



Universidad de Santiago de Compostela

Análisis de polimorfismos de ADN microsatélite de cromosoma Y. Estudio de la población de Galicia y aplicaciones forenses.

M^a Luisa Cal Teba

Tesis de Doctorado

Facultad: Medicina

Director: Dr. Ángel María Carracedo Álvarez

2001

Universidad de Santiago de Compostela
Facultad de Medicina
Instituto de Medicina Legal



BIBLIOTECA VIRTUAL

Análisis de polimorfismos de ADN microsatélite de
cromosoma Y: estudio de la población de Galicia y
aplicaciones forenses

Memoria que presenta para optar al grado de doctor:

M^a Luisa Cal Teba

Santiago de Compostela, Julio de 2000

El Doctor *D. Ángel Carracedo Álvarez*, Catedrático de Medicina Legal de la Facultad de Medicina de la Universidad de Santiago de Compostela y la Doctora *D^a. Carmela Pestoni Porvén*,

CERTIFICAN:

Que la presente Tesis Doctoral, “ANÁLISIS DE POLIMORFISMOS DE ADN MICROSATÉLITE DE CROMOSOMA Y : ESTUDIO DE LA POBLACIÓN DE GALICIA Y APLICACIONES FORENSES” de la Licenciada en Medicina y Cirugía D^a María Luisa Cal Teba, ha sido realizada bajo nuestra dirección en el Instituto de Medicina Legal, considerándola en condiciones para optar al Grado de Doctor y autorizándola para su presentación ante el Tribunal correspondiente.

Y para que así conste, expedimos el presente certificado en Santiago de Compostela, a 24 de octubre de 2000.

Fdo. Prof. Dr. Ángel Carracedo Álvarez

Fdo. Dra. Carmela Pestoni Porvén

A Tino
A Catuxa
A mis padres



“La ciencia será siempre una búsqueda,
jamás un descubrimiento real. Es un viaje,
nunca una llegada”.

Karl Raimund Popper

AGRADECIMIENTOS

Mi gratitud incondicional a todas y cada una de las personas que han colaborado en la realización de esta tesis.

Especialmente a Angel Carracedo y Carmela Pestoni, directores de este trabajo, responsables de mi introducción en el fascinante y complejo mundo de la Genética Forense. Ángel, gracias por permitirme formar parte de tu equipo y por estar siempre accesible para resolver mis dudas y vacilaciones. Carmela, mi agradecimiento por encauzar mi formación científica y compartir conmigo tus conocimientos y experiencia, por esa contagiosa vitalidad para afrontar los constantes desafíos y por tu permanente estímulo.

Al profesor Luis Concheiro, por su amabilidad y por su apoyo, tanto profesional como personal.

A Maviki y Marisol, por asesorarme siempre que necesité sus consejos y por sus palabras de ánimo en las horas bajas. Gracias por vuestra comprensión y generosidad.

A mis compañeros de fatigas del Instituto de Medicina Legal: Ana, Juana, María, Meli, Paula, Sandra, Toño y demás colegas, con los que compartí muchas horas de trabajo e ilusiones y que siempre estuvieron a mi lado para brindarme su ayuda desinteresada. Mi gratitud por los buenos momentos, por vuestro entusiasmo incombustible y vuestra amistad.

A mi familia, por soportar pacientemente mi abandono y por animarme a que este proyecto llegue a ser realidad. Sin vosotros no hubiese sido posible. Gracias por vuestro cariño y comprensión.

Abstract. Some of DNA polymorphisms Y-specifics recently described could be analyzed as very informative haplotypes from a forensic point of view. With the object of including such polymorphisms in the forensic routine it is necessary a study of them that include the attainment of the allelic and haplotypic frequencies of the Galician population (NW Spain), comparatives studies with other populations, sequenced allelic ladders construction and study of possible mutations. The polymorphisms studied are seven Y-STR (Short Tandem Repeats or microsatellites) witch will be analyzed by PCR (Polymerase Chain Reaction). We also include the calculation of some statistic parameters to value the utility of these polymorphisms (DYS19, DYS389-I, DYS389-II, DYS390, DYS393 and DYS385: two loci) in the forensic practice.





ÍNDICE

ÍNDICE	I
ABREVIATURAS	VII
JUSTIFICACIÓN DEL TEMA	XI
I.- INTRODUCCIÓN	1
1.- ADN y Ciencias Forenses	3
1.1.- Estructura y función del ADN	4
1.1.1.- ADN repetido en tándem	6
1.1.2.- ADN repetitivo disperso	8
1.2.- Polimorfismos del ADN	9
1.2.1.- Técnicas de análisis de estos polimorfismos en genética forense	10
2.- Análisis de los polimorfismos de ADN por PCR	13
2.1.- Reacción en cadena de la polimerasa	13
2.2.- Aplicaciones del análisis-PCR en genética forense	14
2.2.1.- Polimorfismos de ADN codificante	14
2.2.2.- Polimorfismos de ADN repetido en tándem	15
2.2.3.- Otros polimorfismos	19
2.3.- Estandarización de los análisis	20
3.- Cromosoma Y	25
3.1.- Introduciéndonos en el cromosoma Y	24
3.1.1.- Su naturaleza	24
3.1.2.- Las regiones pseudoautosómicas	25
3.1.2.1.- Los genes de las regiones pseudoautosómicas	27
3.1.3.- La región específica del cromosoma Y y sus genes	28
3.1.4.- Algunas peculiaridades del cromosoma Y. Analogía con el ADN mitocondrial	29

3.2.- Polimorfismos en el cromosoma Y	33
3.2.1.- Tipos de polimorfismos en el cromosoma Y	33
3.2.2.- STRs de cromosoma Y analizables mediante PCR	41
4.- Contribución de los polimorfismos de ADN al campo médico-legal y de los polimorfismos de cromosoma Y en particular	43
4.1.-Investigación biológica de la paternidad	43
4.2.-Criminalística	44
4.3.-Identificación genética	47
II.- MATERIAL Y MÉTODOS	49
1.- Características de la muestra estudiada	51
2.- Extracción del ADN	51
2.1.- Extracción del ADN a partir de sangre total	51
2.2.- Extracción a partir de manchas de sangre	53
3.- Cuantificación del ADN	54
4.- Amplificación del ADN	55
4.1.- Reacción de PCR	55
4.2.- Condiciones de amplificación de los sistemas	57
5.- Detección de los productos amplificados	59
5.1.- Electroforesis en geles de poliacrilamida y tinción de plata	59
5.2.- Detección automática de fragmentos marcados con fluoresceína	60
6.- Secuenciación	62
6.1.- Elución de los alelos	62
6.2.- Purificación pre-secuenciación.	63
6.3.- Reacción de secuenciación	63
6.4.- Electroforesis y análisis de la secuencia	65
7.- Construcción de ladders alélicos y designación de alelos	66
8.- Análisis estadístico	66

III.- RESULTADOS Y DISCUSIÓN	71
1.- Consideraciones metodológicas.	73
1.1.- Sobre el método de extracción del ADN	73
1.2.- Sistemas electroforéticos	73
2.- Acerca de la secuencia y la nomenclatura empleada	76
3.- Ladders alélicos	82
4.- Resultados genético-poblacionales .	83
4.1.- Caracterización de la población gallega para los sistemas DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS393	83
4.2.- Determinación de haplotipos	92
4.3.- Estudio comparativo con otras poblaciones	95
4.3.1.- Sistema DYS19	95
4.3.2.- Sistema DYS389-I	102
4.3.3.- Sistema DYS389II	108
4.3.4.- Sistema DYS390	114
4.3.5.- Sistema DYS393	118
4.3.6.- Sistema DYS385	126
5.- Parámetros de utilidad forense: eficacia de los sistemas y su validación médico-legal	131
IV.- CONCLUSIONES	133
V.- BIBLIOGRAFIA	137
VI.- ANEXO	157

BIBLIOTECA VIRTUAL



ABREVIATURAS

B A	Adenina
ALF	<i>Automatic Laser Fluorescent DNA Sequencer</i>
ADN/DNA	Ácido Desoxirribonucleico
AMPFLPs	<i>Amplification Fragment Length Polymorphisms</i> (Polimorfismos basados en la longitud de los fragmentos amplificados)
ASO	<i>Allele Specific Oligonucleotides</i>
C	Citosina
CE	Probabilidad de exclusión <i>a priori</i>
cm	Centímetro
D	Diversidad génica
ddNTP	Dideoxinucleótido trifosfato
dNTP	Deoxinucleótido trifosfato
EDNAP	<i>European Dna Profiling Group</i>
EDTA	Ácido etilendiaminotetraacético
et al	y colaboradores
g	Gramo
G	Guanina
GDB	Genome DataBase
H33258	Colorante Hoechst33258
HLA	<i>Human Leukocyte Antigen</i>
ISFG	<i>International Society for Forensic Genetics</i>
Kb	Kilobases
LINEs	<i>Long Interspersed Elements</i> (Elementos dispersos largos)
M	Molar
mA	Miliamperios
Mb	Megabases
mg	Miligramo
ml	Mililitro
MLPs	<i>Multilocus Probes</i> (Sondas de locus múltiple)
mm	Milímetro
mM	Milimolar
MVR	<i>Minisatellite Variant Repeat</i> (Variación en la secuencia de las repeticiones de los minisatélites)
MVR-PCR	<i>Minisatellite Variant Repeat</i> mediante PCR
ng	Nanogramo
nm	Nanómetro
p/v	Peso/Volumen
pb	Pares de bases
PCR	<i>Polymerase Chain Reaction</i> (Reacción en cadena de la polimerasa)
PD	Poder de discriminación

Abreviaturas

pm	Picomoles
RFLPs	<i>Restriction Fragment Length Polymorphism</i> (Polimorfismos basados en la longitud de los fragmentos de restricción)
rpm	Revoluciones por minuto
SDS	Dodecil sulfato sódico
SINEs	<i>Short Interspersed Elements</i> (Elementos dispersos cortos)
SLPs	<i>Single Locus Probes</i> (Sondas de locus único)
SSCP	<i>Single Strand Conformational Polymorphism</i> (Polimorfismo de conformación de cadena única)
STRs	<i>Short Tandem Repeats</i> (Repeticiones en tándem cortas)
SWGDM	<i>Standardization Working Group on DNA Analysis Methods</i>
T	Timina
U	Unidades
V	Voltios
VNTRs	<i>Variable Number of Tandem Repeats</i> (Número variable de repeticiones en tándem (Minisatélites))
W	Wattios
µg	Microgramo
µl	Microlitro
µM	Micromolar

El descubrimiento de las regiones hipervariables de ADN por Jeffreys y colaboradores (1985) y la introducción de la reacción en cadena de la polimerasa (PCR) (Saiki *et al* 1985, Mullis y Faloona 1987), han supuesto una de las mayores innovaciones en las Ciencias Forenses, especialmente en aspectos relacionados con la identificación de personas o restos cadavéricos, análisis de vestigios biológicos en criminalística e investigación biológica de la paternidad.

Para que un marcador analizable por PCR pueda ser considerado de utilidad médico-legal, debe ser muy polimórfico y poseer un alto grado de heterocigosidad genética. Antes de ser aceptado como integrante de la rutina forense es imprescindible disponer de datos fiables acerca de la distribución de sus frecuencias alélicas en la población general, además de cumplir una serie de requisitos y pasar sucesivos controles para su validación como tal.

Los polimorfismos de cromosoma Y son indudablemente de un gran interés forense y permiten arrojar luz sobre problemas de identificación o casos de paternidad de difícil solución con polimorfismos de ADN de cromosomas autosómicos, pero sobre todo, permiten el análisis de vestigios de interés criminal, particularmente restos de esperma mezclados con células femeninas en los que la extracción diferencial no tiene éxito.

De entre todos los polimorfismos del cromosoma Y, en el presente estudio nos hemos centrado en un grupo seleccionado de marcadores microsatélites (STRs), ya que poseen un alto grado de variabilidad (dentro de la tónica general de escaso polimorfismo del cromosoma Y) y permiten el abordaje por técnicas de PCR.

Del conjunto de STRs descubiertos hasta el momento de iniciar este trabajo (algo más de 12) hemos elegido 6 microsatélites: los sistemas DYS19, DYS385, DYS389I, DYS389II, DYS390 y DYS393, todos ellos repeticiones de tetranucleótidos y, en principio, idóneos para los fines perseguidos.

Justificación del tema

La mayor parte de estos STRs no están secuenciados y de la mayoría de ellos no disponemos de escaleras alélicas secuenciadas. Este es el primer objetivo que nos propusimos en el presente trabajo: secuenciar todos los alelos de cada sistema y, en segundo lugar, construir escaleras alélicas secuenciadas para su disposición por toda la comunidad científica. El tercer objetivo consiste en la realización de un estudio genético-poblacional, en una muestra de la población gallega, con esos marcadores y conocer así su utilidad forense.





I. INTRODUCCIÓN

1.- ADN Y CIENCIAS FORENSES

Desde hace casi un siglo, para estudiar las variaciones entre individuos se utilizan los llamados Polimorfismos Genéticos o Marcadores Genético-Moleculares. Son caracteres estables que se transmiten por herencia mendeliana simple y constituyen una expresión de la diversidad genética entre individuos de la misma especie.

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

- Patrón de herencia bien establecido.
- Elevado polimorfismo.
- Alto grado de heterocigosidad.
- Detección fiable de los alelos.
- Datos poblacionales de frecuencias alélicas, fenotípicas y/o genotípicas establecidas.
- Herencia independiente de los otros marcadores usados.
- Tasa de mutación baja.
- Analizable mediante un método simple, rápido y reproducible.
- Precisar poco material para el análisis.

Hace pocos años los análisis forenses estaban basados fundamentalmente en el estudio de marcadores genéticos convencionales (antígenos eritrocitarios y leucocitarios, proteínas séricas y enzimas eritrocitarias). En la década de los 80 se consigue un avance espectacular en el campo de la genética forense a raíz del descubrimiento de las regiones hipervariables del ADN (Jeffreys 1985a). A partir de este momento, la utilidad de los polimorfismos clásicos va disminuyendo conforme se generaliza el estudio de estos nuevos marcadores, mucho más informativos, debido principalmente a la gran variabilidad y estabilidad química del ADN, así como a la alta sensibilidad de las técnicas que lo analizan.

Los polimorfismos de ADN hipervariables poseen tal capacidad identificadora que son considerados, hoy en día, un instrumento de elección en la resolución pericial de casos en Genética Forense. Antes de adentrarnos en el estudio de estos polimorfismos, vamos a referirnos a algunos aspectos básicos del ADN.

1.1.- ESTRUCTURA Y FUNCIÓN DEL ADN

El ADN es la molécula que contiene toda la información genética del individuo. El conjunto de esta información presente en las células se denomina genoma y, según su localización, podemos identificar un genoma complejo nuclear y un genoma mitocondrial simple.

El genoma humano haploide está constituido aproximadamente por 3.3×10^9 pb. En el genoma existen 2 tipos de ADN, según la función biológica que desempeñen:

-ADN expresivo (30%). Este tipo de ADN, también llamado codificante, tiene una función conocida, como por ejemplo la expresión de un gen cuyo producto final es una proteína. Soporta gran presión selectiva, lo que se traduce en una variabilidad de regiones limitada.

-ADN no codificante (70%). Comprende secuencias de ADN transcripcionalmente inactivas de funciones diversas (como por ejemplo, promotores de genes) y, en otros muchos casos, de función desconocida o sin función aparente. Este tipo de ADN, por ser altamente polimórfico, tiene un gran interés de cara a la identificación de individuos. Lo podemos clasificar en :

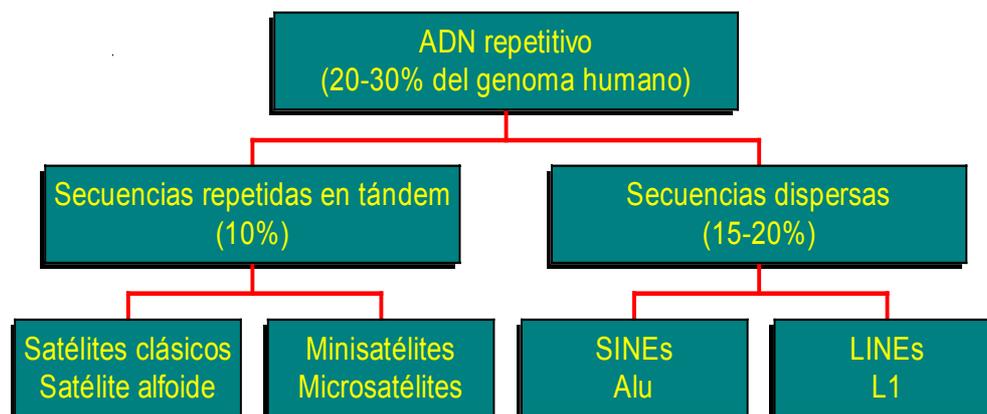
• **ADN de copia única.** Está compuesto por secuencias que se encuentran representadas una o muy pocas veces en el genoma. Se cree que puede actuar como espaciador entre regiones codificantes de ADN.

•**ADN de copia múltiple.** Las secuencias de este tipo de ADN, también denominado ADN Repetitivo, pueden ser altamente repetitivas, moderadamente o poco repetitivas. Podemos clasificarlas en base a sus dos características más importantes: su disposición a lo largo del genoma y el tamaño de la unidad de repetición (Tabla 1). Se pueden reconocer dos grupos principales:

ADN Repetido en Tándem (10% del genoma). Se compone de bloques de secuencias repetitivas agrupadas en tándem. Según el tamaño de la unidad de repetición se subdivide en 3 tipos: ADN satélite, ADN minisatélite y ADN microsatélite.

ADN Repetitivo Disperso (15-20% del genoma). Las unidades de repetición no se agrupan, sino que aparecen dispersas a lo largo del genoma. Esta representado por 2 familias: SINEs y LINEs.

Tabla 1. Clasificación de las secuencias de ADN repetitivo en el genoma humano (modificado de Fowler *et al* 1988)



1.1.1.- ADN REPETIDO EN TÁNDEM

Los distintos componentes de este tipo de ADN repetitivo adoptan un patrón de distribución cromosómica diferente: el ADN satélite se sitúa en la región centromérica, el ADN minisatélite en los telómeros o en sus proximidades, y el ADN microsatélite aparece disperso por todo el cromosoma.

Según Vogt (1990), las secuencias de ADN repetido en tándem se distribuyen por el genoma de dos maneras:

-Tipo I. Grandes bloques formados por repeticiones de distintas unidades de longitud variable. Se corresponde con los satélites clásicos I-IV y el satélite alfoide.

-Tipo II. Pequeños bloques distribuidos a lo largo de todo el genoma, con un número variable de unidades de repetición de secuencia similar. Pertenecen a este tipo los minisatélites y los microsatélites.

•**ADN Satélite.** Las secuencias repetitivas se disponen en grandes bloques de diversas unidades de complejidad variable, con una longitud desde 100 Kb a varias Mb. Este tipo de ADN no se transcribe y al organizarse de un modo tan complejo es difícil su aplicación al campo forense.

Para separar el ADN Satélite del resto del ADN genómico se somete a centrifugación en gradiente de densidad de Plata-Sulfato de Cesio, obteniéndose 2 bandas. La de menor densidad corresponde al ADN Satélite.

Se distinguen 4 tipos (I, II, III, IV) en función de su densidad, que está en relación con el mayor o menor contenido en GC (Singer 1982a). Representan en conjunto del 2 al 6% del genoma y tienen una secuencia consenso (unidad básica de repetición) de 5 a 170 pb (Prosser *et al* 1986, Frommer *et al* 1982, Hollis y Hindley 1988, Waye y Willard 1987). Algunos satélites poseen una unidad de repetición de pequeño

tamaño, como los satélites II y III, que consisten en la repetición en tándem de la secuencia ATTCC.

Existen otros tipos de ADN satélite que no pueden distinguirse por centrifugación en gradiente de densidad, debido a que su densidad es semejante a la de los satélites II y III. Se consiguieron caracterizar por digestión de ADN genómico mediante endonucleasas de restricción. El satélite alfoide pertenece a esta categoría, habiéndose comprobado que es el único satélite que está presente en todos los cromosomas, constituyendo la mayor parte de la heterocromatina del centrómero. Su secuencia consenso contiene aproximadamente 171 pb, con variaciones individuales (Choo *et al* 1991).

•**ADN Minisatélite.** El término minisatélite se debe a Jeffreys, quien lo usó para designar loci de ADN repetitivo de un tamaño menor que el de los satélites clásicos (Jeffreys *et al* 1985a). Los bloques de secuencias de este ADN poseen un tamaño aproximado entre 0.1 y 40 Kb y la unidad de repetición es de 10-100 pb. Tienen un alto grado de variabilidad, tanto en el número de repeticiones en tándem como en la secuencia de la unidad de repetición, por lo que ante tamaños idénticos podemos estar frente a alelos diferentes.

Nakamura *et al* (1987) denominaron a estos loci VNTRs (*Variable Number of Tandem Repeats*) aludiendo a la variación en el número de unidades de repetición. Tomado en sentido estricto, este término podría aplicarse a cualquier tipo de ADN repetitivo y no exclusivamente al minisatélite, y además, al hacer referencia sólo al número de repeticiones, los minisatélites monomórficos en poblaciones humanas no podrían incluirse (Armour *et al* 1990).

En el genoma humano los minisatélites no se distribuyen al azar. Se localizan preferentemente en las regiones subterminales de los cromosomas (Armour *et al* 1989) que son las implicadas en los fenómenos de recombinación, sobre todo en línea germinal masculina.

•**ADN Microsatélite.** Los microsatélites fueron llamados así por su pequeño tamaño, hasta unos 400 pb, lo que los hace especialmente idóneos para técnicas de PCR. También se conocen como STRs (*Short Tandem Repeats*) pues la unidad de repetición oscila entre 2 y 7 pb (Edwards *et al* 1991). Se encuentran ampliamente repartidos por todo el genoma (Litt y Luty 1989, Weber y May 1989).

Basándose en la longitud de la unidad de repetición y el número de veces que ésta se repite, junto con las posibles variaciones en su secuencia, Urquhart *et al* (1994) los clasifica en STRs simples, compuestos y complejos. Posteriormente, Brinkmann (1996) propone una clasificación alternativa: STRs con baja microvariabilidad, intermedia y alta microvariabilidad.

1.1.2.- ADN REPETITIVO DISPERSO

Está representado por dos familias que se diferencian en el número de nucleótidos que constituyen la unidad de repetición (Singer 1982a). Algunos componentes de estas familias pueden ser considerados como transposones (elementos genéticos móviles), que son fragmentos inestables de DNA con capacidad migratoria (Koremborg y Rykowski 1988).

-**SINEs** (*Short Interspersed Nuclear Elements*). Representan alrededor del 10 % del genoma. Los constituyen repeticiones menores de 500 pb. La familia más representativa y mejor estudiada es la *Alu*, así llamada por presentar un lugar de reconocimiento para el enzima de restricción *Alu* I (Slagel *et al* 1987).

Las secuencias *Alu* son las más abundantes en el genoma humano y tienen una secuencia muy conservada de unos 300 pb relativamente rica en GC. Se localizan en las regiones eucromatínicas, preferentemente en las llamadas “bandas R” o bandas reversas (de replicación precoz), que son las regiones cromosómicas más activas a nivel transcripcional.

Existen otras familias, como la NTS (*Non Transcribed Spacer*) que se localiza en primer lugar en la región homónima del ADN ribosómico (Singer 1982b).

-**LINEs** (*Long Interspersed Nuclear Elements*). Suponen del 2 al 5% del genoma. Están constituidos por secuencias repetitivas mayores de 500 pb. El *LINE-1* o familia *LI*, llamada también *Kpn*, es el LINE más abundante en humanos, presentando un número elevado de repeticiones. Los elementos *LI* se localizan principalmente en la eucromatina, al igual que la familia *Alu*, pero en este caso a nivel de las “bandas G” (de replicación tardía) obtenidas al teñir los cromosomas con Giemsa.

1.2.- POLIMORFISMOS DEL ADN

El término polimorfismo fue empleado por Ford (1940) para designar “la aparición conjunta en un lugar de dos o más formas discontinuas de una especie, de tal manera que la más rara de ellas no se puede mantener simplemente a través de la mutación periódica”. En la práctica, para que un locus sea considerado polimórfico, el alelo más común para ese locus debe tener una frecuencia poblacional menor del 99% y, de acuerdo con la ley de Hardy-Weinberg, al menos un 2% de la población debe ser heterocigota para ese locus.

En el ADN codificante existe poca variabilidad individual, exceptuando la región HLA. El margen de variación permitido es muy bajo y los polimorfismos suelen acompañarse de modificaciones fenotípicas. Por ejemplo, si se producen diferentes formas de una proteína se puede condicionar su función o actividad, bien intrínsecamente, bien por influencia ambiental. El ADN no codificante, por el contrario, al no estar sujeto a presión selectiva intensa, puede soportar generalmente grandes niveles de variabilidad sin que se produzca repercusión fenotípica. Esta característica ha convertido a este tipo de ADN en la mayor fuente de investigación de polimorfismos en Genética Forense.

Los polimorfismos pueden ir desde la modificación de una sola base hasta cambios en número y/o tamaño en la unidad de repetición. Pueden dividirse en 2 tipos:

-Polimorfismos de Secuencia. Se producen por el cambio de uno (mutación puntual) o más nucleótidos en una secuencia de ADN. Suelen ser poco polimórficos y son típicos del ADN expresivo.

