ya que, por una lado, ha contribuido a la disminución de variabilidad entre poblaciones y, por el otro, ha aumentado la variabilidad intrapoblacional.

- **Deriva Genética**. En cada generación se produce un sorteo de genes durante la transmisión de gametos de padres a hijos que se conoce como deriva genética. Durante la formación de los gametos de organismos diploides, el que un gameto lleve un alelo u otro de un determinado gen depende del azar. Así, podemos decir que la formación de gametos y su consiguiente unión para formar los huevos de la siguiente generación es un proceso probabilístico. Es por este motivo por el que en cada generación esperamos una fluctuación al azar de las frecuencias alélicas en las poblaciones. Si en algún momento durante esta conducta fluctuante un tipo de alelos no llega a transmitirse a la siguiente generación, entonces este alelo se habrá perdido para siempre. El resultado de la deriva suele ser la pérdida de variabilidad genética, siendo un proceso que contrarresta la entrada de variabilidad genética por mutaciones.

2.4. PROBLEMAS ÉTICOS Y LEGALES RELACIONADOS CON EL ANÁLISIS DEL ADN

Actualmente no existe en España una legislación especial que regule el uso de las pruebas del ADN para la identificación genética en criminalística. Existe, sin embargo, un "Anteproyecto de Ley Reguladora de las Bases de Datos de ADN" (Boletín de Información: Actividad Prolegislativa del Ministerio de Justicia. Año LIII. Suplemento 1854, de 1 de Octubre de 1999; pp: 355-367). Este documento fue elaborado por una comisión de Trabajo del Ministerio de Justicia entre Septiembre de 1998 y Junio de 1999 y está pendiente de su estudio y posible tramitación por parte de las Cámaras Legislativas españolas.

Mientras tanto, todas las decisiones y actuaciones han de ser desarrolladas en base a normas vigentes y, siempre que sea posible, siguiendo la **Recomendación No. R (92) 1 del Consejo de Ministros del Consejo de Europa** sobre el *"Uso del análisis del ácido desoxirribonucleico (ADN) dentro del marco del sistema judicial penal"*.

Este texto europeo equivale a una recomendación de principios general sobre cómo deberían legislar los diferentes países a este respecto. Así, se vio la necesidad de creación de una serie de normas por parte de los laboratorios españoles que garantizasen la correcta aplicación de las técnicas del ADN a las pericias médico-legales. Estas normas fueron elaboradas en 1991 por el grupo español de la ISFH (International Society of Forensic Hemogenetic), de manera que informes realizados por laboratorios pertenecientes a dicho grupo poseen una garantía de calidad además de la posibilidad de futuros contraperitajes, y básicamente hacen referencia a cinco puntos (Lorente y Lorente, 1995):

1. Cual es el objetivo de dichas normas.

2. Qué requisitos debe cumplir el personal que realiza el análisis.

3. Cuáles deben ser los requisitos de los centros en los que se realiza el análisis.

4. Parámetros genéticos y estadísticos requeridos para el estudio biológico de la paternidad.

5. Normativa a seguir en el caso de análisis de vestigios biológicos de interés criminal.

2.4.1. Problemas ético y legales derivados de la investigación criminal

Básicamente, los problemas éticolegales que pueden presentarse en la investigación criminal por medio del análisis del ADN, están relacionados con dos puntos: la *negativa del consentimiento por parte del sospechoso* a donar una muestra con la que comparar con el resultado del análisis del indicio biológico hallado en el lugar de los hechos, y la *puesta en marcha de bancos de datos genéticos* que faciliten dicha investigación (Lorente et al, 1999).

2.4.1.1.Negativa del consentimiento para la toma de muestras del sospechoso

La ausencia de una legislación especial que regule el uso de las pruebas de ADN lleva a que, en ocasiones, no sea posible realizar dicha prueba debido a la negativa del inculpado a donar una muestra que sea comparada con el perfil genético de un indicio y que, por lo tanto pueda relacionarle con un hecho delictivo. En este caso deberían seguirse las normas o derechos recogidos en la Constitución española, pero encontramos una contradicción ya que por un lado, los artículos 15 y 18 proclaman el derecho a la integridad física y a la intimidad de las personas como derechos fundamentales, pero por otro lado, según el artículo 24 toda víctima de un delito tiene derecho a conocer al autor/res del mismo (Ruiz-Jimenez Aguilar, 1993).

Además, el problema del consentimiento para la realización de análisis de identificación genética en los peritajes médico-legales se contempla también desde la perspectiva médico-legal, haciendo referencia a la actuación médica dentro del ámbito de la investigación judicial.

El consentimiento informado es un requisito para la práctica de cualquier actuación médica. Las especiales circunstancias de la prueba en el campo del derecho penal implican por una lado, el supuesto de la falta de colaboración y por otro, la necesidad de informar de las consecuencias jurídicas que puedan derivarse de su realización. Así, la prueba médico-legal que se lleva a cabo dentro de una investigación judicial y que en muchas ocasiones es imprescindible para averiguar la identidad del autor de los hechos, conlleva necesariamente la existencia de otras pruebas o indicios que indiquen la relación del encausado con los hechos, pudiendo por tanto el Juez sustituir el consentimiento del encausado por medio de una resolución motivada para que el acto médico sea lícito (Lorente y Lorente, 1995).

Desde el punto de vista jurídico, los derechos de la Constitución española que pueden lesionarse al realizar una prueba de ADN sin consentimiento son:

1. Derecho a la libertad de movimientos (art. 17.1)
2. Derecho a la integridad física (art. 15)
3. Derecho a no declarar contra sí mismo (art. 17.3)
4. Derecho a no declararse culpable (art. 24.2)
5. Derecho a la presunción de inocencia (art. 24)

En relación al **derecho a la libertad de movimientos**, la Comisión Europea (Decreto 8278/78 de 13 de diciembre de 1979) afirma que "*la ejecución forzosa de exámenes de sangre a una persona constituye una privación de libertad, incluso en el caso de que dicha privación sea de corta duración*". El Tribunal Constitucional español puntualiza los controles de alcoholemia no suponen un atentado contra el derecho a la libertad de movimientos.

En el caso concreto de una investigación criminal siempre que existan indicios y elementos que puedan hacer sospechar de un determinado individuo podría establecerse una privación de libertad como fase previa a la recogida de la muestra.

Respecto al **derecho a la integridad física** de la persona, la Declaración Universal de los Derechos Humanos establece que nadie puede sufrir una lesión en contra de su voluntad por leve que sea. Sin embargo, para el análisis del ADN no es necesario partir de muestras que impliquen una lesión corporal, bastaría con otro tipo de muestras como saliva o pelos recogidos por cepillado cuya toma no requiere lesión alguna. Además, una sentencia del tribunal europeo de Derechos Humanos del 13 de diciembre de 1979 consideró que una prueba tan banal como un análisis de sangre no puede ser considerada como un atentado contra la integridad física.

Sobre los derechos a no declarar contra sí mismo, a no confesarse culpable y a la presunción de inocencia, la Comisión europea afirmó en el Decreto 8239/78, del 4 de diciembre de 1979, lo siguiente: *"la posibilidad ofrecida al inculpado de probar un elemento que le disculpa no equivale a establecer una presunción de culpabilidad contraria a la presunción de inocencia, puesto que, si puede parecer evidente que, siendo positivo el resultado de la prueba, puede derivarse una sentencia condenatoria, tampoco lo es menos que este mismo examen, si fuere negativo, puede exculpar al imputado".*

2.4.1.2. Bancos de datos genéticos

Las bases de datos que serán de uso policial y judicial son realmente *bases de datos de identificación genética*, queriendo significar con ello que los datos que se contienen son simplemente números y letras que identifican a las personas a sus características del ADN no codificante o del ADN mitocondrial, por comparación con la secuencia de Anderson.

Cabría por lo tanto, clasificar las bases de datos de acuerdo con su contenido en:

1) Bases de datos de identificación genética: corresponden a las que acabamos de describir, y contienen números y letras asociados al código de identificación de una persona. Aparecen en los ejemplos de las tablas siguientes (ADN nuclear en la tabla 1.2. y ADN mitocondrial en la tabla 1.1.). Son bases virtuales, o sea, datos que existen simplemente en archivos informáticos o impresos en papeles. El acceso a los datos puede (debe) estar perfectamente controlado, y las conclusiones que se pueden obtener de los mismos va a depender de los programas informáticos que se autoricen por las diferentes legislaciones.

Tabla 1.1. Ejemplo de cómo quedan registrados los datos de ADN mitocondrial.

ISFH 1/00	16024	16365	73	340	16093C	16224C	16311C	73G	152C	263G	315.1C
ISFH 2/00	16024	16365	73	340	16093C	16224C	16311C	73G	152C	263G	315.1C
ISFH 3/00	16024	16365	73	340	16298C	195C	228A	263G	309.1C	315.1C	
ISFH 4/00	16024	16365	73	340	16298C	195C	228A	263G	309.1C	315.1C	

Tabla 1.2. Ejemplo de cómo quedan registrados los datos de ADN nuclear.

Sample Info	D3S1358		vWA		FGA		AMEL		D8S1179		D21S11	
K-	0	0	0	0	0	0	0	0	0	0	0	0
K+	14	15	17	18	23	24	X	X	13	13	30	30
LADDER	12	13	11	12	18	19	X	Y	8	9	24.2	25
M01/99 E.S	15	15	16	17	22	25	X	X	13	13	29	34.2
H01/99 E.S	15	18	16	17	22	25	X	X	13	15	28	34.2
P01/99 E.S	17	18	17	18	22	26	X	Y	14	15	28	29
M02/99 E.S	14	15	17	19	20	23	X	X	13	14	29	31.2
H02/99 E.S	14	15	17	19	23	24	X	Y	13	14	31.2	32.2

Aunque no es objeto de la presente tesis, cabe hacer la reflexión siguiente: algunos de los usos de los datos de ADN no codificante, que entendemos como no perteneciente a la intimidad de la persona, podrían pasar a expresar datos íntimos si se utilizan erróneamente. Por ejemplo, buscar directamente una relación entre un indicio y un sospechoso no plantea mayores problemas éticos ni legales, pero deducir la implicación de una persona en base a muestras biológicas donadas voluntariamente por familiares suyos, o deducir paternidades u otras relaciones de filiación biológica, o establecer correlaciones entre ciertos alelos o genotipos no codificantes con patologías específicas con los que pudieren estar correlacionados, puede tener otra calificación totalmente diferente, y en su momento habrá que plantearse esta situación.

2) Archivos de ADN: no son bases de datos como tales, en tanto en cuanto no acumulan "datos" a los que se puede acceder de modo automatizado, sino que contienen muestras de ADN ya extraído del núcleo celular y de las mitocondrias, dispuesto para ser analizado (en sus regiones codificantes y no codificantes) con la técnica o técnicas adecuadas. Suele confundirse, por parte de personas inexpertas, pero también lamentablemente por algunos profesionales, este tipo de archivos con las bases de datos, y se advierte de la especialísima y restrictiva regulación que las bases de datos necesitan porque, suele argumentarse, podrían obtenerse datos críticos de las mismas.

Este es un hecho cierto e irrefutable, pero que no se deriva de la posibilidad de que existan bases de datos, sino que es consecuencia directa del análisis de miles de muestras por los laboratorios especializados (forenses, hospitales, sanatorios, etc.). Obviamente, se dispone, en nuestra opinión, de instrumentos punitivos suficientes en el Código Penal y en otras disposiciones como para plantear problemas específicos derivados del almacenamiento o acúmulo de material genético.

Finalmente, y por lo que pudiere referirse a los laboratorios forenses, querríamos destacar que muchas veces la cantidad y calidad de las muestras de ADN que se almacenan son tan mínimas, que no permitirían los análisis de ciertos fragmentos de ADN, especialmente de los grandes fragmentos (kilobases) que habitualmente conforman los exones.

3) Archivos de muestras biológicas: tampoco son "bases de datos" en si, ya que representan una especia de "escalafón inferior" en comparación con los archivos de ADN. Cualquier laboratorio forense o de análisis clínicos, hospital, sanatorio, etc., almacena muestras biológicas (sangre, orina, biopsias) de modo más o menos controlado. A las mismas se puede o podría acceder y proceder a obtener ADN para realizar posteriores análisis de uno u otro tipo.

A este apartado aplicaríamos exactamente los mismos comentarios que al anterior, ya que cumple las mismas características. Llegados al absurdo, muestras biológicas de las más diversas personas quedan depositadas todos los días en lugares públicos (bares, restaurantes, peluquerías), y si se hiciese una legislación restrictiva al efecto, no debería de descartarse la posibilidad de incluir esas posibles fuentes de muestras entre los lugares específicamente considerados, otorgando a los profesionales que trabajan en los mismos un status especial, con sus obligaciones pertinentes.

Creemos, por lo tanto, que extender el concepto de "base de datos" de ADN, más allá de lo que realmente es no sólo no dificultaría la puesta en marcha de las mismas, sino que conllevaría la aparición de numerosos problemas científicos, éticos y jurídicos que a nada ni nadie beneficiarían.

Los bancos o bases de datos genéticos, de modo general y con las salvedades que se acaban de considerar, hacen pues referencia al archivo sistemático de material genético o muestras biológicas de determinados grupos de población para ser analizadas en determinadas circunstancias. Según el grupo de personas que abarque pueden clasificarse en:

1. **Generales**
2. **Profesiones de riesgo**
3. **De interés judicial:**
   a) Personas desaparecidas
   b) Criminales:
      - Convictos
      - Sospechosos
      - Víctimas
      - Indicios dubitados obtenidos en el lugar de los hechos.

Los *bancos genéticos generales* que incluyan a toda la población no han sido planteados aún debido a la serie de problemas de tipo técnico, económico, social y científico que supone su creación. Solamente existe un antecedente en Islandia, donde el Parlamento estudia la posibilidad de realizar una base de datos genéticos de interés médico en la que podrían incluirse datos referidos a la identificación. También en Estonia se ha planteado recientemente esta posibilidad.

En cuanto a los *bancos genéticos para profesiones de riesgo* ya se están elaborando en algunas profesiones donde los interesados donan voluntariamente una muestra de saliva o sangre para que sea analizada en caso de accidente. El mayor archivo de estas características se está elaborando en el Armed Forces Institute of Pathology (AFIP) de las Fuerzas Armadas de Estados Unidos. También algunas compañías aéreas norteamericanas y viajeros frecuentes han donado muestras para este fin. Estas son bases de datos de muestras biológicas, no de ADN extraído, y aún menos de datos obtenidos del ADN.

Dentro de los bancos genéticos de interés judicial son los de tipo criminal los que, en España, plantean problemas tanto desde el punto de vista legal como ético.

Entre los países que poseen bases de datos genéticas de utilización rutinaria en criminalística destacamos a EE.UU., Gran Bretaña, Holanda, Alemania y Austria.

Sin duda alguna la mayor de ellas es la de Gran Bretaña, pero algunos estados de EE.UU. como Virginia y Florida, poseen ya (a través del sistema CODIS), cientos de miles de datos de personas envueltas en casos criminales.

España ha sido pionera en la creación de una base de datos genética de tipo civil de personas desaparecidas, con fines exclusivamente de identificación civil y que cuenta con la colaboración voluntaria de las personas (familiares) de las que provienen las muestras de referencia, lo que lo diferencia de las bases de datos criminales, donde se podría obligar, contra la voluntad, a una persona a ceder una muestra biológica.

Dicho proyecto ha sido bautizado con el nombre de "Proyecto Fénix" y surgió como consecuencia de un convenio de colaboración entre la Guardia Civil, la Universidad de Granada y el Ministerio de Interior. Los objetivos primordiales del Proyecto Fénix son la creación de dos bases de datos. Una base de datos indubitados *o de referencia*, en la que se incluyan los perfiles genéticos del ADN mitocondrial de los familiares de los desaparecidos; otra, *dubitada,* con los perfiles del ADN mitocondrial de los restos óseos encontrados en cualquier punto de España y que no hayan sido identificados (en España se estima que existen casi 2.000 restos óseos no identificados). Una vez que se tenga la base de datos de los familiares se procede a enfrentar, a través de un programa informático cedido por el F.B.I. (denominado MitoSearch), el perfil genético de las muestras no identificadas con todos los perfiles genéticos de los familiares de la base de datos para buscar posibles compatibilidades.

En las fechas en las que se está escribiendo esta Tesis (marzo de 2000), se ha conseguido ya 2 identificaciones plenas (corroboradas con otros datos de la investigación policial), y se cuenta con solicitudes de más de 100 familias y con contactos de más de 200 casos, repartidos por toda la geografía española, en los que existen restos óseos sin identificar.

Se estima que para el mes de Diciembre del año 2.001 se habrá completado la primera fase del Proyecto Fénix, consistente en tener analizados todos los restos óseos (humanos en general) a los que se les dio sepultura sin identificar y que aún permanecen anónimos, y todas las muestras de las familias que las hayan donado voluntaria y desinteresadamente.

A partir del año 2.002 se pretende que todas los nuevos casos que vayan apareciendo vayan engrosando la base de datos de modo automático y de la manera más rápida posible. Este Proyecto está sirviendo de referencia para otros países, como Estados Unidos (FBI: Missing Child Genetic Identification Program), Colombia (Fiscalía del Estado) y Chile (Servicio Médico-Legal del Ministerio de Justicia).

3. MATERIAL Y MÉTODOS

3.1. MUESTRAS

3.1.1. Muestras utilizadas para el estudio poblacional

Se han analizado un total de 177 individuos no emparentados de ambos sexos, escogidos aleatoriamente entre aquellos residentes en el sur de España (Andalucía). A cada uno de ellos se le extrajo por venipunción entre 5 y 10 mililitros de sangre periférica que fue depositada en tubos estériles con anticoagulante EDTA-K_3 . Los tubos fueron marcados con números por lo que se desconoce la identidad de cada uno de los donantes. Una vez recibidas las muestras se guardó 1 mililitro de cada una en forma de mancha de sangre sobre un soporte de algodón de 2 x 2 cm y se conservaron en sobres individuales correctamente identificados a 4ºC. Para la extracción de ADN se tomaron unos 200 µl de la muestra original.

3.1.2. Muestras utilizadas para el estudio de validación

Con el objeto de determinar la aplicabilidad forense del triplex SilverSTR™ III, se realizó un estudio de validación con muestras forenses de las analizadas en nuestro laboratorio durante los años 1996, 1997 y 1998, así como controles certificados (K562, High Molecular Weight, Promega y AmpFlSTR Control DNA 9947A, Perkin Elmer). Dicho estudio constó de los siguientes experimentos:

A) Estudio de sensibilidad
B) Estudio de mezclas de ADN
C) Estudio en indicios biológicos de interés forense

A) Estudio de sensibilidad

El objetivo de este estudio fue determinar la mínima cantidad de ADN que se logra amplificar para los tres marcadores siguiendo las condiciones estándar que recomendadas por el fabricante. Para ello se utilizó por un lado, ADN de la línea celular denominada K562 a una concentración de 10 ng/µl y, por otro, ADN de la línea 9947-A a una concentración de 0.1 ng/µl. En el primer caso, se hicieron diluciones seriadas de manera que al tomar 5 µl se añadieran las siguientes cantidades de ADN: 5ng, 2.5ng, 2ng, 1ng, 0.5ng, 0.4ng, 0.3ng, 0.25ng y 0.15ng. Mientras que del otro control, se añadieron 1 ng, 0.5ng, 0.4ng, 0.3ng, 0.2ng y 0.1ng de ADN.

B) Estudio de mezclas

Se hicieron mezclas artificiales utilizando 3 muestras de ADN extraído (A, B y C) perfectamente cuantificado. La muestra A se correspondía con ADN procedente de la línea celular K562 a una concentración de 10 ng/µl, la muestra B se obtuvo a partir de sangre fresca y el ADN extraído se llevó hasta una concentración de 10 ng/µl, y la muestra C correspondía al ADN control denominado 9947-A comercializado a una concentración de 0.1 ng/µl. Para el estudio se hicieron dos bloques de mezclas utilizando por un lado las muestras A y B y por el otro, la A y la C. Los genotipos de las muestras del primer bloque fueron los siguientes:

Muestra	D16S539	D7S820	D13S317
A	11,12	9,11	8,8
B	12,13	10,10	9,12

Para el segundo estudio los genotipos fueron:

Muestra	D16S539	D7S820	D13S317
A	11,12	9,11	8,8
C	11,12	10,11	11,11

Para garantizar al máximo la proporción de los diversos tipos de ADN, las mezclas se hicieron una vez las muestras fueron extraídas y cuantificadas por Slot-blot, como se dijo anteriormente. El estudio se llevó a cabo tomando cantidades constantes de la muestra A frente a cantidades decrecientes de las otras muestras (la B o la C) y amplificando cada una de las mezclas. Así, se llevó a cabo un estudio en paralelo de mezclas con las muestras **A y B** por un lado, y **A y C** por el otro. Las cantidades de cada muestra así como las proporciones en las que se hallaban son mostradas de manera esquemática en las siguientes tablas:

Tabla 3.1. Mezclas A + B

ADN Muestra A (ng)	ADN Muestra B (ng)	Proporciones A : B
5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:20, 1:10, 1:4, 1:2, 1:1, 1:0.5, 1:0.2, 1:0.1, 1:0.05, 1:0.02
2.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:40, 1:20, 1:8, 1:4, 1:2, 1:1, 1:0.4, 1:0.2, 1:0.1, 1:0.04
1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:100, 1:50, 1:20, 1:10, 1:5, 1:2.5, 1:1, 1:0.5, 1:0.25, 1:0.1
0.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:200, 1:100, 1:40, 1:20, 1:10, 1:5, 1:2, 1:1, 1:0.5, 1:0.2
0.25	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:400, 1:200, 1:80, 1:40, 1:20, 1:10, 1:4, 1:2, 1:1, 1:0.4
0.1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:1000, 1:500, 1:200, 1:100, 1:50, 1:25, 1:10, 1:5, 1:2.5, 1:1

Tabla 3.2. Mezclas A + C

ADN Muestra A (ng)	ADN Muestra C (ng)	Proporciones A : C
1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:100, 1:50, 1:20, 1:10, 1:5, 1:2.5, 1:1, 1:0.5, 1:0.25, 1:0.1
0.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:200, 1:100, 1:40, 1:20, 1:10, 1:5, 1:2, 1:1, 1:0.5, 1:0.2
0.25	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:400, 1:200, 1:80, 1:40, 1:20, 1:10, 1:4, 1:2, 1:1, 1:0.4
0.1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:1000, 1:500, 1:200, 1:100, 1:50, 1:25, 1:10, 1:5, 1:2.5, 1:1

D) Estudio de muestras forenses

El número total de muestras utilizadas para esta parte del estudio de validación fue de 45. Para la mayoría de las muestras, el protocolo de extracción utilizado fue el de extracción orgánica con fenol/cloroformo, y los extractos, que habían permanecido congelados a –20ºC hasta el mes de febrero de 1999, fueron descongelados para llevar a cabo esta fase del estudio de validación. Una vez las muestras fueron descongeladas se procedió a su cuantificación en geles de agarosa al 0.8 % (Sambrook et al, 1989) para comprobar la cantidad y calidad de las mismas (si existía degradación o no).

A continuación se muestra una tabla en la que se especifica el tipo de muestra utilizada, el número de muestras de cada tipo y su porcentaje con respecto al total.

Tabla 3.3.

Tipo de muestra	Número de muestras	Porcentaje (%)
Manchas de sangre	17	39,53
Manchas de semen	17	39,53
Caspa	7	16,28
Saliva	1	2,32
Hueso	1	2,32
Total	43	100

3.2. EXTRACCIÓN DEL ADN

3.2.1. Extracción orgánica mediante fenol/cloroformo-isoamilalcohol

3.2.1.1. Extracción de ADN a partir de sangre total o de manchas de sangre

Protocolo:

- Se añade a la muestra (sangre total o si la sangre se depositó en un soporte de algodón, se corta con un bisturí estéril un trozo de aproximadamente 2 x 2 cm): 300 µl de buffer de extracción (10 mM Tris, 100 mM NaCl, 39 mM DTT, 10 mM EDTA, 2%SDS), 7.5 µl de proteinasa K (10 mg/ml) y 12 µl de DTT (1M). Agitar a baja velocidad durante un segundo y centrifugar en una nanofuga durante 2 segundos. A continuación incubar a 56ºC toda la noche (18-24 horas).

- Centrifugar en nanofuga 2 segundos y eliminar el trozo de tela. En una campaña extractora, añadir 300 µl de fenol/cloroformo/isoamilalcohol 24:25:1 (saturado con 10 mM Tris, pH 8.0, 1 mM EDTA, Sigma) agitar hasta conseguir una emulsión lechosa y centrifugar en microcentrífuga 3 minutos a máximas revoluciones (13.000 rpm).

- Una vez separada la fase fenólica y clorofórmica se recupera la fase acuosa o fenólica que es la que contiene el ADN y se transfiere a un tubo Microcon-100 (Microcon YM-100, Millipore) al cual previamente se le ha añadido 100 µl de TE^{-4} buffer.

- Centrifugar a 500 x g durante 10 minutos.

- Eliminar el filtrado y añadir 200 µl de TE^{-4} buffer para limpiar la membrana.

- Centrifugar a 500 x g durante 10 minutos.

- Añadir el volumen de recuperación deseado (entre 40 y 200 µl) y en un tubo limpio invertir la membrana y centrifugar a 500 x g 5 minutos. Descartar la membrana y cerrar el tubo.

3.2.1.2. Extracción de ADN a partir de manchas de semen o hisopos vaginales

En los laboratorios de Criminalística es frecuente recibir hisopos vaginales tomados a mujeres que han sido víctimas de una agresión sexual. En este tipo de muestras suele encontrarse material biológico de al menos dos individuos diferentes, por un lado, espermatozoides y por el otro, células vaginales. Es necesaria por lo tanto una separación de ambos tipos celulares antes de proceder a la extracción de ADN, con lo que, en el caso ideal, se obtienen dos partes o fracciones:

1. Fracción Femenina ("***Female Fraction***" o FF) procedente del ADN de las células epiteliales.
2. Fracción Masculina ("***Male Fraction***" o MF) procedente del ADN de las células espermáticas.

Es sabido que las células vaginales ofrecen menor resistencia a la lisis con proteinasa K que las células espermáticas, las cuales son más resistentes gracias al gran número de proteínas tiólicas que poseen las membranas de las cabezas espermáticas. Por ello el método utilizado para la extracción del ADN se basa en dicha característica. El protocolo que se siguió para la lisis diferencial fue el siguiente:

- Añadir a la muestra (trozo de tela de 2x2 cm o, si es un hisopo se corta la cabeza de algodón) 400 µl de buffer TEN (10mM Tris-ClH, 100mM ClNa, 1mM EDTA, pH=8), 25 µl de sarcosil al 20%, 75 µl de agua estéril y 5 µl de proteinasa K (10 mg/ml). Agitar 1 segundo y centrifugar 2 segundos en nanofuga.

- Incubar a 37ºC durante 2 horas.

- Eliminar el trozo de tela o hisopo utilizando tubos con canasta ("basket tube") y centrifugando a máximas revoluciones durante 5 minutos. Tirar la "canasta" y con mucho cuidado de no tocar el sedimento, tomar el sobrenadante y transferirlo a un tubo Eppendorf nuevo. Este sobrenadante contiene las células vaginales y, por lo tanto, la FF. Esta fase se guarda mientras llevamos a cabo la lisis de las células espermáticas o fracción MF que se encuentran en el sedimento.

- Lavar la MF añadiendo 500 µl de buffer de lavado (10 mM Tris-HCl, 10mM EDTA, 50 mM NaCl, SDS 2%, pH=7.5), agitar brevemente y centrifugar a máximas revoluciones durante 5 min.

- Descartar el sobrenadante y repetir el lavado 2 veces.

- Añadir al sedimento lavado 150 µl de buffer TEN, 50 µl de sarcosil al 20%, 40 µl DTT 0,39M, 150 µl de agua estéril y 10 µl de proteinasa K (10 mg/ml).

- Agitar 1 segundo y centrifugar 2 segundos.

- Incubar a 37ºC durante 2 horas.

- Añadir 400 µl de fenol/cloroformo/isoamilalcohol (24:25:1) y agitar hasta conseguir una emulsión lechosa (unir a este paso de la extracción el tubo que se había guardado con la fracción femenina y seguir con los dos tubos).

- Centrifugar 3 minutos a máximas revoluciones.

- Transferir la fase acuosa a un tubo Microcon-100 al cual previamente se le ha añadido 100 µl de TE^{-4} buffer.

- Centrifugar a 500 x g durante 10 minutos.

- Eliminar el filtrado y añadir 200 μl de TE^{-4} buffer para limpiar la membrana.

- Centrifugar a 500 x g durante 10 minutos.

- Añadir el volumen de recuperación deseado (entre 40 y 200 μl) y en un tubo limpio invertir la membrana y centrifugar a 500 x g 5 minutos. Descartar la membrana y cerrar el tubo.

3.2.1.3. Extracción de ADN a partir de saliva depositada hisopos de algodón y en colillas

El proceso de extracción en sí es exactamente el mismo que el utilizado para la extracción de manchas de sangre lo único que varía es la toma de la muestra. En el caso de los hisopos se corta con un escalpelo estéril la mitad o la cabeza completa de algodón mientras que en el caso de las colillas de cigarros se cortan unos 5 mm del papel que envuelve al filtro del cigarrillo y que ha estado en contacto con la boca. A partir de ahí se sigue el proceso descrito en el apartado 3.2.1.1.

3.2.1.4. Extracción de ADN a partir de hueso

Ante la identificación de restos cadavéricos o en avanzado estado de putrefacción, la extracción de sangre o músculo no es posible y, por lo tanto, se recurre a la extracción de ADN a partir de tejidos duros o resistentes en cuyo interior la agresión celular por parte de microorganismos es menor. El tejido óseo compacto es el más utilizado en estos casos y, en concreto, es preferente utilizar un hueso compacto largo como es el fémur.

En primer lugar se corta el segmento del hueso compacto en el que el diámetro sea mayor para que la probabilidad de que las células hayan sido atacadas por microorganismos y sustancias contaminantes sea menor. Se limpia tanto la superficie del hueso como la parte interna o medular, entonces, en una campana extractora se procede a la extracción del pulverizado óseo. Existen diversos métodos para el pulverizado del hueso, en nuestro laboratorio usamos una lima nueva y papel de aluminio totalmente estériles (esto se consigue tratando el material con luz ultravioleta durante una hora). Es necesario obtener unos 3 gr de pulverizado que se recogen en tres tubos Falcon (1 gr en cada tubo) de 15 ml.

En otra campana se comienza con el proceso de extracción del ADN para lo cual se añade a cada tubo Falcon:

- 1.5 ml de EDTA 0.5M a pH 7-8
- 75 µl de proteinasa K (10mg/ml)
- 105 µl de SDS (10mg/ml)
- 60 µl de DTT

Es imprescindible utilizar un *control negativo* de extracción para posteriormente poder descartar una posible contaminación durante el proceso de extracción.

Una vez añadidos los reactivos se agitan los tubos y se colocan en un baño a 56ºC durante 4 o 5 horas tras lo cual se vuelve a añadir a cada tubo 60-75 μl de proteinasa K y se incuban toda la noche.

Al siguiente día se centrifugan los tubos a 3.000 r.p.m. durante 10 in y se recoge la fracción clara y se guarda para ser purificada. Lavar el precipitado con 1 ml de EDTA, agitar y volver a centrifugar. Recoger de nuevo la fracción clara y añadir 200 μl de fenol/cloroformo-isoamilalcohol, agitar y centrifugar a 14000 rpm durante 4 min. Se recoge el sobrenadante y se purifica con Centricon-100 (Amicon), en un tubo se añaden todas las primeras fracciones claras recogidas, en otro las segundas y en un tercero las obtenidas del control negativo. El ADN obtenido es entonces cuantificado.

3.2.2. Extracción de ADN con Chelex

El "Chelex 100" es una resina quelante captadora de iones polivalentes que permite extraer el ADN a partir de diferentes sustratos de manera rápida y sencilla (Walsh et al, 1991). Además presenta la ventaja de poder extraer ADN a partir de muestras pequeñas puesto que al ser el número de pasos reducido la probabilidad de pérdida de ADN por manipulación es extremadamente baja. Sin embargo presenta el inconveniente de que el ADN obtenido está menos purificado que el obtenido en la extracción orgánica con fenol-cloroformo. No obstante es un método muy efectivo para extraer ADN a partir de muestras de sangre total que se encuentra en buen estado.

METODO:

- El primer paso consiste en un choque hipotónico que hace que las membranas citoplasmáticas de los eritrocitos y linfocitos se lisen para ello se toman unos 10 µl de sangre total y se añade 1 ml de agua estéril o buffer TE^{-4}, se agita durante 2 ó 3 segundos y se deja en incubación media hora a temperatura ambiente.

- Agitar de nuevo y centrifugar a máximas revoluciones durante 3 minutos tras lo cual se formará un sedimento constituido mayoritariamente por los núcleos de los linfocitos.

- Añadir 200 µl de Chelex al 5% (Chelex –100 Resin, BioRad) asegurándonos de tomar bolitas de resina. Agitar e incubar a 56ºC media hora.

- Sacar las muestras del baño, agitar e incubar a 100ºC durante 8 minutos.

- Volver a agitar y centrifugar a máximas revoluciones 3 minutos.

- Guardar los tubos en el frigorífico a 4ºC.

3.3. CUANTIFICACIÓN DEL ADN

La ponderación de la cantidad de ADN fue realizada con el método semicuantitativo del SLOT-BLOT que utiliza una sonda marcada que permite su revelado por un método quimioluminescente. Esta sonda es complementaria al locus D17Z1 el cual se localiza en el centrómero del cromosoma 17 y consiste en una secuencia compuesta por 500 a 1.000 copias de un fragmento de 2,7 Kb el cual a su vez consta de unidades de repetición de 171 pb (Waye, 1986). Este sistema presenta la ventaja de que es altamente sensible para pequeñas cantidades de ADN e incluso para material degradado.

El kit utilizado ha sido el "**ACES™ 2.0⁺ Human DNA Quantitation System**" (Gibco BRL).

MÉTODO:

La cuantificación se realiza por comparación con diluciones de ADN estándar (ADN de la línea celular K562) de concentración conocida. Los patrones que se utilizan tienen las siguientes concentraciones (en ng/µl): 4, 2, 1, 0.4, 0.2, 0.1, 0.04, 0.02, 0.01 y 0.004.

- El primer paso consiste en desnaturalizar los patrones y las muestras para lo cual se mezclan en tubos 10 µl de cada uno de ellos con 100 µl de solución desnaturalizante (0.5 M NaOH, 0.5 M NaCl) y se incuban 5 minutos a temperatura ambiente. Mientras se sumerge la membrana de nylon en 2xSSC durante 5 min. tras lo cual se coloca en el aparato donde se va a llevar a cabo la transferencia del ADN a la membrana ("The Convertible® Filtration Manifold System") y que va conectado a una bomba de vacío.

- Cargar las muestras desnaturalizadas en el aparato y hacer el vacío (10-12 mm Hg) durante 5 min. Una vez filtradas las muestras añadir en cada pocillo 400 µl de solución desnaturalizante y volver a hacer el vacío hasta que toda la solución haya sido absorbida.

- Sacar la membrana del "Manifold" y neutralizarla con una solución 0.2M Tris (pH=7.5), 2xSSC durante 5 min. A continuación se fija el ADN a la membrana exponiendo la misma a luz ultravioleta (254 nm) durante 90 seg. Una vez fijado el ADN a la membrana ésta puede ser conservada a 4ºC hasta el momento en que vaya a procederse a la hibridación.

- Para la hibridación del ADN fijado a la membrana con la sonda el primer paso consiste en una prehibridación con 0.1 ml/cm^2 de solución de hibridación (ACES 2.0 Hybrydization Solution) 50ºC durante 20 min. en agitación. La hibridación en sí se realiza añadiendo 0.05 ml/ cm^2 de solución de hibridación y 0.5 µl/ cm^2 de sonda (ACES 2.0 Human DNA Quantitation Probe Plus) a 50ºC durante 20 min. en agitación. Tras la hibridación la membrana se lava dos veces con 1ml/ cm^2 con una dilución 1:5 de buffer de lavado concentrado (ACES 2.0 Wash Buffer I Concentrate) durante 10 min. cada vez, a 50ºC y en agitación.

- Se vuelve a lavar la membrana dos veces pero esta vez con 0.5ml/ cm^2 de buffer de lavado final 1X (ACES 2.0 10X Final Wash Buffer) a temperatura ambiente y durante 5 min.

- El paso final consiste en la detección del ADN hibridado utilizando una sustancia quimioluminiscente que se une al ADN y que se denomina Lumi-Phos^{tm}530. Antes de añadir el Lumi-phos debe quitarse el exceso de líquido pero sin dejar que la membrana llegue a secarse. A continuación se añaden 5 ml de Lumi-phos y se dejan 5 min. en agitación para permitir que toda la membrana esté cubierta. Pasados los 5

min. la membrana se coloca entre dos láminas de plástico, se sella y se expone dentro de un cassette a una placa radiográfica durante un tiempo comprendido entre media y una hora.

3.4. AMPLIFICACIÓN

Hemos analizado un total de diez loci tipo STR y un locus indicador de sexo (amelogenina) incluidos en dos kit tipo multiplex: uno de ellos fue el kit **SilverSTR™ III Triplex** de Promega cuya detección se hizo con nitrato de plata, y el otro el kit **AmpFISTR Profiler Plus™** de Perkin Elmer Applied Biosystem preparado para ser detectado por electroforesis capilar con emisión de fluorescencia. La amplificación de todos los loci se realizó en un termociclador de la casa Perkin- Elmer (Thermo Cycler 9600). A continuación se muestra una tabla con los loci incluidos en cada kit y algunas características de cada uno de ellos:

Tabla 3.4.

Locus	Localización	Rango de tamaños	Bibliografia
D16S539	16q24-qter	264-304	Schmutzler et al, 1995
D3S1358	3p	114-142	Li et al, 1993
VWA	12p12-pter	157-197	Kimpton et al, 1992
FGA	4q28	219-267	Mills et al, 1992
D8S1179	8	128-168	Oldroyd et al, 1995
D21S11	21	189-243	Sharma and Litt, 1992
D18S51	18q21.3	273-341	Urquhart et al, 1995
D5S818	5q21-31	135-171	Hudson et al, 1995
D13S317	13q22-31	206-234	Hudson et al, 1995
D7S820	7q11.21-22	258-294	Green et al, 1991
Amelogenina	X:p22.22.3 Y:p11.2	107 113	Sullivan et al, 1993

3.4.1. Kit SilverSTR™ III

La amplificación se llevó a cabo siguiendo las recomendaciones del manual (Technical Manual Geneprint™ STR System- Silver Stain Detection. Part # TMD004. Promega Corp., Madison, WI, EEUU).

Mezcla de reacción:

STR 10X Buffer	2.5 µl
Multiplex 10X	
Primer Pair Mix	2.5 µl
Taq polimerasa(5u/µl)	0.2 µl
ADN	.X µl
H_2O	...X µl
Volumen final	.25 µl

Condiciones de amplificación:

96ºC	1'	
94ºC	30"	10 ciclos
60ºC	30"	
70ºC	45"	
90ºC	30"	20 ciclos
60ºC	30"	
70ºC	45"	
60ºC	30'	

3.4.2. Kit AmpFISTR Profiler Plus

Este kit permite el análisis simultáneo de diez loci, nueve STR y el locus amilogenina. Uno de los primers de cada pareja está marcado con un fluorocromo (5-FAM, JOE o NED) que confiere a cada marcador un color (verde, amarillo o azul) cuando son analizados en el aparato de electroforesis capilar. En la siguiente tabla pueden observarse los nombres de los fluorocromos usados, así como las longitudes de onda de excitación y emisión de cada uno de ellos.

Tabla 3.5.

Fluorocromo	$\lambda_{máx}$ de Emisión(nm)	$\lambda_{máx}$ de Excitación	Color	Marcadores
5-FAM	522	493	Azul	D3S1358, VWA, FGA
JOE	554	528	Verde	Amelogenina, d8S1179, D21S11, D18S51
NED	575	553	Amarillo	D5S818, D13S317, D7S820
ROX	607	587	Rojo	Estándar interno de tamaño

Tabla 3.6. Componentes del kit de amplificación:

Reactivo	Composición	Volumen para PCR
AmpFISTR®PCR Reaction Mix	$MgCl_2$, deoxinucleósidos trifosfatos (dATP, dCTP, dTTP, dGTP), BSA, azida sódica (NaN_3)	10.5 µl
AmpFISTR Profiler Plus Primer Set	Parejas de primers para los 9 loci STR mas la amilogenina	5.5 µl
AmpliTaq Gold® DNA Polymerase	Enzima polimerasa	0.5 µl (2.5 u)
AmpFISTR®Control DNA 9947A	ADN de una línea celular humana a una concentración de 0.10 ng/µl en 0.05%NaN_3 y buffer.	10 µl

La cantidad de ADN de cada muestra osciló entre 0.5 y 1.25 ng y el volumen final fue de 25 µl.

Condiciones de amplificación:

95ºC	11'	
94ºC	1'	28 ciclos
59ºC	1'	
72ºC	1'	
60ºC	45'	
10ºC	∞	

En ambos tipos de amplificaciones además de las muestras se amplificó sistemáticamente dos controles: uno *positivo*, para comprobar que la reacción de amplificación ha transcurrido correctamente, y otro *negativo,* para descartar una posible contaminación por ADN extraño.

3.4.3. Comprobación de la amplificación.

Una vez finalizada la reacción de amplificación y antes de comenzar con el tipado de las muestras es recomendable, para ahorrar tiempo y reactivos, comprobar los productos de la amplificación mediante la migración electroforética de los mismos en geles de agarosa al 2% en solución TBE 1X y con 50 μl de bromuro de etidio (0.3 mg/ml). En cada pocillo se depositan 5 μl de producto amplificado mas 2 μl de un buffer de carga compuesto por glicerol y azul de bromofenol. Las condiciones electroforéticas que se aplican son: 10 Watios, 100 mA y 100 voltios. Tras unos 30 min. las muestras son visualizadas iluminando el gel con luz UV. De esta manera es posible saber si la amplificación ha sido positiva y, en caso afirmativo puede observarse la intensidad de amplificación indicándonos si debemos aumentar o disminuir la cantidad de producto amplificado que se va a utilizar para el tipado de las muestras (ya sea en gel vertical desnaturalizante de poliacrilamida o en electroforesis capilar).

3.5. ELECTROFORESIS

3.5.1. Sistema de electroforesis vertical

Mediante este sistema se separan los alelos de las muestras que habían sido amplificadas con el kit SilverSTR™ III. El soporte de la electroforesis es un gel de poliacrilamida desnaturalizante al 6% (Acryl-a-Mix®6, 6% Sequencing Solution, Promega).

Preparación del gel

Para la preparación del gel se utilizaron dos tipos de cristales:

- Uno de dimensiones 34 cm x 19.7 cm x 0,4 cm (cristal corto), tratado con 3 µl de Bind-Silane (Pharmacia Biotech) y 1 ml de ácido acético al 0.5% en etanol al 95%. Esta solución hace que el gel quede adherido al cristal una vez que ha polimerizado.

- Otro de dimensiones 36 cm x 19.7 cm x 0.4 cm (cristal largo) que es tratado con 3 ml de una solución repelente, Repel-Silane Silane (Pharmacia Biotech).

Los cristales se colocan en horizontal con las caras que han sido tratadas enfrentadas y con unos separadores de 0.4 mm. Además, para que los cristales no se muevan y al verter la solución de poliacrilamida ésta no se derrame, se ponen pinzas a lo largo de todo el gel. En un vaso de precipitado se prepara una mezcla compuesta por 30 ml de poliacrilamida al 6% más 230 µl de persulfato amónico al 10%. Esta mezcla es vertida entre los dos cristales de manera que no se formen burbujas; finalmente se pone el peine que ha de conformar los diferentes pocillos y se deja que el gel polimerize durante unos 45 min. como mínimo.

Electroforesis y detección

Cuando el gel ya ha polimerizado se quita el peine, con mucho cuidado de no romper los pocillos, y las pinzas; se limpian los restos de acrilamida y se coloca en la cubeta de electroforesis vertical (GIBCO BRL SA32, Life Technologies). La electroforesis se lleva a cabo en buffer TBE 1X que se deposita en ambos reservorios (tanto anódico como catódico). Es importante eliminar muy bien la urea de los pocillos ya que puede interferir en la migración de las muestras haciendo imposible su lectura.

Antes de colocar las muestras en los pocillos se realiza una premigración ("pre-run") a 2.000 V durante unos 30 min. para que el gel alcance una temperatura de 50ºC.

Preparación de las muestras

Mientras el gel está en el "pre-run" se preparan las muestras mezclando 2.5 µl de producto amplificado y 2.5 µl de buffer de carga (STR 2X Loading Solution) incluido en el kit. Además de las muestras amplificadas se preparan de la misma manera el marcador de peso molecular (pGEM®DNA Markers) y los ladders (SilverSTR^{TM}III Allelic Ladder Mix, Promega) donde van incluidos todos los alelos existentes en la población para cada marcador. A continuación se desnaturalizan las muestras mediante calentamiento a 95ºC durante 5 min. e inmediatamente se colocan en hielo para evitar que el ADN renaturalice.

Colocación de las muestras en el gel y electroforesis

Con una jeringa de 10 ml llena de buffer se eliminan los restos de urea que queden en los pocillos antes de proceder a la carga de las muestras en los mismos. Tras la carga de muestras se aplican las mismas condiciones de electroforesis que en el "pre-run": 2.000 V, 40 W y máximos mA.

El tiempo de electroforesis depende del tamaño del producto amplificado; en nuestro caso este tiempo era de aproximadamente 2 horas. El control de migración se hace tomando como referencia los colorantes incorporados en el buffer de carga. En un gel de acrilamida al 6% el azul de bromofenol migra a la altura de 25 bases mientras que el xilanocianol lo hace a la altura de 105 bases.

Una vez finalizada la electroforesis se quitan los separadores y se separan los cristales, el gel quedará adherido al cristal corto y es en este soporte donde se lleva a cabo el proceso de revelado.

Revelado del gel

Para la visualización del ADN tras la electroforesis seguimos el protocolo de la tinción con plata (***"silver staining"***), exactamente tal y como lo describieron BUDOWLE y ALLEN en 1991 pero ajustando los volúmenes al tamaño del gel.

<u>Método:</u>

- Introducir el gel en 1% HNO_3, manteniéndolo 10' en agitación.

- Lavar brevemente con agua.

- Realizar la tinción con $AgNO_3$ al 0'2% (1.4 gr de $AgNO_3$ en 700 ml de agua destilada) y mantener 15' en agitación.

- Lavar brevemente con agua.

- Introducir varias veces en Na_2CO_3 0.28M con formaldehído (250 ulμ de formaldehído por 500 ml de Na_2CO_3), hasta que aparezcan las bandas con la intensidad deseada.

- Lavar con agua.

- Introducir el gel en Acido Acético 10% durante 5 - 10'.

- Lavar con agua.

Debido a la dificultad para conservar los geles una vez que han sido revelados nosotros decidimos guardar los resultados escaneándolos (ScanJet 3p, Hewlett Packard) y grabándolos debidamente identificados en un soporte informático.

3.5.2. Electroforesis capilar

Las muestras amplificadas con el kit AmpFISTR Profiler Plus fueron tipadas por electroforesis capilar en un analizador genético ABI Prism 310 de Perkin Elmer.

Antes de cargar las muestras en el 310 son necesarios una serie de pasos para prepararlo adecuadamente y que las muestras puedan ser analizadas sin problema.

1. Establecer comunicación entre el ordenador y el 310

2. Abrir el software de recogida de datos "ABI PRISM™ 310 Collection" y permitir a través de dicho programa que el 310 reconozca la posición del distribuidor de muestras ("autosampler") y de la bomba ("syringe").

3. Colocar el bloque de metacrilato que ha de estar limpio y seco.

4. Colocar el capilar. Antes de colocarlo se limpia la ventana del capilar con un papel embebido en etanol absoluto, a continuación con otro embebido en agua destilada y seguidamente se seca. Tras limpiar la ventana se introduce el capilar en la válvula que le conecta con el bloque de metacrilato hasta llegar a la primera intersección y se aprieta bien la válvula; es importante no rebasar esta primera intersección ya que si no se sitúa el extremo del capilar correctamente puede impedirse que el polímero penetre en él. A continuación se sitúa la zona de lectura del capilar en la ventana del láser y se cierra la puerta del láser, se introduce el otro extremo en el bloque que contiene el electrodo negativo o cátodo y se desliza hasta que queden a la misma altura y separados 2 mm. Por último se sujeta el capilar con una cinta adhesiva contra la placa termoestabilizadora.

5. Preparar el buffer 1X. Hacer una dilución 1:10 con agua bidestilada del buffer 10X (310 Genetic Analyzer buffer) con un volumen final de 15 ml (11ml para el vial del ánodo y 4 ml para el vial del buffer).

6. Relleno del capilar y del bloque de metacrilato con el polímero. El polímero utilizado para el análisis de fragmentos es el POP-4 (Performance Optimized Polymer 4, Perkin Elmer) y debe sacarse media hora antes de su uso del frigorífico para que se atempere. Con la jeringa bien limpia se toma el volumen de polímero necesario en función del número de muestras, se enrosca bien para evitar pérdidas de polímero y se llenan los conductos del bloque de metacrilato con mucho cuidado de que no se formen burbujas de aire. Una vez colocado el capilar y relleno el bloque de metacrilato con el polímero, lo cual equivaldría a la preparación del gel en los métodos convencionales de electroforesis, el siguiente paso es la preparación de las muestras.

Preparación de muestras y ladder

Para preparar las muestras se añadió en un tubo de 0.5 ml los siguientes volúmenes por muestra:

- 24 μl de formamida desionizada (Ultra Pure Grade, Amresco)
- 1 μl del estándar de tamaño (GeneScan-500 [ROX] Size Standard, Perkin Elmer)
- 1.5 μl de producto amplificado (o de ladder AmpFlSTR Profiler Plus Allelic Ladder para los tubos del ladder)

Al llevar cada muestra un estándar de tamaño se evitan errores en la asignación del tamaño de los fragmentos analizados debidos a diferencias en la movilidad electroforética entre las distintas inyecciones (bien por variaciones de temperatura,

de concentración de polímero o en las condiciones de electroforesis). Con el estándar utilizado es posible analizar fragmentos de entre 35 y 500 pares de bases. En total consta de 16 fragmentos de las siguientes longitudes:

35	139	250	400
50	150	300	450
75	160	340	490
100	200	350	500

El pico de 250 se deja sin asignar y su único papel es indicar las posibles variaciones entre los diferentes procesos electroforéticos.

Una vez preparadas las muestras éstas se desnaturalizan a 96ºC durante 3 min., se ponen en hielo y se coloca cada tubo en su posición adecuada dentro del soporte que va acoplado al distribuidor de muestras.

Programación de los parámetros de análisis

Antes de iniciar la electroforesis una vez que las muestras han sido preparadas es necesario establecer una serie de parámetros de análisis en el programa GeneScan. Entre los parámetros de mayor interés y que pueden afectar más a la hora de interpretar los resultados se encuentran, por un lado, el rango de análisis, y por otro la intensidad de los picos. En el rango de análisis es importante especificar el rango de tamaños (en pares de bases o en puntos de datos) de los fragmentos que se van a analizar, por lo tanto este parámetro deberá de ser programado en función de los marcadores que se vayan a analizar.

Para la intensidad de picos, normalmente se recomiendan 150 RFU para evitar picos de menor intensidad que podrían corresponder a un ADN extraño. No obstante, cuando los datos van a ser genotipados con el Genotyper se recomienda bajar la intensidad hasta 50 RFU para que este programa pueda identificar señales bajas de manera automática.

Hoja de muestras ("Sample Sheet")

Para preparar la hoja de muestras es necesario abrir el programa *ABIPrism 310 Collection*. Una vez abierto, se procede a rellenar una nueva hoja con el nombre de las muestras en función de su colocación en el tray. Para cada muestra se requiere especificar el color del estándar, en la columna "Std" de la hoja aparecerá un rombo junto al color elegido para el estándar (en nuestro caso en el que las muestras fueron amplificadas con el kit Profiler Plus el color reservado para el estándar es el rojo). A continuación se completa la columna "Sample Info" en la que se anotan los marcadores analizados y se guarda la hoja de muestras en la carpeta "Sample Sheets".

Hoja de inyección

En esta lista se incorpora la hoja de muestras creada y grabada anteriormente y en ella es posible modificar el orden de análisis de las muestras e incluso repetir algunas de las muestras cambiando alguno de los parámetros de análisis como el tiempo de inyección. En la segunda columna se especifica el módulo de análisis el cual contiene las funciones específicas ejecutadas para el proceso de análisis de las muestras. En el caso de análisis de fragmentos el módulo seleccionado es el *GS STR POP-4 F(1ml)*. Una vez seleccionado dicho módulo se establecen automáticamente unos valores para los siguientes parámetros:

Tiempo de inyección.....5 sg

Voltaje de inyección.....15 kv.

Voltaje de electroforesis .....15 kv.

Temperatura de electroforesis.....60ºC

Tiempo de electroforesis .....24 min.

A pesar de preestablecer unas condiciones de electroforesis es posible modificarlas mientras el instrumento trabaja y siempre que las muestras no hayan sido procesadas.

- **Voltaje y tiempo de inyección** (para inyección electrocinética): ambos parámetros pueden ser variados para regular la cantidad de ADN que entra en el capilar. Así, en el caso de obtener una señal débil en una muestra es posible aumentar el tiempo de inyección para permitir que se inyecte más cantidad de producto amplificado al capilar y que aumente la señal de fluorescencia. No obstante, un tiempo de inyección excesivo puede originar una pérdida de resolución.

- **Voltaje de electroforesis**: para fragmentos analizados utilizando el polímero POP-4 el voltaje estándar es de 319 V/cm. Para un capilar de 47 cm esto supone un voltaje total de 15000 V y la corriente a este voltaje es de 8-9 μA. Bajo estas condiciones electroforéticas el tiempo necesario para que un fragmento de 400 pb llegue a la ventana detectora es de unos 25 min. Si se aumenta el voltaje el tiempo necesario es menor pero se pierde resolución.

- **Temperatura de electroforesis**: para secuenciación esta temperatura debe ser de 50ºC y para fragmentos de 60ºC.

- **Tiempo de electroforesis:** puede acortarse cuando sólo se necesite información de productos de amplificación de pequeño tamaño. Normalmente el tiempo de electroforesis programado es un 10% superior a la media del tiempo necesario para que el fragmento de mayor tamaño migre lo suficiente.

En la primera fila de la lista de inyección se inserta siempre una línea en la que se programa el test de los cuatro colores ("Test CCD 4-Color") cuya función es comprobar que la ventana del capilar está limpia y colocada correctamente. Este test obtiene una lectura con el capilar vacío, tiene una duración de 5 min. y en el debemos comprobar que la línea basal de los cuatro colores esté entre 800 y 2500 en escala vertical (Intensidad de fluorescencia).

Una vez completada la lista de inyección y habiéndose alcanzado la temperatura de inyección la electroforesis se inicia pulsando el comando "run".

Análisis de los resultados

A) Análisis de los resultados con el software GeneScan

Una vez recogido el dato bruto o "*raw data*", el software *GeneScan* procesa la información para proporcionar una cuantificación en pares de bases de todos los picos. Como cada muestra lleva un estándar interno éste migra a la vez que el ADN amplificado y bajo las mismas condiciones, con lo que se evita el problema de las variaciones entre muestras.

El estándar de tamaño se define una vez que la electroforesis ha terminado y antes de que el programa GeneScan comience el análisis de las muestras. Para dicha definición existen dos posibilidades: definir un estándar nuevo o bien utilizar un estándar utilizado previamente en otros análisis. Una vez definido el estándar se

comprueba que en todas las muestras coincidan los picos del mismo, lo cual se hace solapando todos los estándar. Si en alguna muestra los picos no coinciden se define uno nuevo. En nuestro caso utilizamos un estándar que había sido usado en análisis anteriores.

Tras la definición del estándar es necesario definir los parámetros de análisis, que fueron los que el programa tiene asignados automáticamente para el Profiler Plus y se procedió al análisis de las muestras. El método por el que el programa determina los tamaños de los fragmentos se denomina Southern Local ("Local Southern Method"). Dicho método utiliza tres puntos de la curva generada con los puntos del estándar: los dos que están por debajo en tamaño del fragmento a determinar y uno por encima, con dichos valores realiza una interpolación y asigna el valor del tamaño.

Los datos analizados pueden ser visualizados de forma gráfica (electroferograma), de forma numérica o bien de ambas formas. Estos datos fueron exportados al programa *Genotyper.*

Asignación de alelos

El programa **Genotyper** es el encargado de asignar los alelos a los fragmentos analizados en GeneScan por comparación con los alelos del ladder. Cuando se importa un archivo desde GeneScan a Genotyper éste último genera una línea (dye/lane list) para cada fluorocromo en la que están incluidos el tamaño, la cantidad y la información de todos los fragmentos marcados con un mismo fluorocromo.

Según la aplicación Genotyper contiene una serie de "macros" que son secuencias de comandos predefinidos necesarios para un determinado tipo de análisis. Para fragmentos amplificados con el kit Profiler Plus el macro seleccionado fue el "*Kazam*". Una vez ejecutado dicho macro aparecieron las muestras con los correspondientes alelos asignados de manera automática. Aquellos alelos que se encontraron por encima o por debajo de los alelos del ladder en un rango de 0.5 pares de bases no fueron designados por el programa.

3.6. ANÁLISIS ESTADÍSTICO

3.6.1. Cálculo de las frecuencias alélicas y genotípicas

Puesto que todos los loci analizados en nuestro estudio poblacional son codominantes y distinguibles, la obtención de las frecuencias alélicas muestrales se realizó por el método clásico del recuento en base a los alelos observados (Li, 1976). Posteriormente, realizamos una agrupación de todos los genotipos observados, homo y heterocigóticos, así como el número esperado de estos genotipos calculado por medio de un estimador el cual, al ser el tamaño de muestra lo suficientemente grande (igual o mayor de 100), fue insesgado o centrado.

Además, el hecho de que nuestras estimas de frecuencias se basen en una muestra de la población, nos lleva a la asunción de un margen de error o un intervalo de confianza en el cual es probable que estén contenidos nuestros valores de frecuencias. El programa informático utilizado calculó los límites de confianza tanto superior como inferior con un valor de confianza del 95%.

3.6.2. Tests para demostrar el equilibrio de Hardy-Weinberg

Para comprobar si la muestra poblacional analizada se encontraba en equilibrio de Hardy-Weinberg para los diez loci se aplicaron dos tipos de tests estadísticos:

- **Test de bondad de ajuste** (test de razón de verosimilitud o "likelihood test", y test de homocigosidad).

- **Tests exactos** (Test exacto de Guo-Thompson).

3.6.2.1. Tests de bondad de ajuste

En general los tests de bondad de ajuste se emplean para comparar dos hipótesis mutuamente excluyentes, en el caso de los estudios poblacionales nos permiten comprobar si las frecuencias genotípicas observadas en la muestra se desvían de las esperadas bajo la hipótesis de equilibrio Hardy-Weinberg. Como lo que se compara es un número *n* de categorías o clases genotipicas, el estadístico (χ^2) que mide dichas discrepancias vendría dado por la expresión:

$$\chi^2_{ob} = \sum (O_i - E_i)^2/E_i$$

Siendo O_i los valores observados y E_i los esperados.

La hipótesis inicial planteada acerca de si las cantidades observadas se desvían de las esperadas bajo la hipótesis de equilibrio será cierta si el valor de χ^2_{ob} es lo suficientemente alto. El problema que plantea este tipo de tests es que la

muestra debe tener un valor de n grande y los valores esperados no pueden ser inferiores a 1. En el caso de los STRs el alto grado de polimorfismo hace que algunas de las clases genotípicas tengan valores muy pequeños y por lo tanto estos tests no llegan a ser efectivos. El problema de utilizar una distribución continua como es la χ^2 para comprobar una hipótesis con datos discretos se puede reducir usando un término de corrección que normalmente tiene un valor de 0.5 o 0.25 (Emigh, 1980).

Al aplicar este tipo de tests se pueden cometer dos tipos errores:

- **Error de tipo I,** también llamado **nivel de significación** o **error** α, que es el error que cometemos cuando rechazamos la hipótesis nula siendo ésta verdadera.
- **Error de tipo II** o β que es el error cometido al aceptar dicha hipótesis cuando ésta es falsa. El valor 1-β es lo que se denomina poder o potencia de la prueba. Un test es tanto mejor cuanto más potente sea.

El error α es un valor controlable y fijado de antemano, cuanto más pequeño sea dicho valor mas garantías se tendrán de aceptar la hipótesis alternativa siendo ésta cierta.

El **valor P** es el mínimo error α al cual un resultado es significativo. Cuando P es menor o igual que el valor del error α se decide la hipótesis altermativa y se rechaza la nula, en cambio, si es mayor que α se acepta la hipótesis nula (Martin Andrés A, 1988).

- **Test de Homocigosidad**

Compara los resultados obtenidos agrupándolos en dos clases: homocigotos y heterocigotos, entonces realiza una χ^2 con un grado de libertad con los homocigotos y heterocigotos esperados y observados.

$$\chi^2 = [\ (Hom_{obs} - Hom_{esp})^2 / \ Hom_{esp}\ ] + [\ (Het_{obs} - Het_{esp})^2 / Het_{esp}\ ]$$

- **Test de razón de verosimilitudes**

Fue aplicado a la situación que nos ocupa en 1969 por Sokal y Rohlf y compara las frecuencias genotípicas observadas y esperadas sin agruparlas. Calcula un valor G siguiendo el algoritmo:

$$2\sum_{i=1}^{a} f_i \times \ln (f_i/f'_i)$$

Siendo f_i la frecuencia observada para el genotipo i y f'_i la frecuencia esperada para el genotipo i.

Este test es aplicable en casos en los que existen frecuencias genotípicas muy bajas pero, por este mismo motivo, plantea el problema de que su nivel de significación no puede ser evaluado directamente por ninguna distribución estándar. Por ello, se ha empleado un procedimiento empírico que realiza permutaciones de los alelos observados y genera una tabla de frecuencias genotípicas hipotéticas por unión aleatoria de dichos alelos. Siguiendo a Chakraborty y col. (1991) se generaron 1000 muestras aleatorias, manteniendo las frecuencias alélicas, y para cada combinación se estimó el valor de G. El valor P obtenido se refiere al número de

combinaciones en las que el valor G obtenido en la simulación es superior al observado.

3.6.2.1. Test Exacto de Guo-Thompson

El test exacto de Guo-Thompson (1992) es utilizado normalmente cuando la muestra poblacional es pequeña o bien cuando existen múltiples alelos y las frecuencias son bajas e incluso no llegan a observarse algunos genotipos en la muestra. Matemáticamente estos test son bastante complejos, no obstante hoy en día este problema queda solucionado mediante la creación de programas informáticos.

Existen dos métodos para calcular los niveles de significación del test exacto:

- El Método de Monte Carlo (Binder y Heermann, 1989).
- El Método adaptado del algoritmo Metropolis (Metropolis et al, 1953).

El método de Monte Carlo realiza un número N de simulaciones (en nuestro caso 2000) con el mismo número de individuos y la misma proporción de alelos que la muestra original. A cada una de estas simulaciones se le calcula su probabilidad exacta (p) para compararla con la de la muestra original. Si el valor de tal probabilidad es menor o igual al valor de nuestra muestra, se aumenta en un valor de 1 un contador denominado K. Así, el valor total de probabilidad que se calcula viene dado por K / N. Valores muy bajos de este cociente indicarán que nuestra población se encuentra entre las distribuciones más raras y por lo tanto podremos rechazar la hipótesis de equilibrio con un determinado margen de error.

El método de Metropolis se basa en la construcción de las cadenas de Markov con una distribución aleatoria de las probabilidades de los genotipos, cuyas muestras poseen las mismas frecuencias alélicas que la muestra poblacional en estudio.
El test utilizado por nosotros ha sido el Monte Carlo que, por regla general, parece ser el más adecuado para estudios con tamaños de muestra pequeños y un gran número de alelos.

Según Chakraborty y Zhong (Chakraborty, R. et al, 1994) cuanto más apartada del equilibrio esté la población, más fácil será detectar el desequilibrio. Al aumentar el tamaño de la muestra también se facilita la dicha detección. La gran diferencia de los tests exactos con respecto a los de bondad de ajuste es que en los primeros el nivel de detección de las desviaciones aumenta con el número de alelos, por lo que puede decirse que el test exacto es el método más poderoso de detección de desviaciones del equilibrio de Hardy-Weinberg para marcadores genéticos muy polimórficos como es el caso de los STRs.

3.6.3. Test utilizados para demostrar la independencia de los alelos

Para demostrar tanto el equilibrio de ligamiento (o independencia intralocus) como el equilibrio de la fase gamética (independencia interlocus) se emplearon los tests de homocigosidad y de correlación interclase de Karlin respectivamente.

Para el test de homocigosidad el estadístico sigue una distribución χ^2 con un grado de libertad cuando la hipótesis de equilibrio de Hardy-Weinberg es cierta.

El test de KARLIN realiza una comparación dos a dos de todos los loci estudiados, determinando la correlación existente entre ambos (***rho***) y sobre ese valor determina una cantidad experimental. Para obtener la probabilidad (valor P) realiza una serie de

permutaciones (1000 en nuestro caso) sobre los datos poblacionales obtenidos para ver como de extraño es la correlación obtenida, lo cual nos indica que los valores de P altos nos dirán que el resultado obtenido es un valor frecuente en nuestros datos y por lo tanto que no se ha obtenido por azar. El test aplicado es de dos colas, ya que nos indica la correlación tanto en sentido positivo como negativo, es decir, si la presencia de un determinado alelo lleva aparejada la de otro, o si por el contrario el que aparezca ese alelo implica que el otro no lo hará.

El planteamiento general del test será:

Ho*:* rho = 0

H1: valores altos o bajos de X suelen ir acompañados de valores altos o bajos de Y.

El hecho de que se obtenga un valor de rho diferente de cero no quiere decir que exista una correlación, ya que podría ser que ese valor se debiera al azar o como consecuencia de la muestra tomada y que realmente no exista una correlación en la población, de ahí la necesidad de hacer el test y ver si el resultado es o no significativo.

3.6.4. Comparaciones poblacionales

Los resultados obtenidos en nuestro estudio poblacional para los diez loci fueron comparados con los obtenidos en estudios de otras poblaciones como fueron: Italia, Suiza y EEUU.

Los resultados fueron comparados dos a dos (para cada marcador) y se obtuvieron los valores de los estadísticos χ^2 y G. A partir de dichos valores se calculó el valor de significancia (p) de manera que valores de p mayores de 0.05 indicaban diferencias significativas entre dos poblaciones determinadas y valores menores de 0.05 eran indicativos de que entre las dos poblaciones comparadas no existían diferencias significativas para el marcador estudiado.

3.6.5. Parámetros estadísticos de interés práctico.

Aunque la mayoría de estos parámetros se utilizan fundamentalmente en la investigación biológica de la paternidad (IBP), algunos se pueden aplicar al campo de la criminalística.

3.6.5.1. Probabilidad de exclusión a priori

Es el porcentaje de presuntos individuos cuya paternidad biológica quedaría excluida en base a este marcador. Para calcularla nos basamos en el análisis estadístico de los sucesos de incompatibilidad genética, el cual nos aporta una fórmula sencilla para el caso de que existan dos alelos codominantes:

$$Pex = p_1\, p_2\, (1\text{-}p_1\, p_2)$$

Conforme el número de alelos es mayor el cálculo se complica enormemente por lo que se recurren a programas computerizados. La probabilidad de exclusión a priori es un valor porcentual que es función directa del polimorfismo de un marcador.

La PROBABILIDAD DE EXCLUSION A PRIORI ACUMULATIVA de varios sistemas, determina la eficacia que posee un laboratorio, la cual debería ser como mínimo del 99% para poder llevar a cabo los estudios de criminalística y de IBP.

3.6.5.2. Eficacia bioestadística de un sistema

Nos indica la eficacia media real en términos del valor ESSEN-MÖLER de un marcador en la prueba positiva de paternidad en base a los valores obtenidos de las posibles combinaciones padre-madre-hijo. La eficiencia también depende del polimorfismo de un sistema y del conjunto de marcadores. Para su cálculo se parte de los valores Essen-Möller (EM) de cada triplete (presunto padre-madre-hijo) teniendo en cuenta su probabilidad relativa de ocurrencia.

$$\text{Valor EM} = \log Y/X + 10$$

Un marcador con un valor EM igual a diez significa que X e Y son iguales lo cual indicaría que ese sistema es poco informativo. Por el contrario, un valor igual a cero indicaría una eficacia bioestadística máxima. Para calcular la eficacia bioestadística acumulada de varios sistemas se multiplican las eficacias bioestadísticas medias de cada marcador y se divide por diez elevado al número de sistemas menos 1.

Tanto la eficia bioestadística de un sistema como la probabilidad de exclusión a priori dependen del polimorfismo del sistema y sirven para evaluar la utilidad a priori de un conjunto de marcadores.

3.6.5.3. Probabilidad a posteriori

Tradicionalmente, la no exclusión de una serie de marcadores era suficiente para establecer una relación entre dos muestras. Hoy en día, gracias a la utilización de un elevado número de marcadores genéticos de ADN, es posible dar unos valores de probabilidad de coincidencia (en criminalística) o de paternidad bastante seguros.
En el caso de paternidades el cálculo de probabilidad se basa en el teorema de Bayes. La fórmula que expresa dicho teorema es algo complicada pero, considerando que la

probabilidad de paternidad a priori es 0,5 (ya que siendo conservadores un supuesto padre tiene la misma probabilidad de serlo que de no serlo)quedaría reducida a lo siguiente:

$$W = X/(X+Y)$$

Siendo X la probabilidad de transmisión del alelo paterno e Y la frecuencia de dicho alelo en la población general.

Otras formas de expresar este resultado son el valor de Essen-Möller y el índice de paternidad.

El valor del **índice de paternidad** es el inverso de la probabilidad de paternidad (W) y relaciona la probabilidad de paternidad del presunto padre con la probabilidad de paternidad de un hombre escogido al azar.

3.6.5.4. Probabilidad de discriminación

Se trata de una medida relativa de la eficacia de los sistemas utilizados. Se define como la probabilidad que tendrían dos individuos elegidos al azar de no coincidir en una serie de genotipos dados. La probabilidad de discriminación vendría dada por:

- P discr. = 1 – P coincidencia

- P discr. = 1 - Q

La PROBABILIDAD DE DISCRIMINACIÓN ACUMULATIVA para varios sistemas (m) viene dada por la fórmula:

- P discr. acumul. = 1 - (Q_1.Q_2. ... Q_m)

Es importante, desde el punto de vista práctico, determinar también la Probabilidad de discriminación teórica máxima y máxima acumulada. La probabilidad de que dos individuos tengan el mismo genotipo para un marcador determinado dependerá de la frecuencia de los genotipos en la población y en general del polimorfismo del marcador. Así, la probabilidad de discriminación máxima se alcanzaría cuando la frecuencia de los alelos fueran iguales ($P_1 = P_2 = ... P_k = 1/k$).

De este modo la probabilidad de discriminación máxima que se puede alcanzar quedaría recogida por la siguiente fórmula:

Prob. discr. teórica = 1 - (2k-1)/k^3

donde *k* es igual al número de alelos del sistema génico del locus en cuestión.

El máximo teórico no se consigue en realidad, pero nos ofrece una base para evaluar la eficacia relativa de los distintos marcadores. Una probabilidad de discriminación de 1 (P de "matching" = 0) sería la mejor situación deseable para establecer una identificación.

Para evaluar la utilidad de un marcador aislado podemos comparar su **P discriminación** con la **P discriminación máxima teórica** y ver la diferencia existente: cuanto más se aleje, menos efectivo será para la individualización en Medicina Legal.

3.6.5.5. Probabilidad de coincidencia

La "identidad genética" o la "evidencia genética" juega un papel importante en las investigaciones médico-legales a la hora de llevar a cabo un estudio comparativo entre un indicio y un sospechoso. Como sucede en los casos de IBP, la exclusión es concluyente, pero la coincidencia puede ser sólo una "evidencia circunstancial". La evaluación de esta evidencia circunstancial no difiere, en principio, del cálculo empleado en la valoración de la identidad génica en las acciones de paternidad.

Debemos diferenciar entre la probabilidad de "matching", referida a los casos en que se comparan las muestras de dos individuos (siendo la probabilidad que tienen estos dos individuos tomados al azar de presentar los mismo genotipos), y la PROBABILIDAD de COINCIDENCIA, que se refiere a la comparación de una serie de marcadores de un individuo con unos genotipos dados, o dicho de otro modo, la probabilidad que tiene una persona concreta de coincidir con otra u otras en una serie de genotipos dados.

La probabilidad de coincidencia es siempre menor en tanto en cuanto es la probabilidad de coincidencia multiplicada por la frecuencia de los individuos que pueden coincidir en los genotipos.

La cifra que nos ofrece directamente la probabilidad de "culpabilidad" o "inocencia" de una persona cuando no ha sido excluida, viene dada por la probabilidad de coincidencia de los distintos genotipos encontrados en el sospechoso y en el indicio.

3.6.5.6. Heterocigosidad y diversidad génica

Es un parámetro indicativo del polimorfismo y eficacia de un marcador genético. En una población que se encuentra en equilibrio Hardy-Weinberg es posible estimar la heterocigosidad media examinando la proporción de heterocigotos. La heterocigosidad de un sistema puede calcularse siguiendo la siguiente ecuación:

$$h_1 = \sum_{u} \sum_{v \neq u} n_{uv}/N$$

donde n_{uv} es el número de individuos heterocigotos para un locus determinado.

La varianza muestral de este estimador viene dada por la fórmula:

$$\mathbf{Var}\ (h_1) = h_1\ (1-h_1)/N$$

Hay casos en los que el uso de este parámetro no es adecuado:

- cuando el tamaño de la muestra es pequeño ya que en este caso las frecuencias de los genotipos pueden desviarse del equilibrio de Hardy-Weinberg por un error de muestreo

- si actuara una fuerte selección en un estadío de desarrollo antes de la observación ya que la proporción de heterocigotos quedaría desproporcionada

- cuando existe endogamia.

En estos tres casos es más correcto medir la variabilidad de la población mediante el parámetro denominado diversidad genética. El valor de este parámetro no depende del polimorfismo del sistema sino de las frecuencias génicas y del tamaño de la muestra. En poblaciones apareadas al azar los valores de heterocigosidad y diversidad genética tienen valores similares.

4. RESULTADOS

En el presente capítulo se procederá a la exposición ordenada y esquematizada de todos los resultados obtenidos. Por un lado, se estudiarán los resultados obtenidos para el estudio poblacional de los loci: **D13S317, D7S820, D16S539, D3S1358, VWA, FGA, D8S1179, D21S11, D18S51** y **D5S818** mostrándose para cada uno de ellos las frecuencias génicas, genotípicas y los resultados de los tests aplicados para demostrar el equilibrio de Hardy-Weinberg, la independencia de los loci así como otros parámetros de interés forense. Por último compararemos nuestros resultados con los obtenidos en otras poblaciones.

Por otro lado, se analizarán los resultados obtenidos en el estudio de validación del triplex de Promega SilverSTR™ III. Se expondrá cual ha sido el límite de sensibilidad hallado por nosotros y los resultados del análisis de mezclas y muestras forenses.

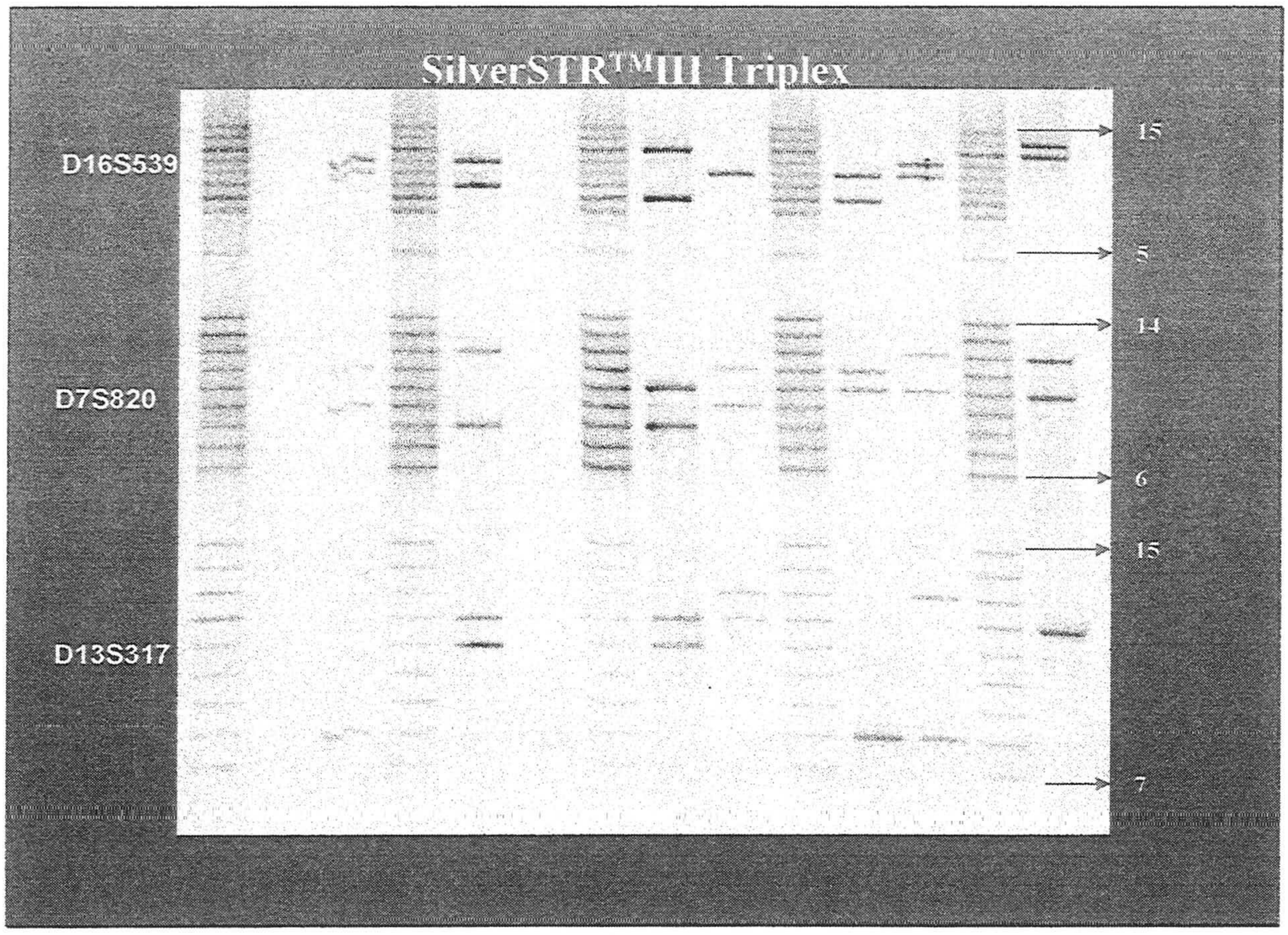
4.1. Resultados del estudio poblacional

4.1.1. Visualización de los resultados obtenidos para el triplex SilverSTR™ III

Como se describió en el apartado de Material y Métodos, la visualización de las muestras de ADN que fueron amplificadas con el kit SilverSTR™ III se realizó en geles de acrilamida al 6%. Tras los procesos de electroforesis y revelado con nitrato de plata, la interpretación de los resultados se llevó a cabo por comparación directa de las muestras con los ladders adyacentes a cada una.

A continuación se muestra una fotografía de un gel vertical de acrilamida, en el que se procesaron muestras del estudio poblacional.

Imagen 4.1.



4.1.2. Visualización de los resultados obtenidos para el kit AmpFISTR Profiler Plus

Tras el análisis de las muestras con el programa Genotyper, éste asignó automaticamente los alelos a cada marcador y los resultados se visualizaron como se muestra a continuación:

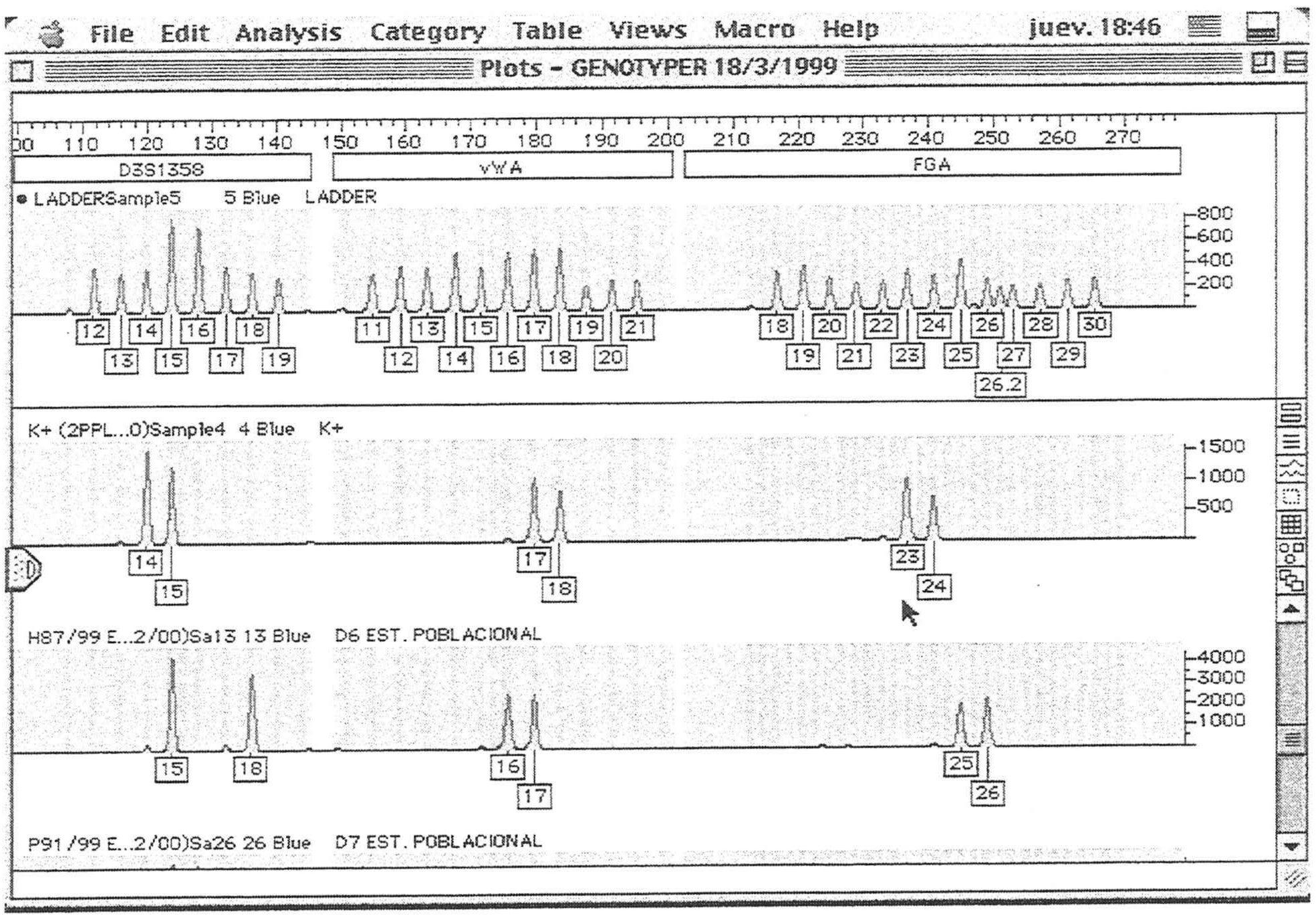


Imagen 4.2. Electroferograma en el que se observan los resultados obtenidos para los loci marcados con el fluorocromo azul (5-FAM): D3S1358, VWA y FGA.

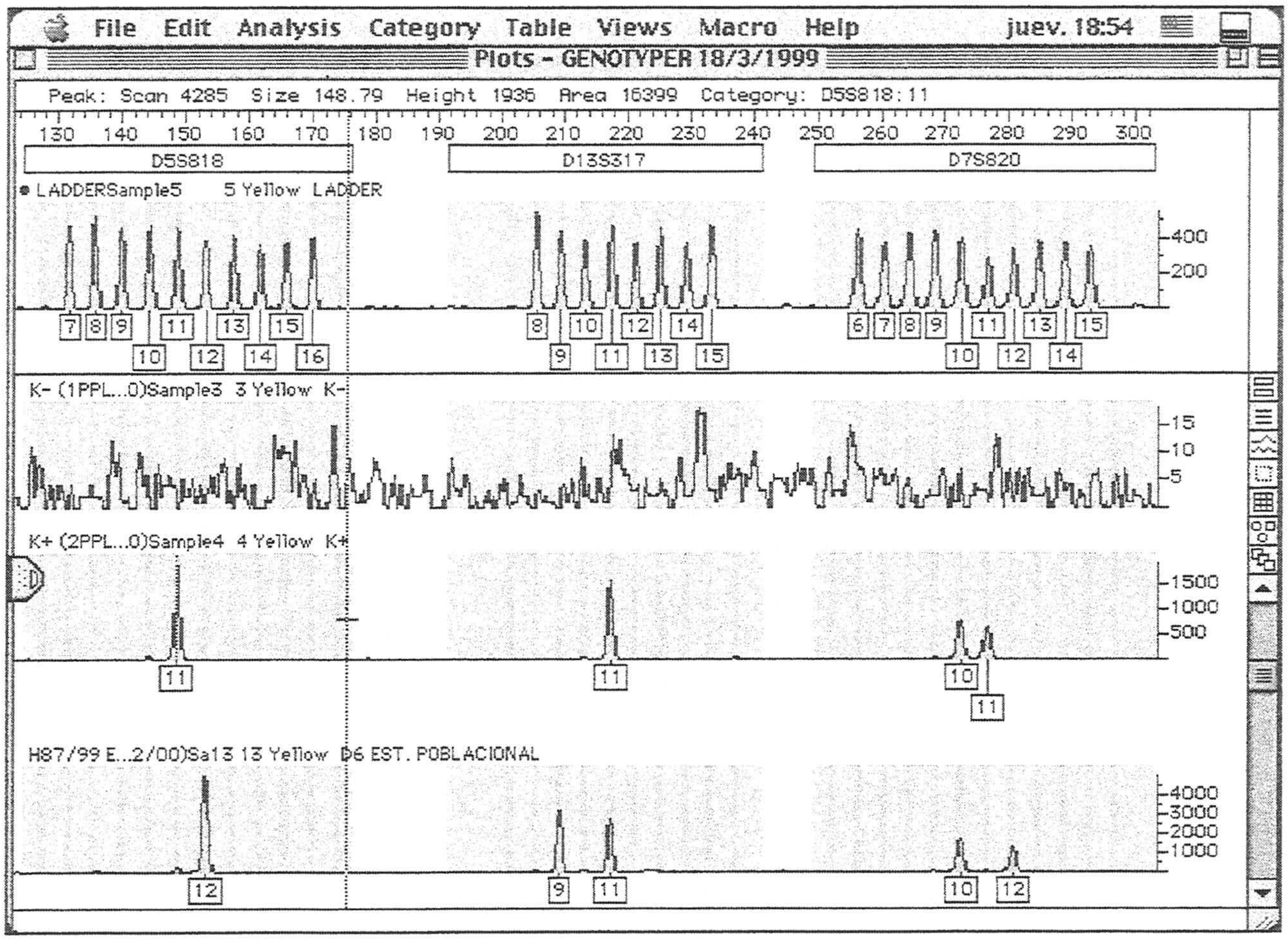


Imagen 4.3. Electroferograma en el que se observan los resultados obtenidos para los loci marcados con el fluorocromo amarillo (NED): D5S818, D13S317 y D7S820.

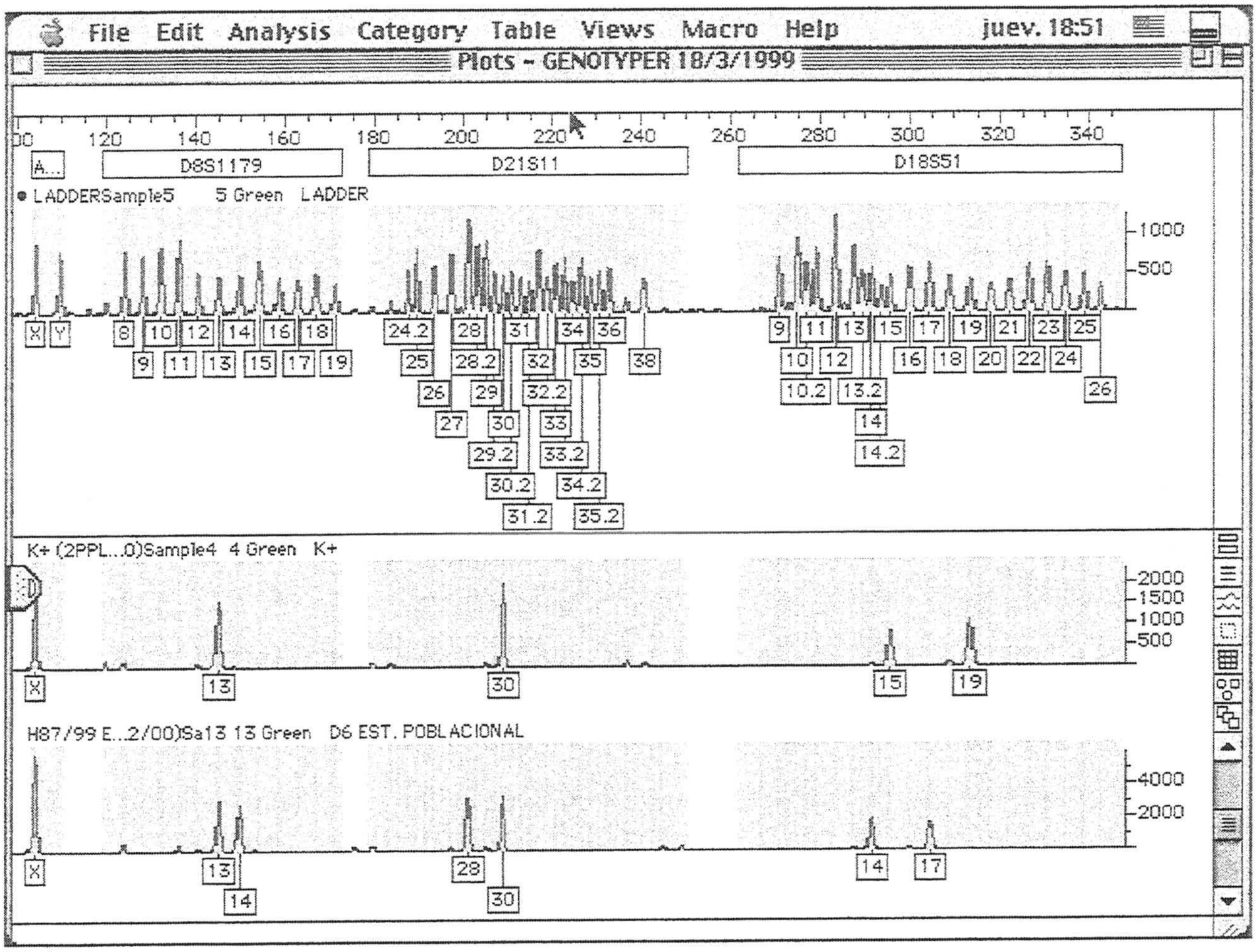


Imagen 4.2. Electroferograma en el que se observan los resultados obtenidos para los loci marcados con el fluorocromo verde (JOE): amelogenina, D8S1179, D21S11 y D18S51.

4.1.3. Resultados obtenidos con el kit SilverSTR™ III

Una vez realizado el recuento de genotipos en la población estudiada, los datos fueron introducidos en formato de WP en el programa informático cedido por el Doctor Budowle del FBI.

Las frecuencias alélicas y genotípicas así como los porcentajes de homocigosidad y heterocigosidad para cada uno de los tres loci se exponen a continuación en las siguientes tablas:

Tabla 4.1. Frecuencias alélicas y genotípicas del locus D13S317

D13S317							
Alelo	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	**Superior**				
8	0,1751	0,1252	0,2395	**8**	**8**	8	0,04519774
9	0,0536	0,0285	0,0986	**8**	**9**	1	0,00564972
10	0,0678	0,0387	0,1161	**8**	**10**	7	0,03954802
11	0,2796	0,2174	0,3517	**8**	**11**	14	0,07909605
12	0,2627	0,2021	0,3339	**8**	**12**	13	0,07344633
13	0,1102	0,0713	0,1663	**8**	**13**	7	0,03954802
14	0,0508	0,0266	0,0951	**8**	**14**	4	0,02259887
				9	**9**	1	0,00564972
				9	**11**	5	0,02824859
				9	**12**	9	0,05084746
				9	**13**	1	0,00564972
				9	**14**	1	0,00564972
				10	**11**	6	0,03389831
				10	**12**	9	0,05084746
				10	**13**	1	0,00564972
				10	**14**	1	0,00564972
				11	**11**	14	0,07909605
				11	**12**	34	0,1920904
				11	**13**	9	0,05084746
				11	**14**	3	0,01694915
				12	**12**	7	0,03954802
				12	**13**	9	0,05084746
				12	**14**	5	0,02824859
				13	**13**	5	0,02824859
				13	**14**	2	0,01129944
				14	**14**	1	0,00564972
Total	0,9998					177	1

Tabla 4.2. Frecuencias alélicas y genotípicas del locus D7S820.

D7S820							
Alelo	Frecuencias alélicas	Intervalos de confianza		Genotipos		Número	Frecuencias
		Inferior	Superior				
7	0.023	0.0087	0.0573	**7**	**8**	2	0,01129944
8	0,1525	0.1065	0.2137	**7**	**9**	2	0,01129944
9	0,1356	0.0924	0.1947	**7**	**11**	2	0,01129944
10	0,3192	0.2542	0.3921	**7**	**12**	2	0,01129944
11	0,1921	0.1403	0.2573	**8**	**8**	3	0,01694915
12	0,1525	0.1065	0.2137	**8**	**9**	12	0,06779661
13	0,023	0.0087	0.0573	**8**	**10**	13	0,07344633
14	0.0028	0.0003	0.003	**8**	**11**	5	0,02824859
				8	**12**	13	0,07344633
				8	**13**	3	0,01694915
				9	**9**	3	0,01694915
				9	**10**	8	0,04519774
				9	**11**	12	0,06779661
				9	**12**	8	0,04519774
				10	**10**	26	0,14689266
				10	**11**	23	0,1299435
				10	**12**	16	0,09039548
				10	**13**	1	0,00564972
				11	**11**	10	0,05649718
				11	**12**	6	0,03389831
				12	**12**	3	0,01694915
				12	**13**	2	0,01129944
				12	**14**	1	0,00564972
				13	**13**	1	0,00564972
Total	1,0007					177	1

Tabla 4.3. Frecuencias alélicas y genotípicas del locus D16S539.