-Polimorfismos de Longitud. Se producen por la inserción o deleción de uno o más nucleótidos. Este tipo es el más abundante en ADN repetitivo, sobre todo en el ADN mini y microsatélite.

Los loci minisatélites y microsatélites están formados, como ya hemos visto, por repeticiones en tándem de secuencias similares que varían en longitud, pudiéndose acompañar simultáneamente de polimorfismos de secuencia. La causa de la variación se supone distinta en ambos casos. Para minisatélites se achaca a un mecanismo de conversión génica, frecuentemente inducida por regiones de flanqueo. En el caso de los microsatélites, según Armour (1996), la causa principal parece debida a un deslizamiento de unidades de repetición durante la replicación (*replication slippage*), aunque Jeffreys *et al* (1997) postulan que también la conversión génica, muchas veces inducida por regiones de flanqueo, puede ser un mecanismo preferente de variabilidad.

1.2.1.- TÉCNICAS DE ANÁLISIS DE ESTOS POLIMORFISMOS EN GENÉTICA FORENSE

Desde hace poco más de una década, la tecnología del ADN ha permitido el estudio de la variabilidad humana mediante el análisis directo del propio material genético y no a través del análisis de las proteínas codificadas por éste.

Wyman y White (1980) dieron el primer paso en la identificación genética por medio del ADN al descubrir un locus polimórfico caracterizado por un número de fragmentos de restricción de longitud variable, aunque el verdadero auge de los polimorfismos de ADN se inicia con Jeffreys y colaboradores, sus introductores en el campo forense.

Los primeros en utilizarse con este fin fueron los denominados RFLPs “*Restriction Fragment Length Polymorphism*” (Jeffreys *et al* 1985a,b), polimorfismos en ADN minisatélite basados en la longitud de los fragmentos de restricción. Se identificaron mediante la digestión de ADN genómico con enzimas de restricción y posterior hibridación con sondas (Botstein *et al* 1980). Un problema legal de inmigración en el Reino Unido (Jeffreys *et al* 1985c) fue el primer caso resuelto satisfactoriamente con ayuda de estos polimorfismos.

En un primer momento se utilizaron las denominadas sondas multilocus (MLPs, *Multilocus Probes*), que detectan múltiples loci minisatélites bajo condiciones poco rigurosas de hibridación, dando lugar a un complejo patrón de bandas. Este patrón de bandas múltiples corresponde a distintos loci con secuencias relacionadas entre sí. Jeffreys y su equipo consideraron que estos patrones serían prácticamente específicos para cada individuo y los denominaron “huellas genéticas” (*DNA fingerprints*).

Poco tiempo después se comienzan a utilizar sondas que permiten detectar un único locus (Wong *et al* 1987, Nakamura *et al* 1987b), las denominadas sondas unilocus (SLPs, *Single Locus Probes*) que proporcionan una o dos bandas según el carácter homocigoto o heterocigoto del individuo para ese locus. De esta manera se obtiene un perfil unilocus de ADN (*DNA profiling*).

Hoy en día sólo las sondas unilocus debidamente caracterizadas (Brinkmann *et al* 1992) están reconocidas para usos forenses. Han supuesto una evidente mejora metodológica en la investigación biológica de la paternidad, pero existen importantes limitaciones para su uso en criminalística. Un método alternativo, y en determinados casos mucho más recomendable que el empleo de sondas, lo constituye la Reacción

en Cadena de la Polimerasa (PCR) (Mullis et Faloona 1987, Saiki *et al* 1988), como veremos más adelante.

En los últimos años estamos asistiendo a un despliegue de nuevas tecnologías, especialmente encaminadas al análisis de STRs, como por ejemplo:

-*LightCycler*®. Esta tecnología permite llevar a cabo, en una misma reacción, la amplificación por PCR de un fragmento de ADN y la detección por fluorescencia, en tiempo real, de la hibridación con dos sondas independientes. Es de una gran especificidad y posibilita la cuantificación exacta de producto amplificado y la detección de posibles mutaciones.

-*Chips* o *Arrays* de ADN. Son sistemas en miniatura que incorporan de cientos a cientos de miles de sondas alelo-específicas (ASO, *Allele Specific Oligonucleotides*). Si el fragmento de ADN diana, marcado mediante la incorporación de un grupo fluorescente durante la PCR, hibrida con la sonda complementaria, se detecta una señal de fluorescencia que es recogida a través de sistemas automatizados. Esta prometedora tecnología puede tener interesantes aplicaciones, fundamentalmente clínicas y forenses.

-*Espectrofotometría de masas*. Permite el análisis de una muestra en cuestión de segundos, sin necesidad de emplear escaleras alélicas, pero de momento sólo es utilizable para fragmentos de ADN menores de 100 pb. Los productos a analizar se cristalizan conjuntamente con una pequeña matriz orgánica, se volatilizan y se someten a un campo eléctrico en fase gaseosa. Se determina el peso molecular del ADN en función del tiempo empleado en el recorrido hasta llegar a un detector, ya que el tiempo es proporcional a la masa del fragmento de ADN.

2.- ANÁLISIS DE LOS POLIMORFISMOS DE ADN POR PCR

La principal limitación que presenta el empleo de sondas es que el análisis debe realizarse con ADN íntegro y en cantidad suficiente. Otros problemas añadidos son la laboriosidad del método, el tiempo de análisis (no menor de dos días) y la difícil estandarización de esta técnica.

La introducción de la Reacción en Cadena de la Polimerasa (PCR) vino a solucionar en gran medida muchas de estas dificultades. Esta técnica ha supuesto un avance metodológico de enorme repercusión en genética forense, ya que nos permite trabajar con cantidades ínfimas de ADN aún estando degradado. Otras ventajas adicionales de la PCR son la rapidez, sencillez, facilidad de interpretación y menor coste de la determinación, además de proporcionar una buena sensibilidad y especificidad.

El poder de exclusión individual de los marcadores analizados por PCR es menor que el de los marcadores detectados por sondas unilocus. Este dato deberá tenerse muy en cuenta aquellos casos en los que esta última determinación pueda ser factible.

2.1.- LA REACCIÓN EN CADENA DE LA POLIMERASA

La PCR es una técnica de amplificación “in vitro” de pequeñas secuencias de ADN que permite sintetizar millones de copias idénticas a partir de una cadena única. Se basa en la facilidad que posee el ADN para desnaturalizarse y renaturalizarse.

Para iniciar la síntesis de ADN se utilizan dos cebadores o *primers* específicos (secuencias de aproximadamente 20-30 nucleótidos de longitud complementarias a

los extremos 5' del fragmento que se desea amplificar) en combinación con una ADN polimerasa termoestable que incorpora dNTPs.

La reacción se lleva a cabo en un termociclador que realiza ciclos automáticos de tres temperaturas (93-95°C para desnaturalizar el ADN, 50-65°C para que tenga lugar el acoplamiento de los *primers* y 72°C para que la polimerasa copie la nueva cadena de ADN), consiguiendo de este modo que el número de copias de la secuencia de ADN que nos interesa aumente exponencialmente.

Experimentalmente, alrededor de 30 ciclos proporcionan una cantidad de ADN más que suficiente para un análisis directo del segmento amplificado sin necesidad de utilizar sondas.

2.2.- APLICACIONES DEL ANÁLISIS-PCR EN GENÉTICA FORENSE

Mientras que en principio varias categorías de polimorfismos de ADN podrían utilizarse con fines de identificación, la mayor parte de los marcadores usados en Genética Forense pertenecen a la categoría del ADN altamente repetitivo.

La amplia lista de polimorfismos amplificables por PCR (Tabla 2) continúa incrementándose a medida que se estudian nuevos loci, se contrastan experiencias y los diferentes grupos de investigación llegan a acuerdos para su validación.

2.2.1.- POLIMORFISMOS DE ADN CODIFICANTE

Aunque muy pocos marcadores de secuencia única son buenos candidatos para usos forenses (Ugozzoli *et al* 1992; Higuchi *et al* 1988), el primer sistema amplificado por PCR con este fin fue un polimorfismo de la región HLA, el locus HLA DQA1 (Saiki *et al* 1986; Erlich *et al* 1986; Higuchi *et al* 1988). Este polimorfismo de secuencia se detecta por *dot-blot* reverso del producto de PCR con

un panel de sondas ASO, inmovilizadas sobre una membrana de nailon (Erlich y Bugawan 1989), que reconocen cada alelo del sistema.

También se han propuesto otros métodos alternativos, más baratos y eficaces, mediante la detección de fragmentos amplificados en geles de poliacrilamida por análisis de heterodúplex y SSCP (*Single Strand Conformational Polimorphism*) (Barros *et al* 1991, 1992, 1994).

2.2.2.- POLIMORFISMOS DE ADN REPETIDO EN TÁNDEM

De todos los polimorfismos analizables por PCR, los de mayor interés médico-legal pertenecen a los tipos minisatélite y, sobre todo, microsatélite. Ambos son particularmente adecuados por ser altamente polimórficos y presentar un alto grado de heterocigosidad. Puesto que estos polimorfismos se deben a diferencias en el número de copias de las unidades repetitivas que presentan los distintos alelos, pueden detectarse tras la PCR mediante electroforesis, generalmente en geles de agarosa o poliacrilamida, como polimorfismos de longitud.

Estos polimorfismos pueden ser separados en clases discretas de alelos y clasificados por el número de repeticiones, lo que no es posible con el empleo de sondas. De este modo se facilita enormemente la estima de frecuencias (Alford *et al* 1994, Schmitt *et al* 1994), constituyendo una ventaja desde el punto de vista estadístico.

Cuando se trata de analizar la secuencia de algunos minisatélites se observa que raramente las unidades de repetición son todas iguales (Wong *et al* 1987). Este polimorfismo de secuencia puede determinarse junto con el polimorfismo de longitud por medio de una técnica denominada MVR (*Minisatellite Variant Repeat*).

- **AMPFLPs.** Los loci minisatélites que pueden ser amplificados por PCR se denominaron AMPFLPs (*AMPlified Fragment Length Polymorphisms*) (Jeffreys *et*

al 1988). Generalmente tienen unidades de repetición largas en múltiples copias y fragmentos alélicos de alto peso molecular. A pesar de los buenos resultados que se obtienen con esta técnica, que conjuga la alta variabilidad de los VNTRs con la sensibilidad y especificidad de la técnica PCR, no debemos olvidar que los productos de amplificación son de un tamaño relativamente grande, lo que puede constituir un factor limitante cuando las muestras se encuentran degradadas.

Los AMPFLPs fueron introducidos en el campo forense por Kasai *et al* (1990) y Budowle *et al* (1991). Los más utilizados son los loci D1S80 (pMCT118) (Kasai *et al* 1990, Budowle *et al* 1990), D17S5 (YNZ22) (Wolff *et al* 1988, Horn *et al* 1989), COL2A1 (Wu *et al* 1990, Rand *et al* 1992), 3'ApoB (Boerwinkle *et al* 1989). De ellos, el más empleado es el sistema el D1S80 (unidad de repetición de 16 pb y 29 alelos), que tiene a su favor una elevada heterocigosidad pero el inconveniente de que en muestras muy deterioradas se puede potenciar la amplificación de los alelos de menor tamaño. Esto nos conduciría al error de tipar como homocigota una muestra en la que no se amplificó el alelo de mayor tamaño.

- **MVR.** Los minisatélites, como ya hemos visto, poseen dos tipos de variación, una debida a la longitud (dependiendo del número de repeticiones, VNTRs) y otra ocasionada por diferencias en la unidad de repetición, lo que se denomina MVR.

En la actualidad este polimorfismo puede determinarse mediante PCR (MVR-PCR), técnica de gran sensibilidad y que posee una elevada capacidad de discriminación, cuyas bases metodológicas fueron descritas por Jeffreys *et al* (1991) y Tamaki *et al* (1992). La introducción de sistemas electroforéticos semiautomáticos y secuenciadores automáticos (Rodríguez-Calvo *et al* 1996) ha supuesto un gran avance en la automatización del análisis MVR-PCR, facilitando su aplicación práctica. Hay que destacar que el análisis no está sujeto a errores de medición en la migración de los fragmentos ni a fenómenos de distorsión de los geles.

Entre los diferentes loci que pueden ser abordados por MVR-PCR, interesantes desde el punto de vista médico-legal, citaremos el MS32 (locus D1S8) (Wong *et al* 1987), MS31A (locus D7S21) (Royle *et al* 1988), YNH24 (locus D2S24) (Holmlund *et al* 1998) y MSY1 (locus DYF155S1), el primer minisatélite descrito específico del cromosoma Y (Jobling *et al* 1994a).

La MVR-PCR es una técnica relativamente sensible con pequeñas cantidades de ADN e incluso se puede aplicar a muestras parcialmente degradadas, pero no se aconseja su uso para análisis de mezclas de muestras ni en casos forenses de paternidad.

- **MICROSATÉLITES.** Descritos inicialmente por Weber y May (1989) son secuencias pequeñas de ADN repetidas en tándem (STRs), constituidas por unidades de 2 a 7 pb que se repiten un número variable de veces y que dan lugar a alelos de tamaño aproximado entre 80 y 400 pb. Su gran utilidad se puso rápidamente de manifiesto debido a su abundancia (Beckman y Weber 1992), distribución bastante regular por el genoma (Stallings *et al* 1990), naturaleza polimórfica (Weber y May 1989, Litt y Luty 1989) y su pequeño tamaño, que los hace idóneos para ser amplificados por PCR. Fueron introducidos en el campo forense a principios de la década de los noventa por Edwards *et al* (1991), Polymeropoulos *et al* (1992a,b) y Tautz (1993), entre otros investigadores.

El polimorfismo de los microsatélites se debe a la variación en el tamaño de los alelos, como propusieron Fowler *et al* (1988) y confirmaron posteriormente Edwards *et al* (1991) y otros autores, pero también pueden presentar variaciones en la secuencia de la unidad de repetición. Urquhart *et al* (1994) clasificaron los STRs en simples, compuestos y complejos. Brinkmann *et al* (1996) proponen denominarlos STRs con microvariabilidad baja, intermedia y alta.

•**STRs simples o de baja microvariación.** Están formados por dos o más unidades de repetición contiguas e idénticas, tanto en longitud como en secuencia. La diferencia en tamaño entre los distintos alelos es de una unidad de repetición. La

presencia de alelos intermedios es excepcional. Son de fácil tipaje, pero tienen el inconveniente de su baja heterocigosidad.

Entre los STRs simples más usados están HUMTHO1 (Edwards *et al* 1991), HUMF13A1 (Polymeropoulos *et al* 1991a), HUMFES/FPS (Polymeropoulos 1991b), que contienen una unidad de repetición tetranucleotídica y HUMCD4, cuya unidad es pentanucleotídica y que presenta importantes diferencias poblacionales (Brinkmann *et al* 1995).

•**STRs compuestos o de microvariación intermedia.** Los sistemas comprenden dos o más unidades de repetición contiguas diferentes, que varían tanto en secuencia como en longitud. El de mayor interés forense es HUMVWA31/A (Kimpton *et al* 1992).

•**STRs Complejos o de alta microvariación.** Pueden contener varios bloques de unidades de repetición de longitud variable, con secuencias intermedias más o menos variables. Son frecuentes los alelos intermedios y las sustituciones simples de bases en algunas unidades, lo que ocasiona importantes variaciones tanto estructurales como en tamaño, produciéndose problemas de designación de alelos. Son los STRs más difíciles de tipar y se acompañan también de una tasa de mutación más alta. Se están utilizando con buenos resultados sistemas muy polimórficos como HUMACTBP2 (SE33) (Moos y Gallwitz 1983), HUMD21S11 (Sharma y Litt 1992) y HUMFIBRA/FGA (Mills *et al* 1992), que presentan una alta tasa de heterocigosidad.

Actualmente parece que se están acumulando evidencias a favor de que incluso los microsatélites más simples presentan algún grado de variabilidad interna en su estructura, por lo que la clasificación anterior sería cuestionable. Existen STRs relativamente simples y extremadamente variables al mismo tiempo, como D12S391 y D1S1656 (Lareu *et al* 1996, 1997), en los que la asociación directa entre variabilidad y complejidad no es sostenible.

Las grandes ventajas de los STRs son su estabilidad y la posibilidad de realizar PCR *multiplex*, amplificando varios loci simultáneamente (Kimpton *et al* 1993). Además su análisis se ha facilitado con el uso de fluorocromos y secuenciadores automáticos de ADN (Ziegle *et al* 1992). Sus aplicaciones son muy diversas, como construcción de mapas genéticos, estudios poblacionales, análisis de ligamiento en enfermedades genéticas, investigación del cáncer, estudios forenses, etc.

Los más abundantes y fáciles de amplificar son los que contienen 2 pb como unidad de repetición (Weber y May 1989) pero ciertos problemas técnicos, como la presencia de bandas “tartamudas” por *slippage* de la polimerasa durante la amplificación, hacen que se utilicen principalmente STRs con repeticiones de 4 pb, más apropiados para fines forenses.

Cuando se parte de ADN altamente degradado (Hagelberg *et al* 1991, Gill *et al* 1992a, Jeffreys *et al* 1992), el tipaje de SLPs e incluso AMPFLPs suele ser negativo, mientras que la amplificación de STRs proporciona generalmente resultados satisfactorios.

En los laboratorios forenses la tendencia actual viene marcada por la utilización de STRs simples e hipervariables (Lareu *et al* 1996, 1997), grandes *multiplexes* (Gill *et al* 1995) y *multiplexes* de STRs de pequeño tamaño, ya que la degradación del ADN es inversamente proporcional al tamaño de los fragmentos (Alvarez *et al* 1996).

2.2.3.- OTROS POLIMORFISMOS DE ADN

En el ADN mitocondrial (ADN_{mt}), la zona más variable se conoce como D-Loop y se han analizado dos regiones de gran interés en Genética de poblaciones (HV1 y HV2). La dotación mitocondrial es haploide y se hereda exclusivamente por vía materna, lo que facilita el análisis de divergencias de secuencias, aunque esta misma propiedad hace que este ADN no sea útil para el estudio biológico de la paternidad.

El ADN mt , mucho más abundante que el genómico, es ideal para el análisis de cabellos y pelos sin bulbo, muestras muy envejecidas como restos óseos de gran antigüedad y, en general, vestigios en los que no sea posible analizar ADN nuclear. La amplificación por PCR de regiones hipervariables del D-Loop y la posterior determinación de esta variabilidad por secuenciación es el método más empleado para la resolución de estos casos (Sullivan *et al* 1994, Gill *et al* 1994a).

El cromosoma Y, de herencia exclusivamente paterna, está suscitando un renovado interés tanto antropológico como médico-legal. Los desalentadores estudios iniciales (Casanova *et al* 1985, Lucotte y Ngo 1985, Jakubizka *et al* 1989, Malaspina 1990) no fueron óbice para que, merced a los avances tecnológicos, el estudio de nuevos polimorfismos de cromosoma Y se encuentre en una fase muy prometedora.

Actualmente se analizan fundamentalmente haplotipos de STRs, secuencia alfoide por análisis de heteroduplex, el minisatélite MSY1 por MVR-PCR, mutaciones puntuales y algunas duplicaciones, inserciones y deleciones.

Los polimorfismos de cromosoma Y, por ser el objeto de esta tesis, se comentarán con detalle más adelante.

2.3.- ESTANDARIZACIÓN DE LOS ANÁLISIS

Para que cualquier polimorfismo pueda ser utilizado en la práctica forense, es preciso conocer unos valores que expresen la frecuencia de aparición de esta variación en la población, es decir, conocer la frecuencia alélica y genotípica en la población implicada y comprobar que está en equilibrio Hardy-Weinberg en la población donde va a ser usado.

Las regiones hipervariables del ADN presentan un gran polimorfismo y un número elevado y no determinado de alelos, lo que unido a variaciones en los procedimientos técnicos empleados podría conllevar dificultades a la hora de interpretar los resultados obtenidos en diferentes laboratorios. Ya que es esencial la comparación de resultados en la práctica forense, es imprescindible que existan unas directrices aceptadas por todos, con lo que se facilitaría la creación de bases de datos comunes y la realización de contrapericias.

Para supervisar este espinoso tema, se constituyeron diversos grupos de trabajo, tanto nacionales como internacionales, con el cometido de estandarizar las técnicas de análisis y su interpretación y garantizar la fiabilidad y reproducibilidad de los resultados obtenidos mediante estrictos controles de calidad.

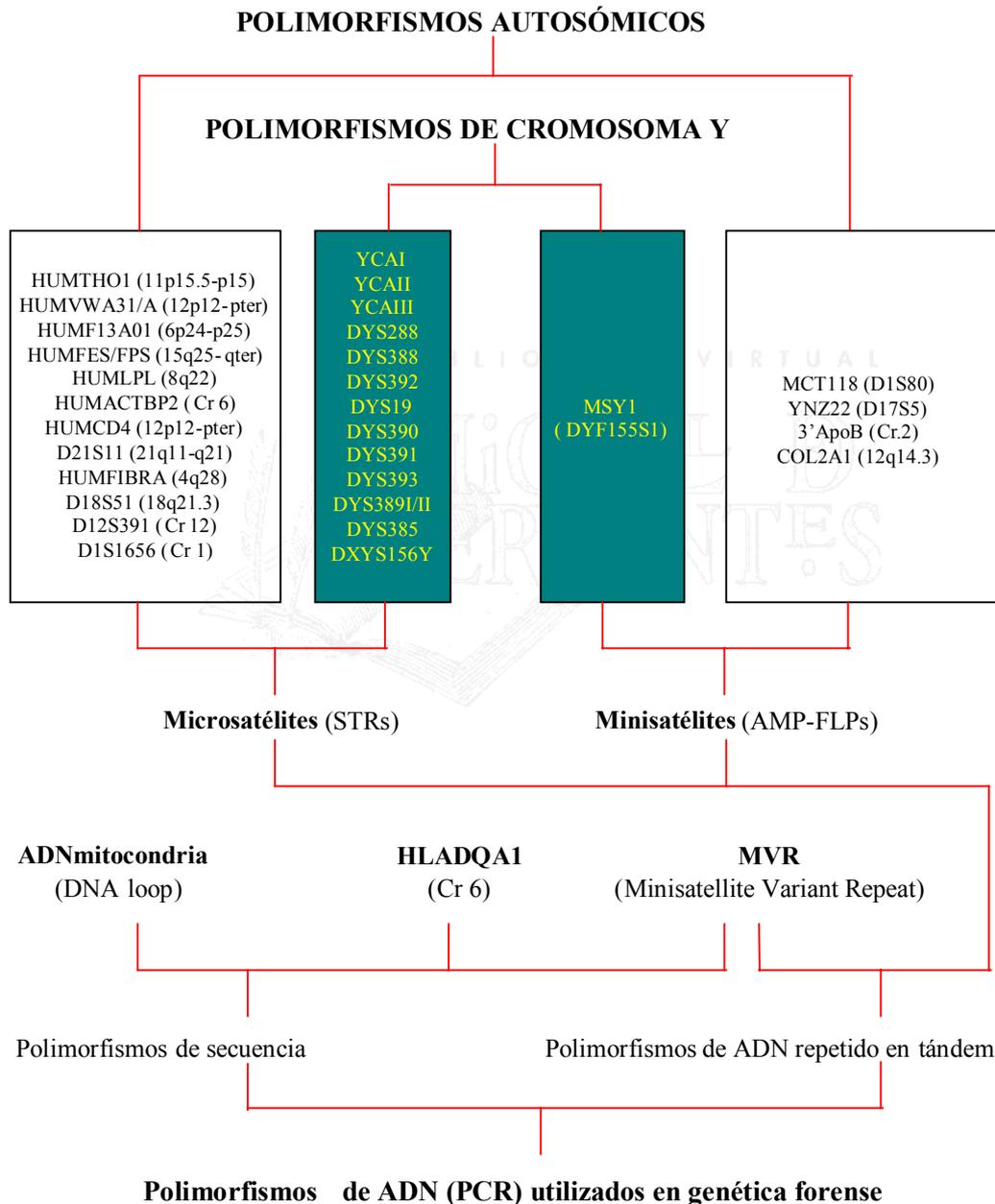
En Europa, en 1989, varios laboratorios de distintos países constituyen el grupo *European DNA Profiling group* (EDNAP) del que forma parte el Instituto de Medicina Legal de Santiago de Compostela. El primer ejercicio de estandarización (Schneider *et al* 1991) consistió en el análisis de una serie de muestras usando el enzima de restricción HinfI y la sonda YNH24, siguiendo cada laboratorio sus propios protocolos, con lo que se evidenció un intervalo de error del 10% en el tamaño de los fragmentos analizados. En el segundo ejercicio (Gill *et al* 1992b) todas las determinaciones se realizaron con un protocolo semejante, con lo que disminuyó enormemente el intervalo de error, demostrándose que la uniformización de protocolos puede permitir la comparación de resultados entre laboratorios. Posteriormente se estandarizaron STRs simples (Gill *et al* 1994b, Kimpton *et al* 1995, Andersen *et al* 1996) y STRs complejos (Gill *et al* 1997).

Otro grupo de estandarización es el grupo norteamericano *Scientific Working Group on DNA Analysis Methods* (SWGDM) que publica periódicamente sus recomendaciones (TWGDAM 1989, 1990, 1991, 1993, 1994, 1995). En Estados Unidos el uso de ADN con fines forenses ha sido específicamente regulado por el National Research Council (NRC 1992, 1997).

Representantes de la EDNAP y de la SWGDAM constituyen la *DNA Comission of the International Society for Forensic Genetics (ISFG)*, que también emite recomendaciones, generalmente anuales, sobre el uso de marcadores de ADN con fines forenses, incluyendo nomenclatura, estadística, gestión y control de calidad (ISFG 1989, 1992a,b, 1995, 1997).



Tabla 2. Algunos polimorfismos de ADN amplificables mediante PCR utilizados en Genética Forense



3.- CROMOSOMA Y

Dentro de los interrogantes que se plantean los seres humanos desde lejanas épocas, la curiosidad acerca de sus orígenes ocupó siempre un lugar muy relevante. Tradicionalmente, el pasado ha sido investigado por historiadores, arqueólogos, y paleontólogos, en tanto que hoy en día se tiende a buscar la evidencia indirecta de poblaciones humanas modernas a través de la lingüística y, cada vez en mayor medida, la biología molecular. Desde que ha sido factible el análisis de nuestro material genético con nuevas técnicas, el estudio del cromosoma Y, por sus especiales características, está despertando grandes expectativas al revelarse como una útil fuente de polimorfismos de ADN, especialmente en el ámbito de la evolución humana y las ciencias forenses.

3.1.- INTRODUCIÉNDONOS EN EL CROMOSOMA Y

Tras la fecundación, la mitad del ADN del cigoto procede de la madre y la otra mitad del padre. La mujer cede a su descendencia obligatoriamente un cromosoma X, el varón en cambio puede contribuir con un cromosoma X o Y determinando, por tanto, el sexo de su descendencia. Este cromosoma en los mamíferos es el responsable del desarrollo testicular, indicando la presencia de un gen responsable que codifica el *testis determining factor* (TDF) que hace que las gónadas indiferenciadas se transformen en testículos en una etapa precoz de la embriogénesis.

3.1.1.- SU NATURALEZA

Es uno de los cromosomas más pequeños del genoma humano, con un tamaño aproximado de 60 Mb (Morton 1991). Desde el punto de vista citológico, el cromosoma Y está formado por regiones de heterocromatina y eucromatina.

La región heterocromatínica se sitúa sobre el brazo largo (Yq) en posición distal. Varía considerablemente en tamaño entre individuos. Se compone de secuencias altamente repetitivas, como por ejemplo DYZ1 y DYZ2, presentando polimorfismos de longitud.

La región de eucromatina se localiza en el brazo corto (Yp), centrómero y en la zona proximal del brazo largo. Su tamaño es bastante constante en varones normales (alrededor de 30 Mb). Es la región de mayor interés genético. Aquí se encuentran algunos genes, así como secuencias que muestran homologías con el cromosoma X y los autosomas, y también secuencias repetitivas específicas del cromosoma Y (DYZ3, DYZ4, DYZ5).

Aunque el grueso de su volumen se considera genéticamente inerte, en el cromosoma Y se pueden diferenciar tres regiones interesantes desde el punto de vista genético: dos pseudoautosómicas (PAR1 y PAR2) y una específica del cromosoma Y (Figura 1).

3.1.2.- LAS REGIONES PSEUDOAUTOSÓMICAS

A pesar de su distinta naturaleza morfológica, los cromosomas X e Y pueden aparearse durante la meiosis e intercambiar información. Este fenómeno ocurre en ciertas pequeñas regiones que contienen secuencias homólogas presentes en ambos cromosomas, denominadas regiones pseudoautosómicas porque en ellas las secuencias de ADN no muestran una herencia ligada al sexo en sentido estricto. En los cromosomas sexuales humanos hay dos de estas regiones, denominadas PAR1 y PAR2, situadas respectivamente en los extremos terminales del brazo corto y brazo largo de dichos cromosomas.

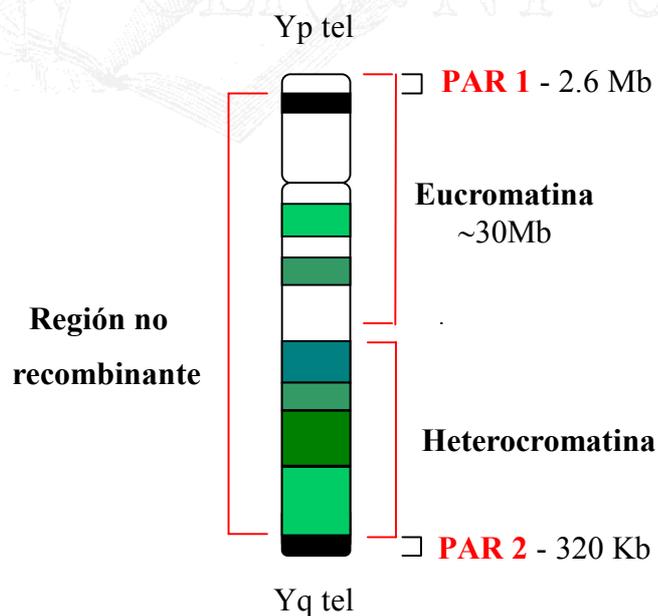
La región pseudoautosómica mayor (PAR1) del cromosoma Y tiene un tamaño de aproximadamente 2,6 Mb. La recombinación a nivel de PAR1 es necesaria para una segregación normal de los cromosomas X e Y en la meiosis (Ellis y Goodfellow

1989a) y por tanto recombina siempre. Pueden existir alteraciones, como en el síndrome de Klinefelter, cuyo cariotipo 47,XXY es producto de la no-disyunción, no produciéndose recombinación a nivel de PAR1 en la gran mayoría de casos (Hassold *et al* 1991).

La región pseudoautosómica menor (PAR2) del cromosoma Y mide sobre 320 Kb y no siempre participa en procesos de recombinación. Además, la actividad de recombinación de PAR2 no puede sustituir a la de PAR1 (Freije *et al* 1992), no siendo, por tanto, ni necesaria ni suficiente para el éxito de la meiosis masculina.

La zona límite de recombinación entre PAR1 y la región específica del cromosoma Y ha sido clonada y secuenciada (Ellis *et al* 1989b) y se conoce como frontera pseudoautosómica.

Figura 1. Idiograma del cromosoma Y. (Modificado de Jobling *et al* 1997)



(PAR1 y PAR2: Regiones pseudoautosómicas)

3.1.2.1.- LOS GENES DE LAS REGIONES PSEUDOAUTOSÓMICAS

Se espera que los genes localizados en estas regiones escapen al mecanismo de inactivación, para asegurar de este modo un reparto equivalente de genes entre los sexos. De los genes que se citan a continuación (Tabla 3), todos tienen homólogo en X.

- El gen **MIC2Y** (PAR1, Yp) ha sido el primer gen pseudoautosómico descrito en el hombre (Goodfellow *et al* 1983). Codifica una proteína estructural del antígeno de superficie celular CD99. Su homólogo en X es **MIC2**.

- El gen **ASMTY** (PAR1, Yp) codifica una enzima que cataliza la última etapa de la síntesis de la hormona melatonina (Yi *et al* 1993). Tiene homólogo en X (**ASMT**).

- El gen **XE7Y** se localiza proximalmente respecto al anterior. Fue aislado en 1992 (Ellison *et al* 1992) y codifica el antígeno XE7, de función desconocida, pero que presenta homología con el CD99. Tiene como homólogo al gen **XE7**.

- **IL3RAY** (PAR1, Yp) es el gen del receptor de la interleucina-3. Se localiza proximal a CSF2RAY (Kremer *et al* 1993). **IL3RA** es su homólogo en X.

- **CSF2RAY** (PAR1, Yp) es el gen más distal de los descritos en PAR1 hasta el momento. Codifica el receptor del factor estimulante de las colonias de granulocitos-macrófagos (GM-CSF) (Goug *et al* 1990). Su homólogo en X es **CSF2RA**.

También han sido identificados algunos genes en la región PAR2, como el gen **SYBL1Y** (*synaptobrevin-like*) que está inactivado en el cromosoma Y (D'Esposito *et al* 1996) y el gen del receptor de la interleucina-9, **IL9RY** (Kermouni *et al* 1995).

3.1.3.- LA REGION ESPECÍFICA DEL CROMOSOMA Y Y SUS GENES

Esta región no recombina con ninguno de los cromosomas, por lo cual sus loci se transmiten inalteradamente por línea paterna de generación en generación (en ausencia de mutaciones). Presenta algunos genes funcionales, de gran importancia para el desarrollo sexual del hombre (Tabla 3).

- El **SMCY**, gen que codifica el antígeno menor de histocompatibilidad H-Y, fue el primer gen identificado en la región específica del cromosoma Y (Wachtel *et al* 1975). **SMCX** es su homólogo en el cromosoma X.

- El gen **ZFY** codifica una proteína que presenta un motivo en dedo de zinc (Page *et al* 1987) cuya función no es conocida por el momento. Su homólogo en el cromosoma X es **ZFX**.

- El gen **SRY** ha sido aislado en el brazo corto del cromosoma Y, en la llamada *Sex-determining Region of the Y chromosome*, a 5kb en dirección proximal a PAR1 (Sinclair *et al* 1990). Presenta un solo exón y carece de homólogo en X. Diversos estudios han demostrado que se expresa únicamente a nivel testicular. En ratones transgénicos hembras, la inserción de un fragmento de ADN conteniendo el homólogo murino de SRY induce la formación de testículos (Koopman 1991). En humanos, las mutaciones en este gen afectan total o parcialmente al desarrollo testicular.

- La familia de genes **YRRM** (*Y chromosome RNA Recognition Motif*) se cree que podría intervenir en la regulación de la espermatogénesis. Como el gen anterior, este grupo carece de homólogo en X.

- El **RPS4Y**, localizado entre SRY y ZFY, es un gen que codifica la proteína ribosómica S4. Posee un homólogo en X que escapa a la inactivación (Fisher *et al* 1990).

• El **AMGY**, gen que codifica un producto *amelogenine-like*, también tiene un homólogo en X (Nakahori *et al* 1991), que es responsable de la formación normal del esmalte dentario.

Tabla 3. Algunos genes conocidos en el cromosoma Y humano

SÍMBOLO	NOMBRE/FUNCIÓN	LOCALIZACIÓN	HOMÓLOGO EN X
CSF2RAY	Receptor del GM-CSF	PAR1, Yp	CSF2RA
IL3RAY	Receptor IL-3	PAR1, Yp	IL3RA
ASMTY	Codifica Acetil serotonin N-metil transferasa	PAR1, Yp	ASMT
XE7Y	Codifica antígeno XE7	PAR1, Yp	XE7
MIC2Y	Codifica antígeno de superficie CD99	PAR1, Yp	MIC2
XGRY	Codifica regulador de grupo sanguíneo	PAR1, Yp	XGR
XGY	Codifica un antígeno de grupo sanguíneo	PAR1, Yp	XG
ZFY	Codifica proteína en dedo de zinc	Distal Yp11.1	ZFX
RSP4Y	Codifica proteína ribosómica S4	Distal Yp11.1	RSP4X
AMGY	Codifica amelogenina	Proximal Yp11.1	AMGX
SMCY	Codifica antígeno H-Y	Proximal Yq	SMCX
SRY	Codifica el TDF	Próximo a PAR1, Yp	No
Familia YRRM	¿Regulación de la espermatogénesis?	Yq11.23	No

3.1.4.- ALGUNAS PECULIARIDADES DEL CROMOSOMA Y. ANALOGÍA CON EL ADN MITOCONDRIAL

El cromosoma Y presenta una compleja secuencia, con grandes fragmentos de ADN no codificante y numerosas familias de secuencias repetitivas dispersas o en tándem. El hecho de que el cromosoma Y sea extraordinariamente pobre en genes (Lahn *et al* 1997) hace pensar en una baja presión selectiva en las zonas no codificantes del cromosoma, por lo que podría acumular más mutaciones y secuencias *junk* que los otros cromosomas. Además, el que permanezca más tiempo en la línea germinal, y que la formación de esperma conlleve un alto número de divisiones celulares, también podría favorecer la aparición de mutaciones (Weber y Wong 1993, Brinkmann *et al* 1998).

Paradójicamente, la frecuencia de polimorfismos en regiones no codificantes del cromosoma Y es mucho menor que en el resto del genoma nuclear. Hay varias posibles explicaciones que intentan justificar este bajo nivel de polimorfismo observado en el cromosoma Y respecto a otros cromosomas (Jobling y Tyler-Smith 1995).

Una de las proposiciones se basa en una razón aritmética, ya que la población efectiva de cromosomas Y (número de cromosomas que pasan a la descendencia) en un momento determinado es claramente inferior a la de otros cromosomas (por cada cromosoma Y están presentes 3 cromosomas X y 4 de cada autosoma), lo que refleja el hecho de que existe correlación entre diversidad y tamaño de la población cromosómica (Kimura 1983). Además, actividades históricamente ligadas al género masculino como guerras y otras expansiones, así como ciertas estructuras sociales o religiosas habituales en el pasado, todavía reducen más el número efectivo de cromosomas Y. Todo esto hace que procesos como la deriva genética hayan tenido un efecto mucho más patente sobre el cromosoma Y que sobre cualquier otro, lo que acrecienta su interés para estudios de evolución humana (Roewer *et al* 1996a, Underhill *et al* 1996).

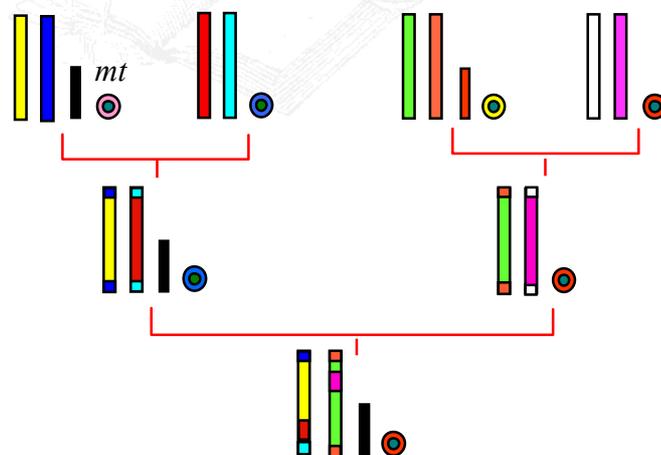
Otro de los argumentos hace referencia a la falta de recombinación del cromosoma Y. Debido a este hecho, si el proceso de selección actuase sobre alguno de los genes del cromosoma, se produciría una selección global de todo el haplotipo que portase dicho cromosoma por efecto “*hitch-hiking*”. El haplotipo en concreto aumentaría mucho de frecuencia y podría llegar a fijarse, lo que provocaría una bajada de los niveles de polimorfismo global del cromosoma Y en la población. De todas maneras, para que un barrido selectivo (“*selective sweep*”) de este tipo se pueda detectar en el presente, tendría que haber sucedido bastante recientemente, ya que si nos remontamos a los comienzos de la historia evolutiva de la especie, la mutación ya habría tenido tiempo de regenerar los niveles de polimorfismo.

El análisis de las primeras secuencias de ADN_{mt} (Anderson *et al* 1981) tuvo una gran repercusión en el estudio de las relaciones filogenéticas entre los diferentes

grupos humanos actuales. El hecho de que una molécula se herede exclusivamente por vía materna y sin sufrir recombinación, permite preservar la información de los sucesos mutacionales que han surgido en un linaje mitocondrial determinado en las generaciones pasadas. Pero el ADN mt no es la única molécula que tiene un gran potencial filogenético. El cromosoma Y (Figura 2), portador de los genes determinantes del sexo masculino, también se hereda uniparentalmente (en este caso por vía paterna) y no padece fenómenos de recombinación en prácticamente toda su extensión (Wolf *et al* 1992).

Figura 2. Esquema de la transmisión diferencial de un cromosoma Y, una molécula de ADN mt , y un par de autosomas*

Herencia de autosomas, cromosoma Y y ADN mt .



(*Las secuencias autosómicas recombinan en cada generación, mientras que la región específica del cromosoma Y y el ADN mt no lo hacen. Por consiguiente, el material autosómico de un individuo procede de múltiples ancestros, pero tanto el cromosoma Y (excluyendo las regiones pseudoautosómicas) como el ADN mt provienen de un único antepasado).

Esta falta de recombinación hace que el cromosoma se comporte como un gran fragmento único de ADN, portador de un haplotipo que se transmite intacto de padres a hijos, a no ser que actúen fenómenos de mutación.

El ADN_{mt} presenta una tasa de sustitución de bases diez veces superior a la del ADN nuclear, lo que significa que existe un gran número de secuencias diferentes de ADN_{mt} en la población, por tanto, una gran diversidad. Al mismo tiempo, esta fuerte tasa de mutación también es causa de inconvenientes: mutaciones puntuales se pueden producir varias veces en el mismo punto o aparecer independientemente en diferentes linajes, haciendo imposible determinar el estado ancestral. Por el contrario, algunas mutaciones puntuales en el cromosoma Y se pueden considerar como acontecimientos únicos que son específicos de un linaje, ya que cuando un grupo de cromosomas Y portan una mutación puntual dada es muy probable que tengan el mismo origen.

A modo de resumen, las principales características de ambos tipos de ADN se muestran a continuación (Tabla 4).

Tabla 4. Analogías/Diferencias entre el ADN de cromosoma Y y el ADN_{mt}

Analogías/Diferencias	ADN Cromosoma Y	ADN mitocondrial
Localización en la célula	Núcleo (ADN genómico)	Mitocondrias (Extranuclear)
Modo de herencia	Monoparental Masculina	Monoparental Femenina
Tipo de Transmisión	En bloque	En bloque
Recombinación	Generalmente, no*	No
Tipo de molécula de ADN	Lineal	Circular
Tamaño molecular aprox.	60 Mb	16.5 Kb
Secuencia conocida	No en su totalidad	Si
Densidad de genes	Actualmente escasa	Alta
ADN no Codificante	Abundante	Escaso
Tasa de mutación	Baja	Alta
Polimorfismos predominantes	Variables**	De secuencia

(* Sólo en las regiones pseudoautosómicas).

(**Pequeños y grandes reordenamientos como inserciones, deleciones, duplicaciones e inversiones. Repeticiones en tándem en ADN satélite, minisatélite y microsatélite).

Las características que hemos comentado han hecho que las aportaciones del cromosoma Y, tanto al campo antropológico (Hammer 1994, Jobling y Tyler-Smith 1995) como al médico-legal (Roewer y Eppelen 1992a, Gomolka *et al* 1994), sean ampliamente reconocidas.

3.2.- POLIMORFISMOS EN EL CROMOSOMA Y

Tras las primeras descripciones de marcadores polimórficos en el cromosoma Y (Casanova *et al* 1985, Lucotte y Ngo 1985), durante casi una década no se van a encontrar nuevos polimorfismos. Inicialmente, una búsqueda sistemática de polimorfismos convencionales (RFLPs) pone de manifiesto su relativa escasez (Jakubicza *et al* 1989, Malaspina *et al* 1990, Spurdle y Jenkins 1992a). A éstos se añaden los detectados por electroforesis en gel de campo pulsante (PFGE) (Oakey y Tyler-Smith 1990, Jobling 1994a) y, posteriormente, un nuevo tipo de polimorfismos detectables por PCR (Roewer *et al* 1992a, Seielstad *et al* 1994). Otra estrategia empleada, la secuenciación de amplias regiones no codificantes del cromosoma Y de varios individuos, apuntó un acusado monomorfismo (Dorit *et al* 1995, Hammer 1995a).

3.2.1.- TIPOS DE POLIMORFISMOS DEL CROMOSOMA Y

Actualmente se conocen diversos tipos de polimorfismos diferentes localizados en este cromosoma (Figura 3), detectables por métodos basados en PCR muchos de ellos, que incluyen entre otros:

- Duplicaciones/deleciones (Casanova *et al* 1985, Jobling *et al* 1996).
- Reordenamientos complejos (Lucotte y Ngo 1985).
- Mutaciones puntuales de cambios de base (Underhill *et al* 1996, 1997).
- Inserciones Alu (Hammer 1994).
- Polimorfismo en ADN repetido en tándem (Roewer *et al* 1992b, Jobling *et al* 1994b).

La mayoría de los polimorfismos de cromosoma Y (Tabla 5) son difíciles de interpretar, pues todavía los mecanismos genéticos que los originan no están bien caracterizados.

Duplicación/Delección

Este tipo de polimorfismo se identifica cuando una variación en el tamaño de los fragmentos es detectada por enzimas de restricción diferentes.

- **12f2 (DYS11)**. Fue el primer RLFP descrito como específico de cromosoma Y (Casanova *et al* 1985). Lo detectan *EcoRI* (5.2 y 3.2 kb) y *TaqI* (10 y 8 kb). Aparentemente es un evento único.

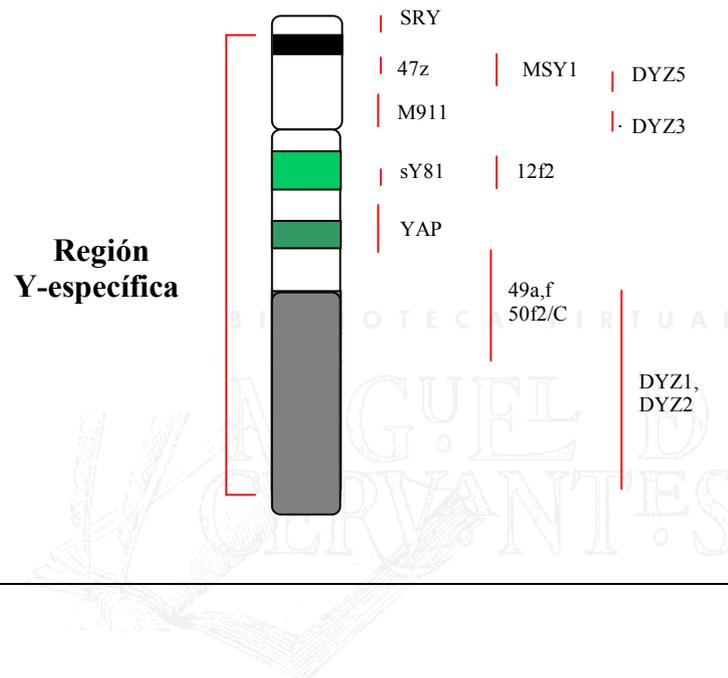
- **50f2(DYS7/C)**. Con *EcoRI* se detectan varios fragmentos (50f2/A, B, C, D y E). La delección de uno de los fragmentos (locus C localizado en Yq) origina un polimorfismo (Disteche *et al* 1986). Este polimorfismo comprende tres categorías: grandes deleciones, pequeñas deleciones y duplicación. Hoy es posible detectar la delección 50f2/C por PCR -producto de 196 pb presente o ausente- (Jobling *et al* 1995). Este locus (DYF155S2) dista 4 kb de 50f2/C (DYS7/C). Los *primers* empleados en la determinación también amplifican el minisatélite MSY1 (DYF155S1).

Reordenamientos complejos

- **49a/f (DYS1)**. Fue descrito inicialmente por Lucotte y Ngo (1985). Con ayuda de *TaqI* detecta una secuencia débilmente repetitiva. También se utilizan otras enzimas (*PvuII*, *BglIII*, *HindIII*, *PstI*, etc.) pudiendo conformarse más de 100 haplotipos diferentes. Las bandas detectadas pueden variar tanto en número (presencia o ausencia) como en tamaño (Ngo *et al* 1986, Spurdle y Jenkins 1992b).

Es el RFLP con mayor variabilidad en el cromosoma Y, lo que dificulta la interpretación de este polimorfismo.

Figura 3. Localización de algunos loci polimórficos en el cromosoma Y



Inserción Alu (YAP)

El YAP (DYS287) o *Y Alu Polimorphism*, es el resultado de una inserción reciente de un miembro de la familia Alu en el brazo largo, Yq11, del cromosoma Y (Hammer 1994). La variación detectada en el tamaño de los fragmentos es idéntica y se identifica con más de una enzima (*TaqI* o *EcoRV*). Por técnicas de PCR, cuando ha tenido lugar la inserción Alu (305 pb) se detecta un producto de 455 pb (YAP+) y si ésta no se ha producido, de 150 pb (YAP-) (Hammer y Horai 1995b).

La secuenciación ha mostrado que en el cromosoma Y el elemento Alu se inserta en la misma posición en diferentes individuos de distintas poblaciones (Hammer 1994) y se considera que la inserción (YAP+) tiene un único origen. El hecho de que esté ausente en la región homóloga del cromosoma Y en chimpancés y gorilas

sugiere que la inserción se ha producido tras la divergencia entre hombres y grandes simios.

La presencia del elemento YAP es muy abundante en la población Sub-Sahariana (sobre todo en bantúes y africanos del Oeste), seguida de la del Norte de África. En Europa es poco frecuente, y nula en la mayoría de las poblaciones de Oceanía y Asia (excepto en Japón, donde muestra una frecuencia bastante elevada). Estos resultados podrían indicar que la distribución de cromosomas Y portando el YAP es posible que ocurriese por migración de varones *Homo sapiens* desde África hacia Europa y Asia (Hammer 1994).

El marcador YAP es particularmente útil en estudios sobre el origen y migración de poblaciones humanas porque representa probablemente un evento único con un estado ancestral conocido (alelo YAP-), aunque es necesario confirmar estas suposiciones con más estudios sobre otras poblaciones y con más marcadores de cromosoma Y.

Mutaciones puntuales

Las sustituciones de una base (Hamer y Horai 1995b, Whitfield *et al* 1995, Underhill *et al* 1996) representan habitualmente un evento mutacional único, probablemente abundante (Jobling *et al* 1997), pudiendo definir un grupo monofilético. La misma sustitución en el mismo punto va a ser rara, excepto que haya ocurrido hace mucho tiempo. Se pueden detectar con una sola enzima o por secuenciación. Estos polimorfismos son poco variables, su tasa de mutación es baja y presentan sólo dos estados alélicos, por lo que son muy adecuados para estudios de evolución humana a lo largo del tiempo.