D16S539							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	**Superior**				
8	0,0113	0.003	0.0413	**8**	**9**	1	0,0056497
9	0,11864	0.0782	0.1761	**8**	**12**	1	0,0056497
10	0,0678	0.0387	0.1161	**8**	**13**	2	0,0112994
11	0,25989	0.1996	0.3308	**9**	**10**	4	0,0225989
12	0,31073	0.2458	0.3841	**9**	**11**	10	0,0564972
13	0,21186	0.1569	0.2795	**9**	**12**	10	0,0564972
14	0,01695	0.0056	0.0497	**9**	**13**	17	0,0960452
15	0,00282	0.00028	0.0275	**10**	**11**	6	0,0338983
				10	**12**	11	0,0621469
				10	**13**	2	0,0112994
				10	**14**	1	0,0056497
				11	**11**	10	0,0564972
				11	**12**	32	0,180791
				11	**13**	22	0,1242938
				11	**14**	2	0,0112994
				12	**12**	17	0,0960452
				12	**13**	21	0,1186441
				12	**14**	1	0,0056497
				13	**13**	4	0,0225989
				13	**14**	2	0,0112994
				13	**15**	1	0,0056497
Total	0,99999					177	1

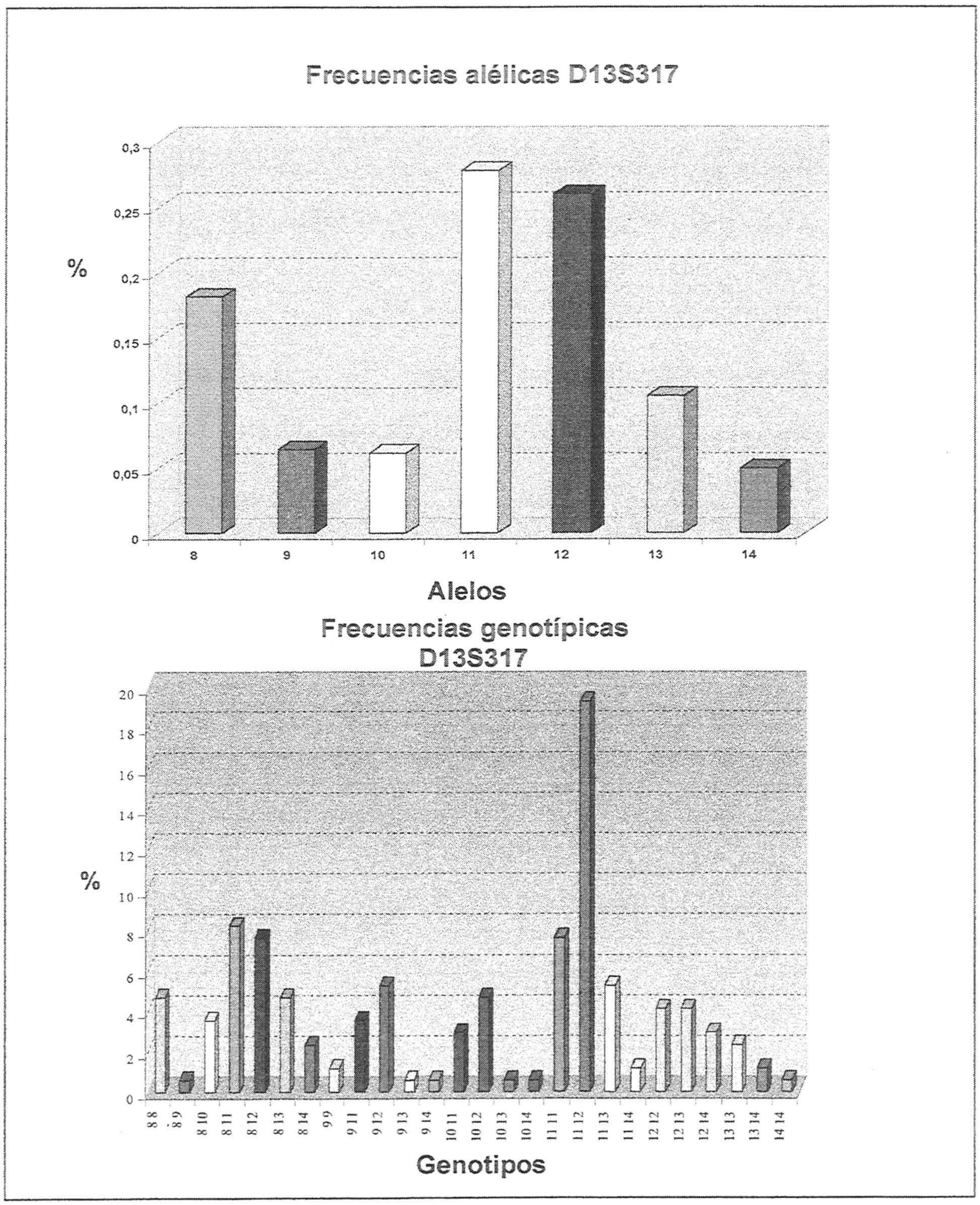
Tabla 4.4. Comparación entre el porcentaje de heterocigotos observado y esperado según un estimador sesgado.

	Heterocigosidad observada (%)	Heterocigosidad esperada (sesgada) (%)	χ^2	Valor p
D13S317	79,66	79,99	0,011953	0,912942
D7S820	74,01	79,53	3,305174	0,069062
D16S539	82,49	77,19	2,817425	0,093246

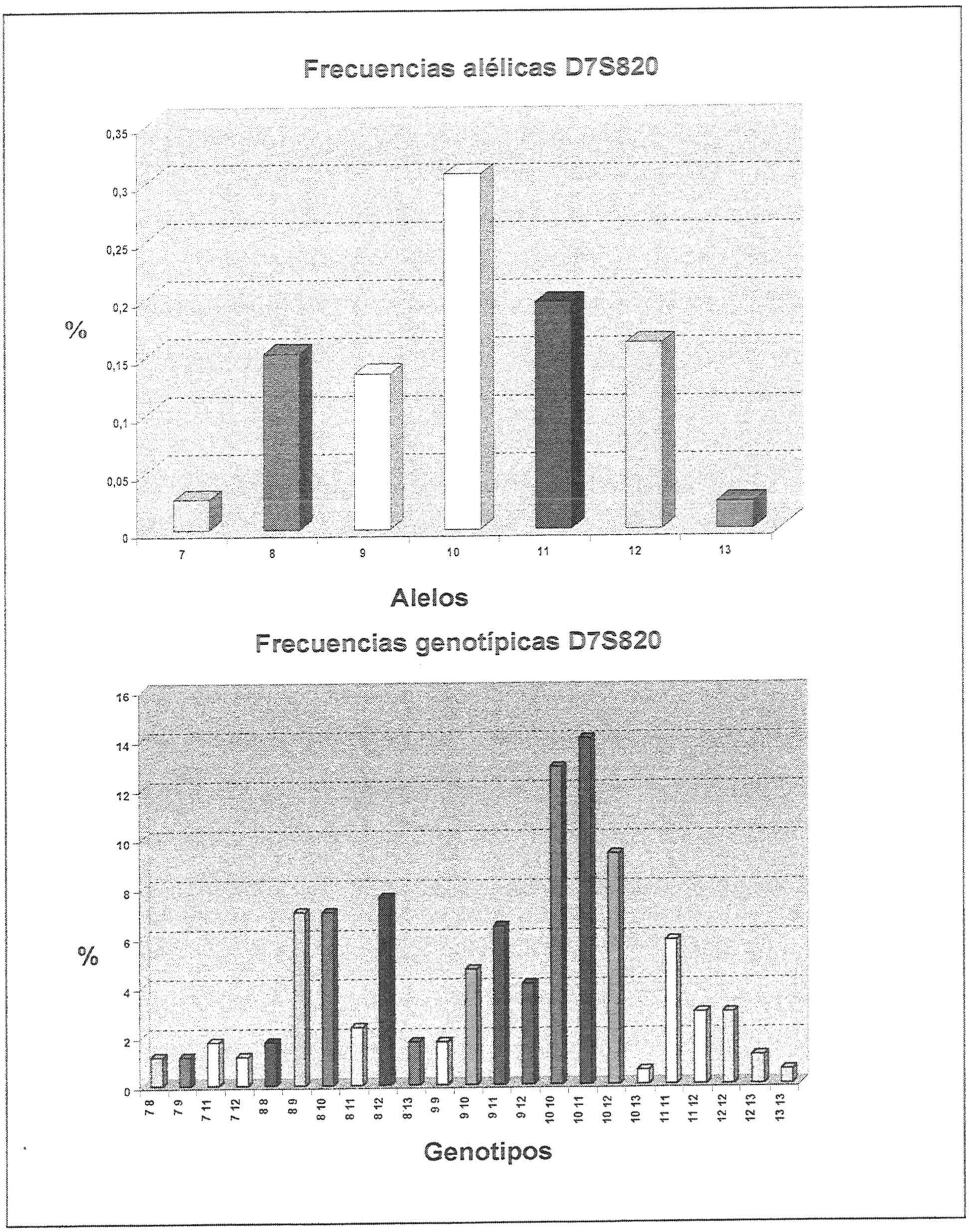
Tabla 4.5. Comparación entre el porcentaje de heterocigotos observado y esperado según un estimador insesgado.

	Heterocigosidad observada (%)	Heterocigosidad esperada (no sesgada) (%)	χ^2	Valor p
D13S317	79,66	80,22	0.034401	0,852857
D7S820	74,01	79,75	3.610376	0,057420
D16S539	82,49	77,41	2,607153	0,106383

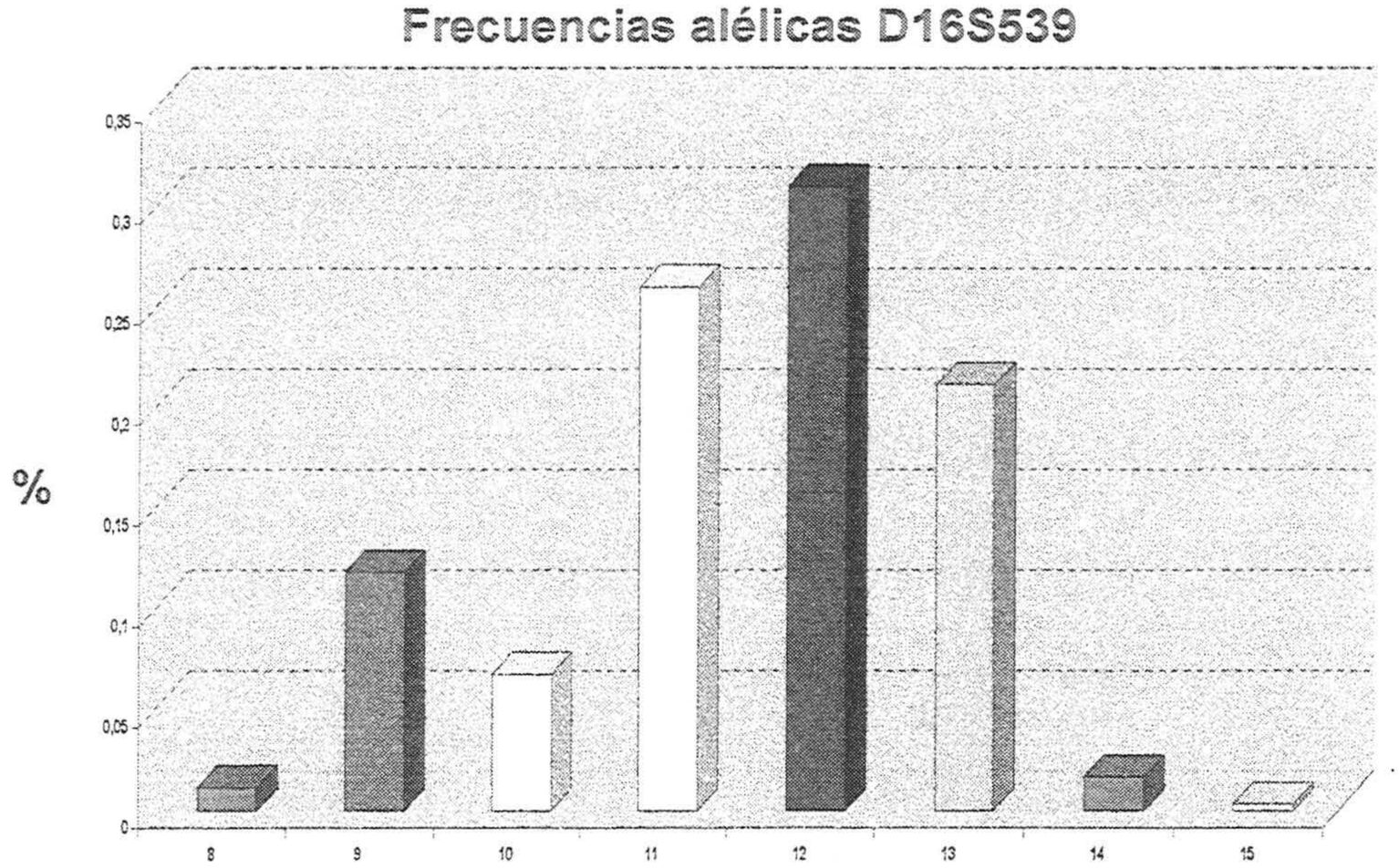
Gráfica 4.1 y 4.2. Frecuencias alélicas y genotípicas D13S317.

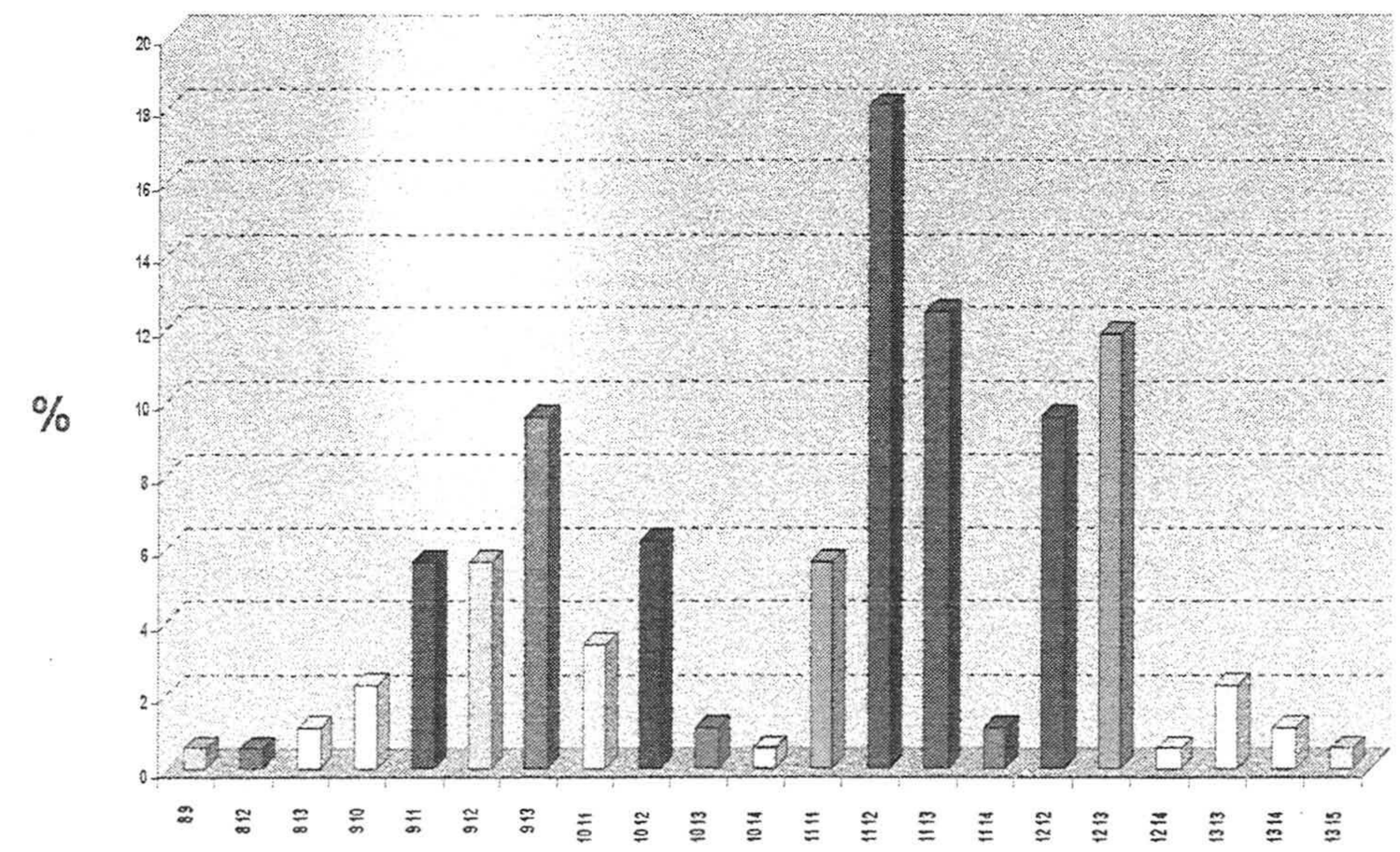


Gráfica 4.3 y 4.4. Frecuencias alélicas y genotípicas D7S280.



Gráfica 4.5 y 4.6. Frecuencias alélicas y genotípicas D16S539.





Como quedó expuesto en el capítulo de material y métodos, los tests aplicados para la demostración de equilibrio Hardy-Weinberg fueron: test de homocigosidad, test exacto de Guo-Thompson y el test de razón de verosimilitudes. Mostramos a continuación los **valores de p** obtenidos con dichos tests para los tres loci:

Tabla 4.6.

	D13S317	**D7S820**	**D16S539**
Homocigosidad observada(%)	20.3	26.0	17.5
Homocigosidad esperada (estimación insesgada)(%)	19.8	20.2	22.6
Test de Homocigosidad	0.853	0.057	0.106
Test de razón de verosimilitud	0.258	0.008	0.097
Test Exacto	0.193	0.003	0.14

Para demostrar la independencia de los alelos el test de correlación de Karlin comparó dos a dos los tres loci obteniendo los siguientes valores de p:

Tabla 4.7.

	D13S317/ D7S820	**D13S317/ D16S539**	**D7S820/ D16S539**
Test de Karlin	0.800	0.168	0.227

También consideramos de interés genético-poblacional, mostrar los valores de la distribución de individuos heterocigotos para uno, dos o los tres loci, asi como los que son homocigotos para los tres loci (ver tabla 4.8.)

Por último, se muestran los valores de probabilidad de discriminación y poder de exclusión de cada uno de los tres loci y del conjunto de los mismos (ver tabla 4.9):

Tabla 4.8.

Nº de loci heterocigotos (x)	Nº individuos con x loci heterocigotos	Frecuencia (%)
0	2	0.01130
1	21	0.11864
2	65	0.36723
3	89	0.50282

Tabla 4.9.

	PD (Observado)	PD (Esperado)	PE
D13S317	0.92342558	0.93192288	0.60999993
D7S820	0.92227649	0.92932327	0.60110246
D16539	0.90306106	0.91207860	0.55789714
Total	0.99942306	0.99957697	0.93122203

4.1.3. Resultados obtenidos con el kit AmpFISTR Profiler Plus

Los resultados de las 171 muestras que fueron analizadas con este kit fueron introducidos de la misma manera que los obtenidos con el triplex de Promega, en el programa informático. Expondremos de manera similar los datos obtenidos a través de dicho programa. En primer lugar se muestran las tablas y gráficos con las frecuencias alélicas y genotípicas obtenidas para cada locus.

Tabla 4.10. Frecuencias alélicas y genotípicas del D3S1358

D3S1358							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
13	0,0029	0.00029	0.00298	**13**	**17**	1	0,005847953
14	0,0999	0.06178	0.15616	**14**	**14**	1	0,005847953
15	0,2164	0.15919	0.28709	**14**	**15**	10	0,058479532
16	0,2895	0.22427	0.36472	**14**	**16**	11	0,064327485
17	0,1901	0.13641	0.2585	**14**	**17**	5	0,029239766
18	0,1696	0.11898	0.23596	**14**	**18**	5	0,029239766
19	0,0292	0.01213	0.06881	**14**	**19**	1	0,005847953
>19	0,0024	0.0003	0.029	**15**	**15**	9	0,052631579
				15	**16**	19	0,111111111
				15	**17**	13	0,076023392
				15	**18**	10	0,058479532
				15	**19**	4	0,023391813
				16	**16**	12	0,070175439
				16	**17**	24	0,140350877
				16	**18**	18	0,105263158
				16	**19**	2	0,011695906
				16	**>19**	1	0,005847953
				17	**17**	4	0,023391813
				17	**18**	13	0,076023392
				17	**19**	1	0,005847953
				18	**18**	5	0,029239766
				18	**19**	2	0,011695906
Total	**1**					**171**	**1**

Tabla 4.11. Frecuencias alélicas y genotípicas del locus VWA

VWA							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
12	0,0029	0.0003	0.0292	**12**	**18**	1	0,005847953
14	0,1199	0.0779	0.1799	**14**	**14**	2	0,011695906
15	0,1257	0.0826	0.1866	**14**	**15**	4	0,023391813
16	0,2368	0.1771	0.3090	**14**	**16**	8	0,046783626
17	0,2456	0.1849	0.3184	**14**	**17**	15	0,087719298
18	0,1754	0.1239	0.2424	**14**	**18**	9	0,052631579
19	0,076	0.0439	0.1283	**14**	**19**	1	0,005847953
20	0,0175	0.0057	0.0523	**15**	**15**	4	0,023391813
				15	**16**	11	0,064327485
				15	**17**	5	0,029239766
				15	**18**	13	0,076023392
				15	**19**	1	0,005847953
				15	**20**	1	0,005847953
				16	**16**	11	0,064327485
				16	**17**	18	0,105263158
				16	**18**	15	0,087719298
				16	**19**	7	0,040935673
				17	**17**	11	0,064327485
				17	**18**	12	0,070175439
				17	**19**	10	0,058479532
				17	**20**	2	0,011695906
				18	**18**	2	0,011695906
				18	**19**	5	0,029239766
				18	**20**	1	0,005847953
				19	**20**	2	0,011695906
Total	**0,9998**					**171**	**1**

Tabla 4.12. Frecuencias alélicas y genotípicas del locus FGA.

FGA							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
18	0,0087	0.0017	0.0432	**18**	**22**	2	0,011695906
19	0,0789	0.0441	0.1372	**18**	**23**	1	0,005847953
20	0,1579	0.1057	0.2291	**19**	**19**	1	0,005847953
21	0,1579	0.1057	0.2291	**19**	**20**	3	0,01754386
22	0,1579	0.1057	0.2291	**19**	**21**	3	0,01754386
22.2	0,0088	0.0017	0.0432	**19**	**22**	4	0,023391813
23	0,1345	0.0867	0.2026	**19**	**23**	5	0,029239766
23.2	0,0029	0.0002	0.0333	**19**	**24**	5	0,029239766
23.3	0,0029	0.0002	0.0333	**19**	**25**	3	0,01754386
24	0,152	0.1009	0.2225	**19**	**26**	2	0,011695906
25	0,0906	0.0527	0.1513	**20**	**20**	4	0,023391813
26	0,0321	0.0129	0.0774	**20**	**21**	8	0,046783626
27	0,0117	0.0027	0.0478	**20**	**22**	15	0,087719298
28	0,0029	0.0002	0.0333	**20**	**23**	4	0,023391813
				20	**24**	8	0,046783626
				20	**24**	8	0,046783626
				20	**25**	8	0,046783626
				21	**21**	6	0,035087719
				21	**22**	2	0,011695906
				21	**22,2**	2	0,011695906
				21	**23**	7	0,040935673
				21	**23,2**	1	0,005847953
				21	**24**	12	0,070175439
				21	**25**	2	0,011695906
				21	**26**	2	0,011695906
				21	**27**	2	0,011695906
				21	**28**	1	0,005847953
				22	**22**	1	0,005847953
				22	**23**	12	0,070175439
				22	**24**	8	0,046783626
				22	**25**	3	0,01754386
				22	**26**	5	0,029239766
				22	**27**	1	0,005847953
				22,2	**24**	1	0,005847953
				23	**23**	2	0,011695906
				23	**24**	5	0,029239766
				23	**25**	7	0,040935673
				23	**26**	1	0,005847953
				23.2	24	1	0.005847953

FGA							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
				24	**24**	3	0.0175438
				24	**25**	6	0.0350877
				25	**26**	1	0.0058479
				25	**27**	1	0.0058479
Total	**0,9997**					**171**	**1**

Tabla 4.13. Frecuencias alélicas y genotípicas del locus D8S1179.

D8S1179							
Alelos	Frecuencias alélicas	Intervalos de confianza		Genotipos		Número	Frecuencias
		Inferior	Superior				
8	0,017	0.0059	0.0509	8	10	1	0,005847953
9	0,015	0.0044	0.0466	8	12	1	0,005847953
10	0,1111	0.0719	0.1678	8	13	3	0,01754386
11	0,0848	0.0513	0.1369	8	15	1	0,005847953
12	0,1199	0.0789	0.1779	9	11	1	0,005847953
13	0,2544	0.1943	0.3255	9	12	1	0,005847953
14	0,2281	0.1709	0.2974	9	15	2	0,011695906
15	0,1462	0.1005	0.2077	9	16	1	0,005847953
16	0,0205	0.0074	0.0055	10	10	5	0,029239766
17	0,0029	0.0003	0.00279	10	11	2	0,011695906
				10	12	5	0,029239766
				10	13	13	0,076023392
				10	14	4	0,023391813
				10	15	2	0,011695906
				10	16	1	0,005847953
				11	11	1	0,005847953
				11	12	3	0,01754386
				11	13	10	0,058479532
				11	14	5	0,029239766
				11	15	5	0,029239766
				11	16	1	0,005847953
				12	12	5	0,029239766
				12	13	5	0,029239766
				12	14	9	0,052631579
				12	15	6	0,035087719
				12	16	1	0,005847953
				13	13	9	0,052631579
				13	14	22	0,128654971
				13	15	14	0,081871345
				13	16	2	0,011695906
				14	14	11	0,064327485
				14	15	15	0,087719298
				14	17	1	0,005847953
				15	15	2	0,011695906
				15	16	1	0,005847953
Total	0,9999					171	1

Tabla 4.14. Frecuencias alélicas y genotípicas del locus D21S11.

D21S11							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
26	0,0029			**26**	**28**	1	0,005847953
27	0,0175			**27**	**28**	2	0,011695906
28	0,1403			**27**	**29**	1	0,005847953
28.2	0,0029			**27**	**30**	1	0,005847953
29	0,1959			**27**	**31,2**	2	0,011695906
30	0,2719			**28**	**28**	2	0,011695906
30.2	0,0351			**28**	**29**	11	0,064327485
31	0,0585			**28**	**30**	12	0,070175439
31.2	0,1374			**28**	**30,2**	2	0,011695906
32	0,0088			**28**	**31**	1	0,005847953
32.2	0,0819			**28**	**31,2**	7	0,040935673
33.2	0,0292			**28**	**32**	2	0,011695906
34.2	0,0088			**28**	**32,2**	4	0,023391813
35	0,0029			**28**	**34,2**	2	0,011695906
35.2	0,0058			**28,2**	**33,2**	1	0,005847953
				29	**29**	4	0,023391813
				29	**30**	17	0,099415205
				29	**30,2**	4	0,023391813
				29	**31**	5	0,029239766
				29	**31,2**	11	0,064327485
				29	**32,2**	8	0,046783626
				29	**33,2**	2	0,011695906
				30	**30**	16	0,093567251
				30	**31**	5	0,029239766
				30	**31,2**	15	0,087719298
				30	**32**	1	0,005847953
				30	**32,2**	5	0,029239766
				30	**33,2**	3	0,01754386
				30	**35**	1	0,005847953
				30	**35,2**	1	0,005847953
				30,2	**30,2**	1	0,005847953
				30,2	**31,2**	2	0,011695906
				30,2	**32,2**	2	0,011695906
				31	**31**	1	0,005847953
				31	**31,2**	3	0,01754386
				31	**32,2**	1	0,005847953
				31	**33,2**	2	0,011695906
				31	**34,2**	1	0,005847953

D21S11							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
				31.2	**31.2**	1	0,005847953
				31.2	**32.2**	4	0,023391813
				31.2	**33.2**	1	0,005847953
				32.2	**32.2**	1	0,005847953
				32.2	**33.2**	1	0,005847953
				32.2	**35.2**	1	0,005847953
Total	**0,9998**					**171**	**1**

Tabla 4.15. Frecuencias alélicas y genotípicas del locus D18S51.

D18S51							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
10	0,0058	0.0009	0.0035	**10**	**12**	1	0,005847953
11	0,0205	0.0071	0.0573	**10**	**15**	1	0,005847953
12	0,1433	0.0964	0.2076	**11**	**13**	1	0,005847953
13	0,0936	0.0568	0.1503	**11**	**14**	3	0,01754386
14	0,1842	0.1307	0.2532	**11**	**15**	1	0,005847953
15	0,1579	0.1085	0.2240	**11**	**16**	1	0,005847953
16	0,114	0.0728	0.1742	**11**	**17**	1	0,005847953
17	0,1403	0.0940	0.2043	**12**	**12**	4	0,023391813
18	0,0789	0.0457	0.1328	**12**	**13**	4	0,023391813
19	0,0322	0.0137	0.0736	**12**	**14**	6	0,035087719
20	0,0234	0.0086	0.0615	**12**	**15**	8	0,046783626
22	0,0058	0.0009	0.0035	**12**	**16**	4	0,023391813
				12	**17**	10	0,058479532
				12	**18**	3	0,01754386
				12	**19**	2	0,011695906
				12	**20**	2	0,011695906
				12	**22**	1	0,005847953
				13	**13**	1	0,005847953
				13	**14**	6	0,035087719
				13	**15**	7	0,040935673
				13	**16**	6	0,035087719
				13	**17**	3	0,01754386
				13	**18**	1	0,005847953
				13	**19**	1	0,005847953
				13	**20**	1	0,005847953
				14	**14**	3	0,01754386
				14	**15**	10	0,058479532
				14	**16**	7	0,040935673
				14	**17**	14	0,081871345
				14	**18**	9	0,052631579
				14	**19**	1	0,005847953
				14	**20**	1	0,005847953
				15	**15**	6	0,035087719
				15	**16**	7	0,040935673
				15	**17**	7	0,040935673
				15	**18**	1	0,005847953
				16	**16**	1	0,005847953

D18S51							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
				16	**17**	4	0,023391813
				16	**18**	4	0,023391813
				16	**19**	3	0,01754386
				16	**22**	1	0,005847953
				17	**17**	2	0,011695906
				17	**18**	1	0,005847953
				17	**19**	1	0,005847953
				17	**20**	3	0,01754386
				18	**18**	3	0,01754386
				18	**19**	2	0,011695906
				19	**20**	1	0,005847953
Total	**0,9999**					**171**	**1**

Tabla 4.16. Frecuencias alélicas y genotípicas del locus D5S518.

D5S818							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
9	0,0234	0.0088	0.0603	**9**	**10**	2	0,011695906
10	0,0819	0.0485	0.1347	**9**	**11**	3	0,01754386
11	0,345	0.2757	0.4215	**9**	**13**	3	0,01754386
12	0,345	0.2757	0.4215	**10**	**11**	10	0,058479532
13	0,1813	0.1293	0.2482	**10**	**12**	10	0,058479532
14	0,0117	0.0031	0.0432	**10**	**13**	6	0,035087719
15	0,0117	0.0031	0.0432	**11**	**11**	22	0,128654971
				11	**12**	40	0,233918129
				11	**13**	18	0,105263158
				11	**14**	1	0,005847953
				11	**15**	2	0,011695906
				12	**12**	24	0,140350877
				12	**13**	19	0,111111111
				12	**15**	1	0,005847953
				13	**13**	6	0,035087719
				13	**14**	3	0,01754386
				13	**15**	1	0,005847953
Total	**1**					**171**	**1**

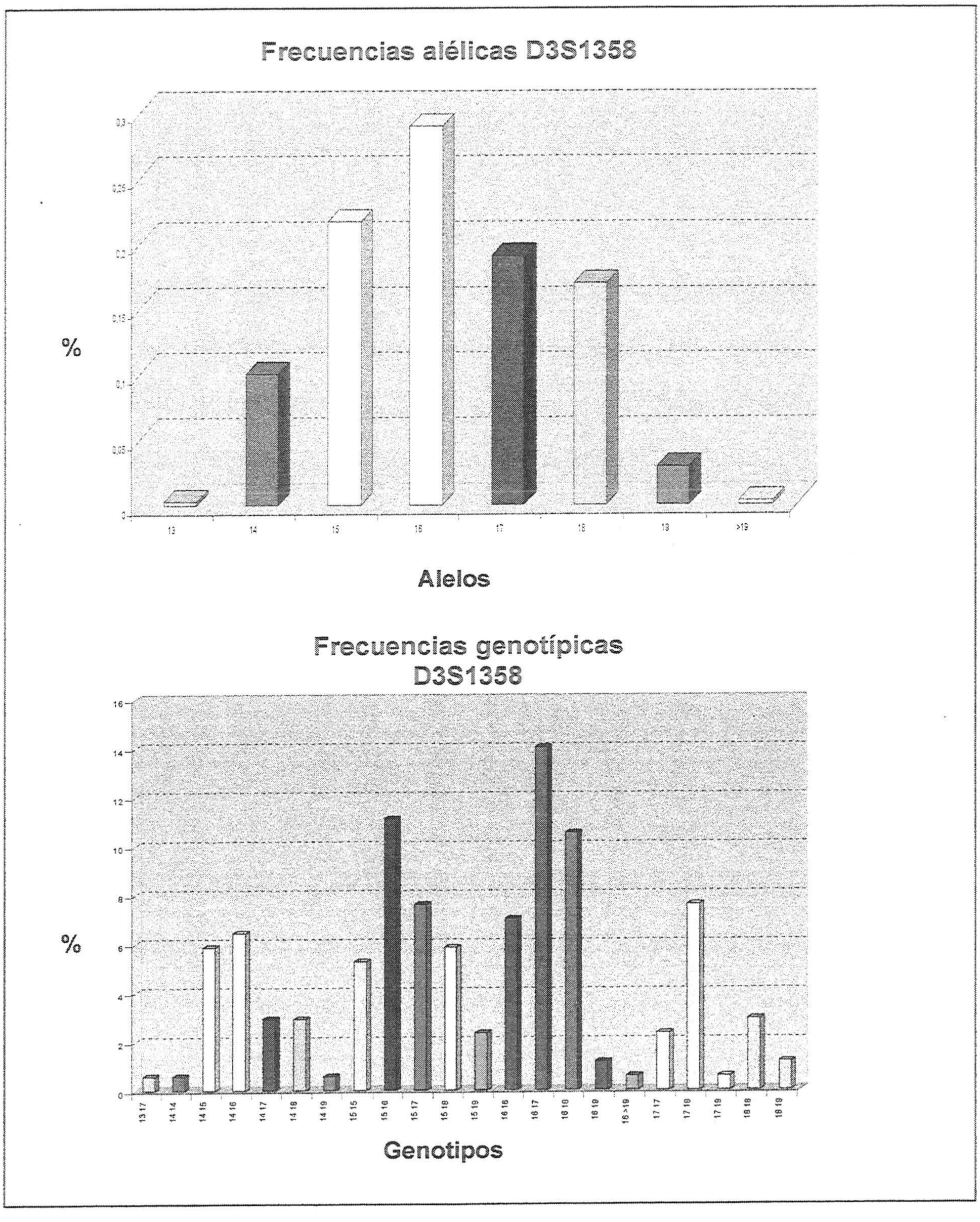
Tabla 4.17. Frecuencias alélicas y genotípicas del locus D13S317.

D13S317							
Alelos	**Frecuencias alélicas**	**Intervalos de confianza**		**Genotipos**		**Número**	**Frecuencias**
		Inferior	Superior				
8	0,1813	0.1297	0.2475	**8**	**8**	8	0,0467836
9	0,0643	0.0358	0.1129	**8**	**9**	1	0,005848
10	0,0614	0.0357	0.1093	**8**	**10**	6	0,0350877
11	0,2778	0.2147	0.3510	**8**	**11**	14	0,0818713
12	0,2602	0.1989	0.3325	**8**	**12**	13	0,0760234
13	0,1053	0.0669	0.1617	**8**	**13**	8	0,0467836
14	0,0497	0.0255	0.0946	**8**	**14**	4	0,0233918
				9	**9**	2	0,0116959
				9	**11**	6	0,0350877
				9	**12**	9	0,0526316
				9	**13**	1	0,005848
				9	**14**	1	0,005848
				10	**11**	5	0,0292398
				10	**12**	8	0,0467836
				10	**13**	1	0,005848
				10	**14**	1	0,005848
				11	**11**	13	0,0760234
				11	**12**	33	0,1929825
				11	**13**	9	0,0526316
				11	**14**	2	0,0116959
				12	**12**	7	0,0409357
				12	**13**	7	0,0409357
				12	**14**	5	0,0292398
				13	**13**	4	0,0233918
				13	**14**	2	0,0116959
				14	**14**	1	0,005848
Total	**1**					**171**	**1**

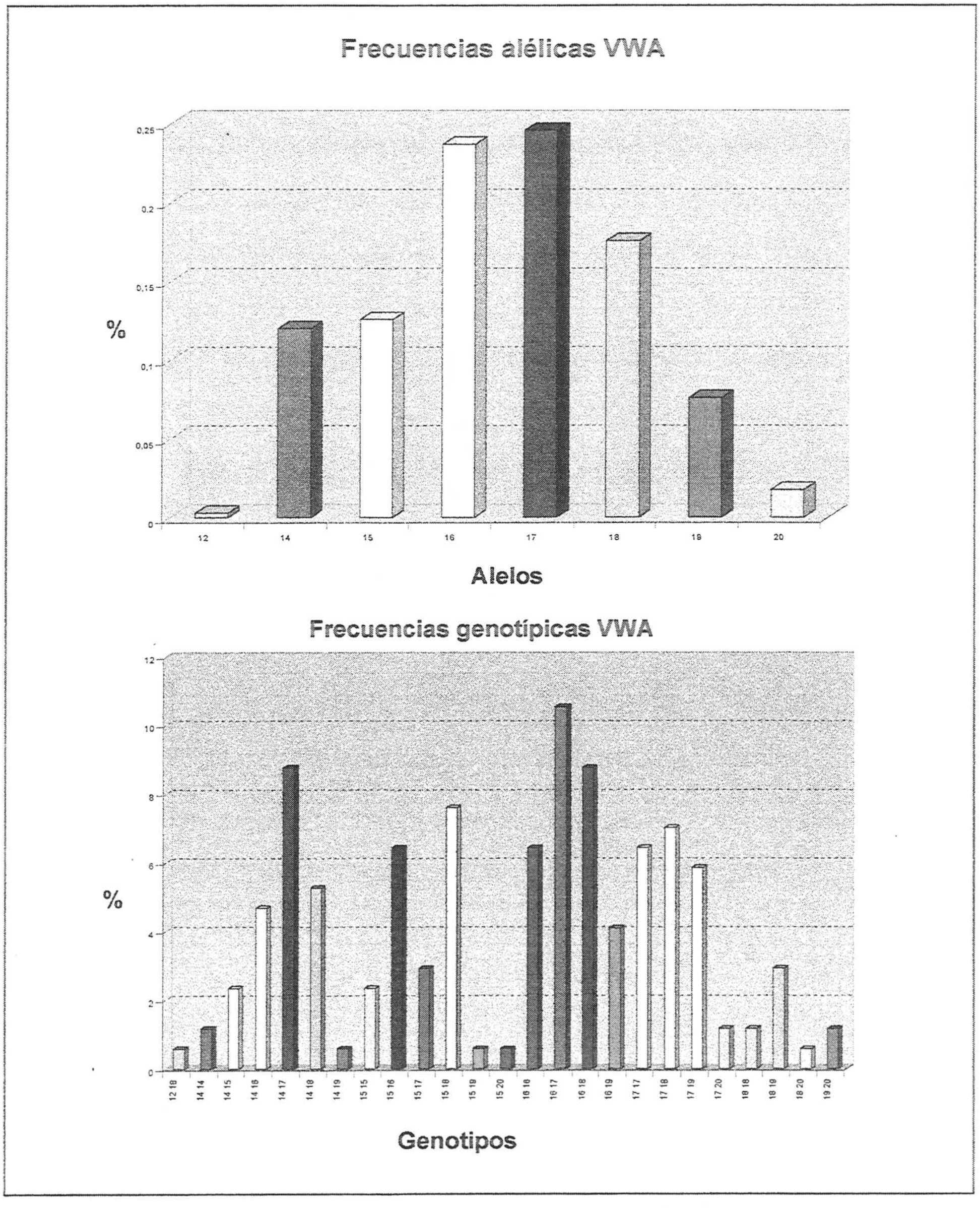
Tabla 4.18. Frecuencias alélicas y genotípicas del locus D7S820.

D7S820							
Alelos	Frecuencias alélicas	Intervalos de confianza		Genotipos		Número	Frecuencias
		Inferior	Superior				
7	0,0263	0.0107	0.0632	**7**	**8**	2	0,0116959
8	0,152	0.1054	0.2143	**7**	**9**	2	0,0116959
9	0,1345	0.0908	0.1946	**7**	**11**	3	0,0175439
10	0,307	0.2420	0.3807	**7**	**12**	2	0,0116959
11	0,1959	0.1428	0.2627	**8**	**8**	3	0,0175439
12	0,1608	0.1128	0.2241	**8**	**9**	12	0,0701754
13	0,0234	0.0090	0.0592	**8**	**10**	12	0,0701754
				8	**11**	4	0,0233918
				8	**12**	13	0,0760234
				8	**13**	3	0,0175439
				9	**9**	3	0,0175439
				9	**10**	8	0,0467836
				9	**11**	11	0,0643275
				9	**12**	7	0,0409357
				10	**10**	22	0,128655
				10	**11**	24	0,1403509
				10	**12**	16	0,0935673
				10	**13**	1	0,005848
				11	**11**	10	0,0584795
				11	**12**	5	0,0292398
				12	**12**	5	0,0292398
				12	**13**	2	0,0116959
				13	**13**	1	0,005848
Total	**0,9999**					**171**	**1**

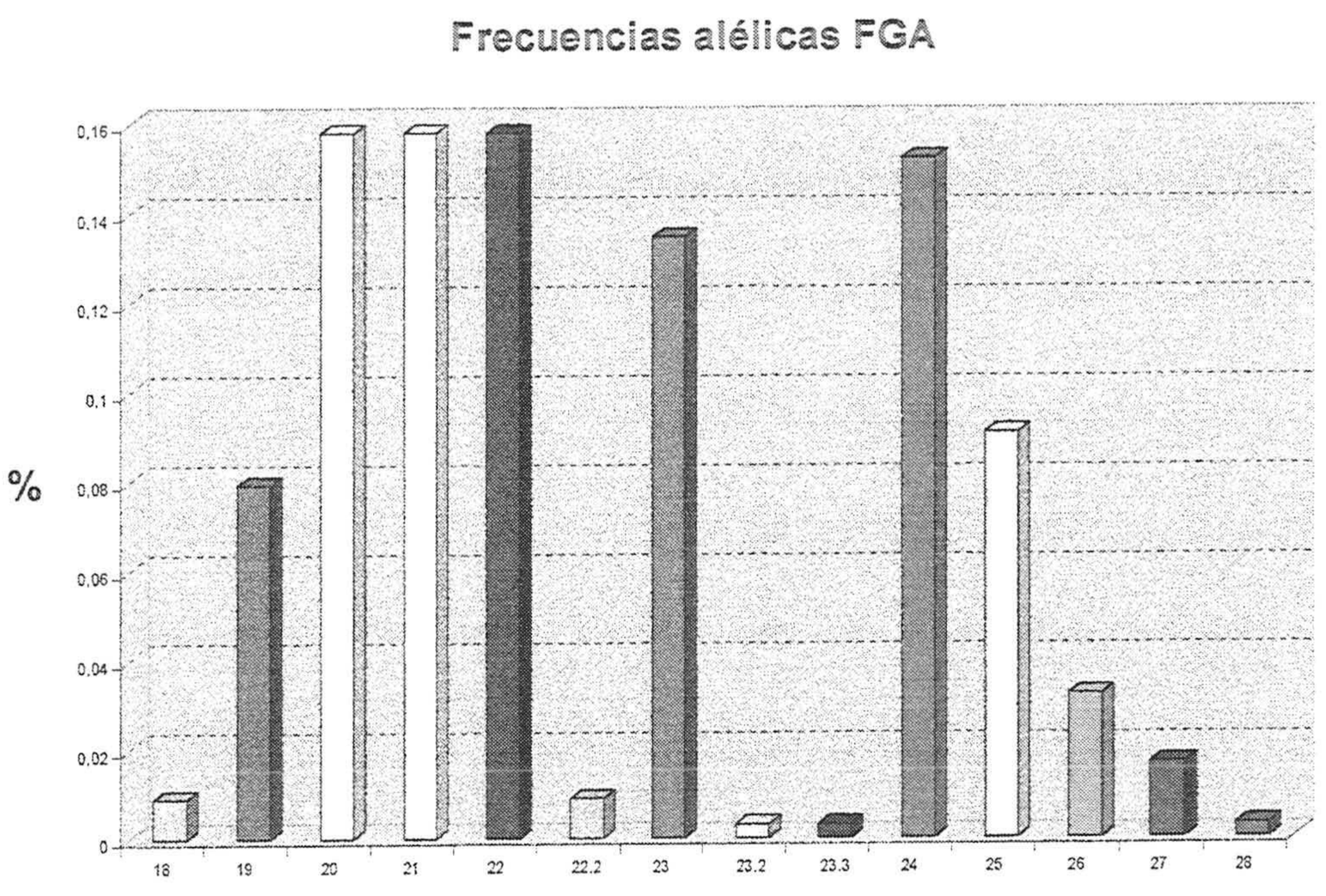
Gráfica 4.7 y 4.8. Frecuencias alélicas y genotípicas del locus D3S1358.