- **47z (DXYS5Y)**. Es un RFLP detectable con la enzima *StuI* (Nakahori *et al* 1989). No es un marcador específico de cromosoma Y puesto que la sonda detecta igualmente un locus en el cromosoma X.

- **92R7** (No tiene locus asignado). El polimorfismo es analizable por PCR. Se identifica una secuencia moderadamente repetitiva. La sonda detecta siete bandas tras digestión con *HindIII* (Mathias *et al* 1994).

- **sY81 (DYS271)**. El polimorfismo convencional sY81, detectado en población Sub-Sahariana YAP+, presenta como sustitución de base la transición A→G (Seielstad *et al* 1994) que origina la pérdida de un lugar de restricción para la enzima *NlaIII*. Este polimorfismo se detecta por PCR-RFLP (Seielstad *et al* 1994), definiéndose dos alelos, A (alelo 0) y G (alelo 1). Es manifiesta la asociación entre el alelo G en sY81 y la inserción YAP. Combinando estos dos polimorfismos se ha intentado inferir el momento en que ocurren estas dos mutaciones. El estudio haplotípico muestra la presencia de un alelo A con YAP- en primates no humanos. Esto hizo sugerir que primero sobrevino la inserción Alu sobre un cromosoma con un alelo A y que posteriormente tuvo lugar la transición A→G sobre un cromosoma YAP+. La transición ha debido producirse una sola vez y, probablemente, en África.

- **SRY-1532** (Jobling y Tyler-Smith 1995, Whittfield *et al* 1995, Kwok *et al* 1996). Es una mutación puntual en la secuencia del gen principal que presenta la transición A→G en la posición 10.831. Es poco frecuente en Europa pero común en el subcontinente indio. Por PCR se obtiene un producto de 167 pb que se analiza por SSCP o usando la enzima *DraIII*. Otras mutaciones puntuales en este gen también detectables por PCR son **SRY-8299**, polimorfismo antes conocido como SRY 4.064 (Whitfield *et al* 1995), que presenta la sustitución de base G→A y **SRY-2627**, denominado primeramente SRY-2628 (Veitia *et al* 1997), que consiste en la sustitución C→T y que es abundante en la Península Ibérica.

- **DYS199** La transición C→T (Underhill *et al* 1996) se detecta con *MunI*.

- Otras mutaciones puntuales como **M9** (Underhill *et al* 1997), **Tat** (Zerjal *et al* 1997), etc.

Tabla 5. Algunos ejemplos de polimorfismos de cromosoma Y

TIPO	LOCUS	NOMBRE	DETECCION	REFERENCIA
Sustitución de una base	DXYS5Y	47z	<i>StuI</i> / hibridación	Nakahori <i>et al</i> (1989)
	–	92R7	<i>HindIII</i> /hibridación PCR	Mathias <i>et al</i> (1994)
	DYS271	sY81	<i>NlaIII</i> /PCR	Seielstad <i>et al</i> (1994)
	SRY	SRY-1532	<i>DraIII</i> /PCR	Witfield <i>et al</i> (1995) Kwok y Hawkins (1996)
	SRY	SRY-8299	<i>BsrBI</i> /PCR	Witfield <i>et al</i> (1995)
	SRY	SRY-2627	HKAI/PCR	Veitia <i>et al</i> (1997)
	DYS199	–	<i>MunI</i>	Underhill <i>et al</i> (1996)
	–	M9	<i>HinfI</i> /PCR	Underhill <i>et al</i> (1997)
	–	Tat	<i>NlaIII</i> /PCR	Zerjal <i>et al</i> (1997)
Duplicación o delección	DYS11	12f2	<i>TaqI</i> o <i>EcoRI</i> / hibridación	Casanova <i>et al</i> (1985)
	DYS7/C	50f2	<i>EcoRI</i> /hibridación	Disteche <i>et al</i> (1986)
	DYF155S2	50f2/C	PCR	Jobling <i>et al</i> (1995)
Inserción	DYS287	YAP	<i>TaqI</i> o <i>EcoRV</i> / hibridación	Hammer (1994)
			PCR	Hammer y Horai (1995)
Reordenamientos complejos	DYSS1	49a/f	<i>TaqI</i> y otras/ hibridación	Lucotte y Ngo (1985)
Repeticiones en tándem				
En ADN satélite	DYZ3	Y α 1	<i>BglII</i> /PFGE-hibridación	Oakey <i>et al</i> (1990)
			<i>HindIII</i> /PCR	Santos <i>et al</i> (1995)
En ADN minisatélite	DYF155S1	MSY1	MVR-PCR	Jobling <i>et al</i> (1994b)
En ADN microsatélite	DYS288	–	PCR	GDB
	–	YCAI	PCR	Mathias <i>et al</i> (1994)
	–	YCAII	PCR	Mathias <i>et al</i> (1994)
	–	YCAIII	PCR	Mathias <i>et al</i> (1994)
	DYS388	–	PCR	GDB
	DYS392	–	PCR	GDB
	DYS19	27H39LR	PCR	Roewer <i>et al</i> (1992a)
	DYS390	–	PCR	GDB
	DYS391	–	PCR	GDB
	DYS393	–	PCR	GDB
	DYS385	–	PCR	GDB
	DYS389	–	PCR	GDB
	DXYS156Y	–	PCR	Chen <i>et al</i> (1994)

Polimorfismo en ADN repetido en tándem

- **En ADN Satélite.** El locus DYZ3 (Satélite Alfoide), de localización centromérica, está formado por subunidades de aproximadamente 170 pb organizadas en unidades de unas 6 kb a lo largo de una región de centenares de kb. Aquí se produce un polimorfismo de longitud ($Y\alpha I$) detectable con diferentes enzimas de restricción (Oakey *et al* 1990) como *BglIII* (*array length polymorphism*), *AvaII* o *EcoO1091*, mediante técnicas electroforéticas en gel de campo pulsante (PFGE) que tienen como principales inconvenientes la lentitud en la determinación y la necesidad de usar sangre fresca o líneas celulares.

Afortunadamente es posible tipar el polimorfismo alfoide por PCR (Santos *et al* 1995), obteniéndose un producto de 285 pb que se corta parcialmente con *HindIII* (lo que indica unidad de 6 kb presente). Se generan fragmentos de 237 y 48 pb, analizándose por heterodúplex la divergencia de subunidades alfoides de la secuencia repetitiva, que se localiza en los extremos del centrómero. Las bases moleculares del polimorfismo son complejas y se basan en la coamplificación de *repeats* alfoides que se diferencian en tamaño y/o secuencia (Santos *et al* 1996). La mayoría de los individuos poseen un locus conservado de 281 pb con una deleción de 4 pb en la rama izquierda de la secuencia alfoide, próxima a Yp, y un número variable de loci (entre 1 y 5) de 285 pb cada uno en la rama derecha, próxima a Yq, cuyas secuencias suelen diferir en una sola base. Combinándolos, se pueden formar desde 1 par hasta 5 pares de heterodúplex por individuo, que asumen diferentes conformaciones según su secuencia y pueden resolverse en geles de poliacrilamida no desnaturizantes. Es un polimorfismo común en Europa, pero no ocurre en un grupo monofilético (Jobling y Tyler-Smith 1995), lo que constituye una limitación para estudios de evolución humana.

También existen otros loci polimórficos como DYZ1/*HaeIII* en el Satélite III y DYZ2/*MspI* en el Satélite I (Mathias *et al* 1994), ambos localizados en el brazo largo del cromosoma Y y que están formados por secuencias altamente repetitivas. En el

brazo corto, DYZ4 y DYZ5 (Y190) (Oakey *et al* 1990) representan loci con un bajo grado de bloques de secuencias repetitivas.

• **En ADN Minisatélite**

Mientras que en la región pseudoautosómica PAR1 se conocen numerosos minisatélites, el único minisatélite descrito hasta la fecha en la región no recombinante es el MSY1 (locus DYF155S1), localizado en brazo corto del cromosoma Y (Jobling *et al* 1994b). Las unidades de repetición, ricas en AT, tienen 25 pb y existen cuatro variedades identificadas por la secuenciación directa de los extremos de los alelos. Los alelos tienen entre 60 y 100 unidades de repetición.

El estudio de MSY1 por medio MVR-PCR (Jobling *et al* 1994b) ha demostrado que este minisatélite posee un alto grado de diversidad alélica y un alto poder de discriminación. Cuando se comparan códigos MSY1 con grupos de haplotipos, se observa una cierta correlación directa entre mayor o menor homogeneidad del grupo y estructuras MSY1. Algunas de las unidades de repetición parecen ser altamente específicas de poblaciones (Jobling *et al* 1994b, Bouzekri *et al* 1995).

El análisis de minisatélites específicos de cromosoma Y, por su polimorfismo, puede resultar muy interesante para estudiar linajes parentales, así como la tasa y proceso de mutación debido a la ausencia de intercambios interalélicos, y, finalmente, como marcadores específicos en procesos de identificación masculina en el campo forense (Jobling *et al* 1994b).

• **En ADN Microsatélite**

Actualmente los microsatélites constituyen una de las clases más utilizadas de marcadores de cromosoma Y (Roewer *et al* 1992a,b; Mathias *et al* 1994, Kayser *et al* 1997, de Knijff *et al* 1997). Son polimorfismos muy variables dentro de una población y analizables por PCR, lo que les hace ser de gran utilidad en genética forense. La facultad de detectar y discriminar ADN de varón hacen de los STRs del

cromosoma Y un eficaz complemento de los establecidos sistemas autosómicos basados también en PCR.

A continuación describiremos algunos de estos sistemas:

3.2.2.- STRs DEL CROMOSOMA Y ANALIZABLES MEDIANTE PCR.

Los STRs del cromosoma Y descritos presentan unidades de repetición que comprenden de dos a cinco nucleótidos. En la Figura 4 se muestra la ubicación de algunos de ellos, de los que haremos una breve descripción:

• *Dinucleótidos*

DYS288 (locus GDB-ID G00-139-631). Este microsatélite de cromosoma Y posee como unidad de repetición el dinucleótido CA y presenta un único locus polimórfico.

YCAI (Mathias *et al* 1994). Su unidad de repetición es el dinucleótido CA Se detectan dos loci polimórficos.

YCAII (Mathias *et al* 1994). También presenta el repeat CA y dos loci polimórficos.

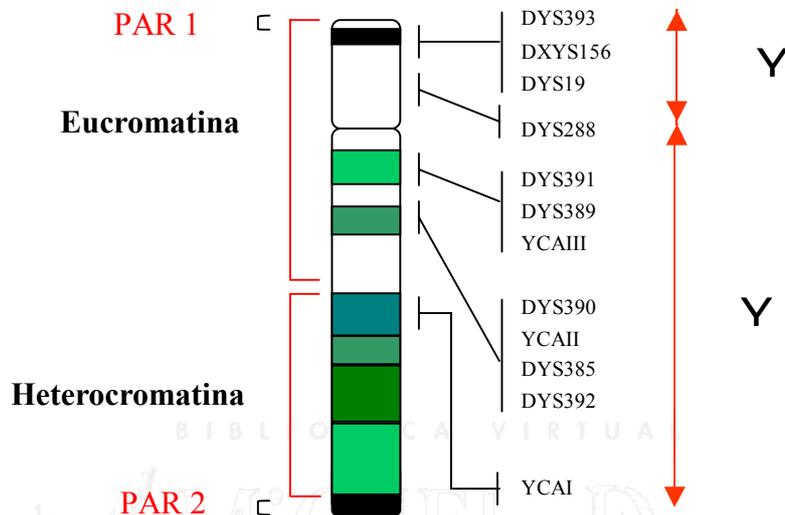
YCAIII (Mathias *et al* 1994). Comparte con los anteriores la misma unidad de repetición, siendo dos los loci polimórficos detectados para el cromosoma Y.

• *Trinucleótidos*

DYS388 (locus GDB-ID G00-365-729). Este STR tiene como unidad de repetición el trinucleótido ATA. Tiene un sólo locus polimórfico.

DYS392 (locus GDB-ID G00-456-509). Se caracteriza por el repeat ATT y es poseedor de un único locus polimórfico.

Figura 4. Esquema de la localización de algunos STRs descritos en el cromosoma Y



• **Tetranucleótidos**

DYS19 (locus GDB-ID G00-121-409). Este STR, primeramente denominado polimorfismo Y27H39, fue el primero que descrito en el cromosoma Y (Roewer *et al* 1992a). La unidad de repetición es el tetranucleótido GATA. Presenta un locus polimórfico.

DYS390 (locus GDB-ID G00-366-115). La unidad de repetición es CTAT y tiene solo un locus polimórfico.

DYS391 (locus GDB-ID G00-366-118). Este sistema tiene como motivo de repetición el tetrámero CTAT y un único locus polimórfico.

DYS393 (locus GDB-ID G00-456-649). La unidad repetitiva básica es el tetranucleótido GATA. Existe un sólo locus polimórfico.

DYS389I/II (locus GDB-ID G00-366-108). La unidad de repetición es tetranucleotídica (CTG/AT) y presenta 2 loci polimórficos.

DYS385 (locus GDB-ID G00-316-257) También en este sistema están implicados dos loci. La unidad de repetición es GAAA.

- **Pentanucleótidos**

DXYS156Y (Chen *et al* 1994). El motivo repetitivo es el pentanucleótido TAAAA. No es un STR específico de cromosoma Y, ya que también se amplifica un locus homólogo en X (DXYS156X).

4.- CONTRIBUCIÓN DE LOS POLIMORFISMOS DE ADN AL CAMPO MEDICO-LEGAL Y DE LOS POLIMORFISMOS DEL CROMOSOMA Y EN PARTICULAR

Desde que Jeffreys y colaboradores consiguieron el reconocimiento de los polimorfismos de ADN para usos forenses, al resolver satisfactoriamente un problema de inmigración en el Reino Unido empleando sondas multilocus (Jeffreys *et al* 1985c), los polimorfismos de ADN, y los de cromosoma Y cada vez en mayor medida, siguen contribuyendo con nuevas aportaciones a ampliar el horizonte médico-legal en las siguientes áreas:

4.1.- INVESTIGACIÓN BIOLÓGICA DE LA PATERNIDAD

La utilidad del análisis de los polimorfismos de ADN es incuestionable, tanto en casos rutinarios, en los que se dispone de todas las muestras necesarias, como en casos especiales de difícil solución por métodos clásicos, por ejemplo, aquellos en los que se carece de muestra perteneciente al presunto padre (Odelberg *et al* 1988), análisis de líquido amniótico para determinación de paternidad en fetos (Ishiyama *et al* 1988), paternidades a realizar sobre restos óseos, en ausencia de madre, con familiares en ausencia de ambos progenitores, etc.

A partir de la introducción de la PCR podemos estudiar diversos tipos de polimorfismos (AMP-FLPs, STRs, polimorfismos de secuencia), que nos van a

permitir, en general, realizar investigaciones biológicas de paternidad con mayor rapidez y con menor coste relativo.

Los polimorfismos de cromosoma Y (STRs) pueden proporcionar información adicional en casos de investigación biológica de paternidad en hijos varones cuando, por fallecimiento del progenitor u otras causas, no estén disponibles datos genéticos relativos al presunto padre (Pena y Chakraborty 1994). Esta falta de datos puede suplirse con los obtenidos de familiares varones por línea paterna para comprobar si su haplotipo coincide con el del supuesto hijo.

También pueden aportar ayuda en la confirmación de exclusiones. Trabetti *et al* (1996) y Weichhold *et al* (1996) han revisado paternidades con exclusión probada y la han corroborado, en un elevado porcentaje de ellas, utilizando únicamente el sistema DYS19.

La inclusión basada únicamente en determinaciones de cromosoma Y no es factible, ya que todos los varones pertenecientes al mismo linaje paterno comparten el mismo haplotipo y se encontrarían inexcusablemente implicados.

Dado el interés de los STRs de cromosoma Y en determinaciones de paternidad, algunos laboratorios forenses acreditados los están incluyendo ya en su sistemática.

4.2.- CRIMINALÍSTICA

Muy frecuentemente los vestigios biológicos de interés en criminalística se encuentran en cantidades ínfimas, sufren deterioro o están contaminados de diversas formas, lo que dificulta enormemente su análisis.

La determinación de polimorfismos de ADN, en especial mediante la técnica de PCR, es un procedimiento óptimo cuando necesitamos analizar manchas de sangre minúsculas o degradadas u otras muestras (como esperma, saliva, pelos o cabellos

sin bulbo, fragmentos óseos, etc) en las que las técnicas convencionales proporcionan resultados muy poco satisfactorios. Están siendo de gran ayuda principalmente en delitos contra la libertad sexual, en los que, ante la negativa del presunto culpable, no suele existir más indicio incriminatorio que el proporcionado por posibles restos de esperma en prendas de vestir y cavidades corporales.

Los polimorfismos de ADN aportan al análisis criminalístico la posibilidad de diferenciar fácilmente muestras compuestas por mezclas de fluidos procedentes de distintos individuos (Evetts *et al* 1991a,b; Wiegand *et al* 1992)) y, en el caso particular de agresiones sexuales, permiten separar el ADN de células vaginales del ADN espermático (Gill *et al* 1987a,b).

El potencial del ADN como medio de identificación hizo que pronto se propusiese la creación de bases de datos de delincuentes para delitos graves y con alta tasa de reincidencia, como el de violación, formadas a partir de las evidencias recogidas con ocasión de hechos delictivos, lo que permitiría la comparación con sospechosos e interrelacionar delitos cometidos por la misma persona. En el Reino Unido y Holanda estas bases de datos ya son una realidad (Werret 1997, Kloosterman y Janssen 1997). En el primero de estos países se almacenan perfiles de ADN de inculpados o simplemente sospechosos y se pretende incluir a toda la población en el futuro. Otros países se inclinan por seguir un criterio más restrictivo, que incluya casos concretos (delitos contra la libertad sexual, por ejemplo). En España, por el momento, no existe legislación al respecto. (En un número especial de la publicación periódica *Forensic Sci Int* (1997) se recogen aspectos legales y éticos sobre este tema).

Los polimorfismos de cromosoma Y (STRs) aplicados al análisis criminalístico (Roewer y Eppelen 1992a), presentan la ventaja de proporcionar información adicional sobre la presencia de ADN de varón en manchas forenses, particularmente en casos de violación y asalto sexual, en los que este análisis ofrece una nueva aproximación.

La principal utilidad de estos sistemas se encuentra en los casos de muestras con mezclas de células masculinas y femeninas. Contrariamente a lo que sucede en los sistemas autosómicos, una elevada cantidad de ADN femenino no inhibe la amplificación de los alelos de cromosoma Y (Prinz *et al* 1997) permitiendo obtener un perfil específico de ADN del asaltante, incluso omitiendo el paso de la lisis diferencial. En aquellas ocasiones en las que esta última determinación se presuma ineficaz o arriesgada, por ejemplo, por estar la muestra muy degradada o por contener una mínima cantidad de esperma, estos STRs pueden incrementar la tasa de éxito en la identificación del componente masculino en fluidos corporales con mezcla de células de ambos sexos. Este abordaje no sólo es rápido, sino que minimiza también la pérdida de ADN espermático y reduce el riesgo de contaminación.

Los STRs de cromosoma Y también pueden ser interesantes en la identificación de otras mezclas, como por ejemplo, sangre-sangre, sangre-saliva o mezcla de tejidos similares, en las que no puede aplicarse la lisis celular diferencial.

Como otras posibles aplicaciones podemos incluir la detección de células epiteliales masculinas procedentes de individuos vasectomizados o el permitir una mejor aproximación al número de contribuyentes de semen en casos de violaciones múltiples, obteniéndose en este caso una mezcla de haplotipos de los agresores. Cuando la víctima de la violación es un hombre, a pesar de la mezcla, también es posible inferir el haplotipo del violador restando el correspondiente a la víctima.

Los STRs de cromosoma Y también pueden usarse como un método de screening rápido para excluir sospechosos, sin tener que emplear tiempo en una extracción diferencial que a veces se muestra infructuosa (Kayser *et al* 1997).

En casos de criminalística, cuando un sospechoso posee un haplotipo de cromosoma Y diferente al del culpable estaríamos ante una buena razón para su exclusión, al igual que ocurre con otros sistemas basados en el tipado de ADN. La inclusión basada solamente en marcadores de cromosoma Y plantea el mismo

problema que hemos comentado al hablar de la investigación de la paternidad. Utilizando sistemas autosómicos, la probabilidad de que dos individuos posean idéntico perfil es habitualmente muy baja (se reduciría al caso de que ambos fuesen gemelos univitelinos), sin embargo, la determinación aislada de polimorfismos de cromosoma Y incluye a todos los familiares varones por vía paterna por compartir idéntico haplotipo, por lo que el veredicto de culpabilidad deberá basarse en otras pruebas adicionales.

Los casos de agresiones sexuales, como vemos, son ideales para la aplicación de este tipo de marcadores de cromosoma Y, lo que va a redundar en un incremento del número de casos resueltos.

Por otro lado, como consecuencia del alto grado de especificidad poblacional de algunos haplotipos de cromosoma Y, sería razonable hacer suposiciones sobre el origen étnico de una muestra de ADN hallada en la escena del delito, caso de portar un haplotipo característico de un grupo particular, lo que permitiría acotar la búsqueda de sospechosos. En algunos casos, cuando las frecuencias haplotípicas sean conocidas, tal información podría utilizarse en investigaciones, siempre y cuando se establezcan normas cautelares que garanticen el uso adecuado de esta información.

4.3.- IDENTIFICACIÓN GENÉTICA

Hasta la aparición de los polimorfismos de ADN la identificación de cadáveres y restos óseos sólo era posible por medio de la antropología física y la odontología forense.

Diversos estudios han demostrado la utilidad de la amplificación del ADN mediante PCR en la identificación de cadáveres humanos, ya que cuando la muestra no se encuentra en condiciones óptimas por efecto de factores adversos como temperatura, humedad y procesos autolíticos y de degradación cadavéricos, la identificación por RFLPs puede no ser posible. Aún en restos muy antiguos es

posible analizar ADN_{mt} o STRs partiendo, por ejemplo, de piezas dentarias (Alvarez-García *et al* 1996) o fragmentos óseos (Gill *et al* 1994a). La comparación con el ADN obtenido de muestras de sangre de algún ascendiente o descendiente indubitado (preferiblemente de los progenitores) suele ser suficiente para establecer una relación de parentesco.

En situaciones de grandes catástrofes (Olaisen *et al* 1997), los polimorfismos de ADN se revelan como un procedimiento de gran utilidad respecto a la asignación de restos cadavéricos a individuos y la identificación de éstos. En circunstancias especiales, en las que el proceso de identificación de cadáveres o restos cadavéricos se prevea dificultoso, puede ser útil contar con fichas genéticas previas de individuos sometidos a conflictos bélicos o con profesiones de riesgo.

Los polimorfismos de cromosoma Y (STRs) en casos forenses de indentificación, especialmente cuando se trata analizar pequeños fragmentos humanos por efecto de desastres en masa, harían posible una rápida determinación del sexo del individuo, contribuyendo además a su identificación (Corach *et al* 1995, 1996).



II. MATERIAL Y MÉTODOS

1.- CARACTERÍSTICAS DE LA MUESTRA ESTUDIADA

En el presente estudio hemos analizado 116 muestras de sangre provenientes de varones tomados al azar pertenecientes a la población gallega, obtenidas de casos rutinarios de paternidad (padres) del Instituto de Medicina Legal de Santiago de Compostela y de donantes anónimos a través del Centro de Transfusión de Galicia. Estas últimas debían cumplir la condición de que no coincidiesen los apellidos de los donantes para obviar posibles relaciones de parentesco, extremo que no pudimos confirmar directamente debido a que las muestras, para preservar el anonimato, las recibimos sin identificación personal.

Adicionalmente incluimos 35 combinaciones padre/hijo para el estudio de mutaciones y 20 muestras pertenecientes a mujeres para comprobar, en este caso, si los sistemas se comportaban como específicos de cromosoma Y bajo las condiciones de nuestro estudio.

A cada persona se le extrajeron 5 ml de sangre periférica por punción venosa, recogiendo en tubos estériles conteniendo EDTA como anticoagulante. Las muestras fueron almacenadas en alícuotas de 700 µl de sangre total a -20°C hasta el momento de la extracción del ADN. (En algunos casos de paternidad las muestras se encontraban almacenadas a temperatura ambiente como manchas de sangre sobre tela).