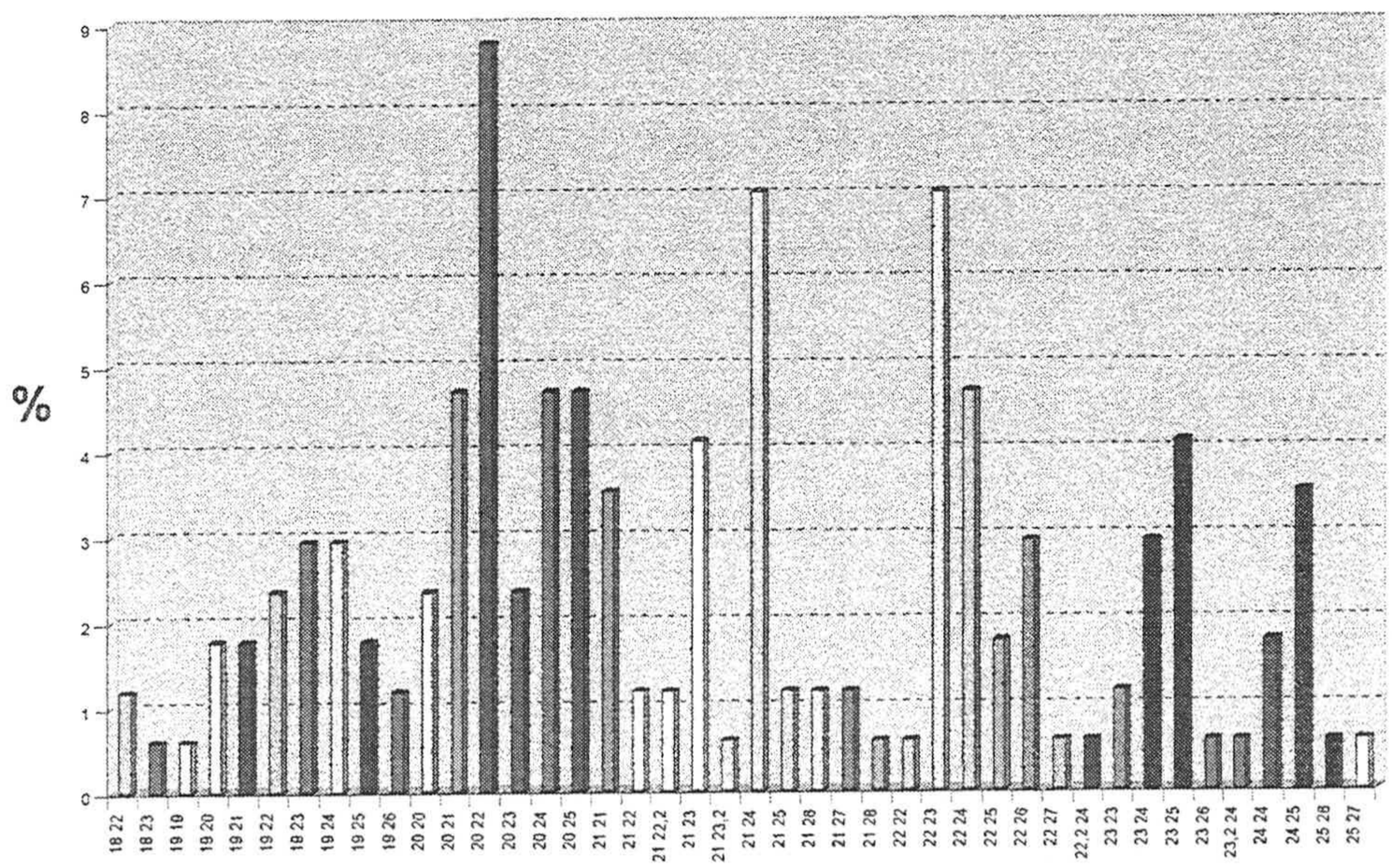


Gráfica 4.9 y 4.10. Frecuencias alélicas y genotípicas del locus VWA.

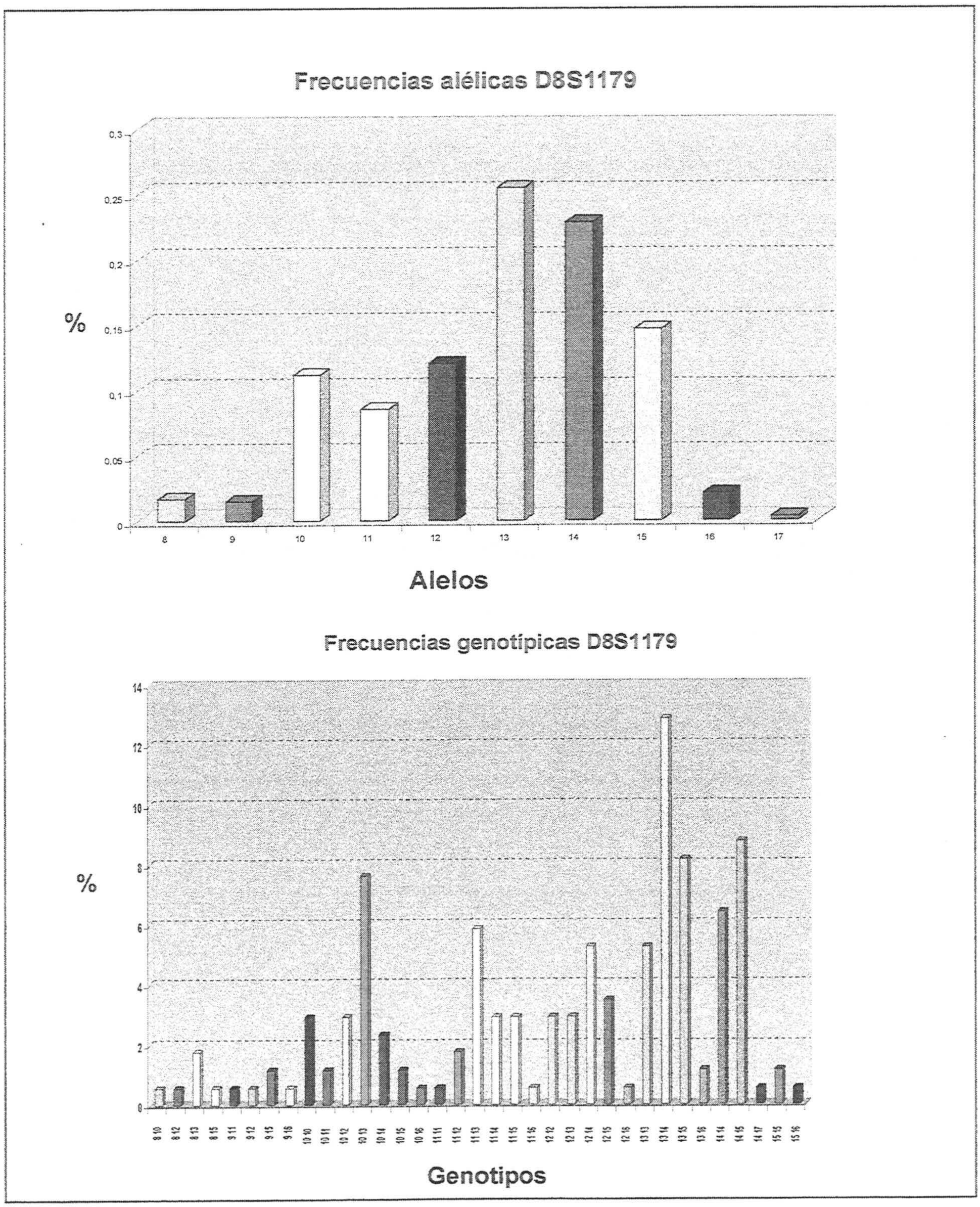


Gráfica 4.11 y 4.12. Frecuencias alélicas y genotípicas del locus FGA.

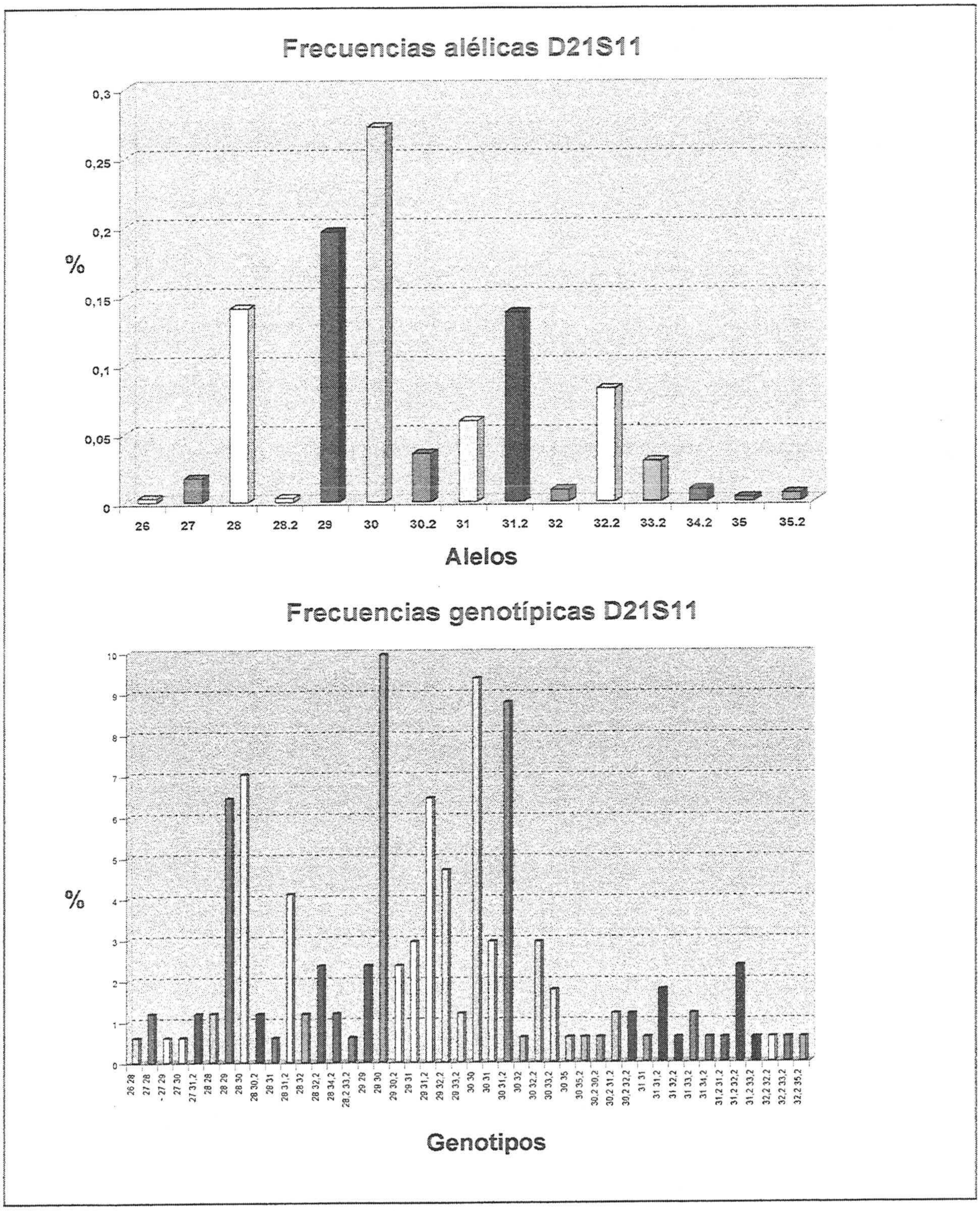




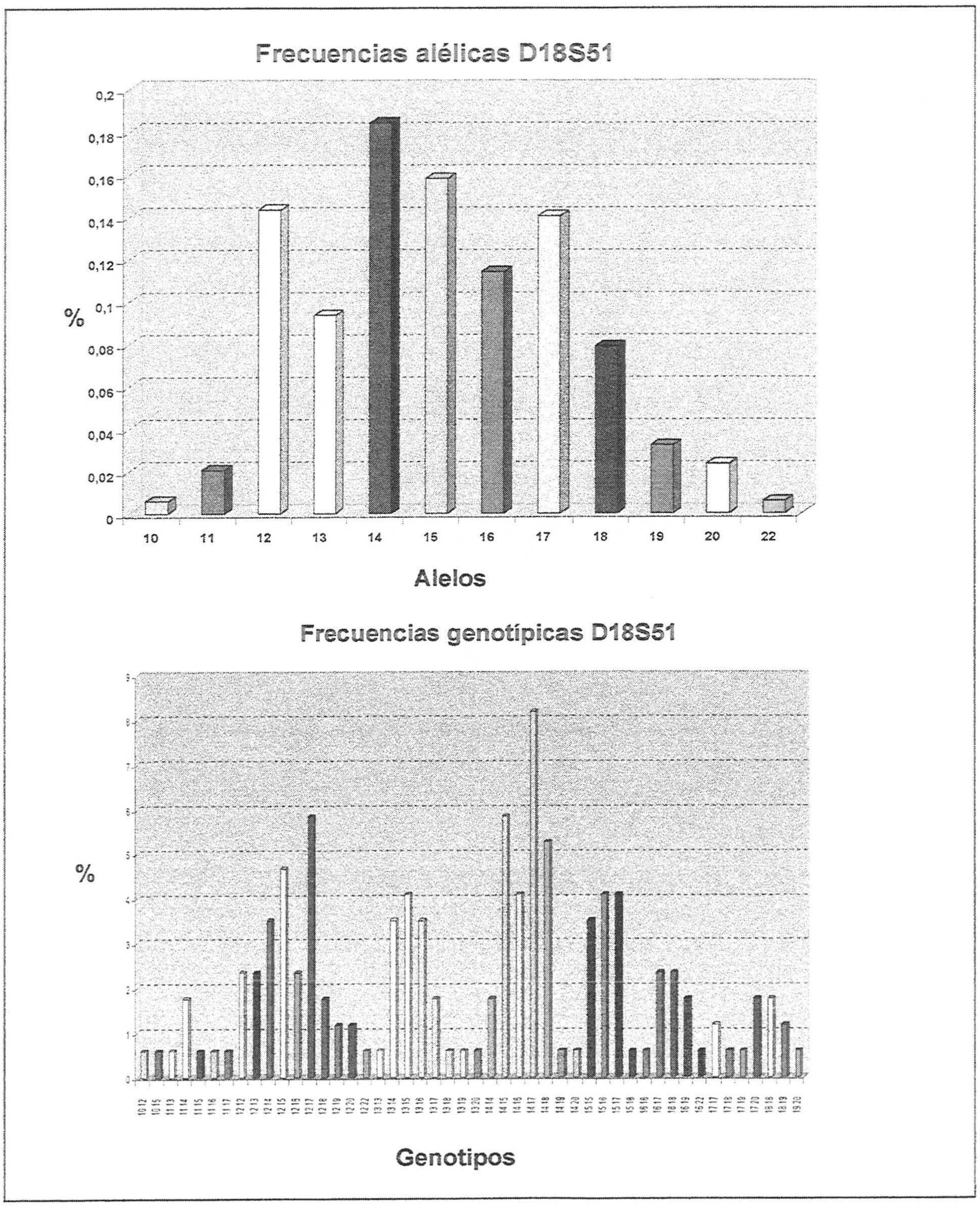
Gráfica 4.13 y 4.14. Frecuencias alélicas y genotípicas del locus D8S1179.



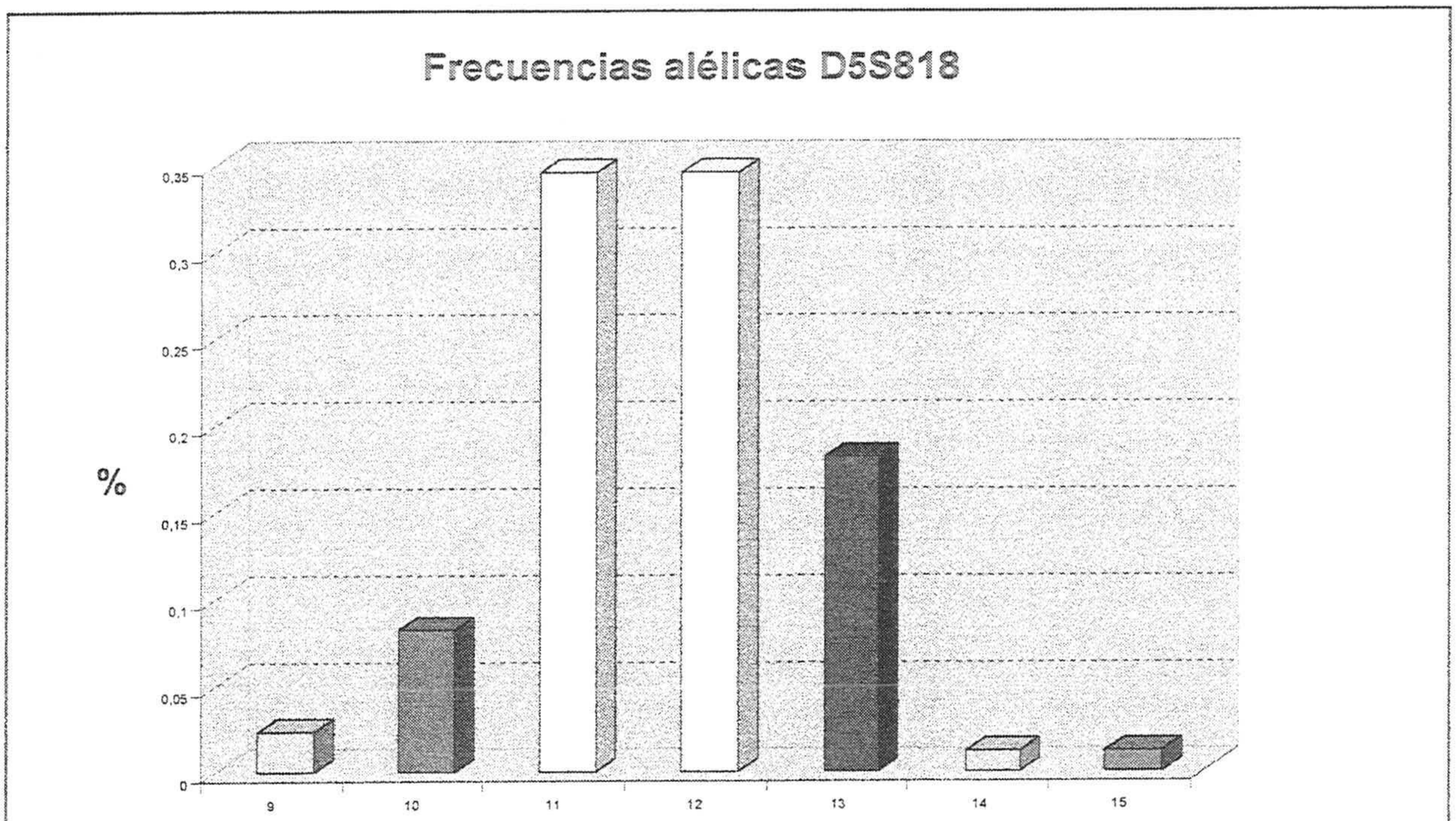
Gráfica 4.15 y 4.16. Frecuencias alélicas y genotípicas del locus D21S11.



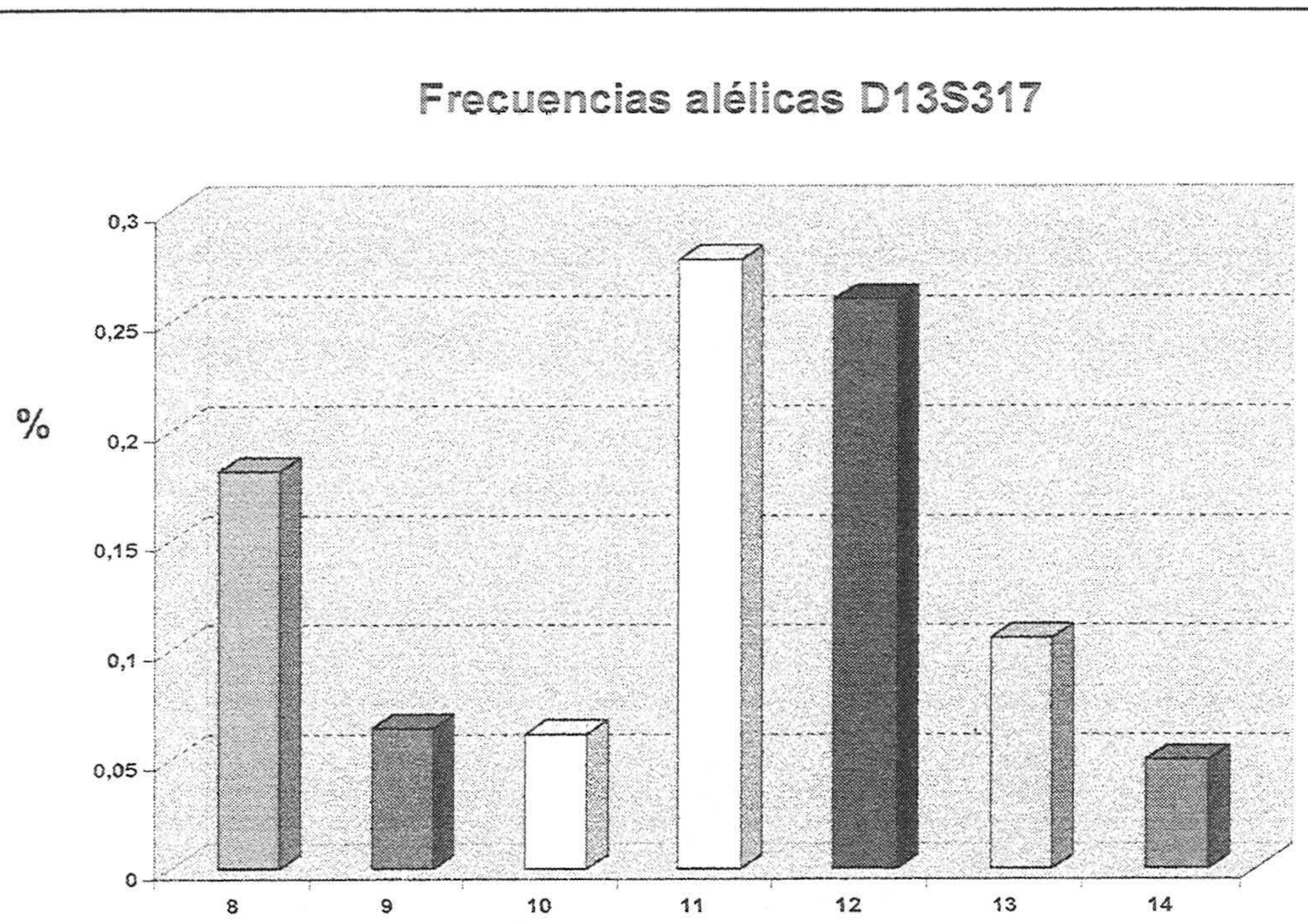
Gráfica 4.17 y 4.18. Frecuencias alélicas y genotípicas del locus D18S51.

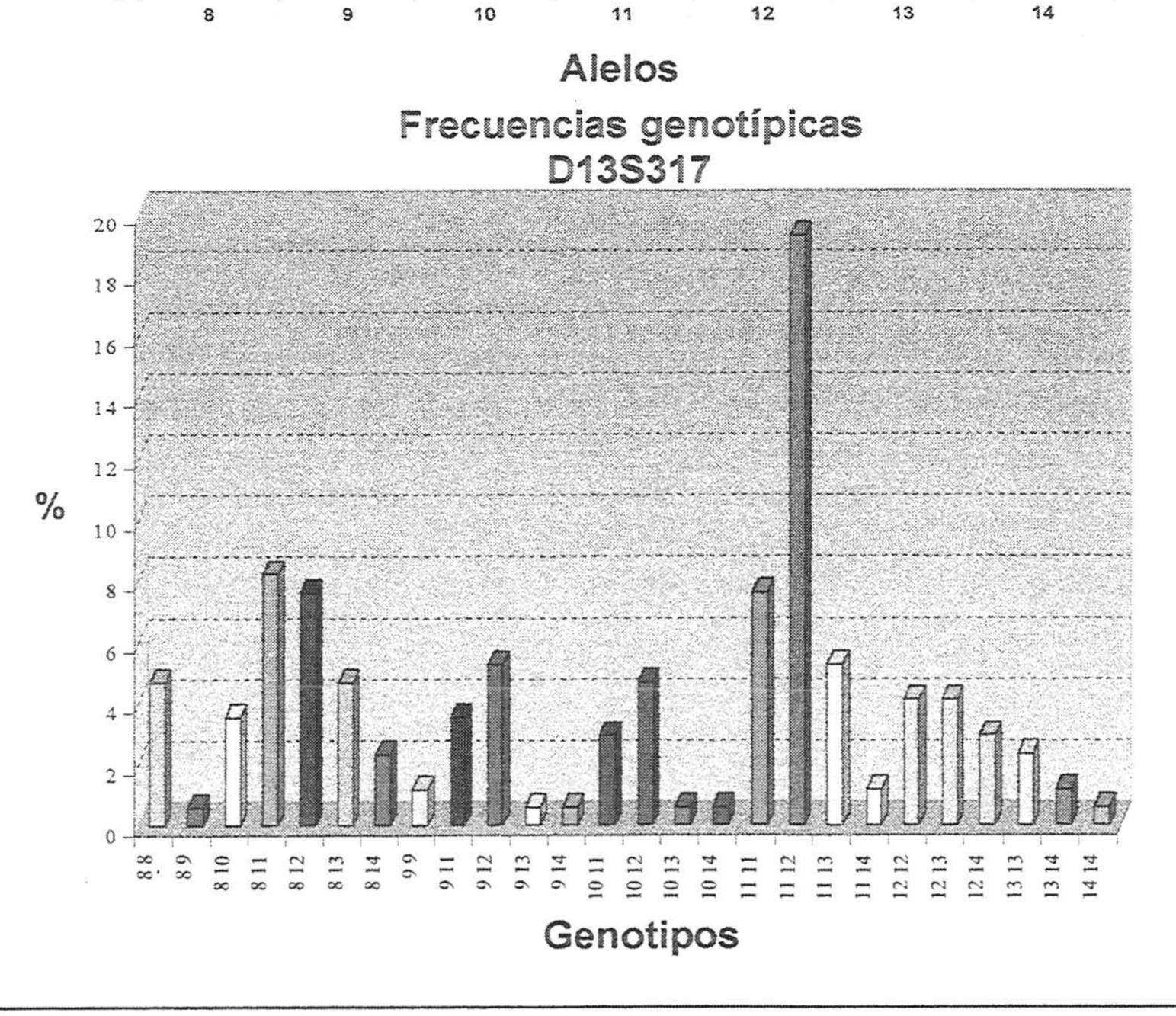


Gráfica 4.19 y 4.20. Frecuencias alélicas y genotípicas del locus D5S818.

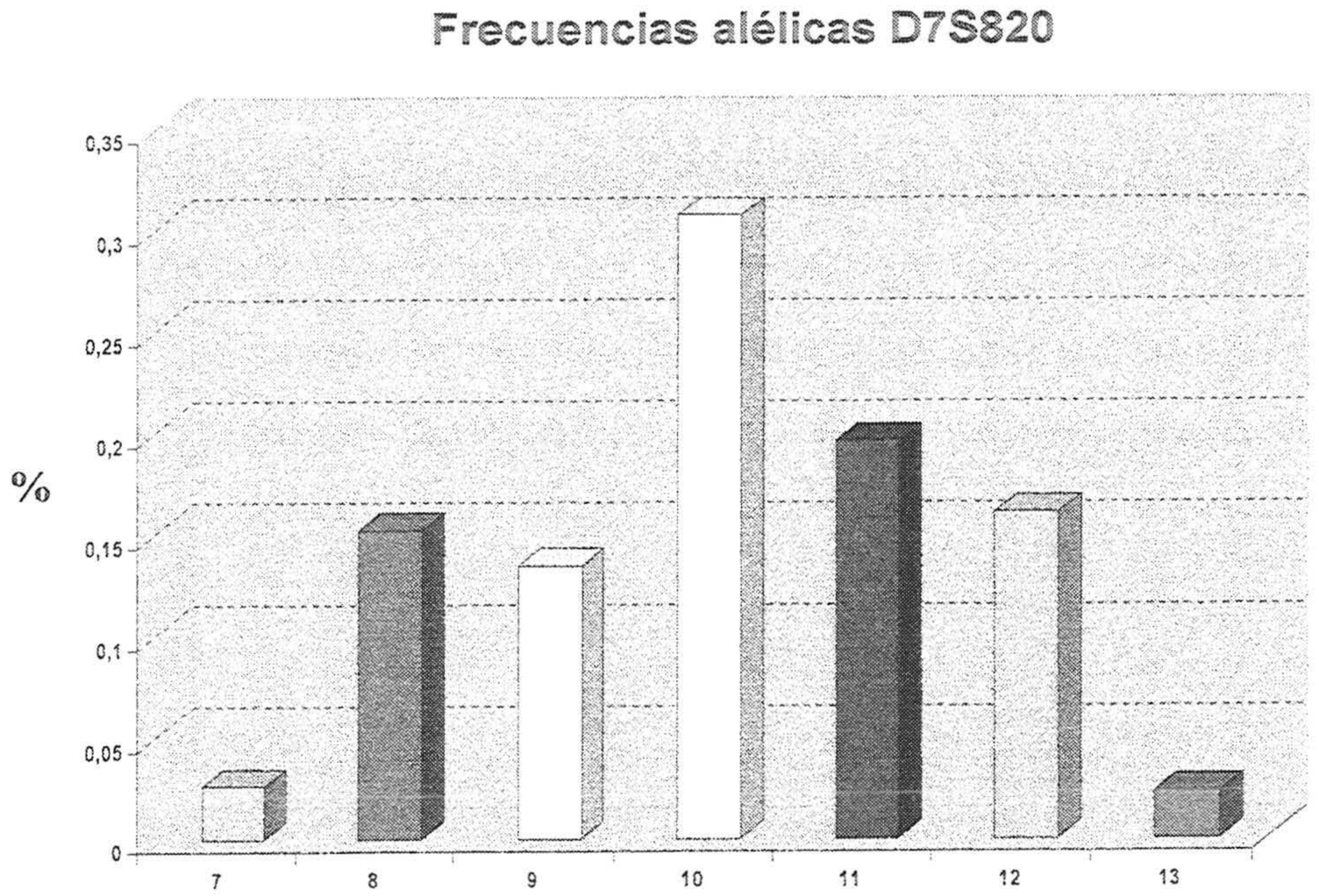


Gráfica 4.21 y 4.22. Frecuencias alélicas y genotípicas del locus D13S317.





Gráfica 4.23 y 4.24. Frecuencias alélicas y genotípicas del locus D7S820.



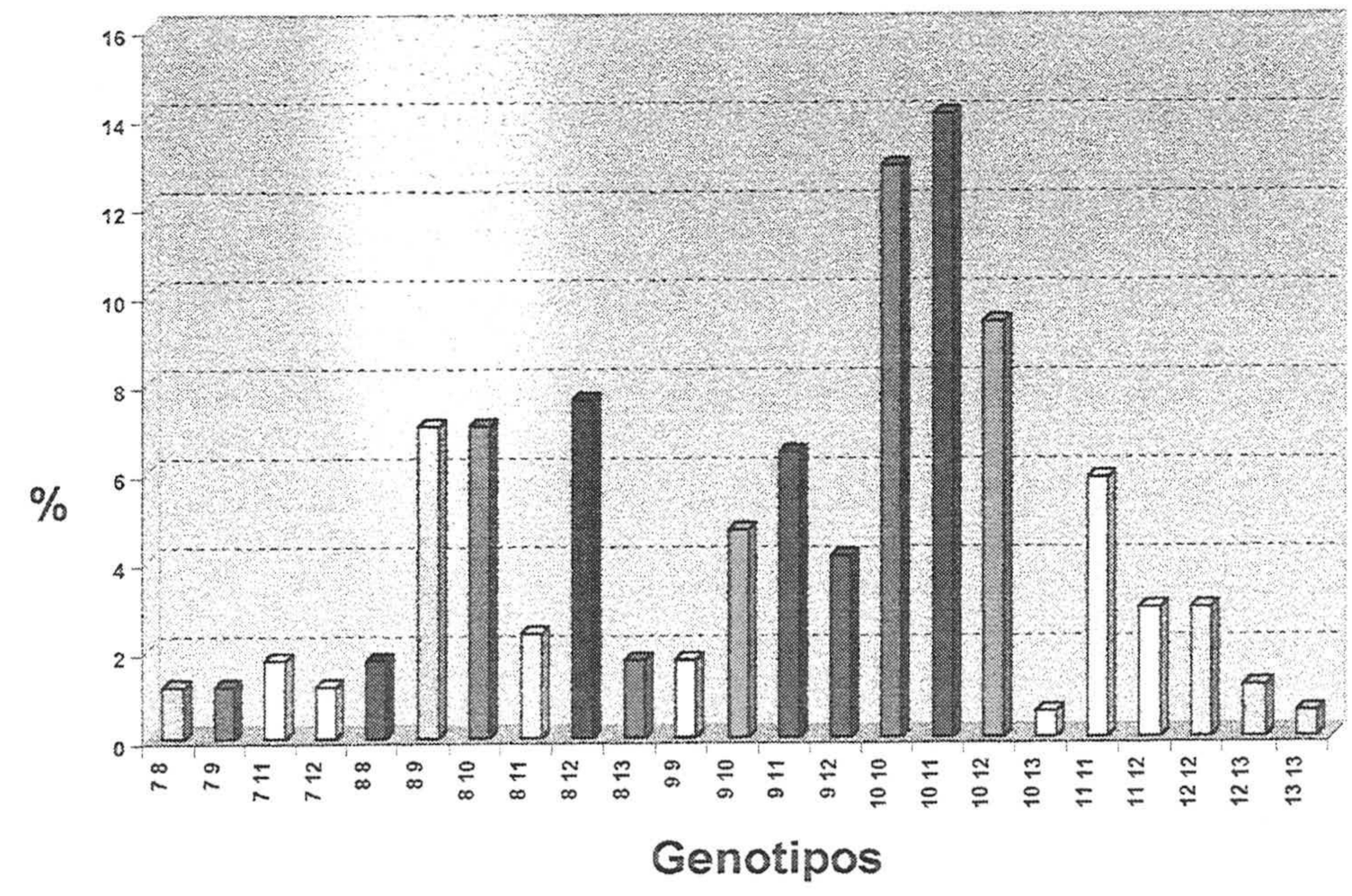


Tabla 4.19. Comparación entre el porcentaje de heterocigotos observado y esperado según un estimador sesgado.

	Heterocigosidad observada (%)	Heterocigosidad esperada (sesgada) (%)	χ^2	Valor p
D3S1358	81.87	79.37	0.650995	0.419758
VWA	82.46	81.65	0.073723	0.785992
FGA	90.06	86.82	1.567375	0.210588
D8S1179	80.70	82.70	0.479286	0.488746
D21S11	84.80	83.64	0.167930	0.681958
D18S51	88.30	87.08	0.225991	0.634513
D5S818	69.59	72.15	0.558385	0.454912
D13S317	79.53	80.08	0.032094	0.857821
D7S820	74.27	79.90	3.382696	0.065884

Tabla 4.20. Comparación entre el porcentaje de heterocigotos observado y esperado según un estimador insesgado.

	Heterocigosidad observada (%)	Heterocigosidad esperada (no sesgada) (%)	χ^2	Valor p
D3S1358	81.87	79.61	0.539785	0.462522
VWA	82.46	81.89	0.036709	0.848059
FGA	90.06	87.07	1.352866	0.244778
D8S1179	80.70	82.95	0.609207	0.435087
D21S11	84.80	83.88	0.105670	0.745128
D18S51	88.30	87.34	0.143658	0.704671
D5S818	69.59	72.36	0.657528	0.417434
D13S317	79.53	80.31	0.066142	0.797039
D7S820	74.27	80.14	3.702270	0.054339

Los valores de probabilidad (p) obtenidos en los tests de equilibrio Hardy-Weinberg para los nueve marcadores así como los valores de probabilidad del test de Karlin, se muestran a continuación:

Tabla 4.21.

	D3S1358	VWA	FGA
Homocigosidad observada(%)	18.1	17.5	9.9
Homocigosidad esperada (estimación insesgada)(%)	20.4	18.1	12.9
Test de Homocigosidad	0.463	0.848	0.245
Test de razón de verosimilitud	0.955	0.110	0.052
Test Exacto	0.955	0.133	0.071

	D8S1179	D21S11	D18S51
Homocigosidad observada(%)	19.3	15.2	11.7
Homocigosidad esperada (estimación insesgada)(%)	17.1	16.1	12.7
Test de Homocigosidad	0.435	0.745	0.705
Test de razón de verosimilitud	0.278	0.531	0.505
Test Exacto	0.130	0.487	0.425

	D5S818	D13S317	D7S820
Homocigosidad observada(%)	30.4	20.5	25.7
Homocigosidad esperada (estimación insesgada)(%)	27.6	19.7	19.9
Test de Homocigosidad	0.417	0.797	0.054
Test de razón de verosimilitud	0.268	0.247	0.006
Test Exacto	0.331	0.151	0.003

Tabla 4.22. Resultados del test de Karlin.

LOCI	p
D3S1358 / VWA	0.773
D3S1358 / FGA	0.518
D3S1358 / D8S1179	0.020
D3S1358 / D21S11	0.463
D3S1358 / D18S51	0.652
D3S1358 / D5S818	0.739
D3S1358 / D13S317	0.118
D3S1358 / D7S820	0.018
VWA / FGA	0.578
VWA / D8S1179	0.327
VWA / D21S11	0.556
VWA / D18S51	0.962
VWA / D5S818	0.002
VWA / D13S317	0.031
VWA / D7S820	0.759
FGA / D8S1179	0.492
FGA / D21S11	0.915

LOCI	P
FGA / D18S51	0.163
FGA / D5S818	0.784
FGA / D13S317	0.024
FGA / D7S820	0.561
D8S1179 / D21S11	0.271
D8S1179 / D18S51	0.571
D8S1179 / D5S818	0.001
D8S1179 / D13S317	0.974
D8S1179 / D7S820	0.844
D21S11 / D18S51	0.263
D21S11 / D5S818	0.048
D21S11 / D13S317	0.548
D21S11 / D7S820	0.699
D18S51 / D5S818	0.899
D18S51 / D13S317	0.585
D18S51 / D7S820	0.563
D5S818 / D13S317	0.373
D5S818 / D7S820	0.313
D13S317 / D7S820	0.954

Tabla 4.23. Distribución de individuos heterocigotos para x loci.

Nº de loci heterocigotos (x)	Nº individuos con x loci heterocigotos	Frecuencia
0	0	0
1	0	0
2	0	0
3	1	0.00585
4	1	0.00585
5	10	0.05848
6	31	0.18129
7	42	0.24561
8	60	0.35088
9	26	0.15205

Tabla 4.24. Valores de Probabilidad de Discriminación (PD) y Poder de Exclusión (PE).

	PD (Observado)	PD (Esperado)	PE
D3S1358	0.92254027	0.92636458	0.59289446
VWA	0.93683527	0.94089676	0.63496255
FGA	0.95981670	0.96809069	0.73259085
D8S1179	0.94155467	0.94793140	0.65780974
D21S11	0.94928354	0.95418849	0.67941225
D18S51	0.96378373	0.96950870	0.73871568
D5S818	0.87500427	0.87436793	0.47969234
D13S317	0.92315584	0.93241375	0.61137733
D7S820	0.92383981	0.93112689	0.60605462

- Comparación poblacional:

Los resultados obtenidos para los nueve loci STR incluidos en el Profiler Plus fueron comparados con otras tres poblaciones caucasianas: Italia, Suiza y EEUU, mediante tablas de contingencia R x C. Las discrepancias entre los datos comparados se midieron por dos estadísticos, χ^2 y G.

En las siguientes tablas mostramos los resultados obtenidos para cada loci:

Tabla 4.25.

D3S1358				
	χ^2	p	G	p
Italia / España	15.4184	0.0225±0.0053	14.9192	0.0340±0.0065
Italia / Suiza	11.0638	0.083±0.0091	10.9202	0.111±0.0104
Italia / EEUU	11.6826	0.0645±0.0085	11.4998	0.0865±0.0097
España / Suiza	2.3024	0.973±0.0057	2.4846	0.973±0.0057
España / EEUU	13.6302	0.0435±0.0075	13.5843	0.0485±0.0079
Suiza / EEUU	9.1541	0.153±0.0123	9.36	0.158±0.0125

Tabla 4.26.

VWA				
	χ^2	p	G	p
Italia / España	7.3412	0.504±0.0178	7.3974	0.512±0.0178
Italia / Suiza	9.4909	0.2865±0.015	9.4633	0.2885±0.015
Italia / EEUU	3.2169	0.858±0.0121	3.1685	0.859±0.012
España / Suiza	3.4915	0.956±0.0072	3.8631	0.9565±0.0072
España / EEUU	7.7992	0.4745±0.0184	8.3735	0.44±0.0183
Suiza / EEUU	9.5323	0.2865±0.0154	9.9898	0.262±0.015

Tabla 4.27.

FGA				
	χ^2	p	G	p
Italia / España	11.07	0.7955±0.014	12.6336	0.7535±0.0154
Italia / Suiza	16.0685	0.4535±0.0165	17.9202	0.4135±0.0163
Italia / EEUU	24.7942	0.0305±0.0059	27.1355	0.0235±0.0052
España / Suiza	11.9443	0.8±0.0141	13.1677	0.783±0.0145
España / EEUU	15.1026	0.3445±0.0175	18.8759	0.3365±0.0174
Suiza / EEUU	11.0369	0.8105±0.0134	12.3825	0.7725±0.0143

Tabla 4.28.

D8S1179				
	χ^2	p	G	p
Italia / España	15.2522	0.08±0.0097	15.3999	0.0845±0.0099
Italia / Suiza	14.5115	0.094±0.0097	14.4168	0.1135±0.0105
Italia / EEUU	23.1053	0.006±0.0027	23.7584	0.0070±0.0029
España / Suiza	6.9058	0.6565±0.0167	6.849	0.6725±0.0165
España / EEUU	10.6722	0.275±0.0165	10.3722	0.323±0.0173
Suiza / EEUU	12.4491	0.167±0.0127	12.6877	0.1805±0.0131

Tabla 4.29.

D21S11				
	χ^2	p	G	p
Italia / España	18.578	0.2950±0.0162	20.0742	0.2895±0.0162
Italia / Suiza	16.0783	0.3705±0.016	17.5488	0.3745±0.016
Italia / EEUU	16.2334	0.266±0.0153	16.7476	0.2995±0.0158
España / Suiza	12.9785	0.642±0.0169	14.2125	0.648±0.0168
España / EEUU	20.8676	0.107±0.0114	22.7881	0.0985±0.011
Suiza / EEUU	15.045	0.434±0.0169	17.3023	0.3625±0.0164

Tabla 4.30.

D18S51				
	χ^2	p	G	p
Italia / España	26.5072	0.008±0.0032	26.9991	0.011±0.0037
Italia / Suiza	16.9702	0.247±0.0143	17.1203	0.2865±0.0150
Italia / EEUU	25.2505	0.013±0.0039	24.9506	0.0225±0.0051
España / Suiza	18.3426	0.129±0.0118	19.5969	0.12±0.0114
España / EEUU	7.6663	0.8370±0.0136	8.3051	0.805±0.0146
Suiza / EEUU	20.2922	0.084±0.0095	20.1323	0.112±0.0108

Tabla 4.31.

D5S818				
	χ^2	p	G	p
Italia / España	3.8516	0.917±0.0098	4.6078	0.8755±0.0118
Italia / Suiza	5.223	0.869±0.0115	5.4279	0.8675±0.0116
Italia / EEUU	8.2654	0.408±0.017	8.781	0.404±0.017
España / Suiza	7.7149	0.4755±0.0182	7.8967	0.467±0.0182
España / EEUU	9.4774	0.1495±0.0132	9.4349	0.1705±0.0139
Suiza / EEUU	10.0174	0.223±0.0147	9.7472	0.268±0.0156

Tabla 4.32.

D13S317				
	χ^2	p	G	p
Italia / España	10.7912	0.1275±0.0119	11.2224	0.1215±0.0116
Italia / Suiza	3.805	0.833±0.0127	4.4628	0.7795±0.0142
Italia / EEUU	4.8939	0.6905±0.016	5.5279	0.6415±0.0166
España / Suiza	8.5091	0.1970±0.0145	8.4583	0.2025±0.0146
España / EEUU	13.0249	0.0405±0.0073	12.9881	0.0425±0.0074
Suiza / EEUU	2.8432	0.836±0.0131	2.8380	0.8355±0.0131

Tabla 4.33.

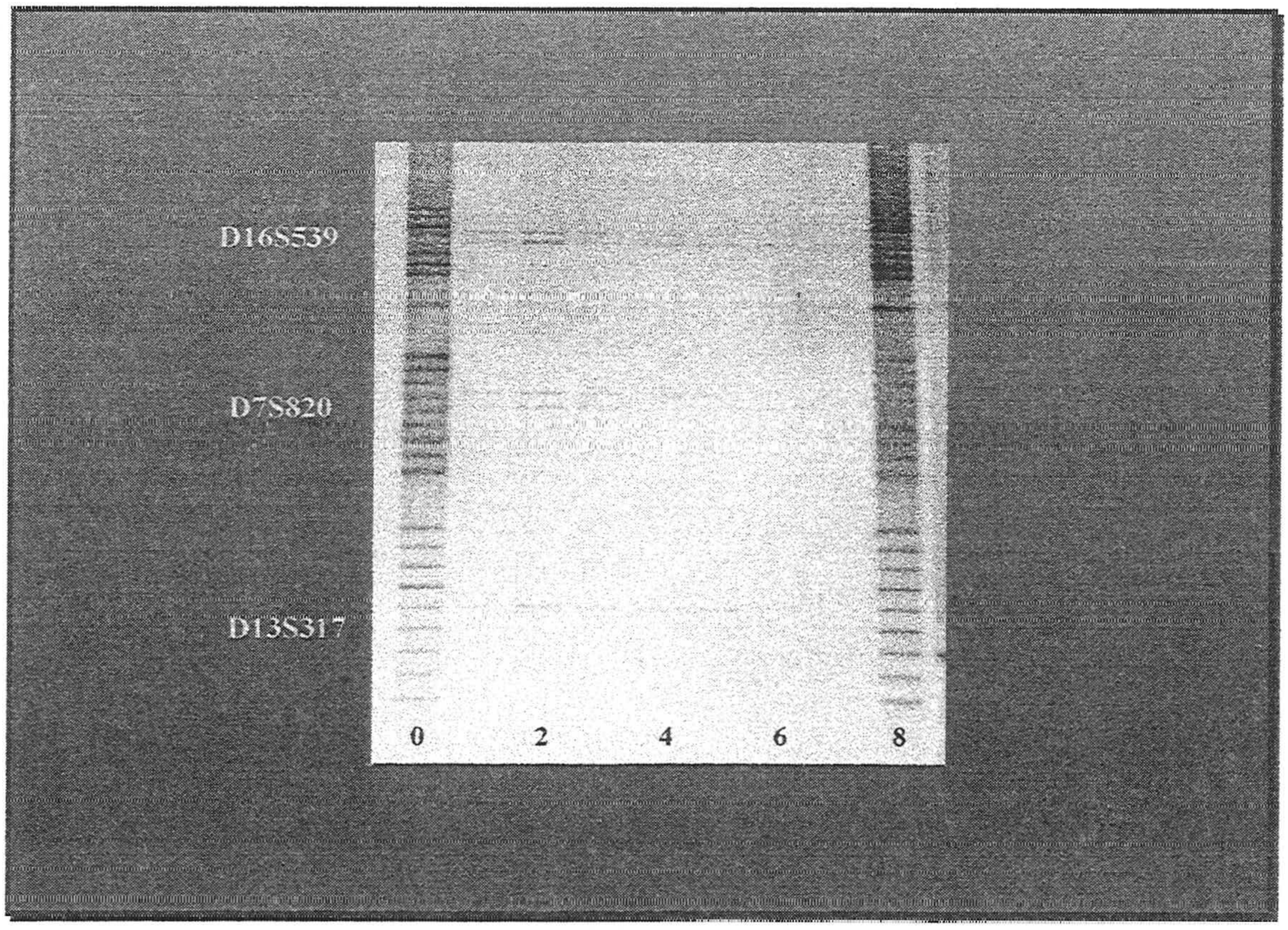
D7S820				
	χ^2	p	G	p
Italia / España	11.7985	0.0965±0.0105	12.7456	0.0855±0.01
Italia / Suiza	21.5726	0.002±0.0015	21.6294	0.003±0.0019
Italia / EEUU	12.7585	0.09±0.0099	12.7128	0.106±0.0106
España / Suiza	10.0278	0.177±0.0139	10.6355	0.1675±0.0136
España / EEUU	6.0224	0.659±0.0175	7.2532	0.5525±0.0184
Suiza / EEUU	9.0497	0.3445±0.0168	9.196	0.3575±0.0169

4.2. Resultados del estudio de validación del triplex SilverIII™

4.2.1. Estudio de sensibilidad

Tal y como se expuso en el capítulo 3.1.2. de Material y Métodos, para determinar la mínima cantidad de ADN amplificable con este triplex se hicieron diluciones seriadas del ADN control 9947-A y se amplificaron siguiendo las condiciones estándar descritas igualmente en el capítulo de Material y Métodos. Como puede observarse en la fotografía de abajo, la mínima cantidad de ADN con la que se observaron bandas fácilmente identificables fue 0.4 ng. Con 0.3 ng, 0.2ng y 0.1 ng, a pesar de que se observaban bandas, la intensidad de las mismas fue tan débil que no consideramos esta cantidad como el límite de sensibilidad.

Fotografía 4.1. Gel de SilverSTR III en el que se observa de izquierda a derecha y numerando del 0 al 9: **0)** ladder, **1)** 1 ng , **2)** 0.5 ng, **3)** 0.4 ng, **4)** 0.3 ng, **5)** 0.2 ng, **6)** 0.1 ng, 7) K-, **8)** ladder.



4.2.2. Estudio de mezclas

El resultado de este estudio nos mostró las variaciones en cuanto a intensidad o rendimiento de la amplificación de una cantidad constante de ADN cuando éste es mezclado con diferentes cantidades de otro ADN. Los resultados de las amplificaciones se muestran en las siguientes tablas en las que pueden observarse tanto las cantidades que se mezclaron de cada muestra (A, B o C) como las proporciones en las que se hicieron dichas muestras. Los resultados de las amplificaciones se consideraron positivos para las cantidades de ADN que encabezan las columnas cuando, de un total de al menos tres intentos, se obtuvieron dos o más resultados positivos.

Tabla 4.34. Resultados de las amplificaciones de las mezclas compuestas por las muestras **A y C** en diferentes proporciones para el locus D13S317.

D13S317

		A							
		1 ng		0.5 ng		0.25 ng		0.1 ng	
C	0.1 ng	1:0.1	+	1:0.2	+	1:0.4	+	1:1	-
	0.25 ng	1:0.25	+	1:0.5	+	1:1	+	1:2.5	+
	0.5 ng	1:0.5	+	1:1	+	1:2	+	1:5	+
	1 ng	1:1	+	1:2	+	1:4	+	1:10	+
	2.5 ng	1:2.5	+	1:5	+	1:10	+	1:25	+
	5 ng	1:5	+	1:10	+	1:20	+	1:50	+
	10 ng	1:10	+	1:20	+	1:40	-	1:100	+
	20 ng	1:20	+	1:40	+	1:80	-	1:200	+
	50 ng	1:50	+	1:100	+	1:200	-	1:500	-
	100 ng	1:100	+	1:200	-	1:400	-	1:1000	-

Tabla 4.35. Resultados de las amplificaciones de las mezclas compuestas por las muestras **A y C** en diferentes proporciones para el locus D7S820.

D7S820

		A							
		1 ng		0.5 ng		0.25 ng		0.1 ng	
C	0.1 ng	1:0.1	+	1:0.2	+	1:0.4	+	1:1	+
	0.25 ng	1:0.25	+	1:0.5	+	1:1	+	1:2.5	+
	0.5 ng	1:0.5	+	1:1	+	1:2	+	1:5	+
	1 ng	1:1	+	1:2	+	1:4	+	1:10	+
	2.5 ng	1:2.5	+	1:5	+	1:10	+	1:25	+
	5 ng	1:5	+	1:10	+	1:20	+	1:50	+
	10 ng	1:10	+	1:20	+	1:40	+	1:100	+
	20 ng	1:20	+	1:40	+	1:80	+	1:200	+
	50 ng	1:50	+	1:100	+	1:200	-	1:500	+
	100 ng	1:100	+	1:200	-	1:400	-	1:1000	-

Tabla 4.36. Resultados de las amplificaciones de las mezclas compuestas por las muestras **A y B** en diferentes proporciones para el locus D13S317.

		A											
		5 ng		2.5 ng		1 ng		0.5 ng		0.25 ng		0.1 ng	
B	0.1 ng	1:0.02	+	1:0.04	+	1:0.1	+	1:0.2	+	1:0.4	-	1:1	-
	0.25ng	1:0.05	+	1:0.1	+	1:0.25	+	1:0.5	+	1:1	+	1:2.5	+
	0.5 ng	1:0.1	+	1:0.2	+	1:0.5	+	1:1	+	1:2	+	1:5	+
	1 ng	1:0.2	+	1:0.4	+	1:1	+	1:2	+	1:4	+	1:10	+
	2.5 ng	1:0.5	+	1:1	+	1:2.5	+	1:5	+	1:10	+	1:25	+
	5 ng	1:1	+	1:2	+	1:5	+	1:10	+	1:20	+	1:50	+
	10 ng	1:2	+	1:4	+	1:10	+	1:20	+	1:40	+	1:100	+
	20 ng	1:4	+	1:8	+	1:20	+	1:40	+	1:80	+	1:200	+
	50 ng	1:10	+	1:20	+	1:50	+	1:100	-	1:200	-	1:500	+
	100 ng	1:20	+	1:40	+	1:100	+	1:200	-	1:400	-	1:100	-

Tabla 4.37. Resultados de las amplificaciones de las mezclas compuestas por las muestras **A y B** en diferentes proporciones para el locus D7S820.

		A											
		5 ng		2.5 ng		1 ng		0.5 ng		0.25 ng		0.1 ng	
B	0.1 ng	1:0.02	+	1:0.04	+	1:0.1	+	1:0.2	+	1:0.4	+	1:1	-
	0.25ng	1:0.05	+	1:0.1	+	1:0.25	+	1:0.5	-	1:1	+	1:2.5	+
	0.5 ng	1:0.1	+	1:0.2	+	1:0.5	+	1:1	+	1:2	+	1:5	+
	1 ng	1:0.2	+	1:0.4	+	1:1	+	1:2	+	1:4	+	1:10	+
	2.5 ng	1:0.5	+	1:1	+	1:2.5	+	1:5	+	1:10	-	1:25	+
	5 ng	1:1	+	1:2	+	1:5	+	1:10	+	1:20	-	1:50	+
	10 ng	1:2	+	1:4	+	1:10	+	1:20	+	1:40	-	1:100	+
	20 ng	1:4	+	1:8	+	1:20	+	1:40	+	1:80	-	1:200	+
	50 ng	1:10	+	1:20	+	1:50	+	1:100	+	1:200	-	1:500	-
	100 ng	1:20	+	1:40	-	1:100	+	1:200	+	1:400	-	1:1000	-

Tabla 4.38. Resultados de las amplificaciones de las mezclas compuestas por las muestras **A y B** en diferentes proporciones para el locus D16S539.

		A											
		5 ng		2.5 ng		1 ng		0.5 ng		0.25 ng		0.1 ng	
B	0.1 ng	1:0.02	+	1:0.04	+	1:0.1	+	1:0.2	+	1:0.4	+	1:1	+
	0.25ng	1:0.05	+	1:0.1	+	1:0.25	+	1:0.5	+	1:1	+	1:2.5	+
	0.5 ng	1:0.1	+	1:0.2	+	1:0.5	+	1:1	+	1:2	+	1:5	+
	1 ng	1:0.2	+	1:0.4	+	1:1	+	1:2	+	1:4	+	1:10	+
	2.5 ng	1:0.5	+	1:1	+	1:2.5	+	1:5	+	1:10	+	1:25	+
	5 ng	1:1	+	1:2	+	1:5	+	1:10	+	1:20	+	1:50	+
	10 ng	1:2	+	1:4	+	1:10	+	1:20	+	1:40	+	1:100	+
	20 ng	1:4	+	1:8	+	1:20	+	1:40	+	1:80	+	1:200	+
	50 ng	1:10	+	1:20	-	1:50	-	1:100	+	1:200	-	1:500	+
	100 ng	1:20	-	1:40	-	1:100	-	1:200	-	1:400	-	1:1000	-

4.2.3. Estudio de muestras forenses

Todas las muestras utilizadas en esta parte del estudio de validación fueron cuantificadas en geles de agarosa al 0.8% para determinar la cantidad y calidad del ADN antes de proceder a la amplificación. Los resultados obtenidos tanto en la cuantificación como en la amplificación se muestran en la siguiente tabla:

Tabla 4.39.

Muestra	Cantidad ADN (ng/µl)	ng ADN amplificado	Calidad ADN	Genotipo D13S317	Genotipo D7S820	Genotipo D16S539
C-1	3	6	B	8,11	9,12	10,12
C-2	1	5	B	11,11	12,12	10,10
C-3	1.6	4.8	B	11,12	11,12	10,11
C-4	1.6	4.8	B	11,12	11,12	10,11
C-5	0.4	4	B	11,12	11,12	9,11
C-6	0.8	4.8	B	12,13	9,10	10,11
C-7	0.8	4.8	B	12,13	9,10	9,9
MSA-1	<1	<10	D	11,12	8,10	11,12
MSA-2	<1	<10	D	11,12	8,10	11,12
MSA-3	•	*	D	11,12	8,10	11,12
MSA-4	4	8	B	10,13	8,10	11,12
MSA-5	<1	<10	B	---	---	---
MSA-6	<1	<10	B	---	---	---
MSA-7	2	4	B	9,10	7,12	10,12

Muestra	Cantidad ADN (ng/μl)	ng ADN amplificado	Calidad ADN	Genotipo D13S317	Genotipo D7S820	Genotipo D16S539
MSA-8	2	6	D	11,12	8,11	9,11
MSA-9	4	8	B	---	---	---
MSA-10	---	*	---	---	---	---
MSA-11	---	*	---	---	---	---
MSA-12	5	10	B	12,13	9,10	11,13
MSA-13	---	*	---	---	---	---
MSA-14	0.1	1.98	B	---	---	---
MSA-15	---	*	---	---	---	---
MSE-1(MF)	<1	<10	B	10,13	8,10	10,12
MSE-1(FF)	5	10	B	8,10,11,13	8,9,10,11	9,11,12,13
MSE-2(MF)	5	10	B	8,11	8,9	9,12
MSE-2(FF)	•	*	MD	8,11,13	8,9,10	9,12
MSE-3(MF)	•	*	MD	13,14	9,9	9,12
MSE-3(FF)	•	*	D	13,14	9,9,	9,12
MSE-4(MF)	•	*	D	10,11	10,11	11,13
MSE-4(FF)	•	*	D	---	---	---
MSE-5(MF)	2	4	D	12,12	9,10	8,9
MSE-5(FF)	•	*	MD	12,12	9,10	8,9
MSE-6(MF)	---	*	---	---	---	---

Muestra	Cantidad ADN (ng/μl)	ng ADN amplificado	Calidad ADN	Genotipo D13S317	Genotipo D7S820	Genotipo D16S539
MSE-6(FF)	---	*	---	---	---	---
MSE-7(MF)	---	*	---	---	---	---
MSE-7(FF)	5	5	B	---	---	---
MSE-8(MF) 1/10	1	2	D	11,11	8,12	10,12
MSE-8(FF) 1/10	1	2	D	11,11	8,12	10,12
MSE-9(MF)	2	4	B	9,12	11,12	14,14
MSE-9(FF)	---	*	---	---	---	---
MSE-10	4	8	D	---	---	---
SAL-1	•	*	D	---	---	---
SAL-2	---	*	---	---	---	---
SAL-3	0.04	0.792	B	---	---	---
H-1	---	*	---	---	---	---

• Notas referentes a las abreviaturas de la tabla:

- **C** = Muestras de ADN extraído a partir de caspa

- **MSA** = Muestras de ADN extraído a partir de manchas de sangre

- **MSE** = Muestras de ADN extraído a partir de manchas de semen

- **SAL** = Muestras de ADN extraído a partir de saliva

- **H** = Muestras de ADN extraído a partir de tejido óseo

- **B** = Calidad buena, ADN sin degradar

- • = Muestras en las que fue imposible cuantificar el ADN debido a que lo único que se observaba en el gel de agarosa era toda la calle teñida, esto es indicativo de una alta degradación del ADN de la muestra.

- **D** = ADN degradado

- **MD** = ADN muy degradado

- * = para la amplificación se utilizó el máximo volumen de ADN que se permitía según el volumen final de amplificación, no se ponen cantidades absolutas de ADN porque son muestras mínimas en las que no se observó una banda clara en geles de agarosa, o bien lo único que se observó fue toda la calle teñida con lo cual la cuantificación fue imposible como se explicó anteriormente.

5. DISCUSIÓN

5.1. ESTUDIO GENÉTICO-POBLACIONAL

5.1.1. Triplex SilverIII STR™

5.1.1.1. Discusión de los resultados

Los tres loci analizados conjuntamente en este triplex comercial de Promega mostraron, en la población analizada, alelos discretos y distinguibles siendo posible así el cálculo de las frecuencias alélicas por recuento de los mismos. Todos los alelos encontrados dentro de cada marcador se diferenciaron en el número de repeticiones en tándem de una secuencia de cuatro pares de bases, no encontrándose microvariantes como las que se observan en algunos loci STR (como el HUMTH01 y e HUMF13A01) originados por delección o adición de pares de bases en la unidad de repetición (Puers et al, 1993; Urquhart et al, 1994).

Para el **D13S317** fueron observados siete alelos numerados desde el 8 hasta el 14. Los dos alelos más frecuentes fueron el 11 y el 12, cuyas frecuencias alélicas sumadas constituyeron más del 50% del total de frecuencias (53,80%).

El número total de genotipos diferentes observados fue de 26, frente a los 28 teóricos que se esperan por combinación de siete alelos. Los genotipos no encontrados en nuestra muestra poblacional fueron el 9,10 y el 10,10 lo cual no es extraño si se tiene en cuenta que las frecuencias de los alelos 9 y 10 son bastante bajas respecto al resto de alelos (6,43 y 6,14% respectivamente). En cuanto al genotipo más frecuente fue el 11,12 con una frecuencia del 19,29%.

Nuestra población mostró similitudes en cuanto a frecuencias génicas con otras poblaciones como Italia (Garofano et al, 1998), Suiza (Gehrig et al, 1999), Alemania (Hanstchel et al, 1999), EEUU (Kupferschmid et al, 1999) y Austria (Neuhuber et al,

1999). En todas ellas los alelos 11 y 12 constituían más del 50% del total de los alelos. Ver Gráfico 5.1 en página 192.

Para este marcador la lectura de las bandas correspondientes a los alelos en geles de acrilamida al 6% no planteó problema alguno puesto que las muestras de ADN utilizadas para el estudio poblacional eran de buena calidad. No podemos decir lo mismo en aquellos casos en los que tanto cantidad como calidad de ADN no era la idónea ya que entonces a menudo aparecían bandas sombra o "stutter bands" que interferían en la interpretación de los resultados, no obstante este problema será comentado más adelante en el apartado del estudio de validación.

El D13S317 junto al D7S820 fue analizado doblemente en nuestra población al estar también incluido en el multiplex de Perkin Elmer (AmpFISTR Profiler Plus). Las frecuencias génicas obtenidas en ambos casos fueron casi idénticas, las pequeñas diferencias que se observaron se debieron a que seis de las 177 muestras analizadas para el SilverIII STR™ no fueron analizadas para el AmpFISTR Profiler Plus, lo cual confirma la fiabilidad de las técnicas utilizadas.

El **D7S820** mostró siete alelos, siendo los más frecuentes el 10 y el 11, cuya suma de frecuencias alélicas constituía aproximadamente un 50% del total. Una distribución de alelos similar mostraron otras poblaciones como Italia (Garofano et al, 1998), Suiza (Gehrig et al, 1999), Alemania (Hanstchel et al, 1999), Austria (Neuhuber et al, 1999) y EEUU (Kupferschmid et al, 1999). Gráfica 5.2. en página 192

Se observaron 23 genotipos para este marcador, los genotipos no encontrados en nuestra población fueron: 7,7; 7,10; 7,13 y 9,13. Dado que las frecuencias alélicas más bajas son las de los alelos 7 y 13 habría que aumentar el tamaño de la muestra para que la probabilidad de encontrar estos genotipos fuera mayor. Tal y como

ocurrió con el D13S317, las diferencias en las frecuencias alélicas encontradas al analizar este locus con dos kits diferentes fueron mínimas y en cualquier caso se debieron al tamaño de muestra analizada en cada caso.

Este locus no presentó ningún problema de los manifestados por algunos STRs a la hora de la interpretación de los resultados.

El tercer marcador incluido en este triplex es el **D13S359**. El número de alelos encontrados fue de 8, siendo los más frecuentes el 11 y el 12, lo cual coincide con los datos encontrados en publicaciones recientes de estudios poblacionales en Suiza (Gehrig, C et al, 1999) e Italia (Garofano, L et al, 1998) (en el presente momento no se han encontrado datos de otras poblaciones para este marcador). Gráfica 5.3. en página 193.

El número de genotipos encontrados fue de 21 frente a los 36 teóricos que debían encontrarse. La existencia de cuatro alelos con frecuencias muy bajas (por debajo de 0.02%) justifica que en una muestra de tamaño no excesivamente grande (n=177), no sean observados genotipos que contengan alguno de estos alelos, por lo tanto, al igual que ocurría con los otros dos loci, sería necesario un tamaño de muestra mayor para que pudieran ser encontrados.

A la hora de interpretar los resultados, ha sido este marcador el que más problemas ha causado debido al desdoblamiento que sufrían las dos cadenas de ADN de cada alelo. Este hecho viene descrito en el manual de Promega (Technical Manual, GeneprintTM STR Systems, Promega, 1997) según el cual este problema aparece cuando la electroforesis se lleva a cabo en geles de acrilamida al 4%. Para evitarlo, proponen el uso de geles de acrilamida al 6%. Siguiendo las recomendaciones de Promega llevamos a cabo las electroforesis en geles al 6% (tal

y como queda expuesto en el apartado de Material y Métodos) pero no conseguimos mejorar el problema.

Este hecho ha sido descrito previamente en otros marcadores de ADN (Saitoh et al, 1998). El desdoblamiento es producido por una diferencia de movilidad electroforética entre las dos bandas de ADN que componen un alelo.

Desde hace algún tiempo, diversos autores han observado una relación entre la diferencia de movilidad electroforética y la composición de bases de las cadenas complementarias del alelo amplificado.

Hayward y colaboradores (1972) describieron que cuando se sometían a electroforesis en geles de agarosa, las bandas de ADN ricas en bases purínicas migraban más rápido que las ricas en pirimidínicas.

En 1989 Weber y cols. (Weber JL et al, 1989) observaron que aquellos alelos STR que contenían repeticiones ricas en citosinas y adeninas migraban más rápido que las que contenían guaninas y timidinas cuando se separaban en geles de acrilamida desnaturalizantes.

Publicaciones más recientes siguen confirmando este hecho. Saitoh (Saitoh, H et al, 1998) determinó la relación entre el cociente AC/GT en alelos de STRs y la diferencia de movilidad electroforética en geles desnaturalizantes de urea. Cuando este cociente es cercano a 1 la diferencia de movilidad decrece. Este autor afirma también que la diferencia de movilidad se da independientemente de la variación en la concentración de acrilamida.

Para las muestras que utilizamos en el estudio poblacional este problema no tuvo mayor repercusión puesto que no fue difícil diferenciar las dos cadenas de cada

alelo, en cambio, en el caso de mezclas de ADN la interpretación fue difícil, a veces imposible (como se verá más adelante) al existir solapamiento de cadenas de alelos distintos.

Para concluir, debemos decir que de manera global los tres marcadores fueron amplificados con éxito en las 177 muestras que componían nuestro estudio poblacional.

5.1.1.2. Discusión del análisis estadístico

Como quedó expuesto en el apartado de Introducción, bajo condiciones de equilibrio Hardy-Weinberg las frecuencias genotípicas de una población vienen determinadas por las frecuencias alélicas, manteniéndose éstas constantes de generación en generación.

Cuando en Criminalística, por ejemplo, se analiza el ADN de una mancha de sangre encontrada en el lugar del delito y se quiere comparar con el ADN de un sospechoso es necesario conocer cuál es la frecuencia en la población de los genotipos encontrados para que, en caso de que coincidan, pueda darse un valor de probabilidad de que esa mancha de sangre procede de dicho sospechoso.

Obviamente, estudiar todos los genotipos de la población es imposible, por lo que se toma una muestra representativa (n>100) y se analizan en ella las frecuencias de los alelos de los diferentes marcadores que serán utilizados para la identificación genética. Por lo tanto, las frecuencias génicas obtenidas en un estudio poblacional no son valores exactos sino estimas con un margen de error.

Una vez calculadas las frecuencias alélicas en la muestra el siguiente paso es comprobar que la población estudiada se encuentra en equilibrio Hardy-Weinberg

para el locus estudiado. Para ello se aplican una serie de tests estadísticos cuyo objetivo es asegurar dicho equilibrio con un grado de fiabilidad del 95%.

Además, para que un marcador genético pueda ser utilizado en la identificación de individuos debe demostrarse que no se encuentra ligado con cualquier otro marcador de los utilizados con el mismo fin, tanto si se encuentran en el mismo o en diferentes cromosomas.

Una vez aplicados todos los tests estadísticos y comprobado que la población se encuentra en equilibrio y que no existe ligamiento entre los distintos loci, es posible utilizar los datos poblacionales obtenidos sin necesidad de repetir los muestreos cada pocos años (Carracedo et al, 1996).

En nuestro caso las heterocigosidades para los tres loci variaron en un rango comprendido entre el 74,01 y 82,49 %, correspondiendo el valor superior al D16S539 y el inferior al D7S820. Ver tabla 4.4 en el apartado 4.1.3.

Los resultados obtenidos tras la aplicación del test de homocigosidad, test de verosimilitud y test exacto de Guo-Thompson indicaron que, de los tres loci, el D7S820 fue el único que dio valores de p por debajo de 0.05 por lo que, en principio, cabría pensar que este marcador se aleja del equilibrio Hardy-Weinberg. No obstante, como se dijo anteriormente, los tests empleados para la demostración de equilibrio no son fiables al 100% y asumen un margen de error o nivel de significancia (en nuestro caso del 0.05%) de aceptar la hipótesis alternativa (nuestros datos se alejan del equilibrio) cuando en realidad es verdadera.

También habría que considerar el tamaño de muestra ya que los tests estadísticos son más exactos cuanto mayor es este valor.

Por otro lado, puesto que los alelos de este marcador fueron fácilmente detectables e interpretables, y no mostraron problemas de "stutter bands" o desdoblamiento de las cadenas de ADN, descartamos posibles errores en la detección de alelos que pudieran haber llevado a tal desequilibrio.

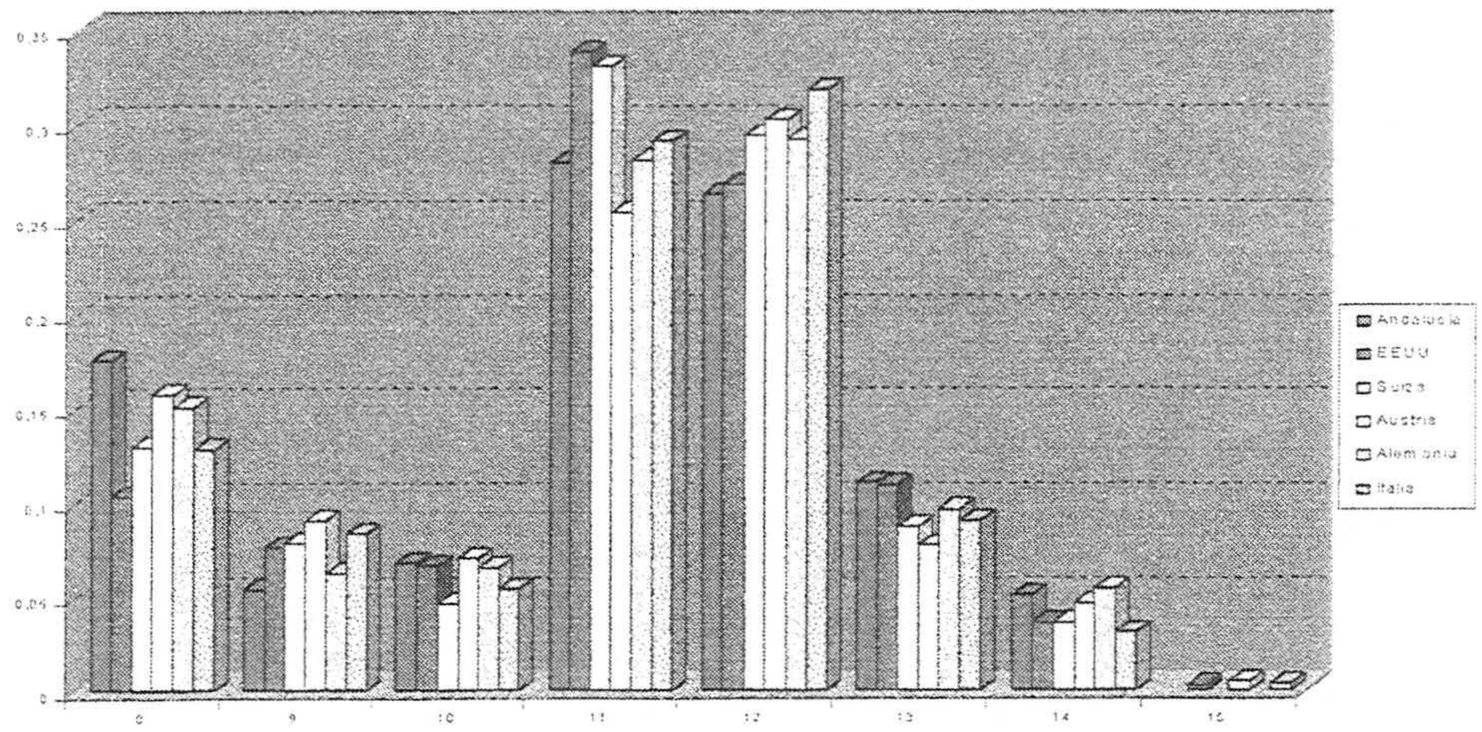
Cuando comparamos nuestros resultados con los obtenidos por otras poblaciones no observamos diferencias significativas por lo que es probable que la causa del desequilibrio fuera un problema de muestreo.

En cuanto al equilibrio de la fase gamética, los resultados obtenidos tras la aplicación del test de correlación de Karlin dieron valores de p por encima de 0.05, demostrándose así que existe independencia de unos loci con respecto a otros y que, por lo tanto se cumple la segunda condición necesaria para la aplicación de dichos marcadores en identificación genética.

Podemos concluir diciendo que el triplex SilverIII STR™ de la casa Promega ha sido aplicado con éxito a las 177 muestras de nuestro estudio poblacional consiguiéndose en conjunto un poder de discriminación (PD) de **0.99942306** y un poder de exclusión (PE) de **0.93122203.**

Gráficos 5.1 y 5.2. Distribución de las frecuencias alélicas de los loci D13S317 y D7S820 en diferentes poblaciones estudiadas.

D13S317



D7S820

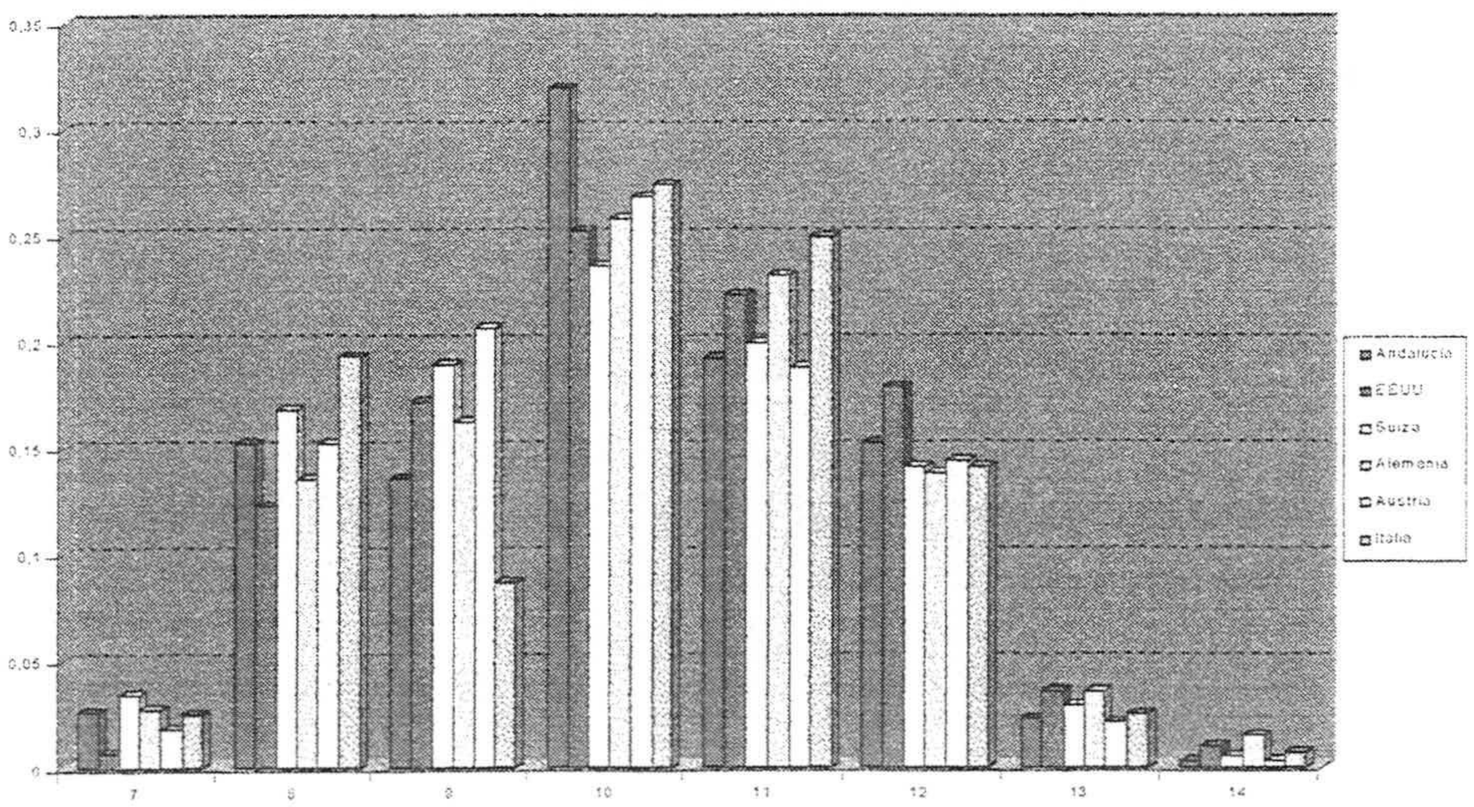
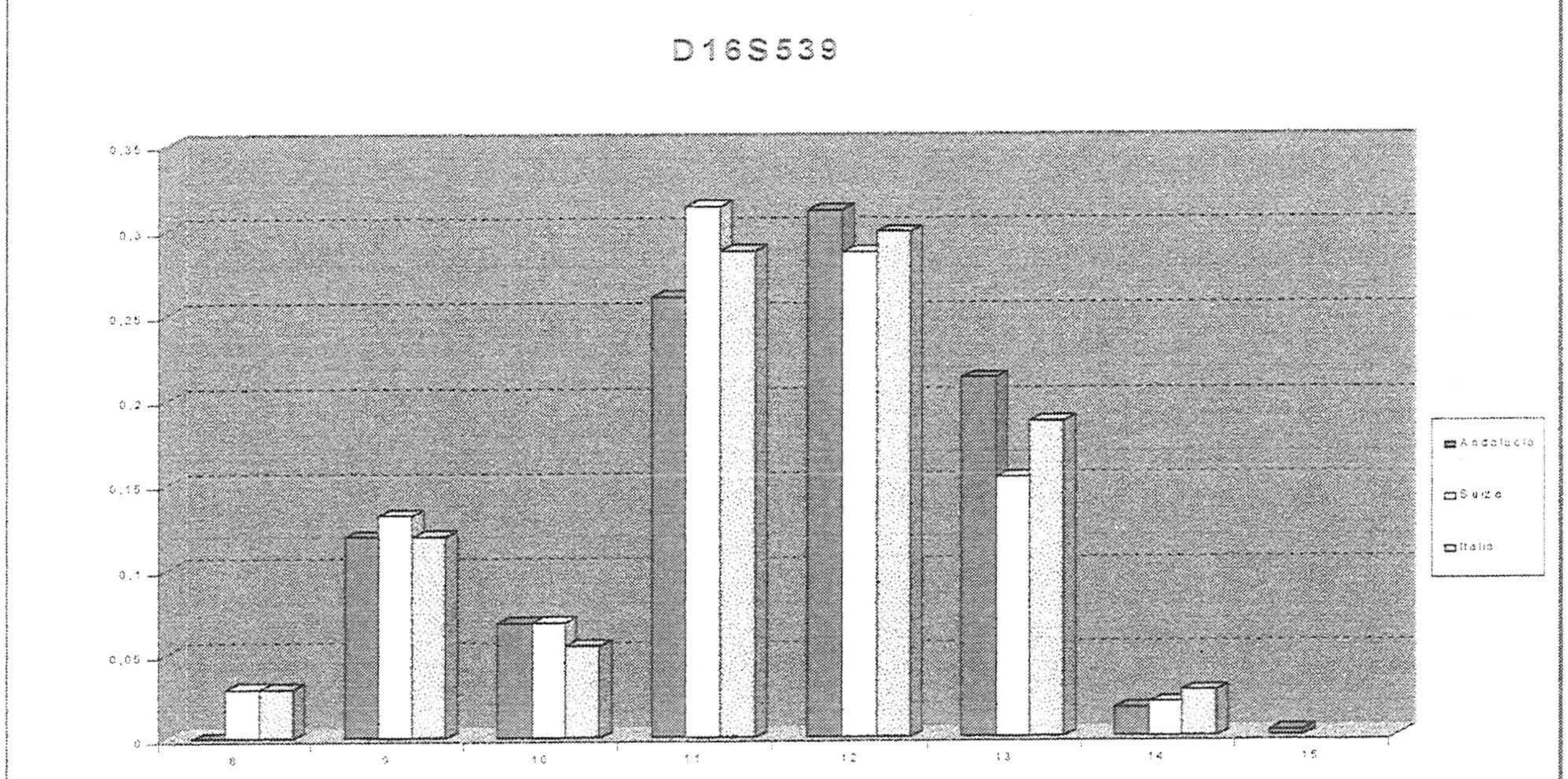


Gráfico 5.3. Distribución de las frecuencias alélicas de los loci D16S539 en diferentes poblaciones estudiadas.



5.1.2. Multiplex AmpFISTR Profiler Plus

5.1.2.1. Discusión de los resultados

Mediante este kit de detección fluorescente se analizaron nueve loci tipo STR (dos de los cuales se encuentran incluidos en el kit de Promega SilverIII STR^{TM}) y un fragmento del gen de la amilogenina localizado en la región homóloga de los cromosomas X e Y (y que es utilizado para el diagnóstico del sexo en humanos).

Para los procesos de amplificación, electroforesis e interpretación de los resultados se siguieron los protocolos recomendados en el manual del fabricante (AmpFISTR Profiler Plus PE Applied Biosystems).

La interpretación de los resultados se hizo por comparación con el ladder alélico que contiene el kit. El estándar de tamaño que se añade a cada muestra amplificada asegura una gran precisión en la determinación del tamaño de los fragmentos, lo cual hace que la comparación con los alelos del ladder sea más exacta, tanto si son alelos compuestos por repeticiones completas de la secuencia consenso, como variantes compuestas por repeticiones incompletas.

Los alelos han sido designados siguiendo las recomendaciones de la Comisión de ADN de la ISFH (DNA Recommendations, 1994, Bär et al, 1997).

El programa Genotyper asignó a los picos obtenidos los alelos correspondientes, siempre y cuando éstos no difiriesen en más de 0.5 pb del pico correspondiente en el ladder.

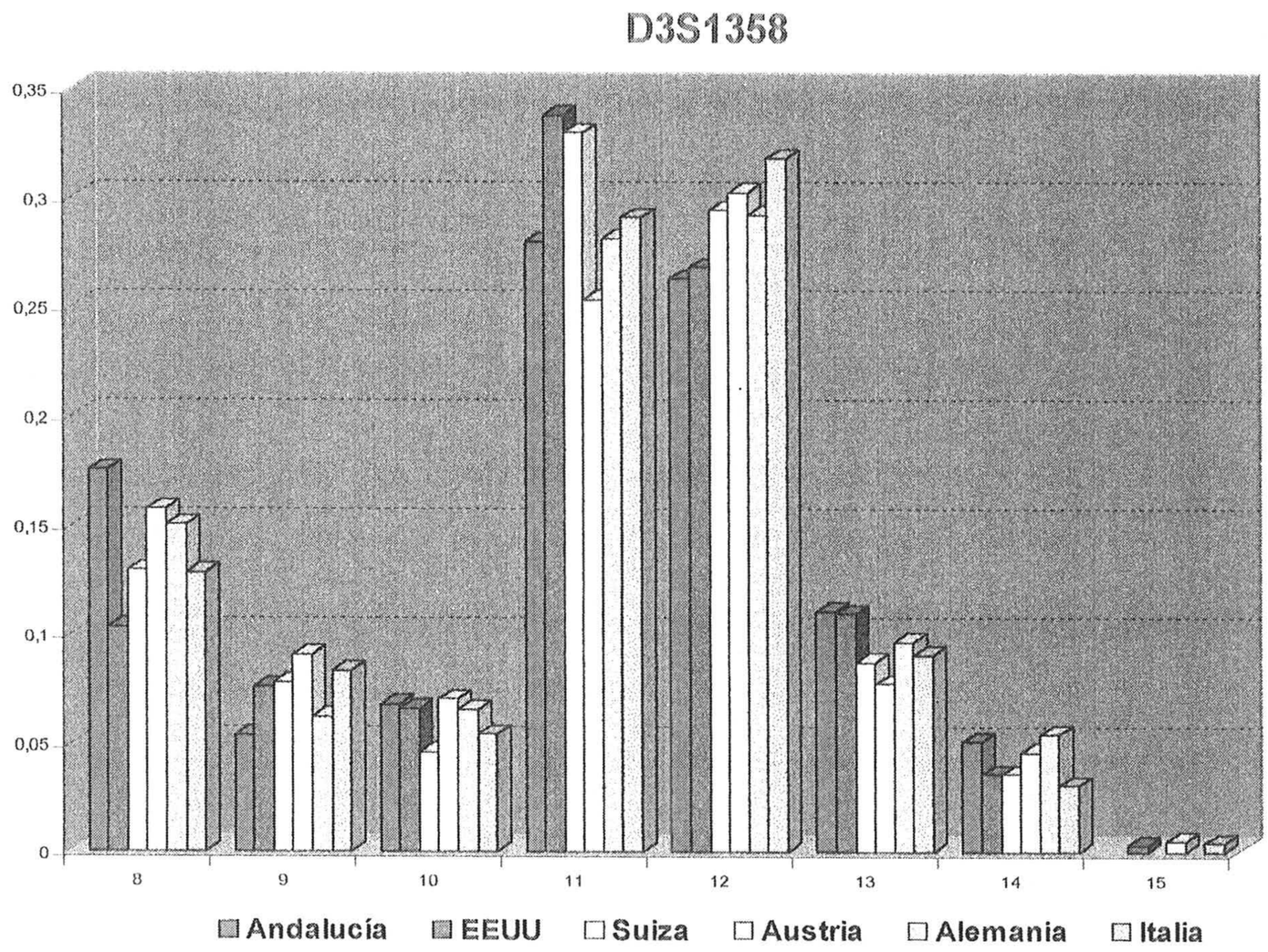
Para la mayoría de las muestras que nosotros utilizamos los alelos fueron asignados correctamente, no obstante, en aquellas en las que se observó algún alelo fuera de ladder ("off-ladder" alleles) volvió a repetirse el proceso de electroforesis.

Tras la asignación de los alelos, los resultados fueron introducidos en el programa informático utilizado para el análisis estadístico.

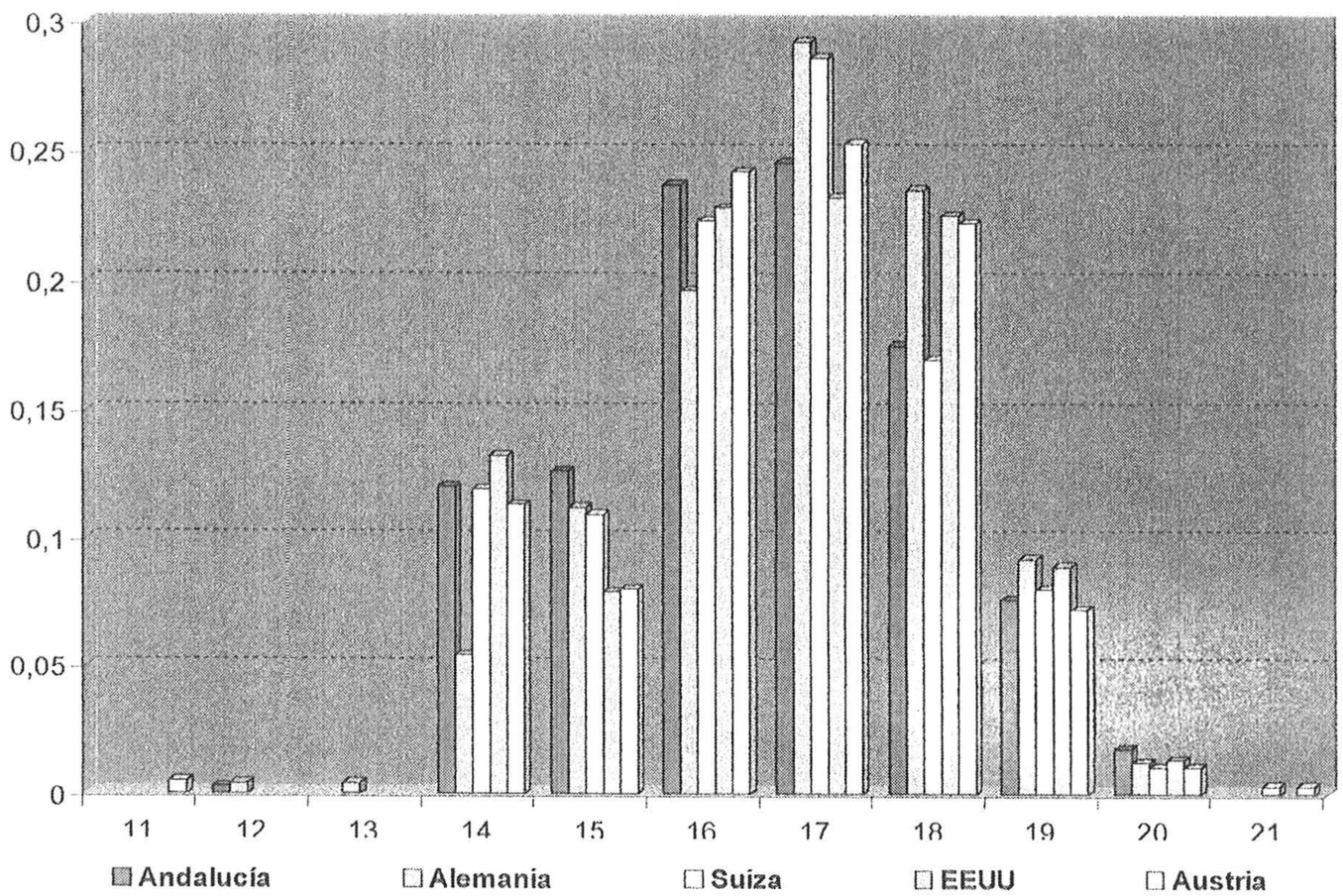
Los alelos más y menos frecuentes, el número de genotipos teóricos esperados según el número de alelos, los genotipos observados en nuestra muestra así como las poblaciones caucasianas con las que se ha encontrado similitud en las frecuencias alélicas (de las publicadas hasta ahora) se muestran a continuación en la siguiente tabla.

Tras la tabla se expondrán las gráficas con las frecuencias alélicas de cada marcador en las diferentes poblaciones caucasianas en las que se han estudiado dichos marcadores en las que, como se ha indicado anteriormente no se encontraron diferencias significativas.