2.- EXTRACCIÓN DEL ADN

2.1.- EXTRACCIÓN DEL ADN A PARTIR DE SANGRE TOTAL

Se utilizó el método fenol/cloroformo-alcohol isoamílico (Valverde *et al* 1993), con arreglo al siguiente protocolo:

Material y métodos

- 1- Descongelar las alícuotas de sangre en baño a 37°C durante 2 minutos.
- 2- Centrifugar a 12.500 rpm durante 2 minutos.
- 3- Retirar el sobrenadante con pipeta, dejando aproximadamente 100 µl de sangre sobre el pellet .
- 4- Resuspender el sedimento de leucocitos en 800 µl de tampón B, elaborado a partir de:
 - Tris. Hcl 50 mM
 - NaCl 150 mM
 - EDTA 100mM
 - pH 8.0
 - Guardar a temperatura ambiente.
- 5- Dividir la mezcla una vez resuspendida en dos tubos Eppendorf de 1.5 ml (450 µl por tubo).
- 6- Añadir 25 µl de SDS 20%. Mezclar invirtiendo varias veces los tubos.
- 7- Añadir 6 µl de proteinasa K (20 mg/ml) y mezclar bien.
- 8- Incubar en baño durante 1 hora a 60°C.
- 9- Centrifugación rápida para llevar todo el líquido al fondo del tubo.
- 10- Añadir 80 µl de perclorato sódico 5M (0.7 g/ml), preparado en el momento.
Mezclar suavemente por inversión
- 11- Añadir 1 volumen (500 µl) de fenol:cloroformo:alcohol isoamílico (25:24:1).
Mezclar bien invirtiendo los tubos.
- 12- Centrifugar a 12.500 rpm durante 10 minutos.
- 13- Recoger el sobrenadante con pipeta y pasarlo a otro tubo.
- 14- Repetir el lavado con el mismo volumen de fenol:cloroformo:alcohol isoamílico y mezclar.
- 15- Centrifugar a 12.500 rpm otros 10 minutos.
- 16- Trasladar la fase superior a un nuevo tubo.
- 17- Añadir 1 volumen de cloroformo:alcohol isoamílico (24:1). Mezclar.
- 18- Centrifugar otra vez a 12.500 rpm durante 10 minutos.
- 19- Pasar el sobrenadante a otro tubo.
- 20- Añadir 2 volúmenes (1 ml) de de etanol absoluto frío (-20°C) y mezclar suavemente. El ovillo de ADN aparecerá rápidamente por precipitación.

- 21- Centrifugar a 12.500 rpm durante 2 minutos. Retirar el etanol por decantación, cuidando de no perder el ovillo.
- 22- Añadir 1 volumen de etanol al 80% frío (4°C).
- 23- Centrifugar a 12.500 rpm durante 5 minutos.
- 24- Retirar inmediatamente el etanol al finalizar el centrifugado, con sumo cuidado para no remover el ADN.
- 25- Dejar secar los tubos a temperatura ambiente 5 ó 10 minutos.
- 26- Resuspender el ADN en 100 µl de agua bidestilada estéril (o EDTA 0.1 mM).
- 27- Dejar los tubos en un baño a 56°C toda la noche.
- 28- Centrifugado rápido para llevar todo al fondo del tubo
- 29- Resuspender con pipeta y mezclar el contenido de ambos tubos.
- 30- Cuantificar en fluorímetro y conservar el ADN extraído a 4°C (si no se va a utilizar seguidamente es preferible congelarlo a -20).

2.2.- EXTRACCIÓN DEL ADN A PARTIR DE MANCHAS DE SANGRE

Se utiliza una resina quelante (Chelex[®] 100 Resin/BioRad Labs, Richmond, CA) con una alta afinidad por iones metálicos polivalentes, según se indica a continuación:

- 1- En un tubo Eppendorf añadir 1 ml de agua destilada a 3 mm² de la mancha.
- 2- Incubar 30 minutos a temperatura ambiente. Mezclar de vez en cuando por inversión
- 3- Centrifugar de 10.000 a 15.000 rpm durante 2-3 minutos.
- 4- Retirar parte del sobrenadante, dejando 20-30 µl y resuspender.
- 5- Añadir Chelex al 5% hasta un volumen final de 200 µl.
- 6- Incubar a 56°C durante 15-30 minutos en un baño.
- 7- Vórtex a alta velocidad durante 5-10 segundos.
- 8- Introducir el tubo en agua hirviendo 8 minutos.
- 9- Vórtex a alta velocidad durante 5-10 segundos.

10- Centrifugar 2-3 minutos a 10.000-15.000 rpm antes de añadir la muestra a la mezcla de PCR.

El ADN extraído por este método se conserva aproximadamente un mes a 4°C o indefinidamente a -20°C

3.- CUANTIFICACIÓN DEL ADN

Cada muestra de ADN se cuantificó por medio de un fluorímetro (*DynaQuant*TM 200 *Hoefler*, Amersham Pharmacia Biotech), utilizando la bisbenzimidida como colorante. Hay que adoptar precauciones al manejar esta sustancia, ya que está considerada como un posible mutágeno.

La presencia de ADN en la solución que contiene la bisbenzimidina modifica las características fluorescentes del colorante (también denominado Hoechst 33258), lo que nos permite calcular con gran exactitud la cantidad de ADN que hay en la solución. En ausencia de ADN el espectro de excitación del colorante es de 356 nm y el de emisión de 492 nm. Cuando el H 33258 se une al ADN, ambos espectros varían. En la cubeta la muestra se expone a luz filtrada (365 ± 7 nm) desde una lámpara de mercurio. Al excitarse el complejo ADN-colorante emite luz a 458 nm. Un filtro de emisión permite registrar la fluorescencia a 460 ± 15 nm. De esta forma, la medida de la fluorescencia se convierte en un indicador directo de la concentración de ADN.

Los fluorímetros miden la fluorescencia en cantidades relativas, más que absolutas.

- En primer lugar se preparan las soluciones control y estándar. Hay dos tipos de solución control, dependiendo de la concentración final que estimemos obtener (A, de bajo rango, adecuada para concentraciones finales entre 10-500 ng/ml y de alto rango, B, si se espera una concentración final más elevada, alrededor de 100-5000

ng/ml). En nuestro caso hemos optado por la solución A preparada con 0.1 µg/ml de H 33258 en TNE 1X (ClNa 0.2 M, Tris-Cl 10 mM, EDTA 1 mM, pH 7.4).

- A continuación se pone el fluorímetro a cero con 2 ml de la solución control A, que sirve como blanco.

- Seguidamente se calibra el fluorímetro. Se añaden 2 µl de solución de ADN estándar (ADN estándar 1mg/ml, TNE 10X) diluída al 1:10 (100 µg/ml) a los 2 ml de solución control. Se coloca la cubeta en el pocillo y se mide hasta que aparezca en pantalla la concentración esperada (100 ng/ml).

- Una vez calibrado el aparato se pone de nuevo a cero con 2 ml de solución control.

- Finalmente se añaden 2 µl de muestra a los 2 ml de solución control, se mezcla bien y se hace la medida de la muestra.

- Cantidad: $2 \mu\text{l} \times 100 \text{ ng}/\mu\text{l} = 200 \text{ ng}$.

- Concentración final en la cubeta: $200 \text{ ng}/2 \text{ ml} = 100 \text{ ng/ml}$.

4.- AMPLIFICACIÓN DEL ADN

Hemos amplificado individualmente mediante PCR los sistemas DYS19, DYS385, DYS389/I-II, DYS390 y DYS393 para su puesta a punto, como paso previo para su integración en multiplexes, con vista a futuros estudios. (Para cada sistema amplifica un sólo locus, excepto DYS389I/II y DYS385, en los que dos loci amplifican simultáneamente con el mismo par de *primers*).

4.1.- REACCIÓN DE PCR

Los oligonucleótidos usados como cebadores en la reacción, han sido suministrados por Amersham Pharmacia Biotech (Suecia). Uno de cada par de *primers* está marcado con fluoresceína en el extremo 5' para su posterior detección automática. La secuencia de cada pareja de *primers* (Tabla 6) es la que figura descrita en el Genome DataBase (GDB).

Para la amplificación de las muestras de cada sistema se utilizaron 1-20 ng de ADN en un volumen final de 25 µl. La concentración de los cebadores y la cantidad de la enzima *Taq* DNA polimerasa varió según el sistema analizado (Tabla 7):

- **DYS19, DYS385:** 0.5 µM de cada *primer* y 2 U de *Taq* DNA polimerasa.
- **DYS389I/II:** 0.5 µM de cada *primer* y 1 U de *Taq* DNA polimerasa.
- **DYS390, DYS393:** 0.2 µM de cada *primer* y 0.65 U de *Taq* DNA polimerasa.

Tabla 6. Secuencia de los cebadores usados para amplificar cada sistema

SISTEMA	SECUENCIA DE LOS CEBADORES
DYS19 (locus GDB:121409)	<i>Primer 1:</i> *5'CTACTGAGTTTCTGTTATAGT3' <i>Primer 2:</i> 5'ATGGCATGTAGTGAGGACA3'
DYS385 (locus GDB:316257)	<i>Primer 1:</i> *5'AGCATGGGTGACAGAGCTA3' <i>Primer 2:</i> 5'GGGATGCTAGGTAAAGCTG3'
DYS389I/II (locus GDB:366108)	<i>Primer 1:</i> *5'CCAACTCTCATCTGTATTATCTATG3' <i>Primer 2:</i> 5'TCTTATCTCCACCCACCAGA3'
DYS390 (locus GDB:366115)	<i>Primer 1:</i> *5'TATATTTTACACATTTTTGGGCC3' <i>Primer 2:</i> 5'TGACAGTAAAATGAACACATTGC3'
DYS393 (locus GDB:456649)	<i>Primer 1:</i> *5'GTGGTCTTCTACTTGTGTCAATAC3' <i>Primer 2:</i> 5'AACTCAAGTCCAAAAAATGAGG3'

* *Primers* marcados con fluoresceína en el extremo 5'

Tabla 7. Componentes de la reacción de PCR

	DYS19	DYS385	DYS389I/II	DYS390	DYS393
ADN	1-20 ng				
Primers	0.5 µM	0.5 µM	0.5 µM	0.2 µM	0.2 µM
dNTPs	0.2 mM				
Buffer 10X	1X	1X	1X	1X	1X
Cl₂ Mg	1.5 mM				
Taq DNA polimerasa	2 U	2 U	1 U	0.65 U	0.65 U
H₂O	Hasta 25µl				

El protocolo para amplificar el sistema DYS385 es el proporcionado por PM Schneider (comunicación personal) y el resto de los protocolos son los utilizados por Kayser *et al* (1997) con ligeras modificaciones.

4.2.- CONDICIONES DE AMPLIFICACIÓN DE LOS SISTEMAS

Los programas de amplificación se llevaron a cabo indistintamente en dos termocicladores de ADN (*DNA Thermal Cycler 9600*, Perkin Elmer Cetus y *PT-100*, MJ Research, Inc) con arreglo a los siguientes protocolos :

- **SISTEMA DYS19** (Tabla 8). Para este sistema se calentaron las muestras a 94°C durante 1 minuto con el fin de desnaturalizar el ADN. Luego se bajó la temperatura a 58°C durante 1 minuto para que tuviese lugar la fase de acoplamiento y, finalmente, la extensión transcurrió a 72°C durante 2 minutos. Se realizaron 30 ciclos.

Tabla 8. Programa de amplificación del Sistema DYS19

DYS19	TEMPERATURA	TIEMPO	Nº CICLOS
Desnaturalización	94°C	1 m	30
Hibridación	58°C	1 m	
Extensión	72°C	2 m	

- **SISTEMA DYS390** (Tabla 9). Realizamos una desnaturalización previa a 94°C durante 2 minutos. Los ciclos repetidos de desnaturalización, hibridación y extensión, dos tandas de 5 y 30 ciclos (94°C-15s, 58°C-15s, 72°C-20s y 94°C-15s, 54°C-15s, 72°C-20s respectivamente), se diferencian únicamente en la temperatura de hibridación. El siguiente paso consistió en una extensión final a 72°C durante 10 minutos.

Tabla 9. Programa de amplificación del Sistema DYS390

DYS390	TEMPERATURA	TIEMPO	Nº CICLOS
Predesnaturalización	94°C	2 m	1
Desnaturalización	94°C	15 s	5
Hibridación	58°C	15 s	
Extensión	72°C	20 s	
Desnaturalización	94°C	15 s	30
Hibridación	54°C	15 s	
Extensión	72°C	20 s	
Extensión final	72°C	10 m	1

• **SISTEMA DYS389I/II Y DYS393** (Tabla 10). A estos sistemas puede aplicárseles el mismo protocolo, que incluye una desnaturalización previa a 94°C durante 2 minutos y una extensión final a 72°C durante 10 minutos. Las etapas de desnaturalización (94°C-15s), hibridación (58°C-15s) y extensión (72°C-20s) se repiten durante 30 ciclos.

Tabla 10. Programa de amplificación de los Sistemas DYS389I/II y DYS393

DYS389I/II y DYS393	TEMPERATURA	TIEMPO	Nº CICLOS
Predesnaturalización	94°C	2 m	1
Desnaturalización	94°C	15 s	30
Hibridación	58°C	15 s	
Extensión	72°C	20 s	
Extensión final	72°C	10 m	1

• **SISTEMA DYS385** (Tabla 11). Previamente se realiza una desnaturalización a 94°C durante 3 minutos. A continuación se suceden las fases de desnaturalización, hibridación y extensión (94°C-30s, 59°C-30s, 72°C-1m durante 3 ciclos; 94°C-30s, 57°C-30s, 72°C-1m durante 2 ciclos; 94°C-30s, 56°C-30s, 72°C-1m durante 29 ciclos). Este sistema no precisa extensión final.

Tabla 11. Programa de amplificación del Sistema DYS385

DYS385	TEMPERATURA	TIEMPO	Nº CICLOS
Predesnaturalización	94°C	3 m	1
Desnaturalización	94°C	30 s	3
Hibridación	59°C	30 s	
Extensión	72°C	1 m	
Desnaturalización	94°C	30 s	2
Hibridación	57°C	30 s	
Extensión	72°C	1 m	
Desnaturalización	94°C	30 s	29
Hibridación	56°C	30 s	
Extensión	72°C	1 m	

Una vez finalizado el proceso de amplificación, las muestras fueron conservadas a 4°C.

5.- DETECCIÓN DE LOS PRODUCTOS AMPLIFICADOS

5.1.- ELECTROFORESIS EN GELES DE POLIACRILAMIDA Y TINCIÓN DE PLATA

Para comprobar la existencia de producto amplificado tras la reacción de PCR se utilizó electroforesis en geles miniaturizados de poliacrilamida (*Phastgeles*, Amersham Pharmacia Biotech, Uppsala, Suecia), de gradiente 8-25.

Estos minigeles (5x5 cm) constan de una zona de concentración de la muestra seguida de una zona de gradiente, en la cual tiene lugar realmente la separación de los fragmentos. El tampón de los geles contiene acetato 0.112 M y Tris 0.112 M; pH

6.4. Los electrodos son de agarosa al 2% y su tampón está compuesto de Tricina 0.20 M, Tris 0.20 M y SDS 0.55%; pH 8.1.

El recorrido electroforético de las muestras (1 μ l de producto amplificado) aplicadas al extremo catódico del gel, se llevó a cabo en la unidad de separación del *PhastSystem* (Amersham Pharmacia Biotech, Suecia), siendo las condiciones electroforéticas 50 V/cm, 10.0 mA y 3.0 W hasta 73 AVh de recorrido. Con las muestras corre también un marcador interno (123).

La detección de los productos amplificados se realizó en la unidad de revelado del mismo aparato, mediante tinción de plata según el método descrito por Heukeshoven y Dernick (1985).

5.2.- DETECCIÓN AUTOMÁTICA DE FRAGMENTOS MARCADOS CON FLUORESCÉINA

Un secuenciador automático de ADN, en nuestro caso el *Automatic Laser Fluorescent (ALF) DNA Sequencer* (Amersham Pharmacia Biotech, Suecia), proporciona otro método para la visualización de productos amplificados. Consta de una unidad de electroforesis y sistema de detección fluorescente asociada a un ordenador, permitiendo controlar el proceso electroforético así como el almacenamiento de datos y su posterior análisis (Figura 5).

Este sistema detecta directamente los fragmentos de ADN marcados con fluoresceína en el transcurso de una electroforesis en geles desnaturizantes de poliacrilamida, permitiendo medirlos y cuantificarlos. Para nuestro trabajo hemos marcado con fluoresceína el extremo 5' de uno de los cebadores de cada sistema.

Los geles desnaturizantes (28x27 cm y 0.35 mm de espesor) se elaboraron con poliacrilamida 6% (p/ v) (acrilamida 5.7%, bisacrilamida 0.3%), urea 7M, en TBE

(Tris-borato 100 mM (pH 8.3), EDTA Na₂ 1 mM) y persulfato amónico al 10% más TEMED 3 mM, dejándolo polimerizar 2 horas.

Figura 5. Secuenciador automático ALF Express (Amersham Pharmacia Biotech, Suecia)



Los productos de PCR que queremos visualizar deben ser preparados previamente. A 1 μ l de producto amplificado se añaden 4 μ l de tampón de carga (5 mg/ml dextran blue/formamida) y 1 μ l de cada marcador interno estándar (114 y 402 pb) y después se desnaturaliza la mezcla a 94°C durante 3 minutos. Para evitar la renaturalización del ADN, las muestras deben mantenerse en hielo hasta el momento de ser cargadas en los pocillos del gel.

El recorrido electroforético se realiza con parámetros constantes: 1600 V, 42 mA, 45W y 45°C de temperatura durante unas 4 horas.

El tamaño de los fragmentos (en pb) se determina automáticamente utilizando el software *Fragment Manager 1.2* (Amersham Pharmacia Biotech, Suecia) y su tipaje

se establece por comparación con los correspondientes ladders alélicos secuenciados que habíamos construido previamente.

6.- SECUENCIACIÓN

El método de secuenciación que hemos empleado se fundamenta en una reacción de elongación mediada por una ADN polimerasa, en presencia de dideoxinucleótidos (ddATP, ddCTP, ddTTP, ddGTP) marcados con fluorocromos de diferente color. La incorporación al azar de un ddNTP en lugar de un dNTP detiene la reacción, con lo que obtenemos múltiples fragmentos de distinta longitud.

6.1.- ELUCIÓN DE LOS ALELOS

Tras seleccionar las muestras de sistemas conteniendo un solo alelo, se realiza una electroforesis en geles de poliacrilamida bajo condiciones nativas, que se tiñen posteriormente con NO_3Ag para identificar las bandas correspondientes a los alelos. Se recortan cuidadosamente del gel y se dejan toda la noche en un baño a 37°C con agitación, centrifugando al finalizar. A continuación se reamplifican y se chequean.

De los sistemas con dos alelos sólo hemos secuenciado el DYS389I/II. Los alelos fueron previamente separados en un gel de tamaño 10.5x8 cm, preparado con agarosa al 3% (*NuSieve*, FMC BioProducts, USA) en tampón TAE 1X (Tris-acetato 44.5 mM, EDTA 1 mM, pH 8) y bromuro de etidio. En cada pocillo del gel se cargaron 15 μl de producto amplificado y 3 μl de tampón de carga. El recorrido electroforético se llevó a cabo durante 3 horas a 96 V. Se cortaron las bandas correspondientes a los alelos y se eluyeron con el kit *SephaglasTM BandPrep Kit* (Amersham Pharmacia Biotech).

6.2.- PURIFICACIÓN PRE-SECUENCIACIÓN

El producto amplificado se purifica mediante columnas, *MicroSpin™ S-300 HR columns* (Amersham Pharmacia Biotech) o *Microcon 100*, (Microconcentrators, Amicon), con el fin de eliminar componentes de la reacción de PCR (*primers*, dNTPs, sales, etc) que puedan interferir en la reacción de secuenciación. (Este paso va implícito cuando se eluyen los alelos con *Sephaglas™*).

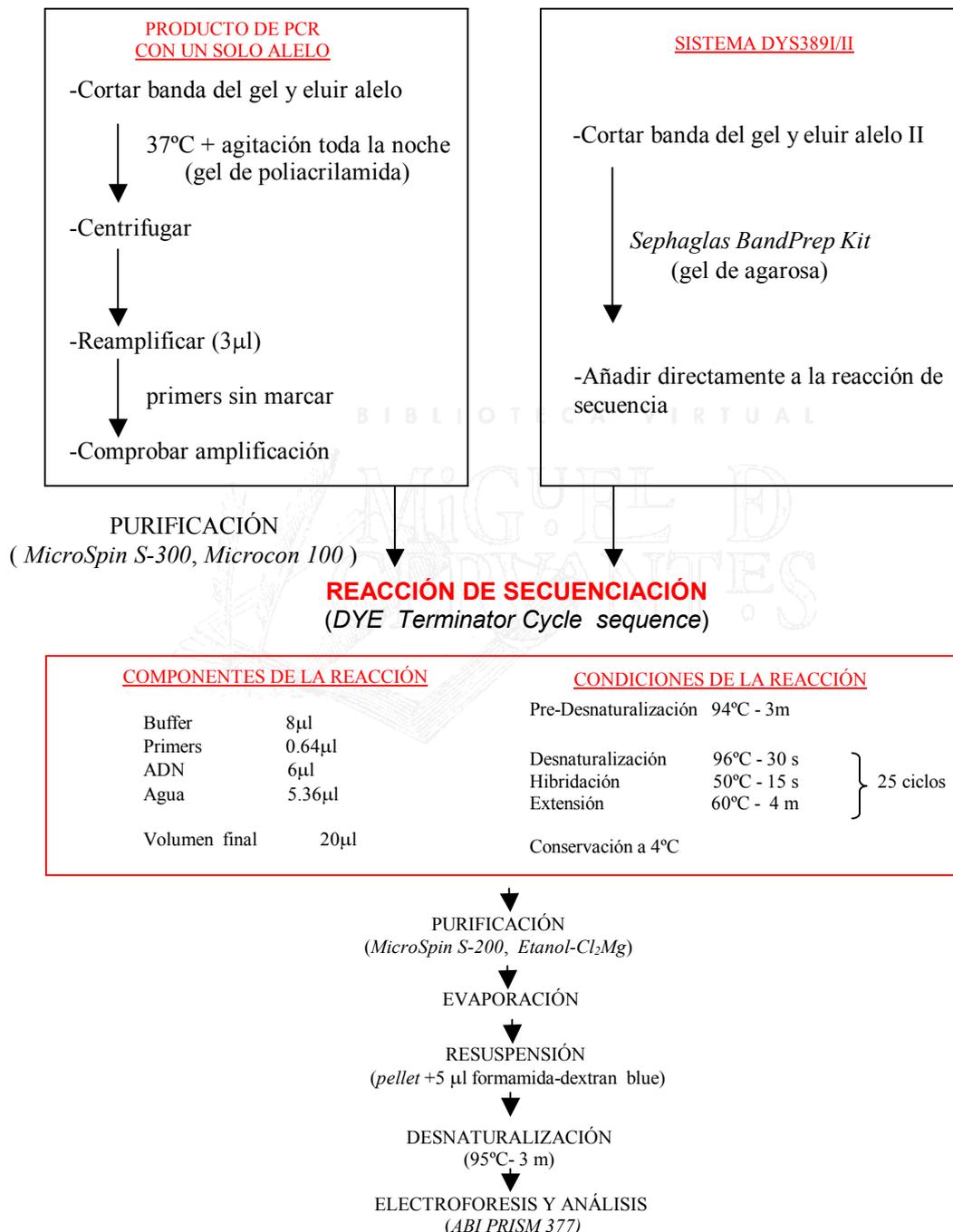
6.3.- REACCIÓN DE SECUENCIACIÓN

Para la reacción de secuenciación de ADN (Tabla 12) hemos empleado el kit *Dye Terminator Cycle Sequencing Ready Reaction* (con AmpliTaq DNA Polymerase, FS) (PE Applied Biosystems) al que se añaden 6µl de ADN purificado y 3.2 pmoles de cada *primer* de secuencia, utilizando para ello uno de los *primers* que intervinieron en la reacción de amplificación del fragmento (el que no estaba marcado con fluoresceína). La reacción tiene lugar en un termociclador programable (*Touch Down*, Hybaid limited, UK). Las condiciones de PCR se muestran en la Tabla 12.

Una vez finalizada la reacción de secuenciación se lleva a cabo una nueva purificación para retirar el exceso de ddNTPs con columnas (*MicroSpin™ S-200 HR columns*, Amersham Pharmacia Biotech) o Etanol-Cl₂Mg, dejando evaporar las muestras.

El "pellet" se resuspende con 4µl de formamida/EDTA-dextran blue (5:1). Las muestras se desnaturalizan durante 2 min a 90 °C y se colocan en hielo hasta ser cargadas en el gel.

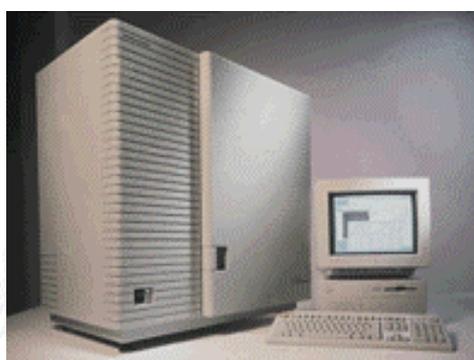
Tabla 12. Esquema del proceso de Secuenciación



6.4.- ELECTROFORESIS Y ANÁLISIS DE LA SECUENCIA

La separación electroforética se llevó a cabo en el secuenciador automático de ADN *ABI Prism™ 377 (ABI) DNA Sequencer* (Perkin Elmer Applied Biosystems), que permite obtener la secuencia completa de cada muestra en una calle (Figura 6).

Figura 6. Secuenciador automático ABI 377 (Perkin Elmer)



El gel (acrilamida 6%, Urea 6M, TBE 1X), de dimensiones 40x24 cm y 0.2 mm de espesor, se elaboró a partir de 7.5 ml de una solución stock de acrilamida al 40% (38 g de acrilamida y 2 g de bisacrilamida en 100 ml de agua), 18 g de urea, 25.5 ml de agua y 0.5 g de *amberlita™* (Amersham Pharmacia Biotech). Se añaden 250 µl de persulfato amónico al 10% y 35 µl de TEMED y se deja gelificar entre 2 y 6 horas.