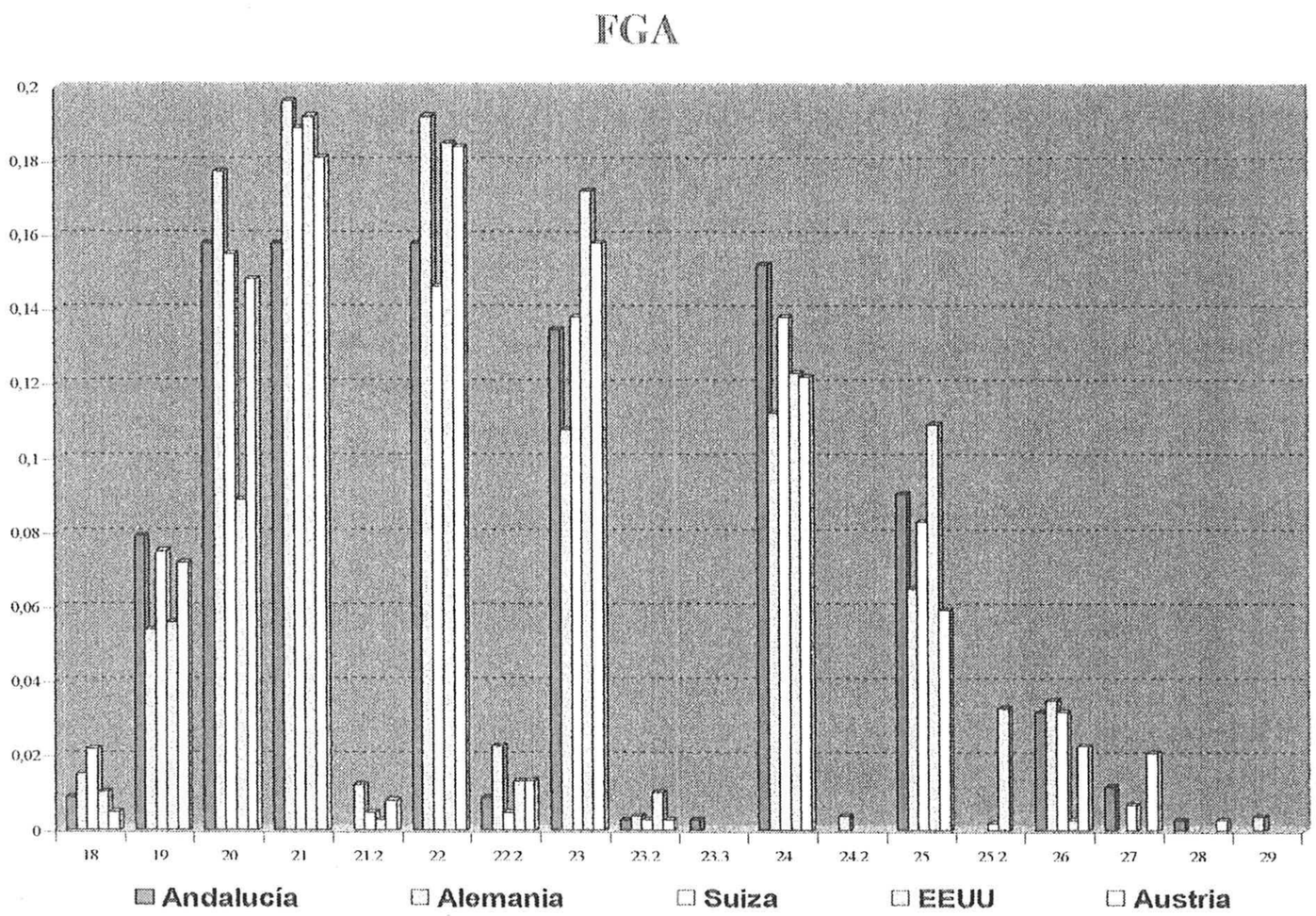
Marcador	Alelos más frecuentes	Alelos menos frecuente	Nº genotipos esperados	Nº genotipos observados	Poblaciones comparadas
D3S1358	15 y 16	13 y >19	36	22	Italia, Alemania Suiza, EEUU, Austria
VWA	16 y 17	12	36	25	Alemania, Suiza, EEUU, Austria
FGA	20, 21 y 22	23.2, 23.3, 28	105	42	Alemania, Suiza, EEUU, Austria
D8S1179	13 y 14	17	55	35	Suiza, EEUU, Austria
D21S11	29 y 30	26, 28 y 35	120	44	Suiza, EEUU, Austria
D18S51	14 y 15	10 y 22	78	48	Italia, Suiza, EEUU, Austria
D5S818	11 y 12	14 y 15	28	17	Italia, Alemania Suiza, EEUU, Austria
D13S317	11 y 12	14	28	26	Italia, Alemania, Suiza, EEUU, Austria
D7S820	10 y 11	13	28	23	Italia, Alemania Suiza, EEUU, Austria



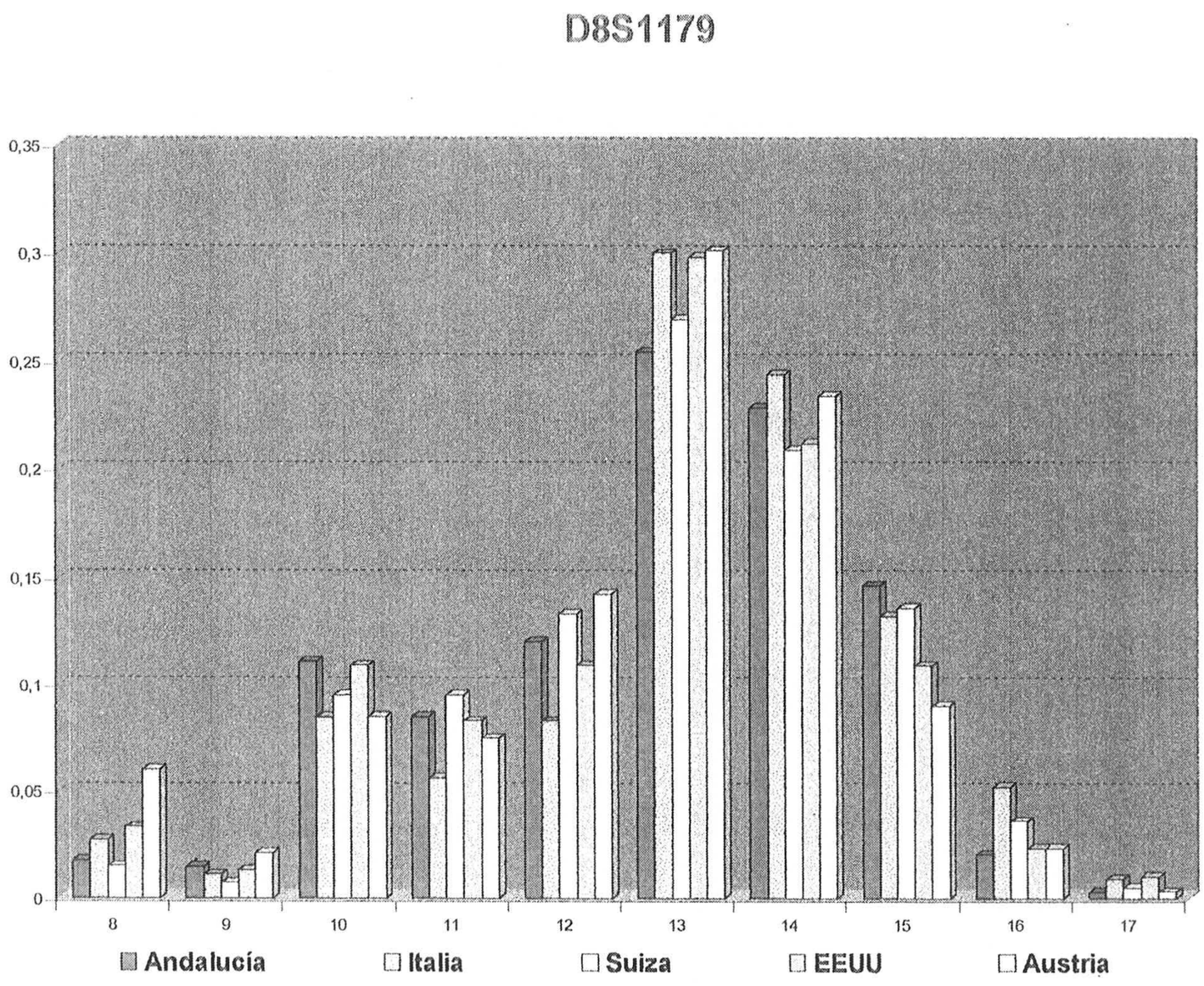
Gráfica 5.4. Distribución de frecuencias alélicas para el locus D3S1358 en diferentes poblaciones analizadas



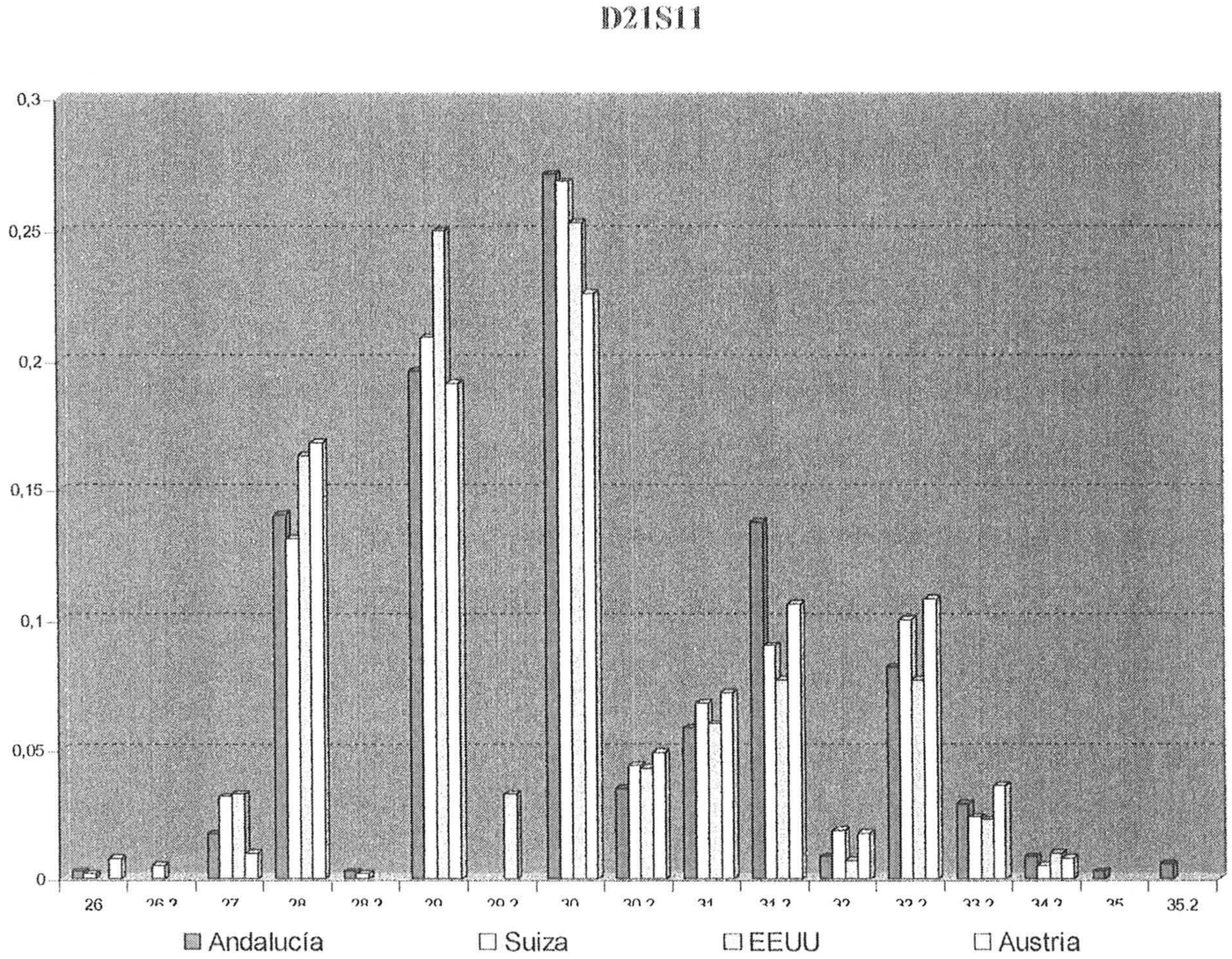
Gráfica 5.5. Distribución de frecuencias alélicas para el locus VWA en diferentes poblaciones analizadas.



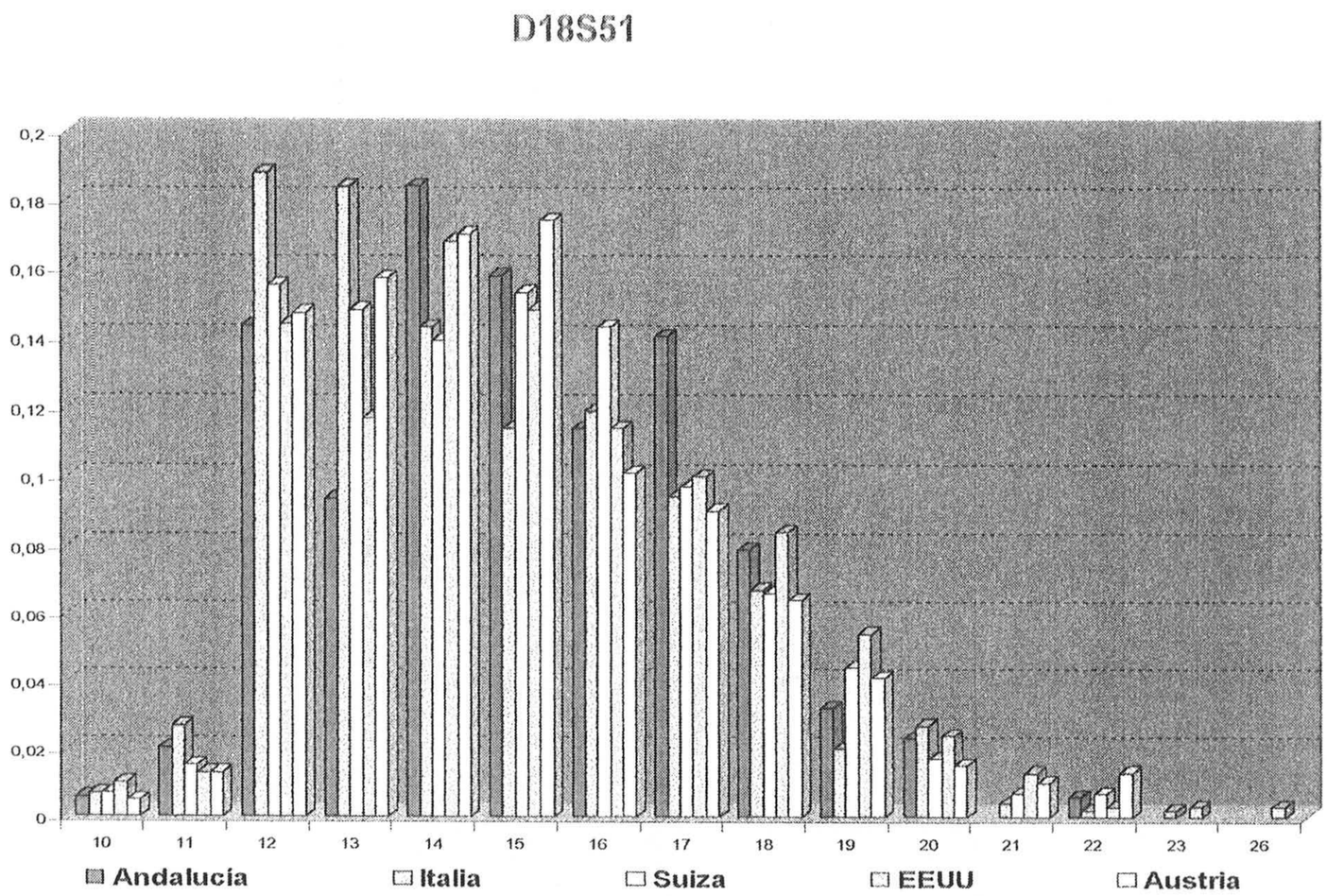
Gráfica 5.6. Distribución de frecuencias alélicas para el locus FGA en diferentes poblaciones analizadas.



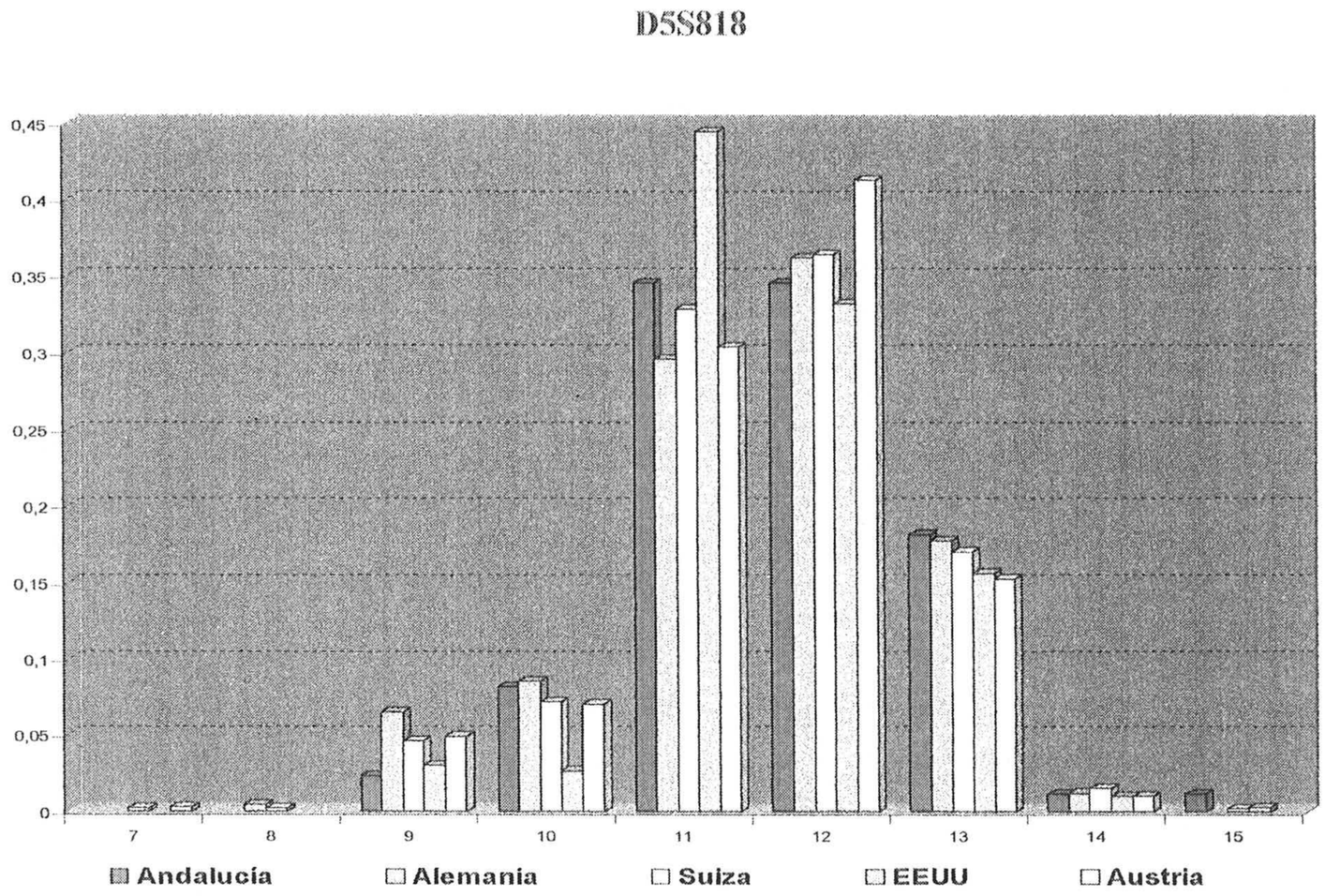
Gráfica 5.7. Distribución de frecuencias alélicas para el locus D8S1179 en diferentes poblaciones analizadas.



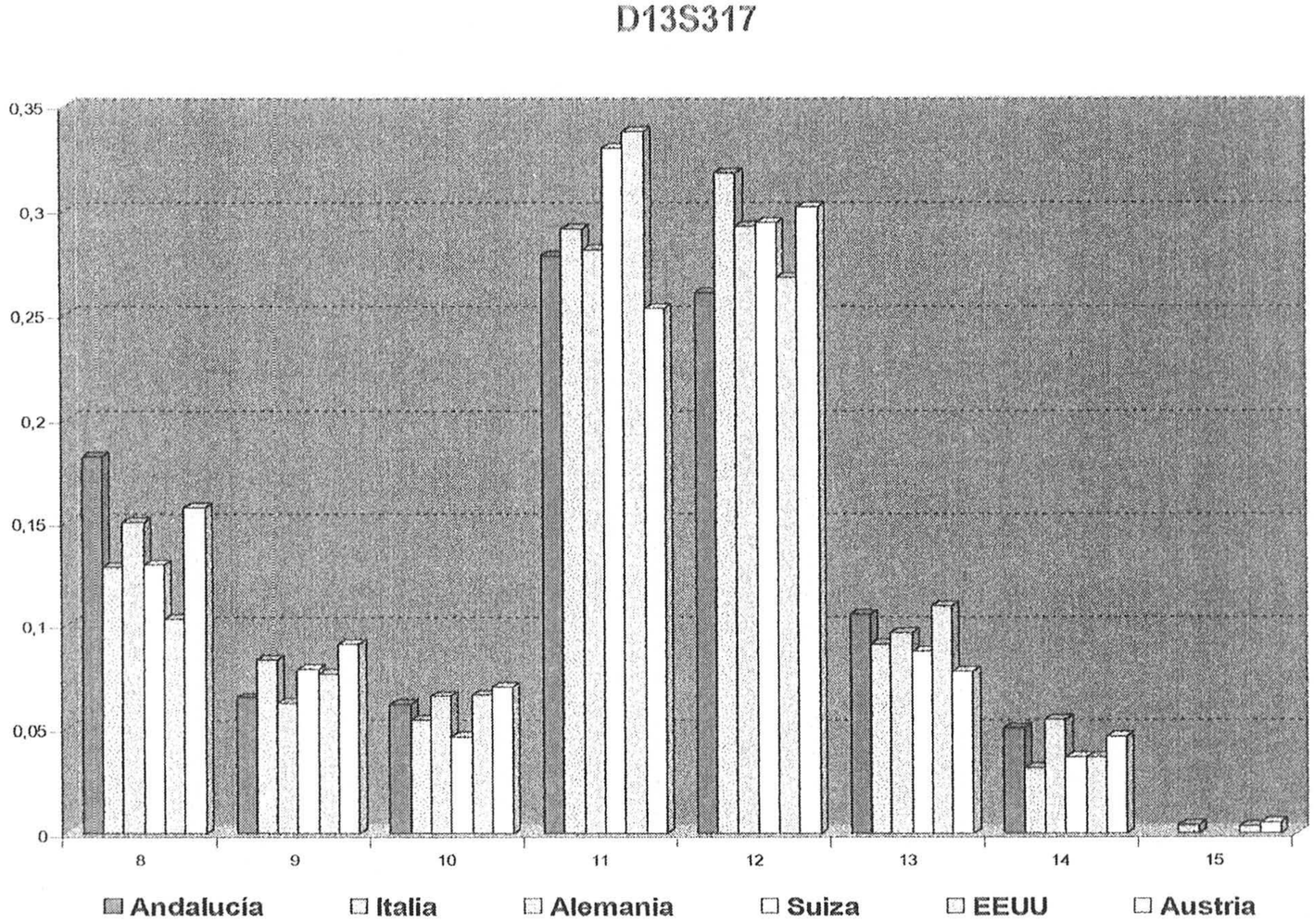
Gráfica 5.8. Distribución de frecuencias alélicas para el locus D21S11 en diferentes poblaciones analizadas.



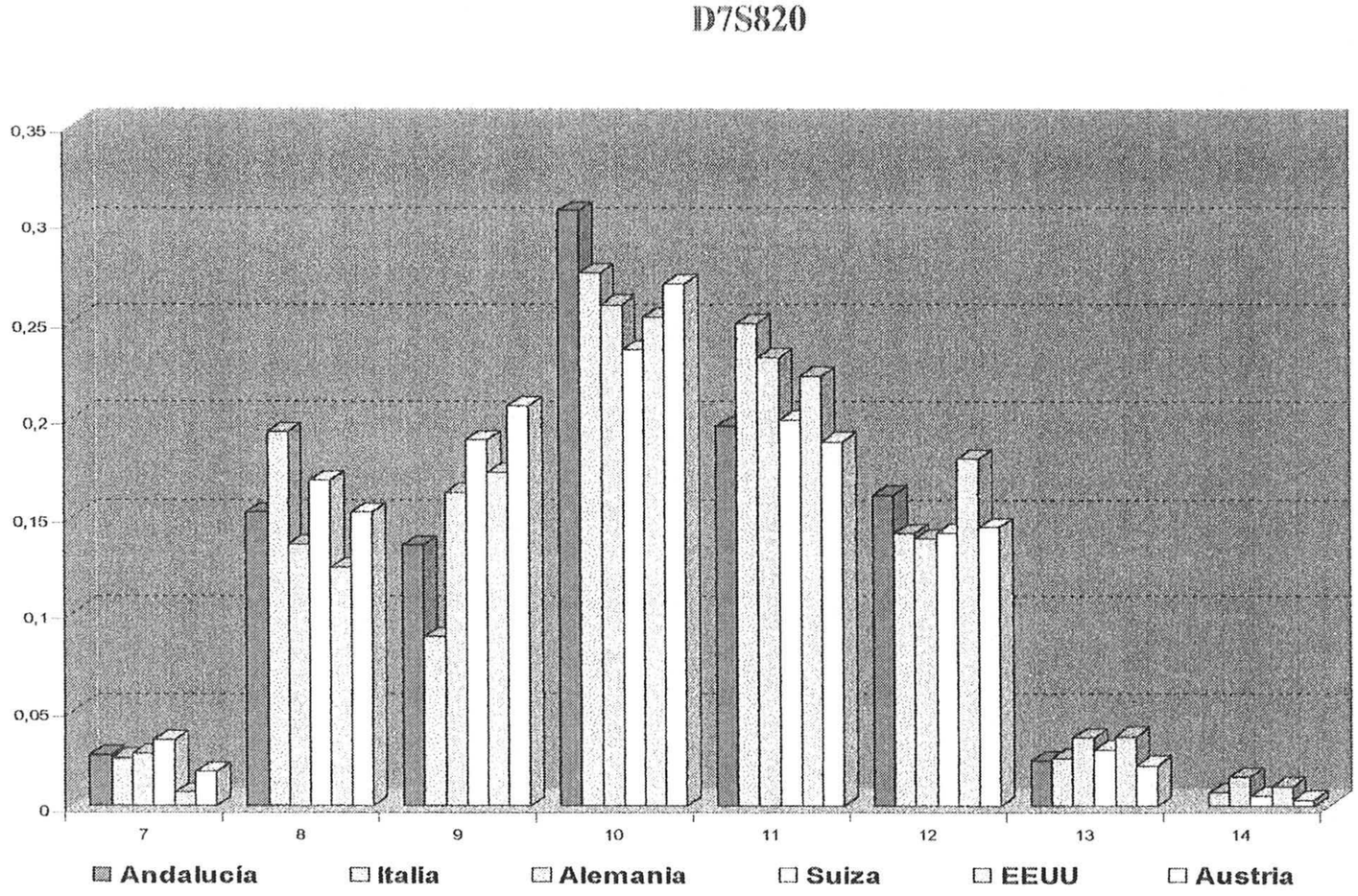
Gráfica 5.9. Distribución de frecuencias alélicas para el locus D18S51 en diferentes poblaciones analizadas.



Gráfica 5.10. Distribución de frecuencias alélicas para el locus d5S818 en diferentes poblaciones analizadas.



Gráfica 5.11. Distribución de frecuencias alélicas para el locus D13S317 en diferentes poblaciones analizadas



Gráfica 5.12. Distribución de frecuencias alélicas para el locus D7S820 en diferentes poblaciones analizadas

5.1.2.2. Discusión del análisis estadístico

Al igual que se hizo con los tres loci del triplex de Promega, aplicamos un análisis estadístico a los nueve loci estudiados para comprobar si cumplían las condiciones necesarias para poder ser aplicados en identificación genética.

Una vez aplicados los tests destinados a comprobar el **equilibrio de Hardy-Weinberg** se observó que, excepto el locus D7S820, los demás loci mostraron valores de p por encima de 0.05. El hecho de que para el D7S820 volviéramos a encontrar un desequilibrio en el test de razón de verosimilitud y en el test exacto de Guo-Thompson nos sugirió que probablemente el desequilibrio se debiese a un error derivado de la estructura interna de la muestra. En la distribución de frecuencias alélicas de este locus observamos que el alelo 10 presenta un valor bastante elevado (31.92%) con respecto al alelo menor representado (0.28%) y que, probablemente, esta sea la causa del desequilibrio. No obstante, como se verá más adelante cuando comentemos los resultados de las tablas de contingencia, no se observaron diferencias significativas con otras poblaciones caucasianas, lo cual hizo que no descartásemos este locus y que las frecuencias alélicas encontradas sean utilizadas actualmente en nuestro laboratorio para el cálculo de probabilidades tanto en criminalística como en la investigación biológica de la paternidad.

En cuanto al **test de correlación de Karlin**, encontramos valores de p por debajo de 0.05 en los loci D3S1358/D8S1179 (0.020), D3S1358/D7S820, VWA/D13S317 (0.018), VWA/D5S818 (0.002), FGA/D13S317 (0.024), D8S1179/D5S818 (0.001) y D21S11/D5S818 (0.048). Estadísticamente, cuando se comparan dos poblaciones (en este caso dos loci) a través de K características medidas en ellas, cada comparación individual ha de hacerse a un error α/K si se desea obtener un valor global de α (Martín Andrés et al, 1989).

Esta corrección se denomina *corrección de Bonferroni* y en nuestro caso daría un error de 0.0014 (0.05/36) con lo que tras su aplicación sólo quedaría significativa la asociación D8S1179/D5S818 cuyo valor de p (p=0.001) sigue quedando por debajo de 0.0014. No obstante, al no dar el programa informático más de tres cifras decimales nos queda la duda de si este valor de p pudiera ser igual (0.0014) o mayor (0.0015) que el factor de corrección.

En caso contrario, nos encontraríamos desde el punto de vista estadístico con una significancia para estos dos loci, aunque pensamos que genéticamente podríamos afirmar que estos dos loci son independientes puesto que así se ha demostrado en otros trabajos publicados recientemente (Kupferschmid et al, 1999; Gehrig et al, 1999; Neuhuber et al, 1999; Hanstschel et al, 1999).

El locus más informativo resultó ser el FGA, con un PD del 96% y un PE del 73,25%.

Con respecto a la comparación de nuestros datos con los de otras poblaciones (Italia, Suiza y EEUU) por medio de tablas de contingencia RxC los resultados obtenidos de probabilidad no muestran, por lo general, diferencias significativas. Se encontraron valores de p significativos: en Italia y EEUU con respecto a España para el locus D3S1358, en EEUU para el locus D13S317 y en Italia para el locus D18S51.

No se encontraron diferencias significativas para el locus D7S820 lo cual nos lleva a pensar, como se explicó anteriormente, que el desequilibrio encontrado en este marcador sea debido a un error interno de muestreo.

5.2. ESTUDIO DE VALIDACIÓN

Todo protocolo analítico debe ser validado antes de proceder a su uso generalizado, esto es, se debe de comprobar que los resultados que se obtienen son consistentes aún cuando las condiciones analíticas y de las muestras varíen para adaptarse a la praxis diaria.

Si lo anterior es importante en cualquier ámbito analítico, no cabe duda que alcanza una de sus máximas expresiones en criminalística forense, y si se quiere, aún más en genética forense.

Las muestras usadas en genética forense son pequeñas, pueden estar degradadas o contaminadas y, en un gran porcentaje, son irrepetibles. Los resultados que se obtengan de lo que a veces puede ser un único análisis son los que constarán formalmente en los informes y los que guiarán, certera o equivocadamente, la investigación judicial.

Entre los loci que se han analizado en esta tesis nos encontramos con el **D13S317**, **D7S820** y **D16S539** que se amplifican y estudian simultáneamente a través de una reacción múltiple (triplex) usando el kit Silver III de la casa Promega (Promega Corporation, Madison, Winsconsin, EE.UU.), diseñado para tinción con nitrato de plata.

La introducción de este kit es reciente (Julio 1998) y la elección de estos tres loci como parte del sistema CODIS (Combined DNA Index System) norteamericano ha hecho que la atención recaiga sobre los mismos por parte de la comunidad forense internacional. En consecuencia, si su uso va a ser generalizado en criminalística, la validación es una necesidad absoluta.

Tanto el Grupo Técnico de Trabajo para los Métodos de Análisis de ADN (Technical Working Group in DNA Methods Analysis, TWGDAM, hoy SWGDAM) como la AABB (American Association of Blood Banks) recomiendan y exigen estudios de validación de los loci genéticos previos al uso forense, civil o penal (Micka et al, 1996).

Nuestro grupo de trabajo tuvo la oportunidad de ser uno de los primeros en comprobar la utilidad del mismo, a través de una serie de kits de prueba cedidos por el fabricante en el mes de febrero de 1998. La compañía Promega estaba especialmente interesada en conocer el rendimiento de los tres loci mencionados tras la detección de los alelos con nitrato de plata, por lo que junto a un estudio poblacional de caucásicos españoles (Entrala et al, 1999) se procedió a validar el kit Silver III.

La validación en el campo forense pretende básicamente comprobar dos extremos:

1. Que la técnica es fiable y funciona adecuadamente con muestras procedentes de condiciones de almacenamiento adversas.
2. Que los resultados que se obtengan sean sólo positivos o negativos, pero que nunca puedan inducir a error por aparición de alelos nuevos.

Obviamente, existen otros aspectos de gran interés en área forense, como comprobar exactamente las cantidades máximas y mínimas con las que se puede trabajar adecuadamente, comprobar el efecto de los contaminantes más comunes, determinar la posible amplificación de ADN de animales domésticos (gatos, perros, aves), etc...

Ello hace que los estudios de validación forense puedan ser verdaderamente amplios, incluso inagotables en todos sus aspectos, tal y como a modo de ejemplo se muestra en la siguiente tabla:

	Fregeau et al 1999	Budowle et al 1997	Micka et al 1999
Tipo de detección	Fluorescente	Nitrato de plata	Fluorescente
Contaminación química	Si	Si	No
Contaminación bacteriológica	Si	Si	Si
Contaminación fúngica	Si	Si	No
Animales	Si	Si	Si
Condiciones ambientales	Si	Si	Si
Mezclas biológicas	Si	Si	Si
Cantidades variables	Si	Si	Si
Casos reales	Si	Si	Si

El gran número de trabajos de validación llevados a cabo por muy diversos autores como los mencionados en la anterior tabla y otros (Micka et al, 1996; Sparkes et al, 1996; van Oorschot et al, 1996; Martínez de Pancorbo et al, 1998; Anderson et al, 1999) han puesto de manifiesto una serie de circunstancias.

El factor común más importante que nosotros podemos resaltar, es que lo verdaderamente intrínseco a cada nuevo conjunto de loci que se describe es la cantidad apropiada de ADN con la que rinden resultados válidos. Al mismo nivel de importancia destacaríamos igualmente, la capacidad de discernir mezclas, o sea, la presencia de ADN de más de una persona en proporciones variables en una misma amplificación.

5.2.1. Estudio de sensibilidad

El estudio de sensibilidad está diseñado para observar cuáles son las cantidades mínimas (principalmente) y máximas que se pueden amplificar en condiciones estándar sin que ello ocasione problemas en la interpretación de los resultados.

En nuestros ensayos hemos utilizado como muestra de ADN el procedente de dos líneas celulares: **K562** y **9947-A** y ello por dos razones básicas.

Primera: porque los genotipos para cada uno de los loci son perfectamente conocidos, ya que ambos son controles positivos habituales en los kits y laboratorios.

Segunda: porque al proceder de productos comerciales de alta calidad para biología molecular, tanto la cantidad (ng/μl) como la calidad se suponen ideales, lo que facilita la preparación de muestras con una concentración de ADN determinada y la interpretación de los resultados.

La fotografía 4.1. del capítulo de Resultados muestra un gel con lo que podría ser un resultado típico de los obtenidos en nuestro estudio. En el mismo se observa que, entre las cantidades por nosotros consideradas, que oscilaban entre 0.1 y 1 ng, aquella que permitió una asignación correcta e inequívoca de los alelos ("tipado" en el lenguaje coloquial en genética forense) fue la de 0.4 ng.

Cantidades menores (0.3 ng) rinden bandas visibles en los geles, pero no son interpretables adecuadamente a no ser que se conozcan a "priori" los alelos que han de aparecer.

Cantidades de 1 a 5 ng rinden resultados que podríamos llamar "ideales", lo que implica que el kit fue optimizado para estas cantidades por el fabricante (Promega Corporation, Madison, Wisconsin, EEUU)

Comparando estos resultados con estudios similares, encontramos que, en conjunto, los resultados publicados por otros autores son similares a los nuestros. Hay que decir, en este momento, que a veces resulta difícil establecer una comparación con otros estudios, y ello por dos razones. En primer lugar porque los modos de mezclar las muestras varían de unos trabajos a otros, y en segundo término porque los sistemas de amplificación (termocicladores, primers, etc.) y de detección (nitrato de plata, fluorescencia y, dentro de la fluorescencia, tipo de aparato utilizado para detectarla) infieren diferencias que hay que valorar con prudencia.

Trabajando con un kit muy similar a éste, el CTT (Promega Corporation, Madinson, Wisconsin, EEUU), Budowle et al (1997) concluyeron que el límite de sensibilidad podía establecerse en 0.20 y 0.25 ng, aunque recomiendan el uso mínimo de 0.40-0.45 ng para evitar problemas en la práctica forense.

Sin embargo, Van Oorschot et al (1996) consiguen amplificar usando tan sólo 0.05 ng de ADN, aunque en este caso no lo hacen con un kit múltiple, sino con el locus HUMTH01, usando condiciones optimizadas para el mismo y un total de 32 ciclos, mientras que los autores anteriormente citados (como Budowle et al, 1997) usan de 28 a 30 ciclos; más aún, Van Oorschot indica que aumentando el número de ciclos de 32 a 35 consigue resultados apropiados a partir de tan sólo 0.01 ng (lo que equivale a 10 pgr, escasamente 2 células).

Un grupo de trabajo español (Martínez de Pancorbo et al, 1998), usando igualmente el kit CTT y detección con nitrato de plata indican que la mínima cantidad que se pudo tipar adecuadamente es de 0.1 ng.

Katherine Micka, liderando un amplio equipo de diversos laboratorios norteamericanos y usando el kit CTT y los loci HUMHPRTB, VWFA31 y FES-FPS concluye que, pese a que existen diferencias dependiendo del locus considerado, la cantidad mínima analizable con el locus CTT es de 0.25 ng (Micka et al, 1996). Los otros tres loci analizados pertenecían a un triplex entonces en desarrollo por la firma Promega, y la cantidad mínima individual variaba desde los 0.1 ng necesarios para los loci HUMHPRTB y HUMVWAF31, hasta los 0.5 ng necesarios para tipar adecuadamente el locus HUMFESFPS.

Usando sistemas de detección basados en la emisión de fluorescencia de los productos amplificados, los resultados vienen a ser similares, pese a que en estos casos influyen muchos más parámetros (modificables por los analistas) en el momento de tomar decisiones.

Un grupo multicéntrico vuelto a encabezar por Micka (Micka et al, 1999) publicó recientemente los estudios de validación llevados a cabo para el kit PowerPlex 1.1 (Promega Corporation, Madison, Wisconsin, EEUU) y el cuádruplex FFFL (F13A01, FESFPS, F13B y LPL). La conclusión de dicho estudio es que la cantidad mínima amplificable con fiabilidad es de 0.2 ng, aunque en algunos casos se pueden usar 0.1 ng con el PowerPlex 1.1 con un total de 30 ciclos de amplificación. En conjunto, se aconseja el uso de una cantidad mínima de 0.2 ng.

Fregeau et al (1999), en un amplísimo trabajo de validación de dos sistemas múltiplex utilizados por la Policía Montada del Canadá, indican que la cantidad mínima que permite resultados positivos consistentes es la de 0.32 ng, aunque se

UNIVERSIDAD DE GRANADA
3 1 MAR. 2000
COMISION DE DOCTORADO

pueden interpretar resultados procedentes de amplificaciones de sólo 0.04-0.16 ng, si bien en estos casos aparecen picos "extra" que podrían invalidar el resultado frente a un tribunal.

Como consideraremos en la discusión global, la interpretación de los resultados en los sistemas multiplex exige la perfección en todos y cada uno de los loci, de tal modo que, sirva el símil, roto o inutilizado el eslabón o locus menos sensible, invalidados los resultados.

El kit validado por nosotros, y de cuyos resultados podemos aconsejar una cantidad mínima de 0.4 ng para el uso forense, muestra dificultades con el locus D16S539 que, en condiciones estándar, presenta graves problemas para la adecuada asignación del valor numérico del alelo, ya que en algunos tipos de geles (dependiendo de la acrilamida que se utilice) aparece una banda desdoblada al ser éstos un medio desnaturalizante. Este desdoblamiento no impide la lectura de los resultados, pero exige una cierta intensidad - difícil de conseguir con cantidades menores a 0.4 ng- para tener la seguridad de asignar la nomenclatura adecuada. Esto se complica aún más como veremos en el siguiente apartado, destinado a la valoración de mezclas.

5.2.2. Estudio de mezclas

La gran sensibilidad derivada del uso de técnicas de PCR conlleva la posibilidad de amplificar cantidades mínimas de ADN ajenas a la muestra que se pretende analizar originariamente. Esta circunstancia conocida como "contaminación biológica", que en cualquier ámbito de la biología molecular correspondería a un error de manipulación tiene, en criminalística, una gran importancia, ya que lejos de ser un artefacto que impida la interpretación de resultados, puede ser un hecho de gran trascendencia.

Las muestras forenses pueden estar, en el momento de llegar al laboratorio, contaminadas con restos biológicos de otras personas. Siguiendo a Lorente y Lorente (1995), aceptamos que la contaminación biológica puede ser:

1. Previa
2. Paralela o concomitante
3. A posteriori
    - 3.1. Accidental
    - 3.2. Criminal
    - 3.3. Imprudente

Cualquiera que sea la circunstancia y el contexto de la investigación, no cabe duda de que los kits y pruebas de uso forense han de ser capaces de detectar pequeñas cantidades de ADN que esté mezclado con un ADN mayoritario o predominante.

A pesar de que en la práctica, cuando se sospecha de la coexistencia de diferentes muestras biológicas (por ejemplo semen y células vaginales; o semen y saliva) se trata de separarlas por medios específicos como la extracción diferencial

REF???, hay otras veces en que esto es imposible (cuando son células del mismo tipo pero de individuos diferentes), o bien lo más interesante es, precisamente, detectar indicios en muy baja concentración para que no pasen desapercibidos y puedan ser adecuadamente valorados en la investigación criminal.

Es, por tanto; parámetro típico en la validación de un nuevo locus polimórfico, el determinar la cantidad mínima de ADN que se amplifica en presencia de proporciones (cantidades) variables de otro ADN diferente.

Como se explica en el apartado correspondiente del capítulo III de Material y Métodos, en nuestra validación se tomaron cantidades constantes de un ADN determinado que se mezclaron con cantidades variables de otro ADN (con genotipos diferentes), como se observa en las tablas 3.1 y 3.2 bis.

Tabla 3.1 bis. **Mezclas A + B**

ADN muestra A (ng)	ADN muestra B (ng)	Proporciones A : B
5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:20, 1:10, 1:4, 1:2, 1:1, 1:0.5, 1:0.2, 1:0.1, 1:0.05, 1:0.02
2.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:40, 1:20, 1:8, 1:4, 1:2, 1:1, 1:0.4, 1:0.2, 1:0.1, 1:0.04
1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:100, 1:50, 1:20, 1:10, 1:5, 1:2.5, 1:1, 1:0.5, 1:0.25, 1:0.1
0.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:200, 1:100, 1:40, 1:20, 1:10, 1:5, 1:2, 1:1, 1:0.5, 1:0.2
0.25	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:400, 1:200, 1:80, 1:40, 1:20, 1:10, 1:4, 1:2, 1:1, 1:0.4
0.1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:1000, 1:500, 1:200, 1:100, 1:50, 1:25, 1:10, 1:5, 1:2.5, 1:1

Tabla 3.2. bis. **Mezclas A + C.**

ADN muestra A (ng)	ADN muestra C (ng)	Proporciones A : C
1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:100, 1:50, 1:20, 1:10, 1:5, 1:2.5, 1:1, 1:0.5, 1:0.25, 1:0.1
0.5	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:200, 1:100, 1:40, 1:20, 1:10, 1:5, 1:2, 1:1, 1:0.5, 1:0.2
0.25	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:400, 1:200, 1:80, 1:40, 1:20, 1:10, 1:4, 1:2, 1:1, 1:0.4
0.1	100, 50, 20, 10, 5, 2.5, 1, 0.5, 0.25, 0.1	1:1000, 1:500, 1:200, 1:100, 1:50, 1:25, 1:10, 1:5, 1:2.5, 1:1

En todos los casos se siguieron los protocolos especificados por el fabricante (Promega Corp., Madison, Wisconsin, EEUU), y no fue motivo de esta investigación el modificar parámetros analíticos de amplificación o electroforesis con objeto de aumentar el rendimiento, ya que son objeto de otros estudios que no conforman la presente tesis.

De acuerdo a nuestros resultados, el conjunto de los tres loci analizados se muestra relativamente robusto, aunque no todos los loci tienen la misma sensibilidad para detectar mezclas. El conjunto de los resultados puede verse en las tablas 4.34 a 4.38, en la que se muestran con un signo (-) las amplificaciones fallidas de las

cantidades que aparecen en la columna de la izquierda tras ser mezcladas con cantidades crecientes de un mismo ADN (fila superior). Se consideraron resultados positivos (+) aquellos en los que se obtuvo dos o más amplificaciones positivas de un mínimo de tres intentos.

En primer lugar, consideraremos la amplificación de K562 (A) frente a cantidades variables de ADN procedente de una muestra de sangre (B); y posteriormente, frente a diversas cantidades del ADN control 9947-A (C). Los comentarios referidos a cada locus individual se referirán a los resultados obtenidos por nosotros ya que hasta el momento no existen trabajos de validación publicados para estos tres loci y, por lo tanto, no podemos establecer comparaciones directas en este apartado. No obstante, al considerar los resultados obtenidos en conjunto si podremos establecer algunas comparaciones con resultados obtenidos por otros autores con kits similares.

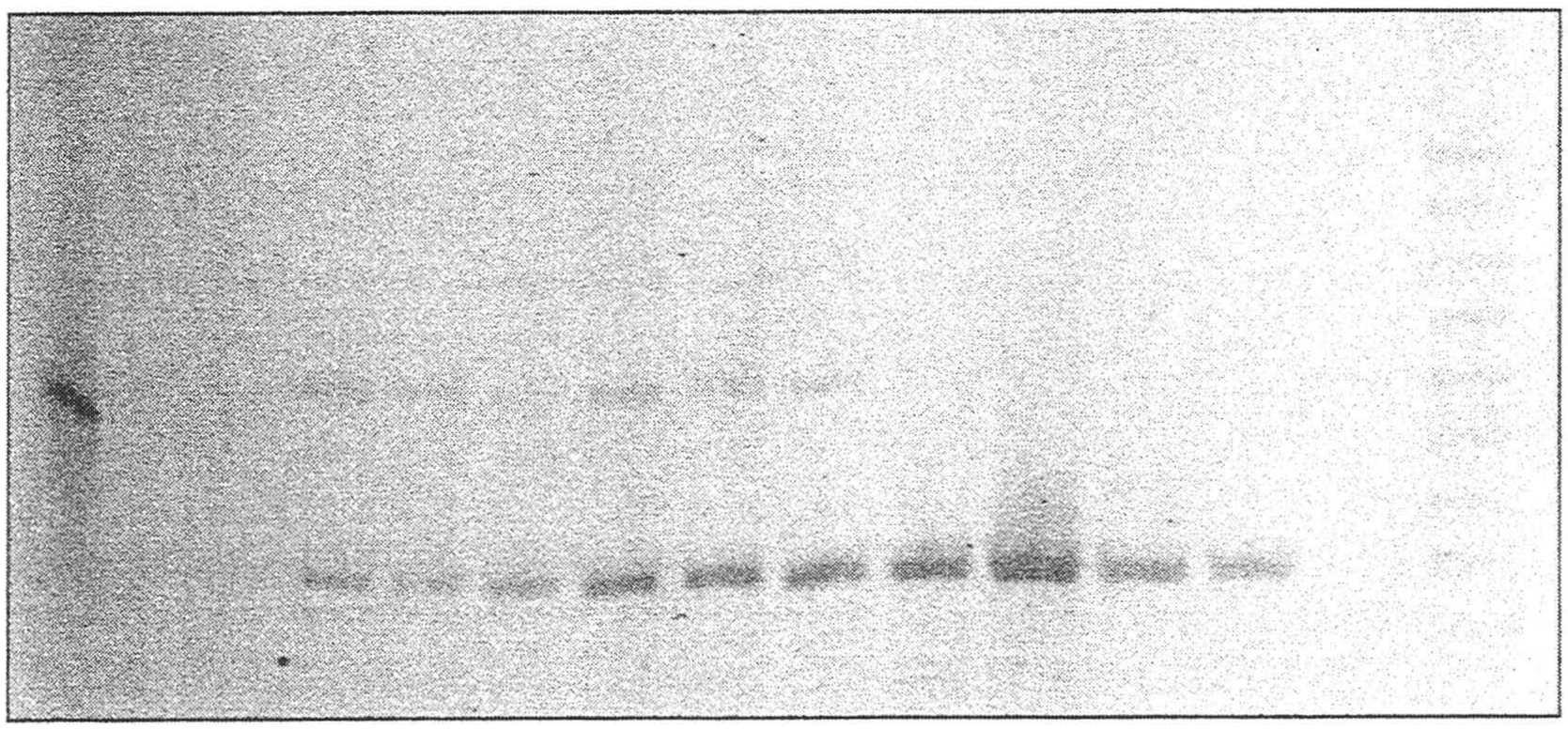
LOCUS D13S317

Tanto en las mezclas A + B (ADN de K562 + ADN procedente de sangre), como en las A + C (ADN de K562 + ADN de 9947-A), los fallos de amplificación aparecen cuando se trata de amplificar cantidades muy pequeñas (0.25 y 0.1 ng) en presencia de cantidades mucho mayores de otro ADN. En conjunto, proporciones de 1:80 (1.25% del ADN presente en la amplificación) o menores son las que muestran problemas.

En el caso de las **mezclas A + B** observamos resultados negativos en las proporciones 1:200 (cuando la cantidad menor es de 0.5 ng), 1:40, 1:80, 1:200, 1:400 (cuando la cantidad menor es de 0.25 ng) y 1:500, 1:1000 (cuando la cantidad menor es de 0.1 ng). Es decir, cantidades grandes de ADN (entre 20 y 100 ng) cuando se mezclan con cantidades muy pequeñas de otro ADN (por debajo del límite de sensibilidad, 0.25 y 0.1 ng) enmascaran la amplificación del ADN que se

encuentra en menor proporción. Debemos señalar que, sin embargo, cantidades comprendidas entre 20 y 0.1 ng permiten la amplificación de la muestra que se encuentra en menor proporción, observándose a veces una intensidad de amplificación mayor en dichas muestras con respecto a la muestra sola (es decir, sin mezclar). El mismo hecho es observado en las mezclas A + C, en las que como se explica en la foto inferior, cantidades pequeñas de ADN (como 0.1 ng en presencia de cantidades mayores (50 y 100 ng) dan resultados negativos en la amplificación.