El recorrido electroforético se realizó bajo las siguientes condiciones: 1680 V, 50mA, 150W y 51°C durante 7 horas.

Para el análisis de las secuencias se utilizó el software *DNA sequencing analysis* (Perkin-Elmer).

7.- CONSTRUCCIÓN DE *LADDERS* ALÉLICOS Y DESIGNACIÓN DE ALELOS

Siguiendo las recomendaciones de la ISFH (1994, 1997), para realizar un tipaje correcto de cada muestra es necesaria su comparación con un *ladder* alélico secuenciado que migre a su lado.

Los *ladders* alélicos que hemos empleado en el presente trabajo (Figura 8), los hemos construido a partir de una mezcla de alelos, previamente secuenciados, conteniendo los más frecuentes para cada marcador en nuestra población, excepto el *ladder* correspondiente al sistema DYS385 (alelos 11, 13, 15, 17 y 19), cedido amablemente por PM Schneider del Institute of Legal Medicine, Mainz, Alemania.

La nomenclatura de los alelos de cada sistema la hemos realizado atendiendo al número de unidades de repetición que poseen y la definición de la unidad repetitiva se ha hecho, en la medida de lo posible, de acuerdo nuevamente con las recomendaciones de la ISFH (1994, 1997). La publicación de los artículos de Kayser *et al* (1997) y de Knijff *et al* (1997) condicionó en gran medida la nomenclatura definitiva, como ya comentaremos en la discusión.

8.- ANÁLISIS ESTADÍSTICO

La frecuencia de un alelo en una población se denomina frecuencia génica o alélica y se puede definir como el cociente resultante de dividir el número de alelos iguales en una población por el número total de alelos.

Para la aplicación de cualquier marcador genético en Genética Forense es fundamental determinar las frecuencias alélicas del sistema polimórfico en la población, realizando una estima de dichas frecuencias a través de un fenotipado muestral representativo, dada la imposibilidad fenotipar todos y cada uno de los integrantes de la población para obtener su valor exacto.

Hemos utilizado los siguientes parámetros estadísticos que nos indican la utilidad a priori de los marcadores:

1-Probabilidad de exclusión a priori (CE) de un marcador en una población determinada. Se define como la probabilidad de exclusión de un individuo como padre en un análisis de paternidad.

Para sistemas autosómicos se calcula mediante la siguiente fórmula (Ohno *et al* 1982):

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

, siendo p_i, p_j = frecuencias génicas respectivas.

Para sistemas de cromosoma Y ésta coincidiría con $1-\sum p_i^2$ siendo $\sum p_i^2$ la probabilidad de inclusión, es decir, que el padre sea efectivamente el padre.

El *CE* depende del número de alelos identificados y de la distribución de las frecuencias alélicas. Es un valor porcentual que es función directa del polimorfismo de un sistema. Cuanto más polimórfico sea éste y más equilibradas estén las frecuencias de sus alelos, tanto mayor será su *CE*, lo que le hace ser un parámetro de gran valor médico-legal, especialmente en determinaciones biológicas de paternidad.

2-Poder de discriminación (PD). Es una medida relativa de la eficacia de un sistema, sobre todo con fines de identificación forense. Se define como la probabilidad de que dos individuos tomados al azar, pertenecientes a la misma población, posean diferentes fenotipos.

Se calcula mediante la siguiente fórmula (Fisher 1951):

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

, siendo p_i = frecuencias genotípicas esperadas, que para cromosoma Y coinciden con las frecuencias alélicas, y siendo $\sum p_i^2$ la probabilidad de coincidencia o "matching".

3-Probabilidad de coincidencia (PM) o matching probability. Se define como la probabilidad de que dos individuos tomados al azar de la misma población coincidan en su genotipo para ese locus. Esta probabilidad fue descrita por Jones (1972) y se calcula a partir de la siguiente fórmula:

$$PM = \sum_{i=1}^n p_i^2$$

, siendo p_i = la frecuencia esperada de cada genotipo (para cromosoma Y, frecuencias alélicas).

Como vemos, el poder de discriminación y la probabilidad de coincidencia son conceptos opuestos ($PD=1-PM$)

4-Diversidad génica y diversidad haplotípica. Para STRs de cromosoma Y, los valores de diversidad génica (D) o haplotípica en su caso, tienen el mismo valor que el poder de discriminación y la probabilidad de exclusión, de tal manera que:

$$PD=CE=1-PM=1-\sum p_i^2$$

, siendo p_i las frecuencias alélicas o haplotípicas.

(La anterior fórmula es equivalente a la heterocigosidad (h) de un marcador autosómico).

Para la comparación de nuestros resultados con los de otras poblaciones estudiadas utilizamos el test χ^2 , considerando como significativos valores de $p < 0.001$.





III. RESULTADOS Y DISCUSIÓN

1.- CONSIDERACIONES METODOLÓGICAS

1.1. SOBRE EL MÉTODO DE EXTRACCIÓN DEL ADN

Hemos elegido para la extracción del ADN el método fenol-cloroformo (Valverde *et al* 1993) por presentar las siguientes ventajas:

-Se parte de un volumen sanguíneo reducido (700 μ l), lo que permite realizar todo el proceso en microtubos desechables, ahorrando en reactivos.

-El proceso se acelera mediante digestión a alta temperatura, obteniéndose un buen rendimiento en la cantidad de ADN.

-El ADN extraído conservado a 4°C, no plantea problemas de degradación. Además puede ser almacenado a -20 °C durante largo tiempo sin perder cualidades.

El método de extracción con resinas quelantes (*Chelex*[®] 100 Resin, Bio Rad Labs, Richmond, CA) también proporciona buenos resultados. Este método es más rápido, sencillo y barato que el anterior y minimiza la pérdida de material por manipulación, pero su principal inconveniente es que el ADN se conserva menos tiempo y contiene más impurezas, por lo que sólo lo hemos utilizado cuando fue preciso extraer el ADN a partir de manchas de sangre.

1.2. SISTEMAS ELECTROFORÉTICOS

Para comprobar el resultado de la amplificación se utilizaron minigeles de poliacrilamida teñidos posteriormente con sales de plata, todo ello analizado en el sistema semiautomático PhastSystem, que permite reproducir fácilmente tanto la separación electroforética como la tinción mediante la programación de variables como temperatura, voltaje, tiempo, etc. El análisis se lleva a cabo en menos de 90 minutos, con un consumo mínimo (1 μ l o algo menos) de producto amplificado por

muestra. Podemos procesar 8 ó 12 muestras por gel, aunque es posible duplicar este número ya que en el PhastSystem pueden colocarse dos geles simultáneamente.

En el caso de que el ADN haya sido extraído por chelex, las bandas aparecen menos nítidas, con más ruido de fondo, aunque esto no será percibido si nos inclinamos por el uso de secuenciadores automáticos para la lectura de las muestras.

La mayoría de los laboratorios forenses utilizan actualmente secuenciadores automáticos de ADN para la lectura de STRs, ya que estos sistemas electroforéticos van a facilitar enormemente la estandarización de los resultados. La electroforesis se realiza en geles desnaturalizantes de poliacrilamida (evitando así diferencias de movilidad debidas a conformaciones distintas de la molécula de ADN motivadas por variaciones en la secuencia, las cuales pueden detectarse en geles nativos de poliacrilamida). Los patrones obtenidos se representan por picos, una vez captada la señal fluorescente por el ordenador en tiempo real. Los resultados se obtienen al finalizar el recorrido electroforético (4 horas aproximadamente).

Para la lectura de los productos amplificados correspondientes a los marcadores de cromosoma Y analizados hemos utilizado el ALF *DNA Sequencer* y para la interpretación de los resultados, el programa informático *Fragment Manager 1.2* (Amersham Pharmacia Biotech, Suecia) como ya hemos citado en el apartado correspondiente de material y métodos. Este programa nos permite obtener información sobre el tamaño en pb de los alelos de cada marcador, tomando como referencia el tamaño en pb de unos marcadores de peso molecular conocido (114 y 402 pb). Para la lectura del fenotipo de cada alelo de un sistema se hace imprescindible su comparación con el *ladder* alélico correspondiente, bien por comparación visual directa, trazando líneas que unen los picos correspondientes a los dos *ladders* entre los que discurren las muestras a estudio, o bien utilizando el programa "*allelix*", que hace una interpolación lineal con un fundamento muy similar. Esta determinación es más exacta que la meramente visual, pero también sujeta a errores.

Es muy importante saber reconocer los picos anómalos (que representan bandas electroforéticas) que no corresponden a alelos y que pueden deberse a distintos artefactos. Entre ellos citaremos las bandas tartamudas (*stutter bands*), frecuentes en los STRs dinucleotídicos, debidas a *slippage* de la Taq polimerasa durante el proceso de PCR; los *pull-up*, picos de pequeña intensidad que aparecen en las calles adyacentes a un pico grande, exactamente en la misma posición, que habitualmente se deben a una muestra sobreamplificada; artefactos inespecíficos, que suelen dar una señal de baja intensidad, producidos en general por hibridaciones inespecíficas de los cebadores; otros artefactos electroforéticos como duplicaciones o distorsiones, etc.

Las características de los STRs utilizados en nuestro estudio, junto con un buen rendimiento de la amplificación, hacen que la incidencia de estos fenómenos haya sido mínima, sin llegar a plantear dificultades de interpretación.

La utilización del ALF para analizar STRs de cromosoma Y nos ha permitido procesar 40 muestras en cada gel utilizando solamente 1 μ l de producto amplificado. Dado que los sistemas estudiados, en general, no coinciden en tamaño (aunque hay que tener presente la gran proximidad entre los sistemas DYS19 y DYS390 y que DYS385 puede solaparse parcialmente con DYS389II), existe la posibilidad de combinar estos STRs y detectar los diferentes productos de PCR en la misma calle, bien mezclando los productos de amplificación previamente, o bien optimizando las condiciones experimentales para realizar reacciones simultáneas (*multiplex*).

Aunque en el ALF hubiésemos podido secuenciar perfectamente, preferimos hacerlo en el ABI, ya que para este tipo de análisis este sistema nos brinda mayores ventajas, entre las que destacamos (Tabla 13):

- La posibilidad de analizar mayor número de secuencias por gel (18 frente a 10), ya que para secuenciar en ALF precisamos 4 calles por muestra.
- La reacción de secuencia se realiza en un único tubo, en lugar de cuatro, con el consiguiente ahorro en material y tiempo.

- Es posible hacer la lectura durante la electroforesis, al estar marcados los dideoxinucleótidos terminadores con cuatro diferentes fluorocromos.

Tabla 13. Algunas características diferenciales entre secuenciadores de ADN: ALF y ABI

SISTEMAS	ALF	ABI
Nº de fragmentos analizados	40 por gel	18 por gel
Nº de secuencias analizadas	10 por gel	18 por gel
Nº de calles para cada secuencia	4	1
Tiempo de análisis de secuencia	5 horas	7 horas
Emisión fluorescente	Monocromática	Policromática
Aplicación principal	Análisis de fragmentos	Análisis de secuencias

La labor de secuenciación se ha visto favorecida por el hecho de que, a diferencia de los sistemas autosómicos, sólo tenemos que analizar un alelo por sistema (excepto en los sistemas DYS389 y DYS385 como ya hemos comentado).

2.- ACERCA DE LA SECUENCIA Y LA NOMENCLATURA EMPLEADA

Los STRs que hemos analizado se pueden catalogar como simples y compuestos de acuerdo con la clasificación propuesta por Urquhart *et al* (1994) y Gill *et al* (1997), pero aún no tratándose de STRs complejos sí existen problemas relacionados con la nomenclatura, debido a la existencia de nomenclaturas previas que en algunos casos no siguen las recomendaciones de la ISFG (1997). Aunque la estructura consensuada es fácil de definir, no ocurre lo mismo a la hora de establecer una nomenclatura coherente para todos los STRs.

La mayoría de los sistemas estudiados comprenden secuencias repetitivas constantes y variables, pudiendo definirse los alelos por el número total de *repeats* (constantes y variables) o bien por el número de *repeats* que varían entre individuos.

También se complica la nomenclatura por la presencia de agrupamientos constantes de dinucleótidos y trinucleótidos. Aunque vamos a comentar diferentes nomenclaturas, hemos seguido en nuestro estudio la propuesta por Kayser *et al* (1997) y de Knijff *et al* (1997) por ser la más utilizada por un amplio número de laboratorios forenses de Europa.

Los sistemas con los que hemos trabajado han sido elegidos en función de su alto grado de variabilidad y su capacidad de adaptación para formar parte de multiplexes, siendo definidos como sigue:

- **SISTEMA DYS19** (Tabla 14) - La nomenclatura de los alelos se definió en un principio por el número de repeticiones del motivo variable **TAGA** (10-14), pero para la nomenclatura definitiva usamos la notación 13-17, que tiene en cuenta los *repeats* constantes y variables, para adecuarla a la propuesta por Kayser *et al* (1997) y de Knijff (1997).

Tabla 14. Secuencia del sistema *DYS19*

Estructura consenso:

P1(21pb)- 8pb (at)₅ agtatt (at)₄ agtggt (at)₅ agtggtt (TAGA)₃ tagg (**TAGA**)_n tata 31pb -P2(19pb)

<u>Alelo (pb)</u>	<u>Secuencia</u>
13 (186): P1(21pb)- 8pb (at) ₅ agtatt (at) ₄ agtggt (at) ₅ agtggtt (TAGA) ₃ tagg (TAGA) ₁₀ tata 31pb -P2(19pb)	
14 (190): P1(21pb)-.....	(TAGA) ₁₁-P2(19pb)
15 (194): P1(21pb)-.....	(TAGA) ₁₂-P2(19pb)
16 (198): P1(21pb)-.....	(TAGA) ₁₃-P2(19pb)
17 (202): P1(21pb)-.....	(TAGA) ₁₄-P2(19pb)

- **SISTEMA DYS385** - Este sistema muestra dos alelos que no pueden asignarse a sus dos loci correspondientes por solaparse en tamaño, así que se analiza como un haplotipo. La secuencia nos fue facilitada por PM Schneider (comunicación personal), siendo el tetranucleótido **GAAA** la unidad repetitiva. La nomenclatura utilizada en nuestro estudio es la propuesta por este autor, que también es la seguida por Kayser *et al* (1997) y de Knijff *et al* (1997).

- **SISTEMA DYS389I** (Tabla 15) - En este sistema hay 2 motivos repetitivos: **TCT(A/G)** pero sólo el *repeat TCTA* es variable. Los alelos los denominamos basándonos únicamente en el número de repeticiones del motivo variable (9-12), coincidiendo con los autores anteriormente citados.

Tabla 15. Secuencia del sistema *DYS389I*

Estructura consenso:

P1(25pb)-2pb-(TCTG)₃(TCTA)_n tcctcccTCTAtcaa(TCTA)₂ ttaTCTA 73pb TGTC 43pb-P2(20pb)

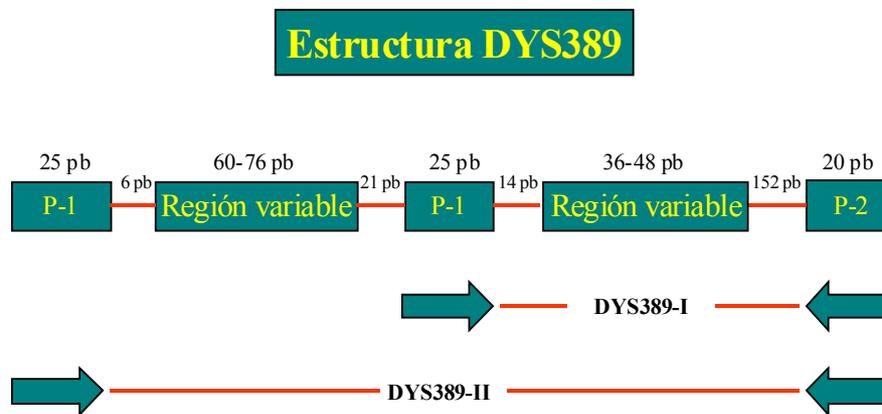
<u>Alelo (pb)</u>	<u>Secuencia</u>
9 (247): P1(25pb)-2pb-(TGTC) ₃ (TCTA) ₉ tcctcccTCTAtcaa(TCTA) ₂ ttaTCTA 73pb TGTC 43pb-P2(20pb)	
10 (251): P1(25pb)-2pb-.....(TCTA) ₁₀-P2(20pb)	
11 (255): P1(25pb)-2pb-.....(TCTA) ₁₁-P2(20pb)	
12 (259): P1(25pb)-2pb-.....(TCTA) ₁₂-P2(20pb)	

- **SISTEMA DYS389II** (Tabla 16) - La secuencia del sistema *DYS389I* está incluida en la secuencia del *DYS389II* (Cooper *et al* 1996). El mismo juego de *primers* amplifica ambos loci simultáneamente porque el primer cebador hibrida en dos lugares diferentes de la misma cadena (Figura 7). La asignación de cada alelo a su locus respectivo es posible porque entre los alelos de cada locus existe una diferencia mínima de 100 pb.

Antes de conocerse la secuencia del sistema, los alelos I y II fueron eluidos en geles de poliacrilamida y reamplificados. No fue posible reamplificar el alelo II en solitario, por lo que recurrimos a otra estrategia que nos permitió solucionar el problema satisfactoriamente: eluyendo el alelo II en un gel de agarosa (cargando 15 µl de producto de PCR), se obtiene suficiente material sin necesidad de reamplificar. Una vez conocida la estructura del sistema *DYS389*, basta con secuenciar el alelo II para obtener las secuencias correspondientes a ambos alelos.

La variación de DYS389II se encuentra a menudo en la región DYS389I, existiendo a veces correlación entre los tamaños de los sistemas I y II. Pero esta circunstancia no siempre está presente y la variación puede originarse en la región específica II, por lo que ambos sistemas pueden incluirse para estudios con haplotipos.

Figura 7. Estructura del sistema DYS389



Considerando la anterior designación alélica de DYS389II, el número de *repeats* de la primera región en los alelos secuenciados es de 15 a 19. Como nos basamos en la nomenclatura propuesta por Kayser *et al* (1997) y de Knijff *et al* (1997), los alelos de DYS389II los designamos como 24-29. Nuevamente estos autores solo tienen en cuenta los *repeats* variables.

Tabla 16. Secuencia del sistema *DYS389II*

Estructura consenso:

P1(25pb)-6pb (TCTG)_n (TCTA)_m 48pb (TCTG)₃ (TCTA)_p 152pb -P2 (20pb)

<u>Alelo (pb)</u>	<u>Secuencia</u>
24 (359): P1(25pb)-6pb (TCTG) ₅ (TCTA) ₁₀ 48pb (TCTG) ₃ (TCTA) ₉ 152pb -P2 (20pb)	
25 (363): P1(25pb)-6pb (TCTG) ₄ (TCTA) ₁₂ 48pb.....(TCTA) ₉ 152pb -P2 (20pb)	
26 (367): P1(25pb)-6pb (TCTG) ₅ (TCTA) ₁₁ 48pb.....(TCTA) ₁₀ 152pb -P2 (20pb)	
27 (371): P1(25pb)-6pb (TCTG) ₅ (TCTA) ₁₁ 48pb.....(TCTA) ₁₁ 152pb -P2 (20pb)	
28 (375): P1(25pb)-6pb (TCTG) ₅ (TCTA) ₁₁ 48pb.....(TCTA) ₁₂ 152pb -P2 (20pb)	
29 (379): P1(25pb)-6pb (TCTG) ₅ (TCTA) ₁₄ 48pb.....(TCTA) ₁₀ 152pb -P2 (20pb)	

BIBLIOTECA VIRTUAL

- ***SISTEMA DYS390*** (Tabla17) - Este STR presenta un motivo de repetición semejante al del sistema *DYS389I/II*, que se debería definir como *TCT(A/G)* siguiendo las directrices de la ISFG (1997). Si nos basamos en el número de *repeats* variables, la designación de alelos iría de 16 a 20. Si se incluye el número de *repeats* invariables habría que emplear otra notación (25.3-29.3) a causa de un *repeat* incompleto, TCA, que es constante en todos los alelos. Sin embargo para coincidir con Kayser *et al* (1997) y de Knijff *et al* (1997), designamos los alelos como 21-25, considerando únicamente los *repeats* que se muestran encuadrados en la siguiente Tabla:

Tabla 17. Secuencia del sistema *DYS390*

Estructura consenso:

P1(23pb)-26pb (TCTA)₂ (TCTG)_n (TCTA)_m TCTG (TCTA)₄ TCA (TCTA)₂ 29pb-P2(23pb)

<u>Alelo (pb)</u>	<u>Secuencia</u>
21 (204): P1(23pb)-26pb (TCTA) ₂	(TCTG) ₈ (TCTA) ₈ TCTG (TCTA) ₄ TCA (TCTA) ₂ 29pb-P2(23pb)
22 (208): P1(23pb)-26pb.....	(TCTG) ₈ (TCTA) ₉29pb-P2(23pb)
23*(212): P1(23pb)-26pb.....	(TCTG) ₈ (TCTA) ₁₀29pb-P2(23pb)
P1(23pb)-26pb.....	(TCTG) ₉ (TCTA) ₉29pb-P2(23pb)
24 (216): P1(23pb)-26pb.....	(TCTG) ₈ (TCTA) ₁₁29pb-P2(23pb)
25 (220): P1(23pb)-26pb.....	(TCTG) ₈ (TCTA) ₁₂29pb-P2(23pb)

*(Al secuenciar el alelo 23 hemos encontrado dos secuencias alternativas)

- **SISTEMA DYS393** (Tabla 18) - Su secuencia muestra 2 motivos de repetición **AGAT-ATGT**, siendo variable únicamente el primero. De acuerdo con el número de *repeats* variables (AGAT), según Kayser *et al* (1997) y de Knijff *et al* (1997), los alelos se denominan 11-15. (Otra posible notación sería 13-17 incluyendo ambos tipos de repeticiones).

Tabla18. Secuencia del sistema *DYS393*

Estructura consenso:

P1(24pb)- (AGAT)_n (ATGT)₂ 17pb -P2(22pb)

Alelo (pb) Secuencia

11 (115): P1(24pb)- (AGAT)₁₁ (ATGT)₂ 17pb -P2(22pb)
 12 (119): P1(24pb)- (AGAT)₁₂.....17pb -P2(22pb)
 13 (123): P1(24pb)- (AGAT)₁₃.....17pb -P2(22pb)
 14 (127): P1(24pb)- (AGAT)₁₄.....17pb -P2(22pb)
 15 (131): P1(24pb)- (AGAT)₁₅.....17pb -P2(22pb)

Como vemos, la nomenclatura más comúnmente aceptada no sigue un criterio unificado, ya que tanto se rige por el número de *repeats* variables como por el de constantes más variables.

Las características de los microsatélites del cromosoma Y empleados en el presente trabajo se resumen en la Tabla19.

Tabla 19. Características de los diferentes sistemas estudiados

Sistema	U. repetición	Localización	Rango en pb ¹	Designación ²	Nº alelos
DYS19	TAGA	Yp	186-202	13-17	5
DYS385	GAAA	Yq	364-402	9-18	24 Haplotipos
DYS389I	TCTA	Yq	247-259	9-12	4
DYS389II	TCT(A/G)	Yq	359-379	24-29	7
DYS390	TCT(A/G)	Yq	204-220	21-25	5
DYS393	AGAT	Yp	115-131	11-15	5

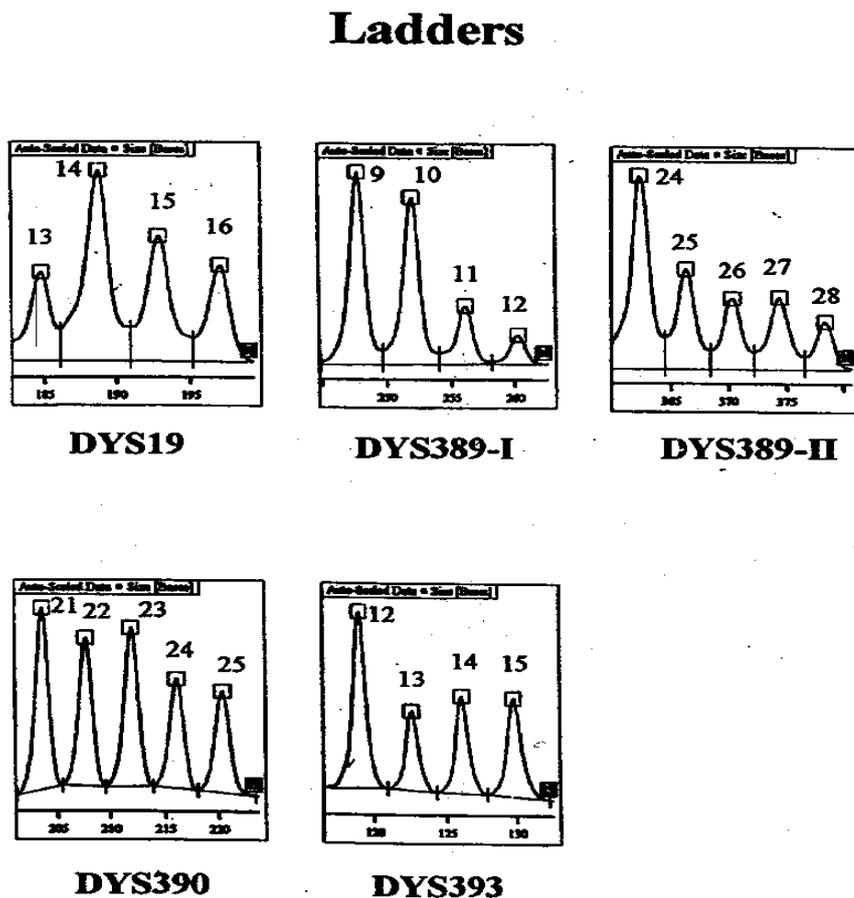
¹ En el rango en pb hemos incluido solamente los alelos de nuestro estudio

² Nomenclatura según Kayser *et al* (1997) y de Knijff *et al* (1997)

3.- LADDERS ALÉLICOS

Para facilitar la estandarización y la correcta lectura de cualquier marcador genético, el uso de escaleras alélicas secuenciadas es un paso inexcusable para garantizar la fiabilidad de los resultados y permitir su comparación entre distintos laboratorios (ISFG 1994, 1997).

Figura 8. Ladders alélicos secuenciados de los STRs estudiados*



*Nomenclatura de acuerdo con Kayser *et al* (1997) y de Knijff *et al* (1997)

Los *ladders* alélicos secuenciados que hemos utilizado para cada sistema (excepto para el DYS385 como ya hemos dicho) los hemos construido con una mezcla de los diferentes alelos encontrados en la población gallega (Figura 8). Empezamos trabajando con un número muy reducido, pero a medida que avanzaba el estudio, cuando aparecía un nuevo alelo procedíamos a secuenciarlo para incluirlo seguidamente en el *ladder*. Nuestro propósito es seguir completando el *ladder* con nuevos alelos, que esperamos identificar en posteriores estudios.

4.- RESULTADOS GENÉTICO-POBLACIONALES

4.1. CARACTERIZACIÓN DE LA POBLACIÓN GALLEGA

Hemos llevado a cabo un análisis genético-poblacional de 116 individuos gallegos no emparentados mediante el estudio de los sistemas DYS19, DYS385, DYS389I/II, DYS390 y DYS393.

En primer lugar hemos determinado las frecuencias alélicas y haplotípicas de estos loci en la citada muestra de población, así como su diversidad génica. En las Tablas 20-25 se muestran las frecuencias calculadas para cada uno de los sistemas, que aparecen representadas gráficamente en las Figuras 9-14.

- **Sistema DYS19.** Hemos encontrado 5 alelos (13-17), siendo el alelo 14 el más frecuente. El valor de diversidad génica calculado para este sistema es de 0.5613 (Figura 9, Tabla 20).

- **Sistema DYS389I.** Con 4 alelos (9-12), tiene como alelo más frecuente el 10 y su diversidad génica, un valor de 0.5553 (Figura 10, Tabla 21).

-Sistema DYS389II. En nuestro estudio está representado por 7 alelos (24-30), con el 26 como alelo modal y 0.6221 como valor de diversidad génica (Figura 11, Tabla 22).

-Sistema DYS390. Presenta 5 alelos (21-25), siendo el 24 el de mayor frecuencia. El valor de la diversidad génica es 0.6221 (Figura 12, Tabla 23).

-Sistema DYS393. Encontramos 5 alelos (11-15), con el 13 como alelo más frecuente. La diversidad génica calculada para este sistema tiene un valor de 0.4331 (Figura 13, Tabla 24).

-Sistema DYS385. Hemos calculado las frecuencias haplotípicas, ya que no es posible separar ambos alelos. Hasta que los datos sobre este sistema no estén más completos sólo podemos abordarlo como haplotipo (Figura 14, Tabla 25).

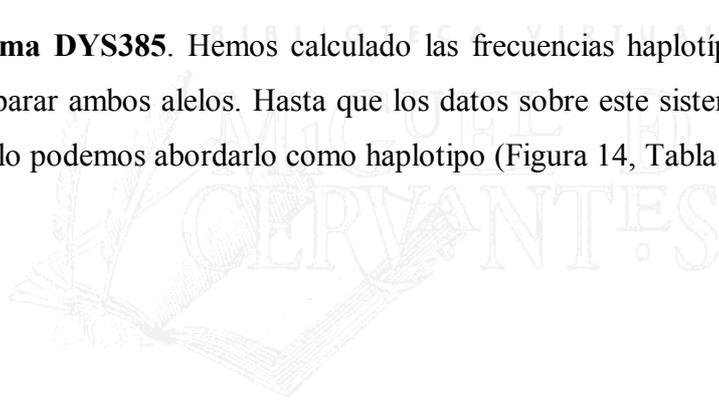
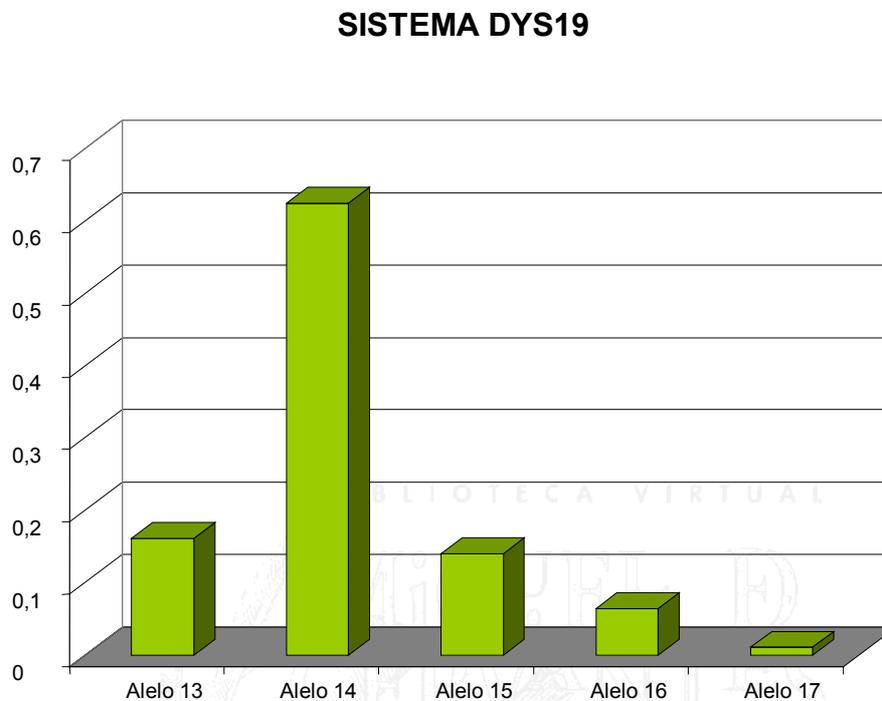


Figura 9. Sistema DYS19: Frecuencias alélicas en población gallega**Tabla 20.** Frecuencias alélicas del sistema DYS19

Locus	Alelos (pb)	Frecuencia	Diversidad. Génica*
DYS19 (n: 93)	13 (186)	0.1613	0.5613
	14 (190)	0.6236	
	15 (194)	0.1398	
	16 (198)	0.0645	
	17 (202)	0.0107	

*Diversidad Génica computada como $1 - \sum p_i^2$ (p_i : Frecuencias alélicas).

Figura 10. Sistema DYS389I: Frecuencias alélicas en población gallega

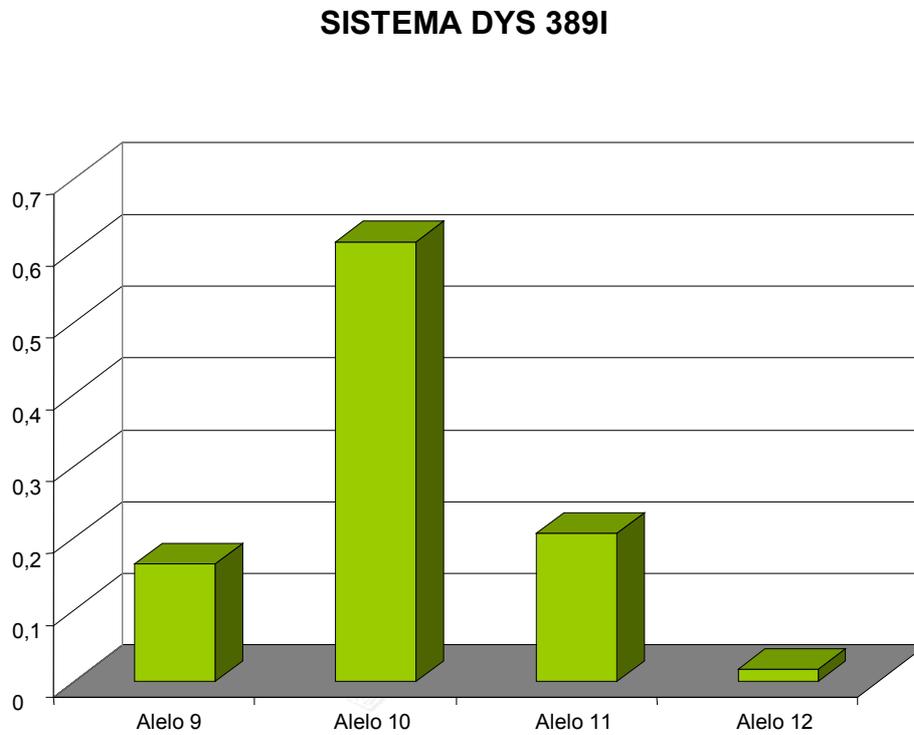
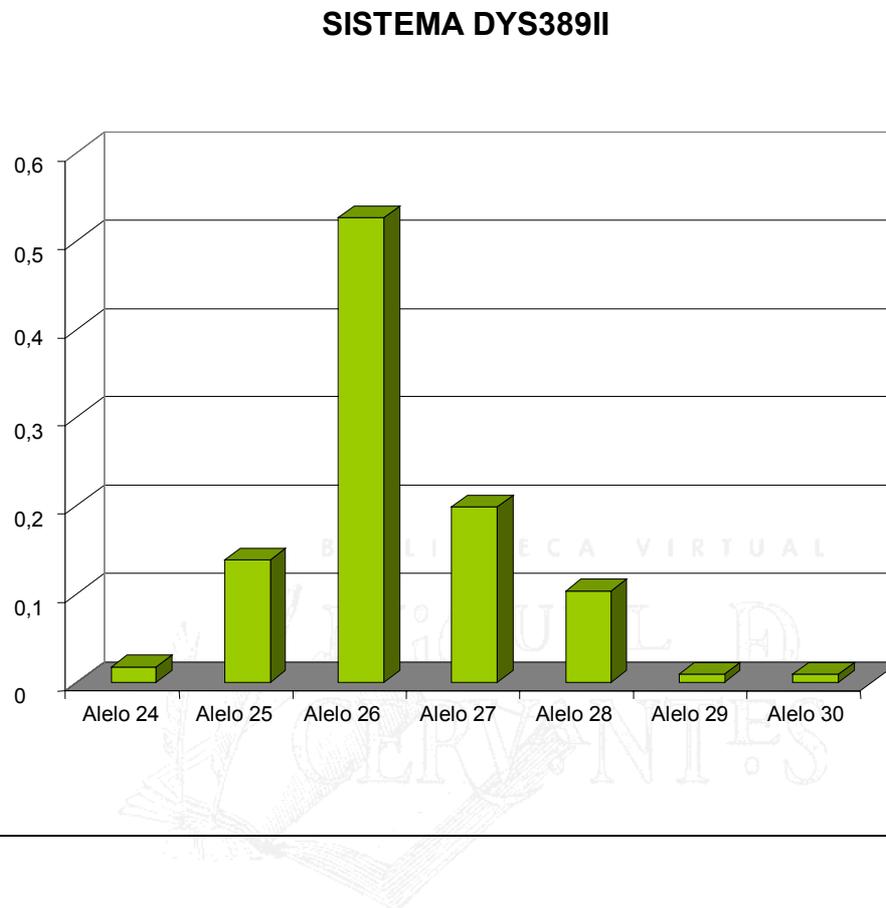


Tabla 21. Frecuencias alélicas del sistema DYS389I

Locus	Alelos (pb)	Frecuencia	Diversidad Génica*
DYS389-I (n: 116)	9 (247)	0.1638	0.5553
	10 (251)	0.6121	
	11 (255)	0.2069	
	12 (259)	0.0172	

*Diversidad Génica computada como $1 - \sum p_i^2$ (p_i : Frecuencias alélicas).

Figura 11. Sistema DYS389II: Frecuencias alélicas en población gallega**Tabla 22.** Frecuencias alélicas del sistema DYS389II

Locus	Alelos (pb)	Frecuencia	Diversidad Génica*
DYS389-II (n: 116)	24 (359)	0.0172	0.6539
	25 (363)	0.1380	
	26 (367)	0.5259	
	27 (371)	0.1983	
	28 (375)	0.1034	
	29 (379)	0.0086	
	30 (383)	0.0086	

*Diversidad Génica computada como $1 - \sum p_i^2$ (p_i : Frecuencias alélicas).

Figura 12. Sistema DYS390: Frecuencias alélicas en población gallega

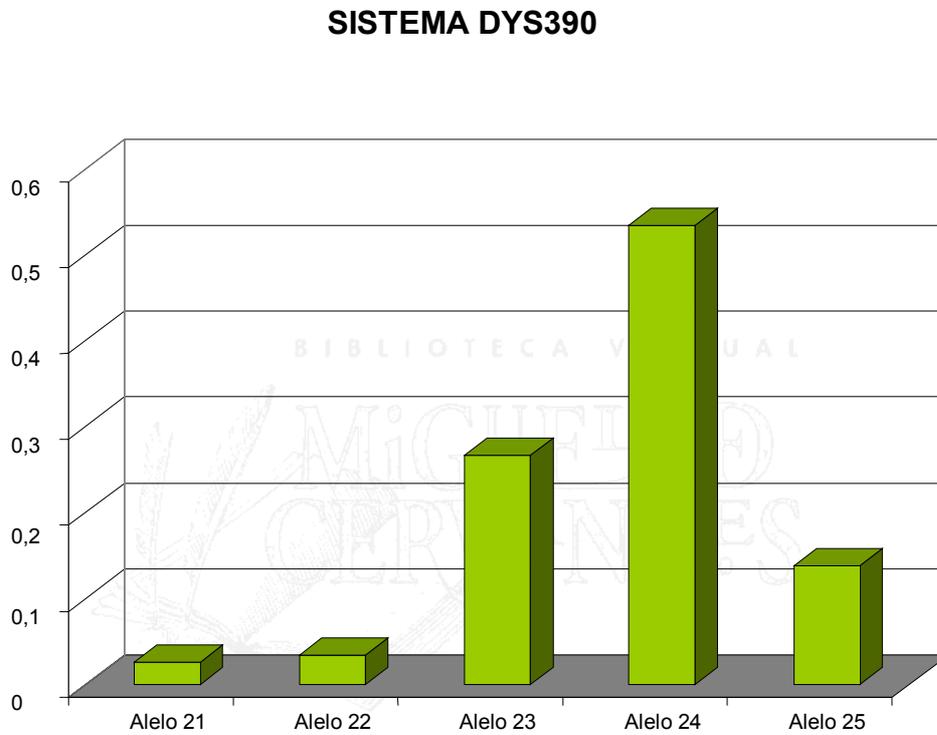
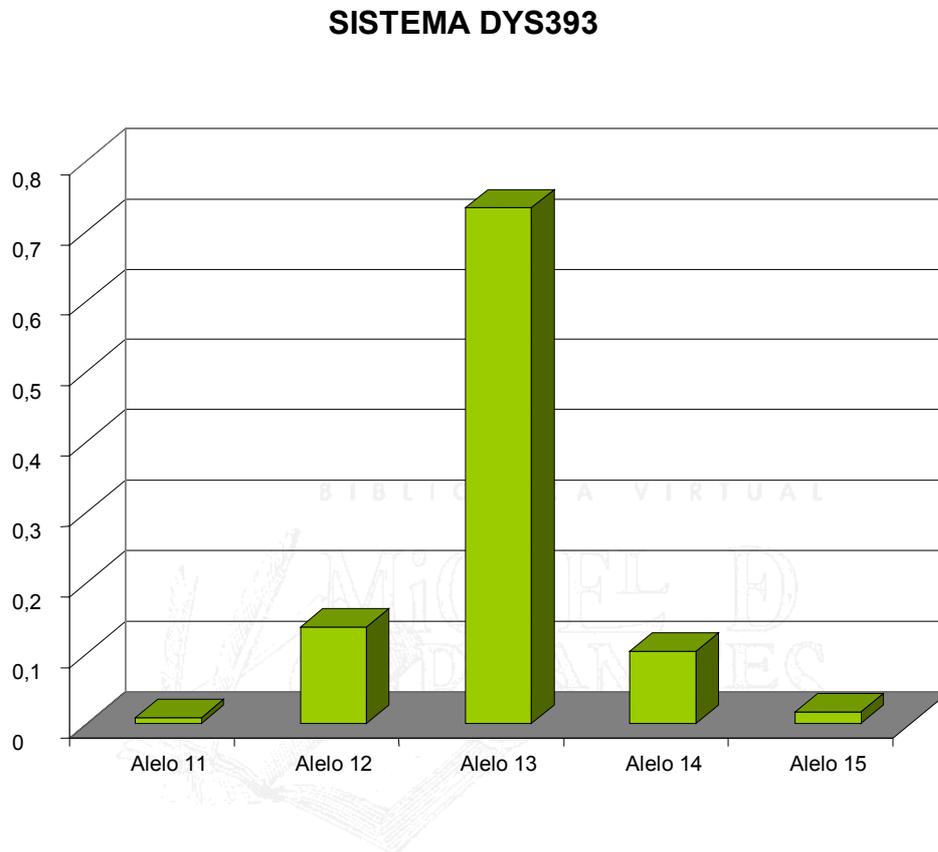


Tabla 23. Frecuencias alélicas del sistema DYS390

Locus	Alelos (pb)	Frecuencia	Diversidad Génica*
DYS390 (n: 116)	21 (204)	0.0259	0.6221
	22 (208)	0.0345	
	23 (212)	0.2672	
	24 (216)	0.5345	
	25 (220)	0.1379	

*Diversidad Génica computada como $1 - \sum p_i^2$ (p_i : Frecuencias alélicas).

Figura 13. Sistema DYS393: Frecuencias alélicas en población gallega**Tabla 24.** Frecuencias alélicas del sistema DYS393

Locus	Alelos (pb)	Frecuencia	Diversidad Génica*
DYS393 (n: 116)	11 (115)	0.0086	0.4331
	12 (119)	0.1380	
	13 (123)	0.7327	
	14 (127)	0.1035	
	15 (131)	0.0172	

*Diversidad Génica computada como $1 - \sum p_i^2$ (p_i : Frecuencias alélicas).

Tabla 25. Frecuencias fenotípicas del sistema DYS385*/**(n=93)

Fenotipo	Frecuencia	Fenotipo	Frecuencia
9-13	0.0108	13-14	0.0322
10-14	0.0108	13-15	0.0860
11-12	0.0108	13-16	0.0215
11-13	0.0322	13-17	0.0108
11-14	0.3226	14-14	0.0322
11-15	0.0215	14-15	0.0215
11-16	0.0430	14-16	0.0430
12-12	0.0430	14-17	0.0215
12-13	0.0215	14-18	0.0108
12-14	0.0645	15-15	0.0322
12-15	0.0645	16-16	0.0108
13-13	0.0215	16-18	0.0108

***Diversidad Haplotípica: 0.8669.**

****Dos loci amplifican simultáneamente.**

DYS385

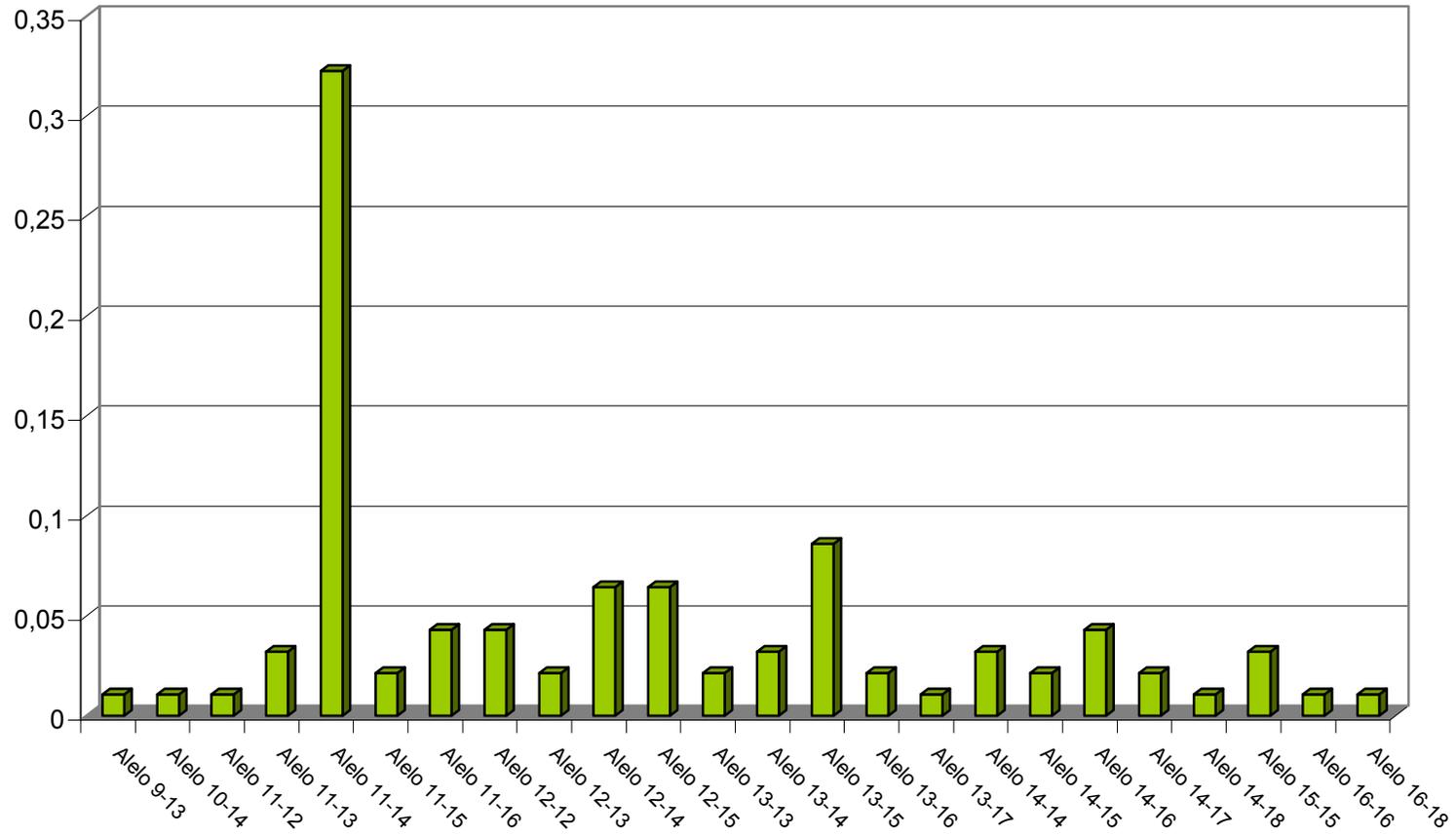


Figura 14. Sistema DYS385: Frecuencias haplotípicas en población gallega

4.2. DETERMINACIÓN DE HAPLOTIPOS

Los anteriores sistemas los hemos agrupado conformando haplotipos en población gallega (Tabla 26), exponiéndolos en orden creciente según el tamaño de los fragmentos de PCR (de izquierda a derecha) con el fin de facilitar el análisis de los datos. Su diversidad haplotípica se muestra en la Tabla 27.