Imagen 5.1. Visualización para el locus D13S317 de una serie de mezclas de las muestras A y C en diferentes proporciones y en las que de la muestra C se pusieron de manera constante 0.1 ng de ADN. De izquierda a derecha y numerando las calles desde el 0 al 14 se observa: **0)** Marcador de peso molecular, **1)** K-, **2)** 1:1, **3)** 1:2.5, **4)** 1:5, **5)** 1:10, **6)** 1:25, **7)** 1:50, **8)** 1:100, **9)** 1:200, **10)** 1:500, **11)** 1:1000, **12)** 5 ng de K562, **13)** 0.1 ng de C, **14)** Ladder D13S317.

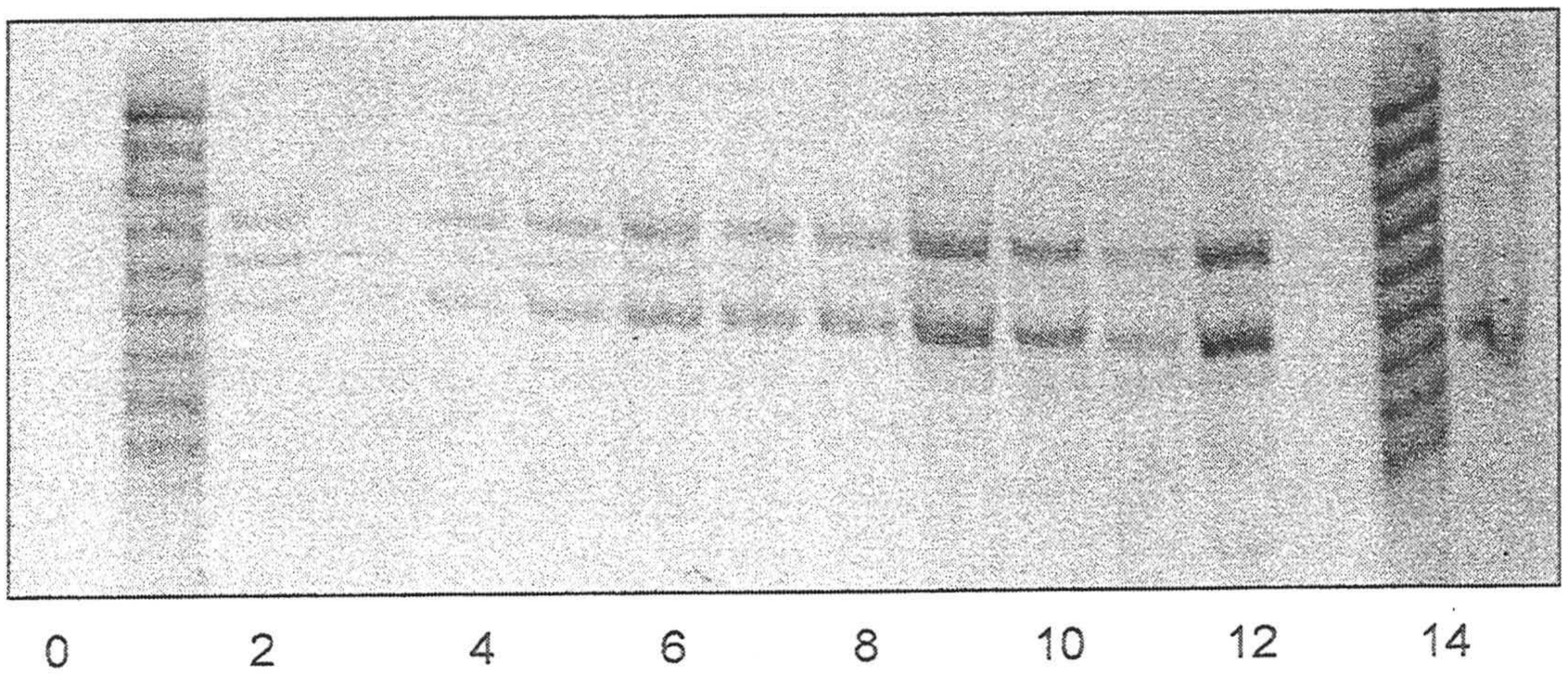


0 2 4 6 8 10 12 14

LOCUS D7S820

Los resultados obtenidos para el locus D7S820 fueron similares a los obtenidos para el locus anterior. En ambos tipos de mezclas se obtienen resultados positivos cuando el ADN que está en menor proporción está por encima del límite de sensibilidad (0.4 ng), en cambio, en aquellas mezclas en las que el ADN en menor proporción está por debajo de dicho límite, se observaron resultados negativos en las proporciones menores. Para este locus seguimos pensando que cantidades grandes de ADN (por encima de 50 ng) enmascaran la amplificación del ADN que está en cantidades muy pequeñas (0.25 y 0.1 ng). Además, como se puede observar en la imagen 2, también se da el hecho de que en ciertas mezclas, la intensidad de amplificación de estas cantidades mínimas de ADN aumenta cuando se amplifican con ciertas cantidades de ADN de la otra muestra.

Imagen 5.2. Visualización para el locus D7S820 de una serie de mezclas de las muestras A y C en diferentes proporciones y en las que de la muestra C se pusieron de manera constante 0.25 ng de ADN. De izquierda a derecha y numerando las calles desde el 0 al 14 se observa: **0)** K-, **1)** Ladder D7S820, **2)** 1:0.4, **3)** 1:1, **4)** 1:2, **5)** 1:4, **6)** 1:10, **7)** 1:20, **8)** 1:40, **9)** 1:80 , **10)** 1:200 **11)** 1:400, **12)** 5 ng de K562, **13)** 0.25 ng de C, **14)** Ladder D7S820.



LOCUS D16S539

Por comparación con los anteriores es, sin duda alguna, el que tiende a ofrecer peores resultados, bien entendido que, pese a esta afirmación, continúa siendo un STR robusto como parte del triplex que hemos considerado.

Efectivamente, como ya se trató parcialmente en el apartado anterior, en nuestras condiciones aparece sistemáticamente con desdoblamiento de las bandas al usar ciertos tipos de acrilamida como la comercializada por Promega (¿??) que complican enormemente la interpretación de los resultados, especialmente si se trata de mezclas. Recientemente se han comercializado un nuevo tipo de acrilamida (Prona Sec 6: acrilamida 6%, TBE 1X, Urea 6M, pH 8.3), fabricada por una empresa española (Pronadisa, España), que evitan el problema del desdoblamiento del locus D16S539, pero que originan problemas en otro locus de otro triplex (HUMvWA, triplex FFv, Promega Corporation). En las imágenes 5.3 y 5.4 Puede apreciarse la diferencia en cuanto a nitidez de las bandas en los tres loci dependiendo del tipo de acrilamida utilizada. En el caso del loci D16S539 el desdoblamiento de bandas desaparece al utilizar la acrilamida de Pronadisa al 6%.

Nosotros, preparando en el laboratorio la acrilamida a la concentración recomendada para el triplex SilverSTR III, no conseguimos paliar el problema mencionado, por lo que damos la bienvenida a esta nueva acrilamida. Desconocemos si existe en la actualidad alguna acrilamida comercializada capaz de rendir resultados perfectos para todos los loci habitualmente incluídos en los triplex (CTT, FFV, SilverSTR III), que obviamente sería lo ideal.

Imagen 5.3. Visualización de un gel realizado con acrilamida de Promega al 6%.

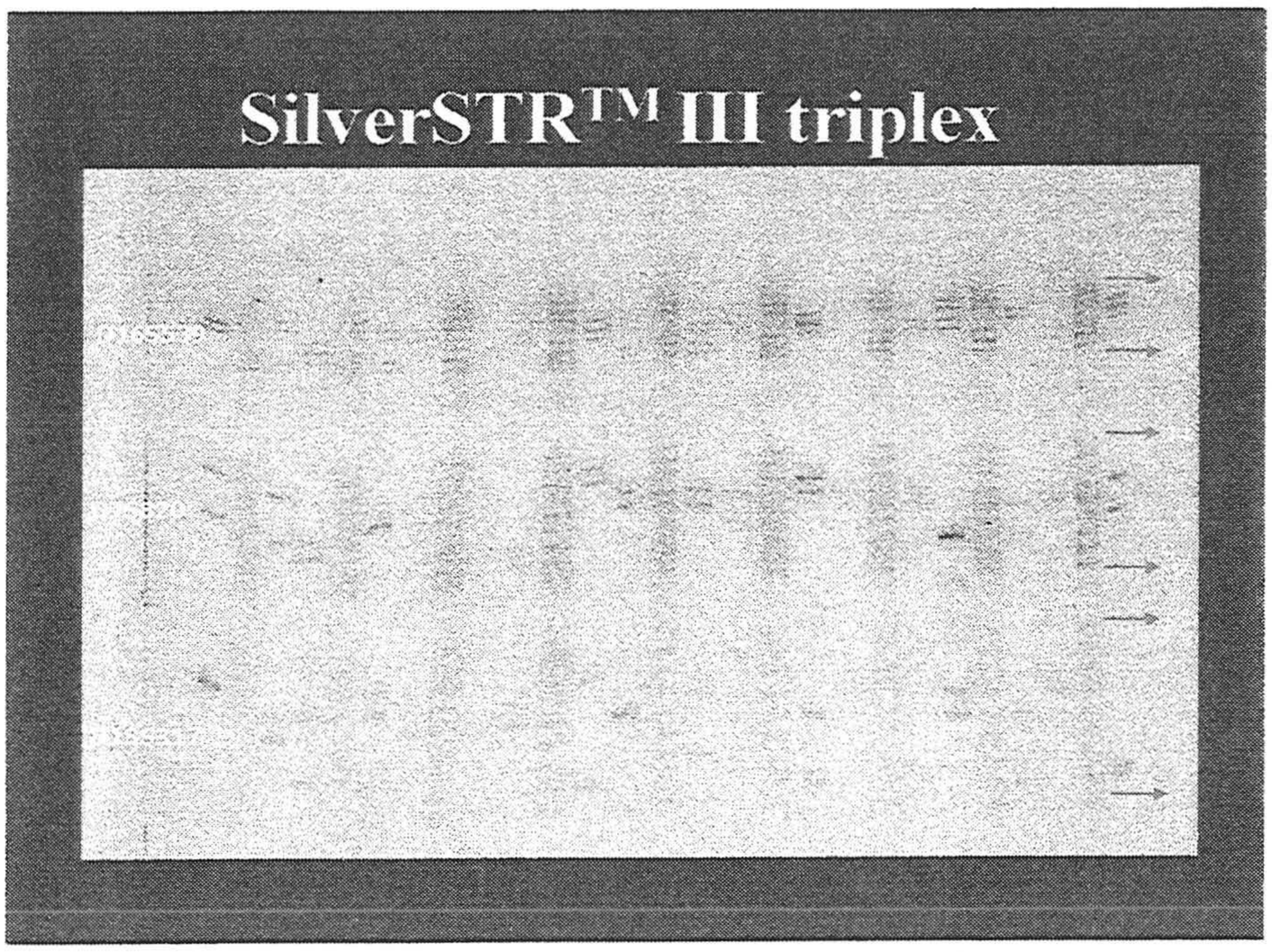
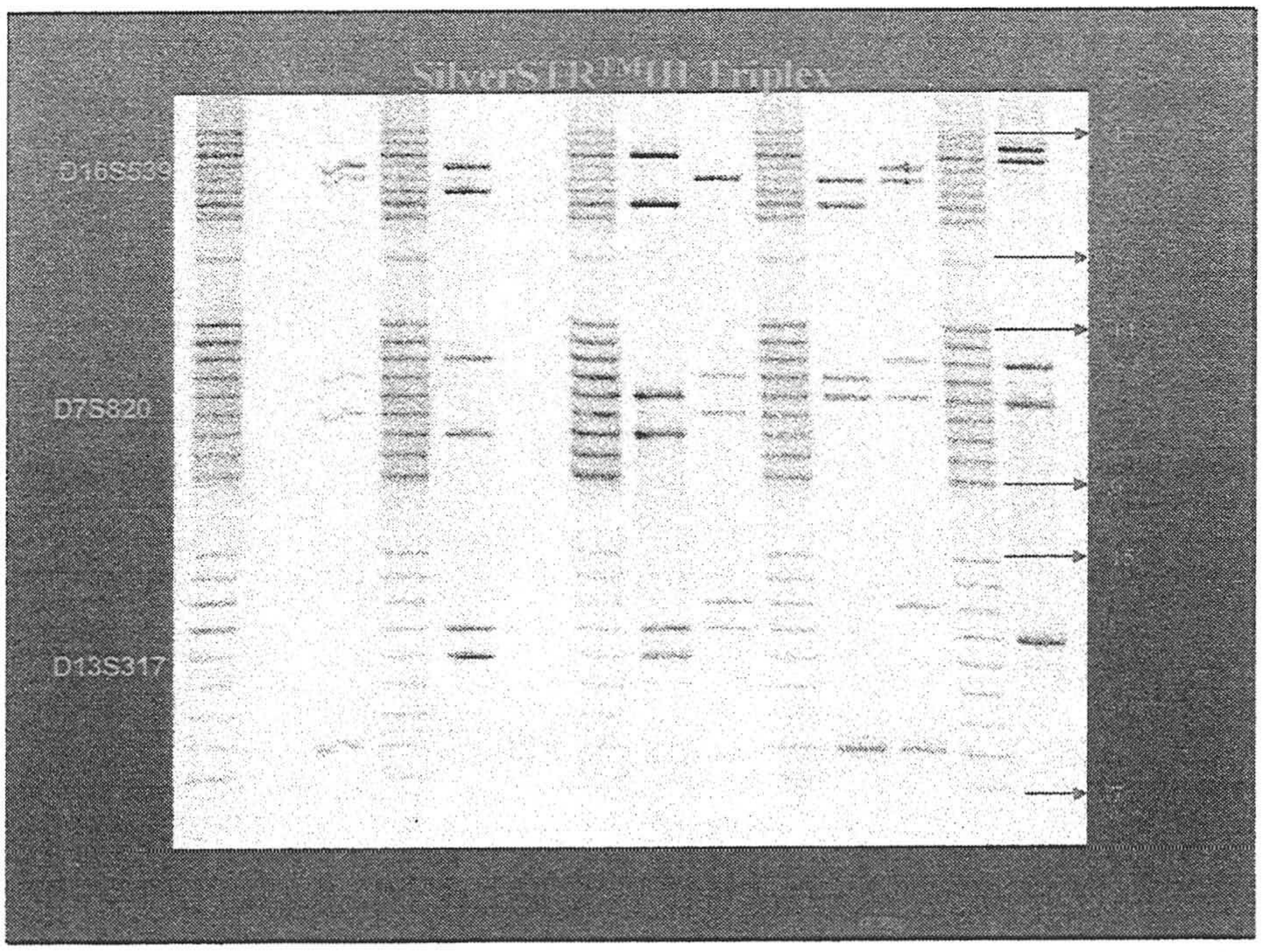
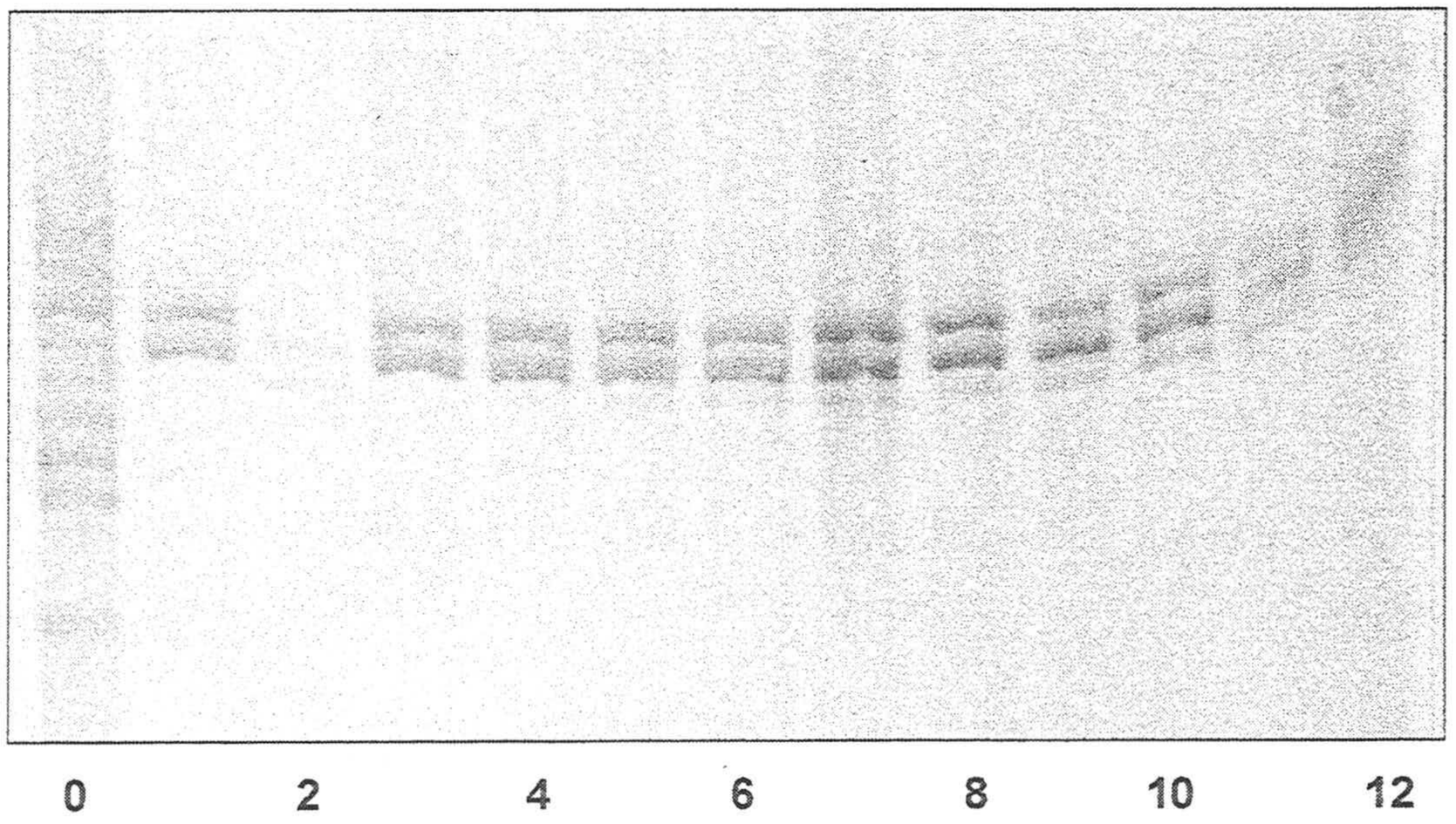


Imagen 5.4. Visualización de un gel realizado con acrilamida de Pronadisa al 6%.



Para el locus D16S539 sólo nos sirvieron los resultados obtenidos con las mezclas A + C ya que las muestras A y B mostraban el mismo genotipo (11,12). Nosotros hemos sido incapaces de interpretar adecuadamente mezclas inferiores al 5% (1:20).

Imagen 3. Visualización para el locus 16 de una serie de mezclas de las muestras A y B en diferentes proporciones y en las que de la muestra B se pusieron de manera constante 0.25 ng de ADN. De izquierda a derecha y numerando las calles desde el 0 al 12 se observa: **0)** Ladder D16S539, **1)** 5 ng A, **2)** 0.25 ng C, **3)** 1:400, **4)** 1:200, **5)** 1:80, **6)** 1:40, **7)** 1:20, **8)** 1:10, **9)** 1:4 , **10)** 1:2 **11)** 1:1, **12)** 1:0.4.



Discusión conjunta de los tres loci

En general, el triplex SilverIII STR, se muestra útil para su uso en casos de muestras forenses en las que se presume o se sabe, coexisten dos o más muestras de ADN ya que, al igual que otros kits comercializados, permite detectar proporciones de 1:20 (5%) e incluso a veces, otras mezclas más desproporcionadas (1:200, 1:500,...).

Los resultados de otras publicaciones referentes a estudios de validación con otros kits muestran, en general resultados similares. Así, un grupo de investigadores liderado por Micka (Micka et al, 1996), usando sistemas de análisis por nitrato de plata, detectan muestras presentes en proporciones de 1:20 y 1:10 (5% y 10%) en el caso del uso de triplex, y llegan a detectar muestras presentes en proporción 1:100 (1%) para algunos loci, como el FESFPS o el VWA.

Sparkes et al (1996), usando detección fluorescente, son capaces de detectar adecuadamente muestras en concentración menor al 10% (proporción 1:10).

Estos resultados son similares a los de Martínez de Pancorbo et al (1998), capaces de detectar 0.1 ng y mezclas presentes en sólo un 5% (1:20), usando un triplex (CTT, Promega Corporation) y posterior detección con nitrato de plata. Van Oorschot et al (1996) fueron capaces de detectar también mezclas de proporciones muy similares a las nuestras (0.1 ng: 5 ng), que equivale a proporciones del 2% (ratio 1:50) del componente menor, usando el locus HUMTH01, uno de los más robustos y clásicos en Genética Forense.

Para finalizar este apartado queremos volver a insistir en la dificultad en este tipo de estudios para obtener resultados reproducibles al 100% por dos motivos:

- Primero por la falta de un método de cuantificación de ADN que nos permitiese añadir cantidades exactas en la amplificación. El margen de error que asumen los métodos actuales de cuantificación puede que no sean muy significativos cuando la cantidad de ADN es relativamente grande, pero a medida que la cantidad va disminuyendo este error aumenta, de manera que es realmente difícil llegar a cuantificar con exactitud muestras mínimas de ADN.

- Segundo porque es sabido que la distribución del ADN en disolución no es homogénea, por lo que hay que asumir otro margen de error a la hora de pipetear los volúmenes de amplificación.

No obstante, pensamos que los resultados obtenidos pueden ser orientativos a la hora de trabajar con muestras en las que se presume la existencia de material biológico de dos o más personas, hecho que, como indicamos anteriormente, es bastante frecuente en los laboratorios de Genética Forense.

5.2.3. ESTUDIO DE MUESTRAS FORENSES

El objetivo perseguido con el análisis de muestras forenses no fue el de evaluar con detalle las cantidades y calidades de ADN necesarios para, según la muestra biológica de origen, conseguir una amplificación adecuada. Así mismo, tampoco se pretendió establecer un estudio de correlación entre tipo de muestras y data de las mismas usando como referencia la calidad de la amplificación.

Por el contrario, lo que se pretendía era observar la capacidad de este kit de rendir resultados con muestras forenses que se sabía que estaban contaminadas, degradadas y con baja concentración de ADN, ya que habían sido analizadas con anterioridad en nuestro laboratorio.

Sin atrevernos a obtener conclusiones estadísticas (puesto que faltan datos para ello), era de nuestro especial interés, en primer lugar, comprobar la capacidad de este nuevo triplex de rendir resultados con muestras reales, difíciles de amplificar, y, de modo simultáneo, pretendíamos poner de manifiesto la falta directa de correlación entre cantidad, calidad y amplificación positiva, hecho ampliamente conocido en la práctica forense.

Así, observamos como muestras en las que se vio un ADN degradado (D) y en baja cantidad, rinden una amplificación positiva; igualmente, muestras muy degradadas han podido ser analizadas adecuadamente con este multiplex. Paralelamente, muestras en las que había ADN de buena calidad (B) e incluso en cantidades suficientes para las exigencias de este kit, no generaron resultados positivos.

Estos resultados vienen a poner de manifiesto, una vez más, las peculiaridades que envuelven al trabajo en los laboratorios de Genética Forense, donde de modo

preliminar, son muy pocos los aspectos que pueden conocerse. Por ello, consideramos que todas las muestras, todos los indicios deben de analizarse, independientemente de la calidad, siendo el único criterio aplicable el de la cantidad, ya que el intento de amplificación de muestras con muy poco ADN puede originar amplificaciones estocásticas, con resultados que pueden llevar a equívocos en el sentido de falsas exclusiones (por detección de falsos homocigotos, por ejemplo) (Budowle,????).

Finalmente, queremos señalar que las cuantificaciones de las muestras fueron hechas con geles de agarosa teñidos con bromuro de etidio porque en el mismo proceso se pretendía observar no sólo la cantidad de ADN, sino la calidad del mismo. Pese a que la cuantificación por técnicas de "slot-blot" e hibridación (Waye et al, 1986) son mucho más sensibles, y de aplicación rutinaria en nuestro laboratorio, consideramos en este caso más útil la cuantificación en geles de agarosa por la razón indicada anteriormente.

Hoy en día, con la existencia y el uso masivo (casi exclusivo) de loci STR (CODIS, AICEF/GITAD, INTERPOL), el problema de la calidad de la muestra pasa a tomar un papel secundario, aunque no por ello hay que dejar de considerarlo. Queremos decir con ello que tanto en los resultados analizados en este estudio como en los derivados de nuestra experiencia, cantidades y calidades mínimas de ADN son capaces de rendir resultados positivos cuando se estudia en el mismo fragmentos muy pequeños de ADN, como son los STR.

En todo caso, problema diferente será, que los resultados, a la par de ser suficientemente claros, sean lo suficientemente "perfectos" como para lograr que sistemas judiciales como el norteamericano (por ejemplo), admita los resultados sin mayor oposición.

5.3. DISCUSION GLOBAL

En su conjunto, el kit desarrollado por la firma Promega (USA) y comercializado bajo el nombre SilverSTR III, se muestra, en nuestras manos, como una herramienta útil en el ámbito de la genética forense. Más útil aún resulta en su aplicación en el estudio biológico de la paternidad, ya que aplicado en conjunción con otros marcadores STR permite llegar a valores de índice de paternidad o probabilidad de paternidad bastante elevados.

En el campo de la criminalística su importancia deriva de varias razones. En primer lugar, porque viene a aumentar la batería de marcadores disponibles para la resolución de casos criminales en términos puramente estadísticos (PD, PE,...). En segundo lugar, por el aspecto metodológico, ya que se usan protocolos similares a otros kits en uso, lo cual significa que es fácil y de rápida implantación en los laboratorios que tienen experiencia con kits similares. Y en tercer lugar, por lo que un kit para detección con nitrato de plata significa en sí, más allá de la facilidad de su implantación práctica. El nitrato de plata, frente a la técnica de detección fluorescente, es muchas veces, para muchos laboratorios, la única alternativa para el análisis de STR.

A pesar de que la técnica estrella del momento sea la detección fluorescente, debido a la serie de ventajas que ofrece respecto a la detección con nitrato de plata, no creemos oportuno dejar de lado las técnicas clásicas que aún se utilizan en muchos laboratorios de genética forense. Las técnicas fluorescentes tienen mayor sensibilidad y el tratamiento informatizado de los resultados puede facilitar la interpretación de los casos complejos (debido a la calidad de la muestra, a la aparición de "stutter bands", etc), y ello son ventajas irrefutables.

Sin embargo, es necesario que los loci analizados actualmente puedan ser estudiados también por técnicas básicas, manuales y baratas, asequibles a cualquier laboratorio o incluso a los grandes laboratorios cuando hay problemas técnicos. En base a estas características, tanto el Grupo Científico de Trabajo para Métodos y Análisis del ADN (SWGDAM) capitaneado por el FBI, como el Grupo Iberoamericano de Trabajo en ADN (GITAD) de la Academia Iberoamericana de Ciencias y Estudios Forenses (AICEF), de espacial interés para nosotros, incluyen a los tres loci de este triplex como loci de interés especial y de uso común entre los diferentes países (Gitad, Profiles in DNA, 2000???) para poder comparar datos, construir bases de datos, etc.

Por comparación con otros loci y kits similares tipo triplex comercializados también por Promega (CTT: CSF1PO, TPOX, TH01; FFV: F13A, FES/FPS, VWA), consideramos que el kit Silver III presenta algunos problemas en ciertas ocasiones (sobre todo con muestras de baja calidad o cantidad de ADN). Problemas que hemos de decir, no siempre son debidos al diseño del kit en cuestión o al estado del ADN (como los derivados de una acrilamida de mala calidad tal y como se comentó anteriormente).

En resumen, y una vez aclarados las ventajas e inconvenientes de este conjunto comercializado de marcadores, podemos afirmar que consideramos que dicho kit es de utilidad en los laboratorios de Genética Forense, tanto en aplicación del estudio biológico de la paternidad como en resolución de casos de criminalística, y que viene a colmar una carencia en el campo de los STR analizados por nitrato de plata, superando las ventajas que supone su uso a los inconvenientes puntuales que pudiera presentar.

6. CONCLUSIONES

PRIMERA: Los loci microsatélites estudiados con los kits Profler Plus (D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 y D7S820) y SilverSTR III (D13S317, D7S820 y D16S539), se encuentran, excepto para el locus D7S820, en equilibrio de Hardy-Weinberg en la población analizada.

El desequilibrio encontrado para el locus D7S820 se debe, probablemente, a un error de muestreo interno, puesto que la comparación de las frecuencias génicas obtenidas con las de otras poblaciones estudiadas no mostró diferencias significativas.

SEGUNDA: Al estudiar la correlación o equilibrio de la fase gamética de los diez loci entre sí, no se observaron valores de significancia relevantes, lo cual nos lleva a afirmar que los loci considerados son independientes entre sí en nuestra población.

TERCERA: De los diez loci estudiados el más informativo resultó ser el FGA, con un poder de discriminación del 96% y una probabilidad de exclusión del 73,25%. En conjunto, los nueve loci analizados en el kit Profiler Plus confieren un poder de discriminación de 0,99999999 y una probabilidad de exclusión de 0,99990927. En cuanto a los tres loci del kit SilverSTR III el poder de discriminación global conseguido es de 0.99942306 mientras que la probabilidad de exclusión es de 0.93122203.

CUARTA: El límite de sensibilidad, en condiciones estándar, observado para el triplex SilverSTR III en el estudio de validación realizado, se establece en torno a 0.4 ng, valor que es similar al obtenido para otros kits similares como el CTT.

QUINTA: A pesar de las dificultades que planteó el estudio de mezclas, podemos afirmar a modo orientativo que, el triplex SilverSTR III puede ser útil en el análisis de indicios forenses en los que coexisten dos o más muestras ya que permite detectar mezclas de proporciones 1:20 (5%) e incluso mezclas más desproporcionadas.

SEXTA: Los resultados conseguidos tras la amplificación de ADN procedente de diversas muestras de tipo forense para el estudio de validación, fueron variables, debiéndose dicha variabilidad a las características de cada muestra en particular. No obstante, obtuvimos resultados positivos en algunas muestras en las que el ADN estaba degradado o en muy baja concentración, por lo que, de manera general, aconsejamos el uso de dicho triplex para el análisis de muestras "difíciles" como algunas de las analizadas habitualmente en casos forenses.

SÉPTIMA: Habida cuenta de la importancia que en Génetica Forense supone la búsqueda de nuevos pollimorfismos de ADN con un alto poder de discriminación y cuyo análisis sea rápido y sencillo, podemos afirmar que los loci analizados en este trabajo de tesis cumplen todos los requisitos necesarios para su aplicación en identificación genética humana.

Las técnicas utilizadas para el análisis de los mismos (tanto manual como automatizada), son el reflejo de la adaptación de la Genética Forense al desarrollo y evolución de la Biología Molecular, desarrollo que está haciendo posible el estudio conjunto de múltiples loci tipo microsatélite de una manera sencilla y rápida, disminuyendo los errores de interpretación de los resultados y permitiendo alcanzar valores de probabilidad con una fiabilidad cada vez más cercana al 100%.

7. BIBLIOGRAFÍA

1. **Alberts, B**. "Biología Molecular de la célula" 2ª edición. Editorial Omega, Barcelona 1992.

2. **Allshire, R. C., M. Dempster, and N. D. Hastie.** 1989. Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Research. **17**(12):4611-27.

3. **Andreassen, R., and B. Olaisen.** 1997. Length and sequence variation in D7S22 (g3) alleles studied by high resolution length measurements and nucleotide sequencing. Electrophoresis. **18**(5):675-81.

4. **Andrés, A., Luna, J.D.** 1989. Bioestadística para las Ciencias de la Salud. 2ª Edición. Ediciones Norma. Madrid.

5. **Armour, J. A., P. C. Harris, and A. J. Jeffreys.** 1993. Allelic diversity at minisatellite MS205 (D16S309): evidence for polarized variability. Human Molecular Genetics. **2**(8):1137-45.

6. **Bär, W., B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, P. Lincoln, W. Mayr, and B. Olaisen.** 1997. DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics [editorial]. International Journal Of Legal Medicine. **110**(4):175-6.

7. **Bell, G., M. Selby, and W. Rutter.** 1982. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. Nature. **295**(5844):31-35.

8. **Bellas, S., Soto, J.L., Carracedo, A., Rodriguez-Calvo, M.S**. 1997. Minisatellite variant repeat coding using semiautomatic system. BioTechniques. **23**:44-46.

9. **Boerwinkle, E., W. J. Xiong, E. Fourest, and L. Chan.** 1989. Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. Proceedings Of The National Academy Of Sciences Of The United States Of America. **86**(1):212-6.

10. **Britten, R. J., and D. E. Kohne.** 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science. **161**(841):529-40.

11. **Budowle, B., R. Chakraborty, A. M. Giusti, A. J. Eisenberg, and R. C. Allen.** 1991. Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. American Journal Of Human Genetics. **48**(1):137-44.

12. **Budowle, B., J. Lindsey, J. DeCou, B. Koons, A. Giusti, and C. Comey.** 1995. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ alpha using a multiplex amplification and typing procedure. J Forensic Sci. **40**(1):45-54.

13. **Budowle, B., T. Moretti, and M. Keys.** 1997. Validation studies of the CTT STR Multiplex System. J Forensic Sci. **42**(2):701-707.

14. **Castro, A.** 1996. Evaluación de los polimorfismos de loci minisatélite y microsatélite en la identificación genética humana. Tesis Doctoral. Universidad del País Vasco.

15. **Capon, D. J., P. H. Seeburg, J. P. McGrath, J. S. Hayflick, U. Edman, A. D. Levinson, and D. V. Goeddel.** 1983. Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations. Nature. **304**(5926):507-13.

16. **Carracedo, A., and F. Barros.** 1996. Problemas bioestadísticos en Genética Forense. Ed. Servicio de Publicaciones e Intercambio Científico, Universidad de Santiago de Compostela.

17. **Chakraborty, R.** 1992. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. Human Biology. **64**(2):141-59.

18. **Chakraborty, R.** 1991. Statistical interpretation of DNA typing data [letter; comment]. American Journal Of Human Genetics. **49**(4):895-7, 899-903.

19. **Chakraborty, R., and K. K. Kidd.** 1991.. The utility of DNA typing in forensic work .Science. **254**(5039):1735-9.

20. **Chakraborty, R., and Y. Zhong.** 1994. Statistical power of an exact test of Hardy-Weinberg proportions of genotypic data at a multiallelic locus. Human Heredity. **44**(1):1-9.

21. **Chakraborty, R., Y. Zhong, L. Jin, and B. Budowle.** 1994. Nondetectability of restriction fragments and independence of DNA fragment sizes within and between loci in RFLP typing of DNA. American Journal Of Human Genetics. **55**(2):391-401.

22. **Cheung, V. G., M. Morley, F. Aguilar, A. Massimi, R. Kucherlapati, and G. Childs.** 1999. Making and reading microarrays. Nature Genetics. **21**(1 Suppl):15-9.

23. **Clayton, T., J. Whitaker, R. Sparkes, and P. Gill.** 1998. Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int. **91**(1):55-70.

24. **Comey, C., and B. Budowle.** 1991. Validation Studies on the Analysis of the HLA DQ alpha Locus Using the Polymerase Chain Reaction. J of Forensic Sciences. **36**(6):1633-1648.

25. **Coyne, V.E., James, M.D., Reid, S.J., Rybicki, E.P.** 1996 Molecular Biology Techniques Manual. Third Edition. Ed. Rybicki.

26. **Devlin, B., N. Risch, and K. Roeder.** 1990. No excess of homozygosity at loci used for DNA fingerprinting. Science. 249:1416-1420

27. **Devlin, B., N. Risch, and K. Roeder.** 1993. Statistical evaluation of DNA fingerprinting: a critique of the NRC's report [see comments]. Science. **259**(5096):748-9, 837.

28. **Dover, G.** 1982. Molecular drive: a cohesive mode of species evolution. Nature. **299**(5879):111-7.

29. **Duggan, D. J., M. Bittner, Y. Chen, P. Meltzer, and J. M. Trent.** 1999. Expression profiling using cDNA microarrays. Nature Genetics. **21**(1 Suppl):10-4.

30. **Edwards, A., A. Civitello, H. Hammond, and C. Caskey.** 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. **49**(4):746-56.

31. **Edwards, A., H. Hammond, L. Jin, C. Caskey, and R. Chakraborty.** 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics. **12**(2):241-53.

32. **Eng, B., P. Ainsworth, and J. Waye.** 1994. Anomalous migration of PCR products using nondenaturing polyacrylamide gel electrophoresis: the amelogenin sex-typing system. J Forensic Sci. **39**(6):1356-9.

33. **Entrala, C., J. A. Lorente, M. Lorente, J. C. Alvarez, B. Budowle, and E. Villanueva.** 1999. Spanish population data on the loci D13S317, D7S820, and D16S539 generated using silver staining (SilverSTR III Multiplex). Journal Of Forensic Sciences. **44**(5):1032-4.

34. **Entrala, C., M. Lorente, J. A. Lorente, J. C. Alvarez, T. Moretti, B. Budowle, and E. Villanueva.** 1998. Fluorescent multiplex analysis of nine STR loci: Spanish population data. Forensic Science International. **98**(3):179-83.

35. **Epstein, N., O. Nahor, and J. Silver.** 1990. The 3' ends of alu repeats are highly polymorphic. Nucleic Acids Research. **18**(15):4634.

36. **Evett, I.W., Weir B.S**.1998. Interpreting DNA Evidence: Statistical Genetics for forensic scientists. Sinaeur Associates. USA.

37. **Fernández, I**. 1998. Polimorfismos de ADN microsatélite y minisatélite en el País Vasco y su aplicación a la identificación ganética en delitos de agresión sexual. Tesis Doctoral. Universidad de Santiago de Compostela.

38. **Fu, Y., A. Pizzuti, and R. Fenwick.** 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science. **255**(5049):1256-1258.

39. **Fu, Y. H., D. P. Kuhl, A. Pizzuti, M. Pieretti, J. S. Sutcliffe, S. Richards, A. J. Verkerk, J. J. Holden, R. G. J. Fenwick, S. T. Warren, and e. al.** 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell. **67**(6):1047-58.

40. **García-Foncillas, J.** 1998. "Genetic Diagnosis in Medicine". Editores: Armengod, M.E., Echevarría, J.M., Blasband, A. Pamplona (España).

41. **Garofano, L., M. Pizzamiglio, C. Vecchio, G. Lago, T. Floris, G. D'Errico, G. Brembilla, A. Romano, and B. Budowle.** 1998. Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWFA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S1358. Forensic Science International. **97**(1):53-60.

42. **Gehrig, C., M. Hochmeister, U. V. Borer, R. Dirnhofer, and B. Budowle.** 1999. Swiss Caucasian population data for 13 STR loci using AmpFISTR profiler plus and cofiler PCR amplification kits. Journal Of Forensic Sciences. **44**(5):1035-8.

43. **Gene, M., E. Huguet, P. Moreno, C. Sanchez, A. Carracedo, and J. Corbella.** 1996. Population study of the STRs HUMTH01 (including a new variant) and HUMVWA31A in Catalonia (northeast Spain). Int J Legal Med. **108**(6):318-20.

44. **Gill, P., R. Sparkes, R. Pinchin, T. Clayton, J. Whitaker, and J. Buckleton.** 1998. Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int. **91**(1):41-53.

45. **Goellner, G. M., D. Tester, S. Thibodeau, E. Almqvist, Y. P. Goldberg, M. R. Hayden, and C. T. McMurray.** 1997. Different mechanisms underlie DNA instability in Huntington disease and colorectal cancer. American Journal Of Human Genetics. **60**(4):879-90.

46. **Guo, S. W., and E. A. Thompson.** 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics. **48**(2):361-72.

47. **Hacia, J. G.** 1999. Resequencing and mutational analysis using oligonucleotide microarrays. Nature Genetics. **21**(1 Suppl):42-7.

48. **Hantschel, M., R. Hausmann, T. Lederer, P. Martus, and P. Betz.** 1999. Population genetics of nine short tandem repeat (STR) loci-DNA typing using the AmpFISTR Profiler PCR amplification kit. International Journal of Legal Medicine. **112**:393-395.

49. **Hayward, G. S., and M. G. Smith.** 1972. The chromosome of bacteriophage T5. I. Analysis of the single-stranded DNA fragments by agarose gel electrophoresis. Journal Of Molecular Biology. **63**(3):383-95.

50. **Hochmeister, M., B. Budowle, J. Jung, U. Borer, C. Comey, and R. Dirnhofer.** 1991. PCR-based typing of DNA extracted from cigarette butts. Int J Legal Med. **104**(4):229-33.

51. **Hopkins, B., N. J. Williams, M. B. Webb, P. G. Debenham, and A. J. Jeffreys.** 1994. The use of minisatellite variant repeat-polymerase chain reaction (MVR-PCR) to determine the source of saliva on a used postage stamp. Journal Of Forensic Sciences. **39**(2):526-31.

52. **Horn, G. T., B. Richards, and K. W. Klinger.** 1989. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nucleic Acids Research. **17**(5):2140.

53. **Huberman, J. A.** 1995. Cell cycle. A licence to replicate [news; comment]. Nature. **375**(6530):360-1.

54. **Hudson, T. J., L. D. Stein, S. S. Gerety, J. Ma, A. B. Castle, J. Silva, D. K. Slonim, R. Baptista, L. Kruglyak, S. H. Xu, and e. al.** 1995. An STS-based map of the human genome [see comments]. Science. **270**(5244):1945-54.

55. **Huguet E., Carracedo, A., Gené M**. 1988. Introducción a la Investigación Biológica de la Paternidad. Promociones y Publicaciones Universitarias, S.A. Barcelona.

56. **Jarjoura, D., J. Jamison, and S. Androulakakis.** 1994. Likelihood ratios for deoxyribonucleic acid (DNA) typing in criminal cases. J Forensic Sci. **39**(1):64-73.

57. **Jarman, A. P., R. D. Nicholls, D. J. Weatherall, J. B. Clegg, and D. R. Higgs.** 1986. Molecular characterisation of a hypervariable region downstream of the human alpha-globin gene cluster. Embo Journal. **5**(8):1857-63.

58. **Jeffreys, A., V. Wilson, and S. Thein.** 1985a. Hypervariable 'minisatellite' regions in human DNA. Nature. **314**(6006):67-73.

59. **Jeffreys, A. J., V. Wilson, and S. L. Thein.** 1985b. Individual-specific 'fingerprints' of human DNA. Nature. **316**(6023):76-9.

60. **Jeffreys, A., A. MacLeod, K. Tamaki, D. Neil, and D. Monckton.** 1991. Minisatellite repeat coding as a digital approach to DNA typing [see comments]. Nature. **354**(6350):204-9.

61. **Jeffreys, A. J., D. G. Monckton, K. Tamaki, D. L. Neil, J. A. Armour, A. MacLeod, A. Collick, M. Allen, and M. Jobling.** 1993. Minisatellite variant repeat mapping: application to DNA typing and mutation analysis. Exs. **67**:125-39.

62. **Jung, J., C. Comey, D. Baer, and B. Budowle.** 1991. Extraction strategy for obtaining DNA from bloodstains for PCR amplification and typing of the HLA-DQ alpha gene. Int J Legal Med. **104**(3):145-8.

63. **Kao, F. T.** 1985. Human genome structure. International Review Of Cytology. **96**:51-88.

64. **Kasai, K., Y. Nakamura, and R. White.** 1990. Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. Journal Of Forensic Sciences. **35**(5):1196-200.

65. **Kimpton, C., A. Walton, and P. Gill.** 1992. A further tetranucleotide repeat polymorphism in the vWF gene. Human Molecular Genetics. **1**(4):287.

66. **Kimpton, C., A. Walton, and P. Gill.** 1992. A further tetranucleotide repeat polymorphism in the vWF gene. Hum Mol Genet. **1**(4):287.

67. **Kit, S.** 1961. Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. J Mol Biol. **3**:711-716.

68. **Krawczack, M. and J.Schmidtke**.1998. "DNA Fingerprinting". 2ªEdición. Ed. BIOS Scientific Publishers Ltd. UK.

69. **Kreike, J., and A. Lehner.** 1995. Sex determination and DNA competition in the analysis of forensic mixed stains by PCR. Int J Legal Med. **107**(5):235-8.

70. **Kupferschmid, T. D., T. Calicchio, and B. Budowle.** 1999. Maine Caucasian population DNA database using twelve short tandem repeat loci. Journal Of Forensic Sciences. **44**(2):392-5.

71. **Lahermo, P., M. L. Savontaus, P. Sistonen, J. Béres, P. de Knijff, P. Aula, and A. Sajantila.** 1999. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. European Journal Of Human Genetics. **7**(4):447-58.

72. **Lareu, M., M. Pestoni, F. Barros, A. Salas, and A. Carracedo.** 1996. Sequence variation of a hypervariable short tandem repeat at the D12S391 locus. Gene. **182**(1-2):151-3.

73. **Li, H., L. Schmidt, M. H. Wei, T. Hustad, M. I. Lerman, B. Zbar, and K. Tory.** 1993. Three tetranucleotide polymorphisms for loci: D3S1352; D3S1358; D3S1359. Human Molecular Genetics. **2**(8):1327.

74. **Lorente, M., J. A. Lorente, M. R. Wilson, B. Budowle, and E. Villanueva.** 1993. Composite PAGE: an alternate method for increased separation of amplified short tandem repeat alleles. International Journal Of Legal Medicine. **106**(2):69-73.

75. **Lorente, M.** 1994. Polimorfismo del ADN e identificación médico-legal: Estudio de siete loci (D1S80, D17S5, HUMTH01, HUMVWA, ACTBP2, D21S11 y HLA-DQ-alfa) mediante la reacción en cadena de la polimerasa (PCR). Tesis Doctoral. Universidad de Granada.

76. **Lorente, J.A., Lorente M.**(1995). El ADN y la identificación en la investigación criminal y en la paternidad biológica. Editorial Comares. Granada.

77. **Lorente, M., J. A. Lorente, M. R. Wilson, B. Budowle, and E. Villanueva.** 1997. Spanish population data on seven loci: D1S80, D17S5, HUMTH01, HUMVWA, ACTBP2, D21S11 and HLA-DQA1. Forensic Science International. **86**(3):163-71.

78. **Lorente, M., C. Entrala, J. A. Lorente, J. C. Alvarez, E. Villanueva, and B. Budowle.** 1998. Dandruff as a potential source of DNA in forensic casework. Journal Of Forensic Sciences. **43**(4):901-2.

79. **Lorente , M.**, Lorente, J.A., Villanueva, E. Identificación humana y Medicina Legal: consideraciones éticas y jurídicas. En "La prueba del ADN en Medicina Forense". (1999) Martínez Jarreta, M.B. Editorial Masson. Barcelona.

80. **Martinez Jarreta, B., P. Diaz Roche, and E. Abecia.** 1999. Genetic variation at six STR loci (HUMTH01, HUMTPOX, HUMCSF1PO, HUMF13A01, HUMFES/FPS,

HUMVWFA31) in Aragon (north Spain). Forensic Science International. **100**(1-2):87-92.

81. **Martínez Jarreta, B.** 1999. "La prueba del ADN en Medicina Forense". Ed. Masson. Barcelona (España).

82. **Mills, K. A., D. Even, and J. C. Murray.** 1992. Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). Human Molecular Genetics. **1**(9):779.

83. **Moller, A., E. Meyer, and B. Brinkmann.** 1994. Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. Int J Legal Med. **106**(6):319-23.

84. **Monckton, D. G., R. Neumann, T. Guram, N. Fretwell, K. Tamaki, A. MacLeod, and A. J. Jeffreys.** 1994. Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. Nature Genetics. **8**(2):162-70.

85. **Mora, J.M**. 1999. La identificación criminal a través de la tecnología del ADN. Tesis Doctoral. Universidad de Granada.

86. **Moreno Suárez, P.**, Gené Badía, M. 1999. Fundamentos de la variabilidad genética de las poblaciones humanas. En "La prueba del ADN en Medicina Forense". (1999) Martínez Jarreta, M.B. Editorial Masson. Barcelona.

87. **Morral, N., and X. Estivill.** 1992. Multiplex PCR amplification of three microsatellites within the CFTR gene. Genomics. **13**(4):1362-4.

88. **Mullis, K., and F. Faloona.** 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. **155:**335-50.

89. **Mullis, K. B.** 1990. Target amplification for DNA analysis by the polymerase chain reaction. Annales De Biologie Clinique. **48**(8)**:**579-82.

90. **Mully, J., A. Gedeon, S. White, E. Haan, and R. Richards.** 1991. Predictive diagnosis of myotonic dystrophy with flanking microsatellite markers. Journal of Medical Genetics. **28:**448-452.

91. **Neil, D. L., and A. J. Jeffreys.** 1993. Digital DNA typing at a second hypervariable locus by minisatellite variant repeat mapping. Human Molecular Genetics. **2**(8)**:**1129-35.

92. **Neuhuber, F., M. Klintschar, and M. Radacher.** 1998. A collaborative genetic study on the STR system FGA in two Austrian population samples. Forensic Sci Int. **91**(1)**:**1-6.

93. **Ohno, S.** 1984. Repeats of base oligomers as the primordial coding sequences of the primeval earth and their vestiges in modern genes. Journal Of Molecular Evolution. **20**(3-4)**:**313-21.

94. **Oldroyd, N. J., A. J. Urquhart, C. P. Kimpton, E. S. Millican, S. K. Watson, T. Downes, and P. D. Gill.** 1995. A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification. Electrophoresis. **16**(3)**:**334-7.

95. **Oste, C.** 1988. Polymerase chain reaction. Biotechniques. **6**(2):162-7.

96. **Oudet, C., R. Heilig, A. Hanauer, and J. L. Mandel.** 1991. Nonradioactive assay for new microsatellite polymorphisms at the 5' end of the dystrophin gene, and estimation of intragenic recombination. American Journal Of Human Genetics. **49**(2):311-9.

97. **Owerbach, D., and L. Aagaard.** 1984. Analysis of a 1963-bp polymorphic region flanking the human insulin gene. Gene. **32**(3):475-9.

98. **Pastinen, T., A. Kurg, A. Metspalu, L. Peltonen, and A. C. Syvänen.** 1997. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. Genome Research. **7**(6):606-14.

99. **Perkin Elmer Corporation**. 1996. GeneScan Analysis Software, P/N 904435.

100. **Perkin Elmer Corporation**. 1996. Genotyper. Applications Tutorials, P/N 904649.

101. **Perkin Elmer Corporation**. 1997. GeneScan Reference Guide, P/N 4303189 rev A.

102. **Perkin Elmer Corporation**. 1998. ABI PRISM^{TM}310 Genetic Analyzer User´s Manual, P/N 903565 rev B.

103. **Persidis, A.** 1998. Biochips. Nature Biotechnology. **16**(10):981-3.

104. **Plummer, S. J., M. J. Paris, J. Myles, R. Tubbs, J. Crowe, and G. Casey.** 1997. Four regions of allelic imbalance on 17q12-qter associated with high-grade breast tumors. Genes, Chromosomes And Cancer. **20**(4):354-62.

105. **Pouchkarev, V., E. Shved, and P. Novikov.** 1998. Sex determination of forensic samples by polymerase chain reaction of the amelogenin gene and analysis by capillary electrophoresis with polymer matrix. Electrophoresis. **19**(1):76-9.

106. **Promega Corporation.** 1998. GenePrint™ STR Systems Technical Manual, # TMD004.

107. **Proudfoot, N. J., A. Gil, and T. Maniatis.** 1982. The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell. **31**(3 Pt 2):553-63.

108. **Puers, C., H. A. Hammond, L. Jin, C. T. Caskey, and J. W. Schumm.** 1993. Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01[AATG]n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. American Journal Of Human Genetics. **53**(4):953-8.

109. **Rand, S., C. Puers, K. Skowasch, P. Wiegand, B. Budowle, and B. Brinkmann.** 1992. Population genetics and forensic efficiency data of 4 AMPFLP's. International Journal Of Legal Medicine. **104**(6):329-33.

110. **Richards, R. I., and G. R. Sutherland.** 1992. Heritable unstable DNA sequences [news]. Nature Genetics. **1**(1):7-9.

111. **Rodriguez-Calvo, M. S., S. Bellas, J. L. Soto, F. Barros, and A. Carracedo.** 1996. Comparison of different electrophoretic methods for digital typing of the MS32 (D1S8) locus. Electrophoresis. **17**(8):1294-8.

112. **Rodriguez-Calvo, M. S**. 1999. Análisis de la variación interna de los loci minisatélites. Un enfoque digital para el análisis del ADN. En "La prueba del ADN en Medicina Forense". 1999. Martínez Jarreta, M.B. Editorial Masson. Barcelona.

113. **Ruiz-Jiménez Aguilar, J**. 1993. Aspectos legales y prácticos del empleo de la información genética en criminalística y pruebas de paternidad. Utilización de material genético en criminalística y pruebas de paternidad: aspectos éticos y legales. (ISFH): 15-27.

114. **Ross, A., and H. Harding.** 1989. DNA typing and forensic science. Forensic Sci Int. **41**(3):197-203.

115. **Saitoh, H., S. Ueda, K. Kurosaki, and M. Kiuchi.** 1998. The different mobility of complementary strands depends on the proportion AC/GT. Forensic Sci Int. **91**(2):81-90.

116. **Sambrook, J., E. Fritsch, and T. Maniatics.** 1989. Molecular Cloning. A Laboratory Manual., Second ed, vol. I, II, III. Cold Spring Harbor Laboratory Press, USA.

117. **Schildkraut, C., J. Marmur, and P. Doty.** 1962. Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl. J Mol Biol. **4**:430-443.

118. **Schmutzler, R. K., A. Homann, E. Bierhoff, O. D. Wiestler, A. von Daimling, and D. Krebs.** 1995. [Detection of genetic alterations in sporadic breast tumors]. Gynakologisch-Geburtshilfliche Rundschau. **35 Suppl 1**:63-7.

119. **Sen, D., and W. Gilbert.** 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature. **334**(6180):364-6.

120. **Sharma, V., and M. Litt.** 1992. Tetranucleotide repeat polymorphism at the D21S11 locus. Human Molecular Genetics. **1**(1):67.

121. **Silva, F., L. Gusmão, C. Alves, R. Seruca, L. David, and A. Amorim.** 1997. Tetra- and pentanucleotide short tandem repeat instability in gastric cancer. Electrophoresis. **18**(9):1633-6.

122. **Singer, M.** 1982a. Highly repeated sequences en mammalian genomes. Int Rev Cytol. **76**:67-112.

123. **Singer, M.** 1982b. SINE and LINE: highly repeated short and long interspersed sequences in mammalian genomes. Cell. **28**(3):433-434.

124. **Smith, G. P.** 1976. Evolution of repeated DNA sequences by unequal crossover. Science. **191**(4227):528-35.

125. **Smith, J., R. Anwar, J. Riley, D. Jenner, A. Markham, and A. Jeffreys.** 1990. Highly polymorphic minisatellite sequences: allele frequencies and mutation rates for five locus-specific probes in a Caucasian population. J Forensic Sci Soc. **30**(1):19-32.

126. **Southern, E. M.** 1979. Measurement of DNA length by gel electrophoresis. Analytical Biochemistry. **100**(2):319-23.

127. **Strachan, T.**, Andrew, P.R. 1999. Genética Molecular Humana. Ediciones Omega. Barcelona.

128. **Sullivan, K. M., A. Mannucci, C. P. Kimpton, and P. Gill.** 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques. **15**(4):636-8, 640-1.

129. **Tamaki, K., D. G. Monckton, A. MacLeod, D. L. Neil, M. Allen, and A. J. Jeffreys.** 1992. Minisatellite variant repeat (MVR) mapping: analysis of 'null' repeat units at D1S8 [published erratum appears in Hum Mol Genet 1992 Oct;1(7):558]. Human Molecular Genetics. **1**(6):401-6.

130. **Tamaki, K., X. L. Huang, T. Yamamoto, R. Uchihi, H. Nozawa, and Y. Katsumata.** 1995. Applications of minisatellite variant repeat (MVR) mapping for maternal identification from remains of an infant and placenta. Journal Of Forensic Sciences. **40**(4):695-700.

131. **Urquhart, A., C. Kimpton, T. Downes, and P. Gill.** 1994. Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers. Int J Legal Med. **107**(1):13-20.

132. **Urquhart, A., N. J. Oldroyd, C. P. Kimpton, and P. Gill.** 1995. Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. Biotechniques. **18**(1):116-8, 120-1.

133. **Valverde, E.** 1992. Análisis del polimorfismo en los loci minisatélite D2S44, D7S21 y D12S11: Aplicaciones Médico-Legales. Tesis Doctoral. Universidad de Santiago de Compostela.

134. **Vogt, P.** 1990. Potential genetic functions of tandem repeated DNA sequence blocks in the human genome are based on a highly conserved chromatin folding code. Human Genetics. **84**(4):301-36.

135. **Vosberg, H. P.** 1989. The polymerase chain reaction: an improved method for the analysis of nucleic acids. Human Genetics. **83**(1):1-15.

136. **Walker, G. J., J. M. Palmer, M. K. Walters, D. J. Nancarrow, P. G. Parsons, and N. K. Hayward.** 1994. Simple tandem repeat allelic deletions confirm the preferential loss of distal chromosome 6q in melanoma. International Journal Of Cancer. **58**(2):203-6.

137. **Walsh, P., D. Metzger, and R. Higuchi.** 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. **10**(4):506-13.

138. **Waring, M., and R. J. Britten.** 1966. Nucleotide sequence repetition: a rapidly reassociating fraction of mouse DNA. Science. **154**(750):791-4.

139. **Weber, J. L.** 1990. Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics. **7**(4):524-30.

140. **Weber, B., O. Riess, G. Wolff, S. Andrew, C. Collins, R. Graham, J. Theilmann, and M. R. Hayden.** 1992. Delineation of a 50 kilobase DNA segment containing the recombination site in a sporadic case of Huntington's disease. Nature Genetics. **2**(3):216-22.

141. **Weir, B.** 1992. Population genetics in the forensic DNA debate. Proc Natl Acad Sci U S A. **89**(24):11654-9.

142. **Willard, H. F., and J. S. Waye.** 1987. Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. Journal Of Molecular Evolution. **25**(3):207-14.

143. **Wong, Z., V. Wilson, I. Patel, S. Povey, and A. J. Jeffreys.** 1987. Characterization of a panel of highly variable minisatellites cloned from human DNA. Annals Of Human Genetics. **51 (Pt 4)**:269-88.

144. **Wooster, R., A. Cleton-Jansen, N. Collins, J. Mangion, R. Cornelis, C. Cooper, B. Gusterson, B. Ponder, A. von Deimling, O. Wiestler, and a. et.** 1994. Instability of short tandem repeats (microsatellites) in human cancers. Nat Genet. **6**(2):152-6.

145. **Wu, S., S. Seino, and G. I. Bell.** 1990. Human collagen, type II, alpha 1, (COL2A1) gene: VNTR polymorphism detected by gene amplification. Nucleic Acids Research. **18**(10):3102.

146. **Wyman, A. R., and R. White.** 1980. A highly polymorphic locus in human DNA. Proceedings Of The National Academy Of Sciences Of The United States Of America. **77**(11):6754-8.

147. **Yamamoto, T., K. Tamaki, T. Kojima, R. Uchihi, and Y. Katsumata.** 1994. Potential forensic applications of minisatellite variant repeat (MVR) mapping using the polymerase chain reaction (PCR) at D1S8. Journal Of Forensic Sciences. **39**(3):743-50.

8. ANEXO

ELSEVIER

Forensic Science International
98 (1998) 179–183

Forensic Science International

Fluorescent multiplex analysis of nine STR loci: Spanish population data

Carmen Entrala[a], Miguel Lorente[a], Jose A. Lorente[a,*], J. Carlos Alvarez[a], Tamyra Moretti[b], Bruce Budowle[b], Enrique Villanueva[a]

[a]*Laboratory of Genetic Identification, Dept. de Medicina Legal, Facultad de Medicina, Universidad de Granada, E-18012 Granada, Spain*

[b]*Forensic Science Research Unit, FSRTC, Laboratory Division, FBI Academy, Quantico, VA 22135, USA*

Received 30 July 1998; received in revised form 14 September 1998; accepted 28 September 1998

Abstract

A total of 171 Caucasians living in Andalucia (southern Spain) have been typed for nine short tandem repeat (STR) loci by multiplex PCR amplification using a commercially available kit (Profiler Plus; Perkin–Elmer, Norwalk, CT, USA) and semi-automatic electrophoresis (ABI Prism 377 DNA Sequencer, Applied Biosystems, Foster City, CA, USA). The kit enables typing of the STR loci D3S1358, VWA, FGA, D5S818, D13S317, D7S820, D8S1179, D21S11, and D18S51. All loci, except D7S820, meet Hardy–Weinberg equilibrium. Because of the large number of loci that can be analyzed, the power of discrimination (PD) is greater than 0.99999, and the probability of exclusion (PE) reaches 0.99991 in our population sample. © 1998 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Spanish population; STR; PCR; Forensic genetics; Fluorescent

1. Introduction

Short tandem repeat (STR) loci are routinely used for paternity and forensic identity testing [1,2]. Furthermore, STR loci are useful components of national DNA databanks, such as those in the United Kingdom [3] and the United States [4]. The typing of STR

*Corresponding author. Address for correspondence: Dept. Medicina Legal - University of Granada, Av. Madrid 11, E-18012 Granada, Spain. Tel.: +34-958-243546; fax: +34-958-246107; e-mail: jlorente@goliat.ugr.es

0379-0738/98/$ – see front matter © 1998 Elsevier Science Ireland Ltd. All rights reserved.
PII: S0379-0738(98)00165-0

loci is facilitated by the ability to amplify several loci simultaneously in a multiplex polymerase chain reaction (PCR). Several commercial kits have been developed for STR multiplex amplification. One kit - Profiler Plus™ (Perkin–Elmer) - is designed for analysis of the amelogenin locus and nine STR loci: D3S1358, VWA, FGA, D5S818, D13S317, D7S820, D8S1179, D21S11, and D18S51. Based on fluorescent detection technology, this multiplex analysis has a high power of discrimination. We have used the Perkin–Elmer Profiler Plus™ kit to type a sample population from Andalucia, Spain.

2. Material and methods

Whole blood was obtained in EDTA vacutainer tubes by venipuncture from 171 unrelated Spanish Caucasians from Andalucia (southern Spain). DNA was extracted organic procedure then concentrated and purified through Microcon-100 filtration cartridges [5]. The quantity of recovered DNA was determined by slot-blot hybridization [6].

The starting PCR reaction on which the Profiler Plus™ kit required 1 to 2 ng of template DNA. Amplification was performed in a Perkin–Elmer 9600. Since Taq Gold™ was used as the DNA polymerase, samples were pre-heated in the thermocycler at 95°C for 11 min. Electrophoresis was carried out in an ABI Prism 377 DNA sequencer; all procedures were carried out in accordance to manufacturer's instructions.

Statistical analysis of the results was performed using the following tests. The relative frequency of each allelic class was obtained from the allele counts. Estimates of expected heterozygosity were computed as described by Edwards et al. [1]. Compliance with the Hardy–Weinberg equilibrium (HWE) was determined by calculating the estimate of the expected homozygote/heterozygote frequencies [7–9], the likelihood ratio test [10,11], and the exact test by Guo and Thompson [12].

3. Results and discussion

The allelic frequencies observed for each of the 9 STR loci are shown in Table 1. A majority of the samples ($n=120$) at some of the loci (i.e., D7S820, D13S317, and VWA) in this Spanish data set were typed previously by our group [13,14]; there were no discrepancies between the two typing results. These data support the reliability of STR typing with the Profiler Plus™ kit in that identical results were obtained even though different primer sets, different PCR conditions, and different detection methods were used.

Statisitical inferences for each locus are shown in Table 2. The most informative loci turned out to be FGA and D18S51, the least informative locus was D5S818. The overall power of discrimination (PD) was 0.999997, the probability of exclusion (PE) reached 0.99991.

All loci, except D7S820, meet Hardy–Weinberg equilibrium (HWE). It was shown previously that the departure from HWE for the D7S820 locus in this sample population may be due to sampling error, and the D7S820 allele frequncies still would be useful for

Table 1
Allele frequencies for the nine loci included in the Profiler Plus™ (Perkin–Elmer) in a sample of Spanish Caucasians from Andalucia ($n=171$)

Allele	D13S358	VWA	FGA	D8S1179	D21S11	D18S51	D5S818	D3S317	D7S820
7									0.0263
8				0.0017				0.1813	0.1520
9				0.0015			0.0234	0.0643	0.1345
10				0.1111		0.0058	0.0819	0.0614	0.3070
11				0.0848		0.0205	0.3450	0.2778	0.1959
12		0.0029		0.1199		0.1433	0.3450	0.2602	0.1608
13	0.0029			0.2544		0.0936	0.1813	0.1053	0.0234
14	0.0099	0.1199		0.2281		0.1842	0.0117	0.0497	
15	0.2164	0.1257		0.1462		0.1579	0.0117		
16	0.2895	0.2368		0.0205		0.1140			
17	0.1901	0.2456		0.0029		0.1403			
18	0.1696	0.1754	0.0087			0.0789			
19	0.0292	0.0760	0.0789			0.0322			
20		0.0175	0.1579			0.0234			
21			0.1579						
22			0.1579			0.0058			
22.2			0.0088						
23			0.1345						
23.2			0.0029						
23.3			0.0029						
24			0.1520						
25			0.0906						
26			0.0322		0.0029				
27			0.0170		0.0175				
28			0.0029		0.1403				
28.2					0.0029				
29					0.1959				
30					0.2719				
30.2					0.0351				
31					0.0585				
31.2					0.1374				
32					0.0088				
32.2					0.0819				
33.2					0.0292				
34.2					0.0088				
35					0.0029				
35.2					0.0058				

estimating DNA profile frequencies [14]. In addition, our southern Spain sample population has now been typed for a total of 15 loci [13–15]. The observation of only one locus departing from HWE out of a total of 15 is no more than would be expected by chance.

In conclusion, a Spanish population database has been established for the loci D3S1358, VWA, FGA, D5S818, D13S317, D7S820, D8S1179, D21S11, and D18S51. The allelic frequencies of these PCR-based loci can be used to estimate the frequency of a multiple locus DNA profile in Spanish population. The data demonstrate that a high

Table 2
Statistical values of interest for the Hardy–Weinberg equilibrium

	D13S358	VWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820
Observed homozygosity	18.1%	17.5%	9.9%	19.3%	15.2%	11.7%	30.4%	20.5%	25.7%
Expected Homozygosity	20.4%	18.1%	12.9%	17.1%	16.1%	12.7%	27.6%	19.7%	19.9%
Homozygosity test (*P*)	0.463	0.848	0.245	0.435	0.745	0.705	0.417	0.797	0.054
Likelihood Ratio test (*P*)	0.955	0.110	0.052	0.278	0.531	0.505	0.268	0.247	0.006[a]
Exact test (Guo–Thompson) (*P*)	0.955	0.133	0.071	0.130	0.487	0.425	0.331	0.151	0.003[a]
Probability of Discrimination (PD)	0.923	0.937	0.960	0.942	0.949	0.964	0.875	0.923	0.924
Probability of Exclusion (PE)	0.593	0.635	0.733	0.658	0.679	0.739	0.480	0.611	0.606

[a]=significant, probability of discrimination (PD) and probability of exclussion (PE); for all the nine loci together, PD=0.999997 and PE=0.99991.

degree of discrimination can be obtained when all nine (or fewer) loci are used to characterize forensic biological evidence.

Acknowledgements

This work was supported by a Proyecto de Investigación (PM97-0175) from the DGES, Spanish Ministry of Education and Science (MEC).

References

[1] A. Edwards, H. Hammond, L. Jin, C.T. Caskey, R. Chakraborty, Genetic variation at five trimeric and tetrameric repeat loci in four human population groups, Genomics 12 (1992) 244–253.
[2] C.J. Frégeau, R.M. Fourney, DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification, BioTechniques 15 (1993) 100–119.
[3] A. Hopwood, J. Brokes, A. Shariff, P. Cage, E. Tatum, R. Mirza, M. Crook, K. Brews, K. Sullivan, A fully integrated robotic system for high sample throughout within a DNA databasing unit, in: Proceedings of the VII Symposium on Human Identification. Promega Corp., Scottsdale, AZ, 1997, pp. 20–24.
[4] T.R. Moretti, B. Budowle, The CODIS STR project: evaluation of fluorescent multiplex STR systems, in: Proceedings of the 50th Meeting of the American Academy Forensic Sciences, San Francisco, CA, 1998, pp. 48–49.
[5] C.T. Comey, B.W. Koons, K.W. Presley, J.B. Smerick, J.B. Smerick, C.A. Sobieralski, D.M. Stanley, F.S. Baechtel, DNA extraction strategies for amplified fragment length polymorphism analysis, J. Forensic Sci. 39 (1994) 1254–1269.
[6] B. Budowle, F.S. Baechtel, C.T. Comey, A.M. Giusti, L. Klevan, Simple protocols for typing forensic biological evidence: chemiluminescent detection for human DNA quantitation and RFLP analysis and manual typing of PCR amplified polymorphisms, Electrophoresis 16 (1995) 1559–1567.
[7] R. Chakraborrty, P.E. Smouse, J.V. Neel, Popualtion amalgamation and genetic variation: observations on artificially agglomerated tribal populations of Central and South America, Am. J. Human Genet. 43 (1988) 709–725.
[8] M. Nei, A.K. Roychoudhvry, Sampling variances of heterozygosity and genetic distance, Genetics 76 (1978) 583–590.
[9] M. Nei, Estimation of average heterozygosity and genetic distance from a small number of individuals, Genetics 89 (1978) 583–590.

CASE REPORT

Miguel Lorente,[1] M.D.; Carmen Entrala,[1] M.S.; Jose A. Lorente,[1] M.D.; J. Carlos Alvarez,[1] M.S.; Enrique Villanueva,[1] M.D.; and Bruce Budowle,[2] Ph.D.

Dandruff as a Potential Source of DNA in Forensic Casework*

REFERENCE: Lorente M, Entrala C, Lorente JA, Alvarez JC, Villanueva E, Budowle, B. Dandruff as a potential source of DNA in forensic casework. J Forensic Sci 1998;43(4):901–902.

ABSTRACT: Dandruff is a clinical alteration of the skin that consists histologically of orthokeratotic clumps with minute parakeratotic foci found in inflammatory pathologies such as seborrheic dermatitis and psoriasis. Therefore, some nucleated cells should be found in dandruff and hence there is a possibility that forensically typeable DNA could be extracted from dandruff.

Because of a particular case in which we were involved, a study was carried out to determine whether or not DNA could be extracted from dandruff, and if the two most widely used extraction techniques (Chelex and organic) would be applicable.

Results show that sufficient quantities of DNA (more than 30 to 40 ng) can be obtained from as little as 1.0 to 1.5 mg of dandruff. Both methods yield DNA, although the organic procedure seems to yield more (72.5 ng Chelex vs. 183.3 ng organic). All the DNA samples extracted were typed correctly for the loci HUMTH01 and HUMvWA.

Therefore, dandruff can be considered a potential source of DNA for forensic identification.

KEYWORDS: forensic science, dandruff, DNA typing, polymerase chain reaction, HUMTH01, HUMvWA

Dandruff is produced as a pathological desquamation of the skin, and histologically consists of orthokeratotic clumps with minute parakeratotic foci. Dandruff is often found in inflammatory pathologies such as seborrheic dermatitis and psoriasis (1).

Based on the etiology, where inflammatory pathologies are involved, it is presumed that some nucleated cells could be found in dandruff; therefore, in some cases, where sufficient quantities of dandruff could be obtained, DNA for the polymerase chain reaction (PCR) might be obtained for DNA analysis.

Because of a missing body case in Granada, Spain, a preliminary set of studies was performed to determine: (i) the potential to obtain typeable DNA from dandruff; and (ii) whether the Chelex-100 and/or organic phenol:chloroform extraction method is useful to recover forensically typeable genetic material from dandruff.

[1] Department of Legal Medicine, University of Granada, Granada, Spain.

[2] Forensic Science Research Unit, FBI Academy, Quantico, VA.

* Presented in part at the 49th Annual Meeting, American Academy of Forensic Sciences, New York, NY, February 1997.

Received 12 August 1997; and in revised form 26 Nov. 1997; accepted 8 Dec. 1997.

Materials and Methods

Dandruff sampling—Dandruff samples, ranging between 1.0 and 2.3 mg of total weight, were obtained from 14 different donors. All the samples were directly recovered from the head of the donor by scratching with the hand protected with sterile gloves. Dandruff flakes were collected on sterile plain white paper and then transferred inside sterilized paper envelopes, where they were stored in a refrigerator (+4°C) until used.

DNA extraction—Since there were no known references for extracting DNA from dandruff, the two most widely used extraction procedures for blood were employed. These are the Chelex-100 and the organic phenol:chloroform methods.

Chelex extraction—The classical procedure (2) with slight modifications was used. Basically, 5% Chelex-100® was directly added to 1.0 to 1.5 mg of dandruff without previous washing/incubations steps. The protocol is as follows: using a sterile brush, 0.5 to 1 mg of dandruff was placed in to a 1.5 mL microcentrifuge tube. To the tube, 50 to 100 μL of 5% Chelex-100® were added directly. The tube was incubated at 56°C for 30 min. The tube was vigorously vortexed for 10 s, then incubated in a boiling water bath for 8 min. The tube was then vortexed vigorously for 10 s and placed in a microcentrifuge and subjected to maximum speed centrifugation (13,000 *g*) for 3 min. The quantity of extracted DNA was estimated using the slot-blot hybridization assay (3).

Phenol-chloroform extraction—The organic extraction protocol previously described (4) was used. Minor variations from this protocol are the following. Using a sterile brush, 1.0 to 1.5 mg of dandruff was added to a 1.5 mL microcentrifuge tube. To the sample were added 200 μL of stain-extraction buffer (10 mM Tris-Hcl, pH 8.9; 10 mM EDTA; 0.1 M Na Cl; 2% SDS), 8 μL 1 M DTT, and 5 μL of 10 mg/mL Proteinase K. The tube was incubated at 56°C overnight (12 to 14 h). Then, an equal volume (215 μL) of phenol/chloroform/isoamyl-alcohol 25:24:1) was added to the extract. The tube was vortexed vigorously for 10 s and then subjected to centrifugation (2500 *g*) for 10 min. The aqueous phase was transferred to a Microcon®-100 concentrator. The DNA was recovered in 100 μL of TE^{-4} (10 mM Tris•Cl, pH 8.0, 0.1 mM EDTA). The quantity of extracted DNA was estimated using the slot-blot hybridization assay (3).

DNA typing—Approximately 3 ng of DNA were successfully

amplified for the loci HUMTH01 and HUMvWA according to previously reported methods (5,6).

Results and Discussion

Apparently DNA extraction methods designed for blood or epithelial cells also are applicable to extracting DNA from dandruff. Both the Chelex-100 and organic methods proved to be useful for extracting a sufficient quality and quantity of DNA. Amounts of dandruff ranging between 1.0 to 1.5 mg of dandruff yielded 183.3 $\pm$28.9 ng (n = 7) with the organic protocol, and 72.5 $\pm$37.1 ng with Chelex (n = 6). All genotypes from dandruff perfectly matched the known controls obtained from the saliva of donors (data not shown).

It is known that the frequency of seborrheic dermatitis ranges between 3% and 5% (7). The adult form is more common in men than in women, and there does not appear to be any racial predilection (7).

In a particular case, it would be very difficult to estimate accurately the amount of DNA (ng) that could be extracted from a given amount of dandruff (mg) because of: (a) the limited amount of this biological material that might be found at a crime scene or serve as an exemplar. We found that the amount of dandruff that could be obtained from our sample of 14 individuals ranged between 0.8 and 6.2 mg, and these amounts are difficult to collect and handle. (b) The drastic influence of inflammatory processes (psoriasis, spongiotic psoriasiform dermatitis, etc.), which could facilitate the release of nucleated cells (basically neutrophils and/or lymphocytes) that adhered to the non-nucleated desquamation tissues. Thus, it has been reported that a large number of parakeratotic cells associated with some infiltrated polymorphonuclear leukocytes (PMNLs) were found in the scales obtained from 11 individuals complaining of dandruff (8), making it possible for these authors to demonstrate the presence of neutrophil chemotactic anaphylotoxins in human dandruff.

Regardless, at least 30 to 40 ng of DNA were recovered from dandruff with either extraction procedure. The data support that dandruff can be a reliable source of biological material in some casework scenarios, and this material could be collected and stored along with hairs, nails, and other biological specimens.

It has been suggested that accidental contamination from foreign sources could be problematic in forensic science (9), and dandruff could be a potential contaminant in some cases. Therefore, depending on the case situation, we would recommend wearing haircaps.

In our particular case, DNA typing from dandruff obtained from clothes in the house of a missing person was compared with the DNA obtained from a set of bones. The results were consistent with the presumed identity. Further, mock case analyses performed in our laboratory gave always positive and conclusive results using dandruff as the only source of DNA.

Further studies are underway to determine the influence of medication and cosmetic products on the recovery of a sufficient quality and quantity of DNA from dandruff and to determine if the relationship between the amount of dandruff recovered and the amount of DNA expected can be roughly predicted.

Acknowledgments

This work was supported by a Proyecto de Investigación (PB93-1155) from the DGICYT-DGES, Spanish Ministry of Education and Science (MEC). We thank Drs. A. Buendia, S. Serrano, and J. J. López-Caballero for providing us with samples and scientific advice.

References

1. Cropley TG. Seborrheic dermatitis. In: Arndt P, Le Boit A, Robinson SA, Wintroub W, editors. Cutaneous medicine and surgery. New York: Saunders, 1996;214–7.
2. Walsh PS, Metzger DA, Higuchi R. Chelex® 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. BioTechniques 1991;4:506–13.
3. Waye JS, Presley LA, Budowle B, Shutler GS, Fourney RM. A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. BioTechniques 1989;7:852–5.
4. Sweet DJ, Lorente M, Valenzuela A, Lorente JA, Alvarez JC. Increasing DNA extraction yield from saliva stains with a modified Chelex method. Forensic Sci Int 1996;83:166–77.
5. Lorente JA, Lorente M, Wilson MR, Budowle B, Villanueva E. Analysis of the HUMTH01 allele frequencies in the Spanish population. J Forensic Sci 1994;39:1270–4.
6. Lorente JA, Lorente M, Wilson MR, Budowle B, Villanueva E. Analysis of the short tandem repeat (S.T.R.) VWA (HUMvWA) in the Spanish population. Forensic Sci Int 1994;65:169–75.
7. Johnson M, Roberts J. Prevalence of dermatological diseases among persons 1–74 years of age. Publication No. 79. Washington, DC: Department of Health, Education and Welfare, 1977.
8. Kikuchi T, Horii I, Sakamoto T, Nakayama Y, Tagami H. Demonstration of neutrophil chemotactic anaphylotoxins in human dandruff. Arch Dermatol Res 1989;218:482–6.
9. Van Oorschot RAH, Jones MK. DNA fingerprints from fingerprints. Nature 1997;387:767.

Additional information and reprint requests:
Prof. Dr. Jose A. Lorente
Department of Legal Medicine
University of Granada
Av. Madrid 11
18012-Granada, Spain
E-mail: jlorente@goliat.ugr.es

TECHNICAL NOTE

Carmen Entrala,[1] M.Sc.; Jose A. Lorente,[1] M.D., Ph.D.; Miguel Lorente,[1] M.D., Ph.D.; J. Carlos Alvarez,[1] M.Sc.; Bruce Budowle,[2] Ph.D.; and Enrique Villanueva,[1] M.D., Ph.D.

Spanish Population Data on the Loci D13S317, D7S820, and D16S539 Generated Using Silver Staining (SilverSTR III™ Multiplex)*

REFERENCE: Entrala C, Lorente JA, Lorente M, Alvarez JC, Budowle B, Villanueva E. Spanish population data on the Loci D13S317, D7S820, and D16S539 generated using silver staining (SilverSTR III™ Multiplex) J Forensic Sci 1999;44(5):1032–1034.

ABSTRACT: A set of 212 samples from unrelated Spanish Caucasians living in Andalucia (southern Spain) were analyzed with a new commercially-available kit for multiplex amplification of 3 STR loci (D13S137, D7S820, and D16S539), manual denaturing polyacrylamide gel electrophoresis and silver staining. These three loci are of special interest for the forensic community since they are a part of the 13 CODIS-core STR loci. The results show that the loci D13S317 and D16S539 meet Hardy-Weinberg expectations (HWE), but the locus D7S820 did not meet HWE ($p = 0.003$). However, there was no detectable departures from independence (i.e., linkage disequilibrium) between any pair-wise combination of loci. The D7S820 data were further investigated. The excess homozygosity was due to an excess of D7S820 10, 10 homozygotes. To determine if the allele frequency data are meaningful and can be applied to forensic identity cases, the Spanish D7S820 allele frequency data were compared with four other Caucasian sample populations. The D7S820 allele frequencies were statistically similar; thus, the results support that the allele frequency data can be used reliably for estimating DNA profile frequencies.

KEYWORDS: forensic science, D13S319, D7S820, D16S539, silver staining, DNA typing, short tandem repeats, population genetics, Spain

Typing short tandem repeats (STR) loci is useful for paternity and forensic identity testing. Furthermore, STR loci have been chosen as core markers for CODIS (1). Thus, STR loci typing figure prominently in forensic identity testing for the forseeable future. A new commercially-available kit enables the simultaneous amplification of three of the CODIS core loci: D7S820, D13S317, and D16S539 (DDD) (SilverSTR III™ Multiplex, Promega Corp., Madison, WI). After electrophoretic separation of the STR amplicons, the DNA profiles can be detected by silver staining. Thus, simple, inexpensive procedures can be used to type the DDD triplex. It is well-established in the literature (1–3) that amplification of STR loci by the polymerase chain reaction (PCR), electrophoresis of the amplified products, and detection by silver staining offer useful and reliable approaches for characterizing DNA derived from forensic biological specimens. This study describes results of a population study of unrelated individuals from Andalucia (southern Spain) using the DDD kit.

Material and Methods

Whole blood was obtained in EDTA vacutainer tubes by venipuncture from 212 unrelated Spanish Caucasians from Andalucia (southern Spain). DNA was extracted organically and concentrated and purified by Microcon-100 filtration (4). The quantity of recovered DNA was determined by slot-blot hybridization (5).

The SilverSTR III™ Multiplex kit (Promega Corp., Madison, WI) was used to amplify the loci D7S820, D13S317, and D16S539. The PCR contained 3 to 5 ng of DNA, 2.5 μL of STR 10X buffer, 2.5 μL Multiplex 10× primer pair mix, and 0.75 units of Taq polymerase. Sterile water was used to adjust to a final volume of 25 μL. Amplification was performed in a Perkin-Elmer 9600 for 30 cycles according to the manufacturer's recommendations (Technical Manual. GenePrint™ STR Systems—Silver Stain Detection. Part #TMD004. Promega Corp.).

Electrophoretic separation of the amplified alleles was performed in vertical, denaturing, polyacrylamide gels (4%) using a BRL SA32 apparatus (Life Technologies, Gaithersburg, MD) (2). The gels were pre-run at 40 W for 30 to 45 min in order to reach an approximate temperature of 56°C. Silver staining was performed as described previously (6).

The alleles of the samples were identified by comparison with the allelic ladders in adjacent lanes in the gel and confirmed by a second independent reader. The ladders contain the alleles 7 to 15 for locus D13S317; 6 to 14 for locus D7S820; and 5, 8 to 14 for locus D16S539.

Statistical analysis of the results was performed using the following tests. The frequency of each allele for each locus was calculated from the numbers of each genotype in the sample test. Unbiased estimates of expected heterozygosity were computed as described by Edwards et al. (7). Possible divergence from Hardy-

[1] Laboratory of Genetic Identification. Dept. de Medicina Legal. Facultad de Medicina. Universidad de Granada. E-18012 Granada, Spain.

[2] Forensic Science Research Unit. FSRTC. Laboratory Division. FBI Academy, Quantico, VA.

* This work was supported by a Proyecto de Investigación (PM97-0175) from the DGES, Spanish Ministry of Education and Science (MEC).

Received 28 July 1998; and in revised from 17 Nov. 1998; accepted 5 Jan. 1999.

Weinberg expectations (HWE) was determined by calculating the unbiased estimate of the expected homozygote/heterozygote frequencies (8–10), the likelihood ratio test (7,11,12), and the exact test by Guo and Thompson (13). An interclass correlation criterion (14) was used for detecting disequilibrium between loci.

Results and Discussion

The results using the SilverSTR™ III kit are easy to obtain and interpret (Fig. 1). For those laboratories familiar with manual gel electrophoresis and silver staining technology, transfer should be relatively easy, and for those laboratories without the more expensive fluorescent detection equipment, the typing of the loci D7S820, D13S317, and D16S539 can still be implemented.

The distribution of observed allelic frequencies, observed and expected homozygosity, tests for independence, probability of discrimination (PD), and probability of exclusion (PE) for the DDD loci are shown in Table 1. The loci D13S317 and D16S539 meet Hardy-Weinberg expectations (HWE), but the locus D7S820 does not meet HWE ($p = 0.003$). The excess homozygosity at the D7S820 locus was due primarily to an excess of observed 10, 10 homozygotes (29 observed versus 21 expected).

This excess of homozygosity could be due to possible population substructure at the locus, selection, typing errors, and/or sampling error. There is no evidence to date to support selection. Typing errors are unlikely because the same samples were typed using the Profiler Plus kit (Perkin-Elmer) and the typing results were the same for the two kits (data not shown). To evaluate if the departure from HWE is due predominately to either substructure or sampling error the Spanish D7S820 allele data were compared with data from three U.S. Caucasian sample populations and one from Italy (all of which meet HWE at the D7S820 locus) using a test for homogeneity, and they were statistically similar (data not shown); thus sampling error at the genotype level appears to be the main

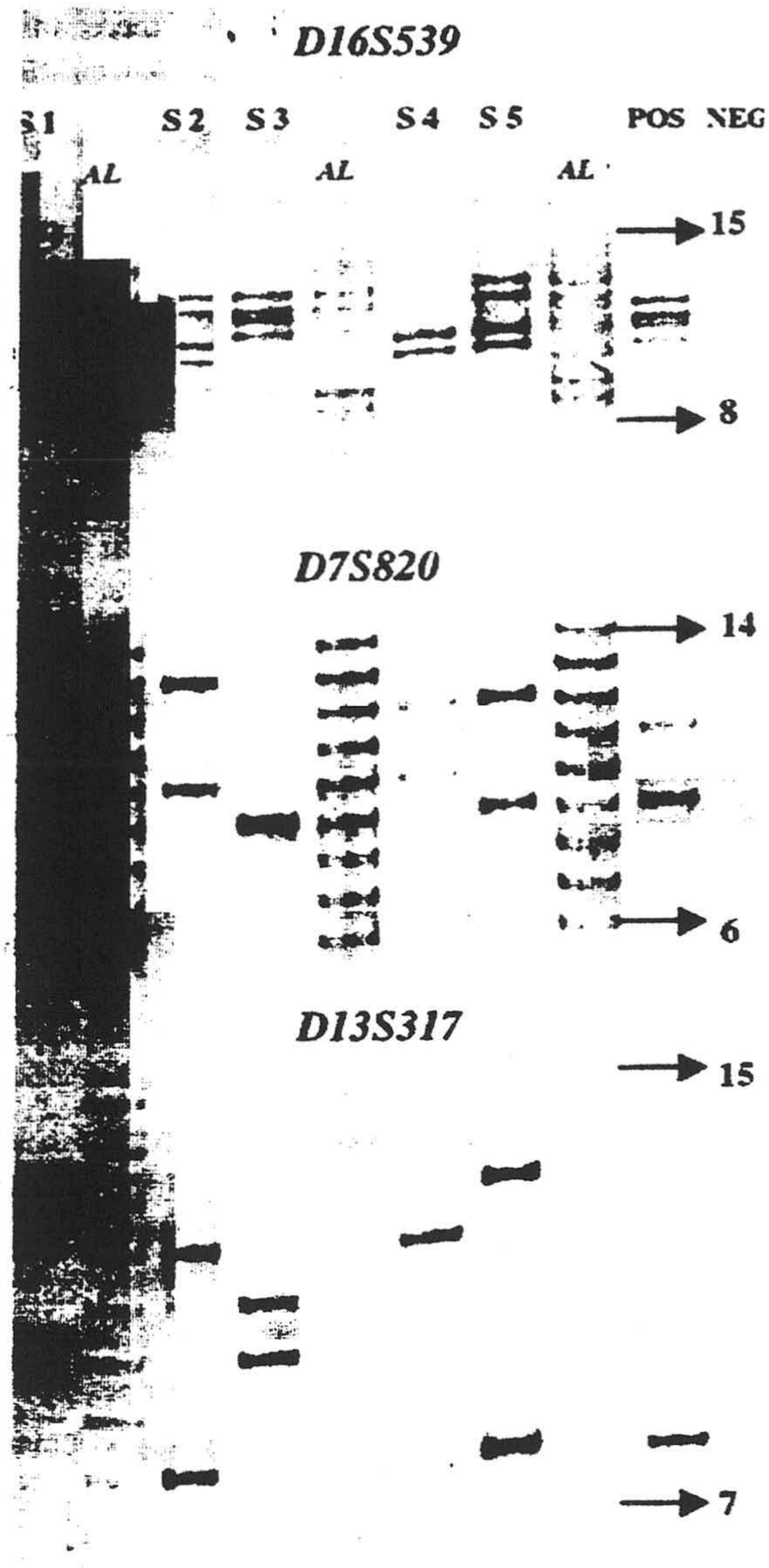


FIG. 1—*SilverSTR™ III genotypes of five samples (S1 to S5), K562 positive control (POS) and negative control (NEG). For each locus, the shortest and largest alleles present in the allelic ladder supplied by the manufacturer are pointed with arrows.*

TABLE 1—*Allele frequencies and parameters of statistical interest to check the Hardy-Weinberg equilibrium of the 3 loci included in the SilverSTR™ III kit. Cummulated PD: 0.9995—Cummulated PE: 0.9333*

Loci	D13S317*		D7S820†		D16S539‡	
Allele	Number	Proportion	Number	Proportion	Number	Proportion
7	—	—	8	0.0226	—	—
8	62	0.1751	54	0.1525	4	0.0113
9	19	0.0536	48	0.1356	42	0.1186
10	24	0.0678	113	0.3192	24	0.0678
11	99	0.2796	68	0.1921	92	0.2599
12	93	0.2627	54	0.1525	110	0.3107
13	39	0.1101	8	0.0226	75	0.2118
14	18	0.0508	1	0.0028	6	0.0169
15	—	—	—	—	1	0.0028

* Observed Homozygosity = 20.8%, Expected Homozygosity = 19.3%, Homozygosity Test ($p = 0.595$); Likelihood Ratio Test ($p = 0.339$); Exact Test ($p = 0.270$); PD = 0.929; PE = 0.618.

† Observed Homozygosity = 25.9%, Expected Homozygosity = 20.0%, Homozygosity Test ($p = 0.030$); Likelihood Ratio Test ($p = 0.006$); Exact Test ($p = 0.003$); PD = 0.927; PE = 0.606.

‡ Observed Homozygosity = 18.4%, Expected Homozygosity = 22.7%, Homozygosity Test ($p = 0.139$); Likelihood Ratio Test ($p = 0.048$); Exact Test ($p = 0.088$); PD = 0.904; PE = 0.556.

cause for the departure of HWE. Furthermore, there was no detectable departures from independence (i.e., linkage disequilibrium) between any pair-wise combination of loci.

In conclusion, the analysis of the DDD loci was facilitated due to the development of a commercial kit. The typing of these STR loci can be performed by most laboratories, since the kit enables manual gel electrophoresis and silver staining to be used to type the markers.

Furthermore, a Spanish population database has been established for the loci D13S317, D7S820, and D16S539. The allelic frequencies of these PCR-based loci can be used to estimate the frequency of a multiple locus PCR-based DNA profile in Spanish population. The data demonstrate that a significant degree of discrimination can be obtained (PD = 0.9995) when all three loci are used to characterize forensic biological evidence; the power of exclussion (PE) reaches 0.9333 for the 3 loci together.

Acknowledgments

We want to thank Randy Nagy and Carol Zabit, (Promega Corporation, Madison, WI), for their continous help and assistance.

References

1. Moretti TR, Budowle B. The CODIS STR project: evaluation of fluorescent multiplex STR systems. Proc 50th Meeting of the Am Acad Forensic Sciences. San Francisco, CA 1998:48.
2. Budowle B, Moretti T, Keys KM, Koons BW, Smerick JS. Validation studies of the CTT STR Multiplex system. J Forensic Sci 1997;42:701–7.
3. Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B. Population genetics and forensic efficiency of 4 AMPFLP's. Int J Legal Med 1992;104:329–33.
4. Comey CT, Koons BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM, et al. DNA extraction strategies for amplified fragment length polymorphism analysis. J Forensic Sci 1994;39:1254–69.
5. Budowle B, Baechtel FS, Comey CT, Giusti AM, Klevan L. Simple protocols for typing forensic biological evidence: chemiluminescent detection for human DNA quantitation and RFLP analysis and manual typing of PCR amplified polymorphisms. Electrophoresis 1995;16: 1559–67.
6. Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC. Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. Am J Hum Genet 1991;48:37–144.
7. Edwards A, Hammond H, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric repeat loci in four human population groups. Genomics 1992;12:244–53.
8. Chakraborty R, Smouse PE, Neel JV. Population amalgamation and genetic variation: observations on artificially aglomerated tribal populations of Central and South America. Am J Human Genet 1988;43:709–25.
9. Nei M, Roychoudhvry AK. Sampling variances of heterozygosity and genetic distance. Genetics 1978;76:583–90.
10. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 1978;89:583–90.
11. Chakraborrty R, Fornage M, Guegue R, Boerwinkle E. Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population. In: Burke T, Dolf G, Jeffreys AJ, Wolff R, editors: DNA fingerprinting: approaches and applications. Birkhauser Verlag, Berlin 1991; 127–43.
12. Weir BS. Independence of VNTR alleles defined by fixed bins. Genetics 1992;130:873–8.
13. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:371–2.
14. Karlin S, Cameron EC, Williams PT. Sibling and parent-offspring correlation estimation with variable family size. Proc Nat Acad Sci 1981; 78:2664–8.

Additional information and reprint requests:
Prof. Jose A. Lorente, M.D., Ph.D.
Dept. Medicina Legal—University of Granada
Av. Madrid 11
E-18012 Granada, Spain
Tel: + 34 958 243546
Fax: + 34 958 246107
E-mail: jlorente@goliat.ugr.es