Tabla 26. Haplotipos de STRs de cromosoma Y de muestras de población gallega (n=93)

Haplotipo	n	DYS393	DYS19	DYS390	DYS389-I	DYS389-II	DYS385
1	1	11	13	22	9	25	11-14
2	1	12	13	24	10	26	11-14
3	1	12	13	24	10	26	16-18
4	1	12	14	22	10	26	11-14
5	1	12	14	23	10	26	14-17
6	1	12	14	23	10	27	13-15
7	1	12	14	23	10	27	13-16
8	1	12	14	23	10	29	14-14
9	1	12	14	23	11	28	12-15
10	1	12	14	24	10	25	11-14
11	1	12	14	24	10	27	12-15
12	1	12	15	23	10	26	13-15
13	1	12	15	24	10	26	12-14
14	1	12	15	24	10	28	13-15
15	1	13	13	23	10	27	11-14
16	1	13	13	24	10	28	16-16
17	1	13	13	24	11	27	13-13
18	2	13	13	24	11	27	13-14
19	1	13	13	24	11	27	13-15
20	1	13	13	24	11	27	14-14
21	1	13	13	24	12	28	11-16
22	1	13	13	25	10	26	12-14
23	1	13	13	25	11	28	13-15
24	1	13	13	25	12	28	11-16
25	1	13	14	22	9	26	14-15
26	1	13	14	23	9	25	13-15
27	2	13	14	23	10	26	11-14
28	1	13	14	23	10	26	12-12
29	1	13	14	23	10	26	12-14
30	1	13	14	23	10	26	13-17
31	1	13	14	23	11	27	10-14
32	1	13	14	23	11	27	12-13
33	1	13	14	23	11	27	12-14

Haplotipo	n	DYS393	DYS19	DYS390	DYS389-I	DYS389-II	DYS385
34	1	13	13	24	9	25	11-14
35	1	13	14	24	9	26	11-14
36	1	13	14	24	10	24	11-14
37	1	13	14	24	10	26	11-12
38	2	13	14	24	10	26	11-13
39	9	13	14	24	10	26	11-14
40	1	13	14	24	10	26	11-16
41	1	13	14	24	10	26	12-13
42	1	13	14	24	10	26	12-14
43	1	13	14	24	10	26	13-15
44	1	13	14	24	10	26	14-15
45	1	13	14	24	10	26	14-16
46	1	13	14	24	10	27	11-14
47	1	13	14	24	11	26	11-14
48	2	13	14	24	11	27	11-14
49	1	13	14	24	11	27	11-15
50	1	13	14	24	11	28	11-14
51	2	13	14	25	9	25	11-14
52	1	13	14	25	9	25	12-14
53	1	13	14	25	10	26	9-13
54	1	13	14	25	10	26	11-14
55	1	13	14	25	10	26	11-16
56	1	13	14	25	10	26	12-15
57	1	13	14	25	10	26	13-15
58	1	13	15	22	10	25	12-12
59	1	13	15	23	9	26	13-13
60	1	13	15	23	9	26	14-18
61	1	13	15	23	10	26	13-16
62	1	13	15	24	10	26	11-13
63	1	13	15	24	10	26	11-14
64	1	13	15	24	11	26	11-15
65	1	13	16	21	9	26	14-16
66	1	13	16	23	10	25	12-12
67	1	13	16	25	9	25	14-16
68	1	13	16	25	10	28	11-14
69	1	13	17	24	10	25	12-12
70	1	14	13	24	10	28	13-14
71	1	14	14	23	9	26	14-14
72	1	14	14	24	10	26	11-14
73	1	14	14	24	10	27	12-15
74	1	14	15	23	10	26	14-16
75	1	14	15	23	11	28	15-15
76	1	14	15	23	11	30	15-15
77	1	14	16	21	9	26	14-17
78	1	14	16	23	10	26	15-15
79	2	15	14	24	11	27	12-15
Total	93						

De los sistemas estudiados (excepto DYS385), el que presenta mayor diversidad génica es el DYS389II, correspondiendo la menor al DYS393. Pero dado que estos sistemas se transmiten en bloque, no pueden ser estudiados independientemente, debiendo ser abordados forzosamente como haplotipos cuando se proceda al análisis de dos o más de ellos.

Como vemos en la Tabla 27, a medida que aumenta el número de sistemas que conforman un haplotipo, el valor de PD también se incrementa.

Tabla 27. Valores de Diversidad Haplotípica

	Número de individuos	Haplotipos diferentes	Diversidad haplotípica
Haplotipo 1 (DYS389-I (PD=0.5553), DYS389-II.(PD=0.6539), DYS390 (PD=0.6221) y DYS393 (PD=0.4331))	116	50	0.9214
Haplotipo 1 + DYS385 (PD=0.8669) + DYS19 (PD=0.5613)	93	79	0.9794

Como estrategia para diferenciar haplotipos idénticos en STRs de cromosoma Y puede ser útil el tipado de sustituciones de base. Esta determinación puede generar adicionalmente una importante información para estudios antropológicos.

Por otro lado, hemos analizado 20 muestras femeninas en cada sistema. El resultado de la amplificación ha sido negativo en todos los casos.

También hemos contribuido al estudio de mutaciones analizando 35 parejas padre/hijo con paternidad probada, sin encontrar mutaciones en ninguno de los sistemas. La presencia de mutaciones es poco habitual, detectándose ocasionalmente en algún sistema aislado, como han puesto de manifiesto algunos estudios (Kayser *et al* 1997, Roewer *et al* 1998, Lessig y Edelman 1998).

4.3.- ESTUDIO COMPARATIVO CON OTRAS POBLACIONES

Los resultados obtenidos para los seis STRs de cromosoma Y estudiados en nuestra población, los hemos comparado con los datos publicados referentes a otras poblaciones.

Para facilitar el análisis comparativo hemos tratado de agrupar las poblaciones, en la medida de lo posible, en función de características raciales y de localización geográfica. Las hemos distribuido en tres grupos: poblaciones españolas, poblaciones europeas y un conjunto misceláneo de poblaciones no europeas.

La comparación se ha llevado a cabo mediante el test de contingencia χ^2 , comparando las poblaciones dos a dos para comprobar la homogeneidad de sus frecuencias alélicas o haplotípicas (consideramos significativos aquellos valores en los que $p < 0.01$).

4.3.1.- SISTEMA DYS19

- **Comparación entre la población gallega y otras poblaciones españolas**

En un estudio comparativo de frecuencias alélicas entre muestras de población gallega y otras procedentes de población catalana y vasca (Figura 15, Tabla 28), no hemos apreciado diferencias significativas para este sistema (Tabla 43).

Hemos detectado un mayor número de alelos en nuestra población siendo el más común el alelo 14, al igual que en catalanes y vascos, aunque para estas poblaciones la frecuencia alélica es más baja. El alelo menos representado, el 17, sólo apareció en una ocasión en nuestra casuística, mientras que en la población catalana y vasca no ha sido descrito. Posiblemente la diferente distribución de alelos en nuestra población, respecto a las citadas, habría que buscarla en el tamaño de la muestra, sobre todo para población catalana, en la que sólo hay descritos tres alelos.

Tabla 28. Sistema DYS19: Comparación de frecuencias alélicas Galicia-Población Española

	n	Alelo 13	Alelo 14	Alelo 15	Alelo 16	Alelo 17
Galicia	93	0.1613	0.6236	0.1398	0.0645	0.0107
Cataluña¹	25	0.0400	0.8000	0.1600	0.0000	0.0000
P.Vasco¹	52	0.0770	0.7890	0.0960	0.0380	0.0000

(Las celdillas sombreadas contienen el alelo modal)

• **Comparación entre población gallega y población europea**

Al comparar los datos de nuestra población con los obtenidos de poblaciones europeas (Figura 16, Tabla 29), sí se aprecian diferencias significativas en algún caso (Tabla 43).

Respecto a la población alemana, vemos que se encuentran representados los mismos alelos. Referente a la distribución de sus frecuencias alélicas, los alelos 13 y 14 (alelo modal) presentan una cifra menor en la población alemana, mientras que los alelos 15, 16 y 17 aparecen en general con unas frecuencias más altas.

A pesar de la aparente homogeneidad de los datos en esta población, a la hora de interpretar los resultados encontramos que existen diferencias, significativas o no, según el estudio al que hagamos referencia (Tabla 43). La explicación a esta contradicción podría encontrarse, tal vez, en variaciones relativas a la selección de la muestra o a la metodología empleada en los diferentes estudios. Por ejemplo, Junge *et al* (1998) emplean geles nativos, aunque Lessig y Edelmann (1998), que estudian este sistema bajo condiciones desnaturalizantes y no desnaturalizantes, no observan variación en los resultados .

Al contrastar nuestros datos con los de varios estudios realizados en población holandesa e italiana, no hemos encontrado diferencias significativas en ningún caso (Tabla 43). Con estas poblaciones compartimos los mismos alelos, con la salvedad de que el alelo 10 únicamente aparece descrito en población italiana (caso aislado).

DYS19

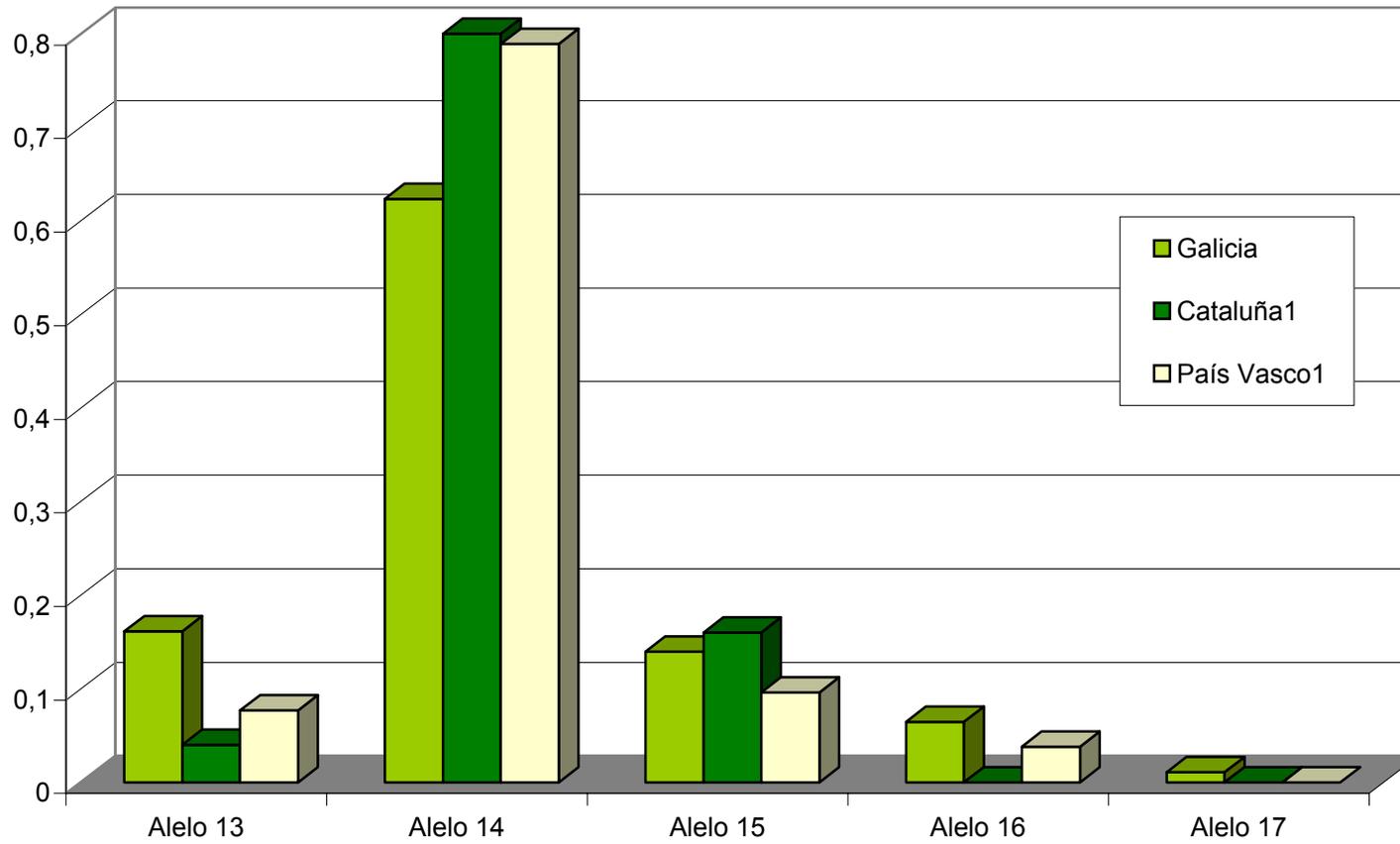


Figura 15. Sistema DYS 19: Comparación de frecuencias alélicas Galicia-Población Española

Tabla 29. Sistema DYS19: Comparación de frecuencias alélicas Galicia-Población Europea.

	n	Alelo 10	Alelo 13	Alelo 14	Alelo 15	Alelo 16	Alelo 17
Galicia	93	0.0000	0.1613	0.6236	0.1398	0.0645	0.0170
Alemania²	109	0.0000	0.0450	0.5960	0.2470	0.0640	0.0450
Alemania³	250	0.0000	0.0560	0.4600	0.2240	0.1720	0.0880
Alemania⁴	53	0.0000	0.0750	0.4910	0.2260	0.1320	0.0750
Alemania⁵	120	0.0000	0.0600	0.5300	0.2300	0.1200	0.0700
Alemania⁶	233	0.0000	0.0700	0.3900	0.2700	0.2100	0.0600
Alemania⁷	70	0.0000	0.1000	0.3710	0.2570	0.2430	0.0290
Holanda⁶	88	0.0000	0.0400	0.7000	0.1900	0.0300	0.0200
Holanda⁷	70	0.0000	0.0450	0.7080	0.1910	0.0340	0.0220
Holanda⁸	110	0.0000	0.0290	0.7280	0.1650	0.0490	0.0290
Noruega⁹	150	0.0000	0.0330	0.5530	0.2470	0.1400	0.0270
Italia⁶	100	0.0100	0.1300	0.5200	0.2600	0.0500	0.0300
Italia¹⁰	114	0.0000	0.0800	0.6100	0.2000	0.0800	0.0300
Italia¹¹	102	0.0000	0.1580	0.5840	0.1980	0.0500	0.0100
Italia¹²	100	0.0000	0.1000	0.6100	0.2200	0.0700	0.0000
Austria¹³	221	0.0000	0.0810	0.4200	0.2200	0.2130	0.0630

(Las celdillas sombreadas contienen el alelo modal)

Al comparar nuestra población con la noruega y austríaca, encontramos en éstas una mayor representación de los alelos 15, 16 y 17 (muy notoria para este último alelo en población austríaca) y una menor frecuencia de los alelos 13 y 14. En esta ocasión sí se producen diferencias significativas con nuestra población (Tabla 43).

Queremos hacer notar que el estudio en población austríaca se ha realizado utilizando condiciones nativas (Glock *et al* 1998). De cualquier manera, sería deseable ampliar el número de estudios para confirmar que existen realmente estas variaciones poblacionales.

• Comparación entre población gallega y poblaciones no europeas

Al comparar nuestra población con poblaciones de otros continentes (Figura 17, Tabla 30), encontramos diferencias sobre todo con población no caucasoide (Tabla 43).

En población caucasoide hallamos diferencias significativas con el pueblo marroquí. En éste aparece como alelo más frecuente el 13, representando casi el 60%

DYS19

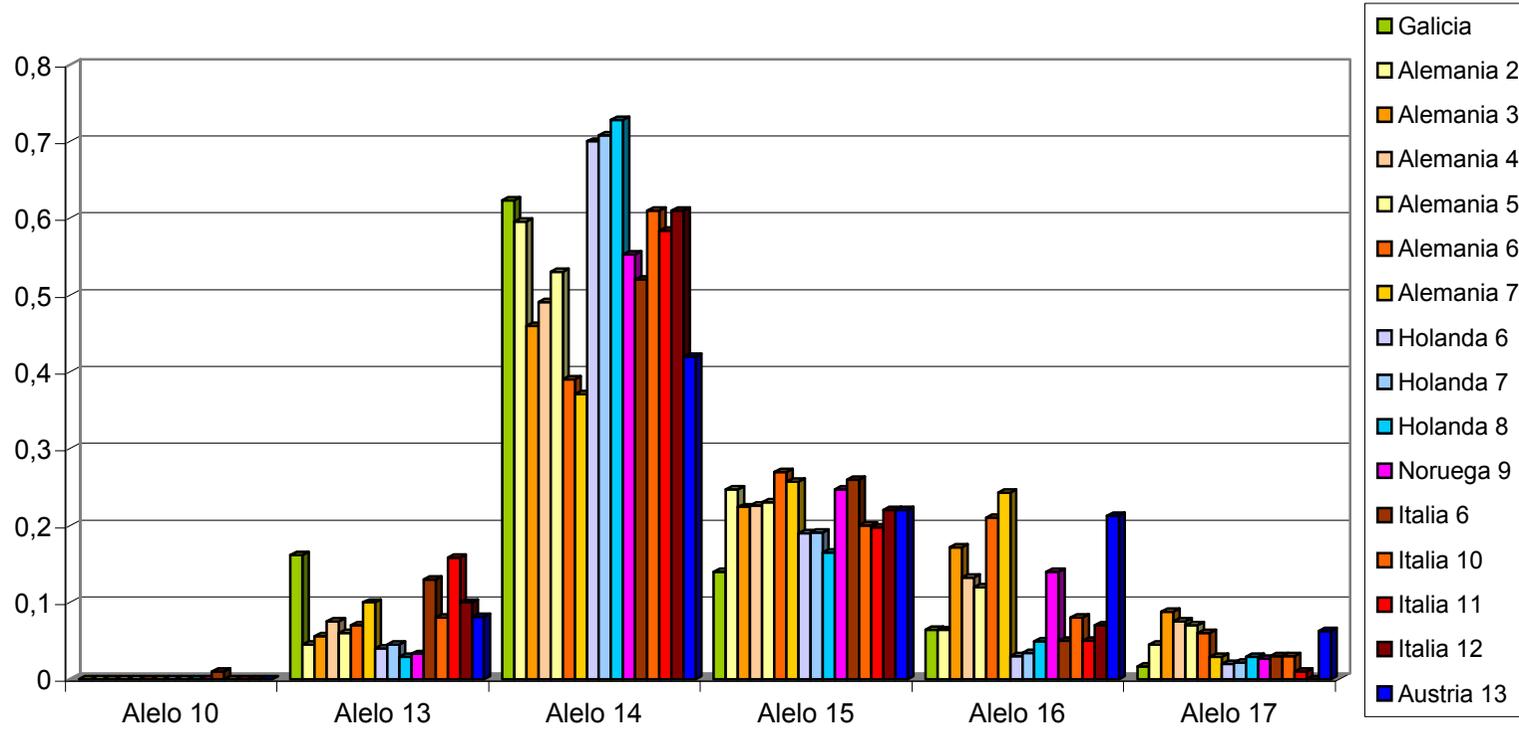


Figura 16. Sistema DYS19: Comparación de frecuencias alélicas Galicia-Población Europe:

del total. Para el resto del grupo, el más representado es el 14 como en nuestra población, aunque con frecuencias más bajas.

Tabla 30. Sistema DYS19: Comparación de frecuencias alélicas Galicia-Población no Europea

	n	Alelo 12	Alelo 13	Alelo 14	Alelo 15	Alelo 16	Alelo 17	Alelo 18
Galicia	93	0.0000	0.1613	0.6236	0.1398	0.0645	0.0170	0.0000
Brasil¹⁴	100	0.0000	0.1900	0.4900	0.2400	0.0700	0.0100	0.0000
B. Aires⁶	100	0.0000	0.2200	0.5600	0.1900	0.0300	0.0000	0.0000
N. York⁶	88	0.0100	0.1800	0.4800	0.1800	0.0900	0.0600	0.0000
Marruecos⁶	41	0.0000	0.5800	0.2400	0.1600	0.0200	0.0000	0.0000
China⁶	36	0.0000	0.0000	0.2500	0.2800	0.3900	0.0500	0.0300
Mongolia⁶	40	0.0000	0.0000	0.2700	0.3200	0.3000	0.0050	0.0050
Japón¹⁵	131	0.0000	0.0530	0.0760	0.4050	0.2290	0.2370	0.0000
Japón¹⁵	120	0.0000	0.0330	0.0750	0.5420	0.2500	0.1000	0.0000
Japón¹⁶	184	0.0000	0.0540	0.0710	0.4570	0.1900	0.2230	0.0050
Taiwan¹⁷	53	0.0000	0.0750	0.3580	0.5470	0.0190	0.00000	0.0000
Inuit⁶	62	0.0000	0.7700	0.1600	0.0300	0.0300	0.0000	0.0000
Surinam⁶	54	0.0200	0.1100	0.2200	0.5000	0.1100	0.0400	0.0000
Yanomami⁶	27	0.0000	0.8100	0.1800	0.0000	0.0000	0.0000	0.0000
Papua⁶	23	0.0000	0.0000	0.1800	0.7700	0.0000	0.0000	0.0400
Trobriands⁶	63	0.0000	0.0600	0.0800	0.5100	0.3300	0.0200	0.0000
Pigmea⁶	31	0.0000	0.0000	0.0600	0.4200	0.2600	0.2600	0.0000

(Las celdillas sombreadas contienen el alelo modal)

Con respecto a poblaciones no caucasoides, como algunas asiáticas y otras pertenecientes a diversas localizaciones geográficas, al compararlas con nuestra población hemos encontrado diferencias significativas en todos los casos (Tabla 43).

El alelo 15 es el más difundido en la población japonesa, pueblo mongol y algunos pueblos indígenas. En China y Taiwan el alelo más frecuente es el 16, mientras que en poblaciones amerindias, como la inuit y la yanomami, predomina el alelo 13, compartido por más del 75% de la población, lo que va en detrimento de su diversidad. El alelo 18 es muy infrecuente, observándose principalmente en pueblos asiáticos, aunque también está descrito en población papua.

DYS19

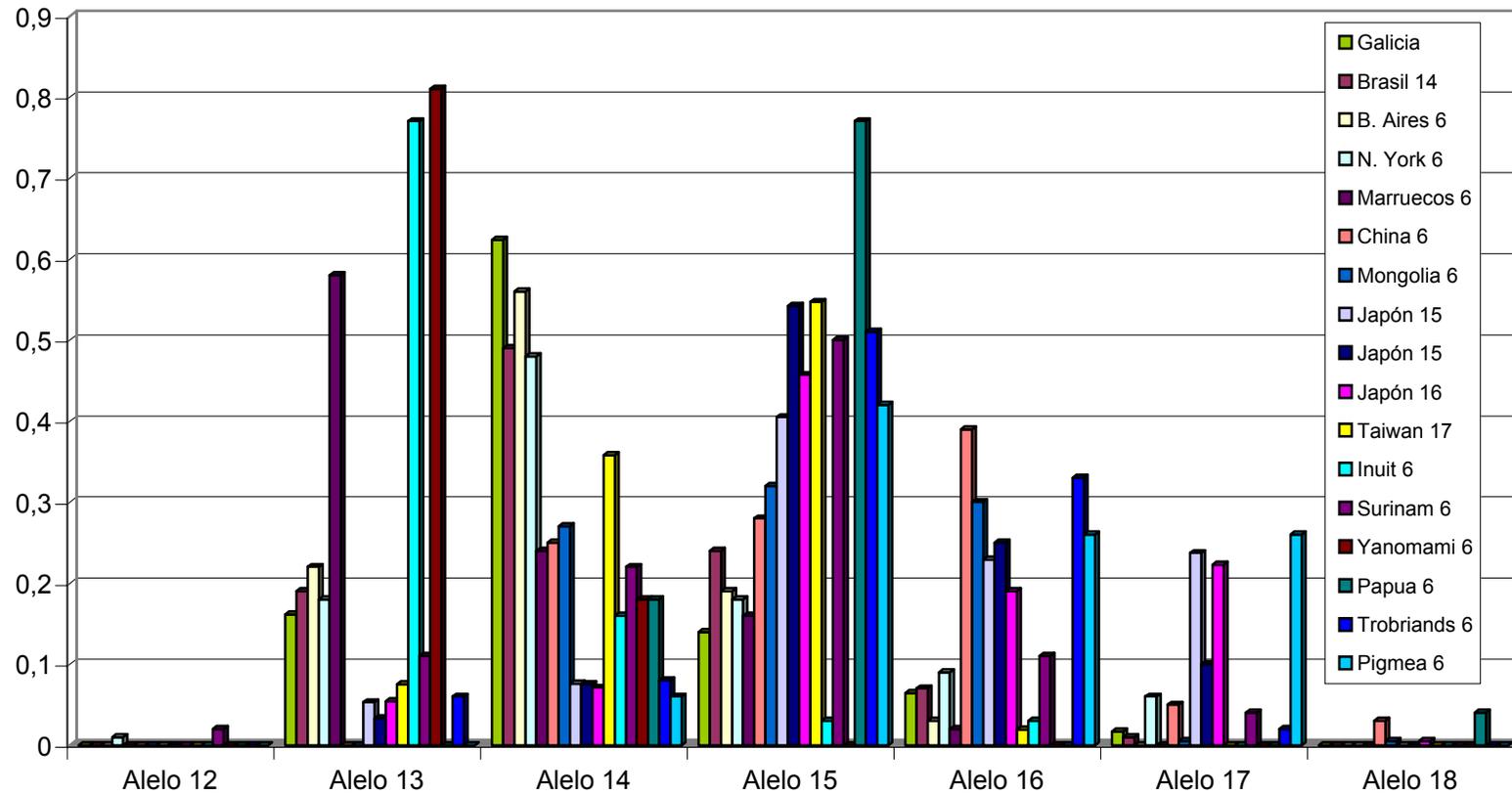


Figura 17. Sistema DYS19: Comparación de frecuencias alélicas Galicia-Población no Europea

4.3.2.- SISTEMA DYS389I

- **Comparación entre población gallega y otras poblaciones españolas**

Hemos encontrado 4 alelos en nuestra población mientras que en la población de Cataluña han sido descritos 3 alelos y 5 en la del País Vasco. En todas ellas el alelo 10 es el más frecuente (Figura 18, Tabla 31). El alelo 8 sólo se ha detectado en población vasca, en la cual está presente el alelo 12 con una frecuencia muy similar a la calculada para nuestra población. Este último alelo no aparece descrito en población catalana, aunque el tamaño de la muestra estudiada es muy pequeño.

Para este sistema no hemos encontrado diferencias significativas entre estas dos poblaciones y la nuestra (Tabla 43).

Tabla 31. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población Española.

	n	Alelo 8	Alelo 9	Alelo 10	Alelo 11	Alelo 12
Galicia	116	0.0000	0.1638	0.6121	0.2069	0.0172
Cataluña¹	31	0.0000	0.0970	0.7420	0.1610	0.0000
P. Vasco¹	56	0.0180	0.1250	0.5340	0.3050	0.0180

(La celdillas sombreadas contienen el alelo modal)

- **Comparación entre población gallega y población europea**

Los alelos 9, 10 y 11 están representados en todas las poblaciones estudiadas (Figura 19, Tabla 32). El alelo 10 es el más frecuente para todas ellas. El alelo 8 está descrito en población alemana y holandesa (caso aislado), mientras que el alelo 12 es compartido por nuestra población y la italiana.

En este sistema no hemos encontrado diferencias significativas (Tabla 43) respecto a la población europea estudiada.

DYS389I

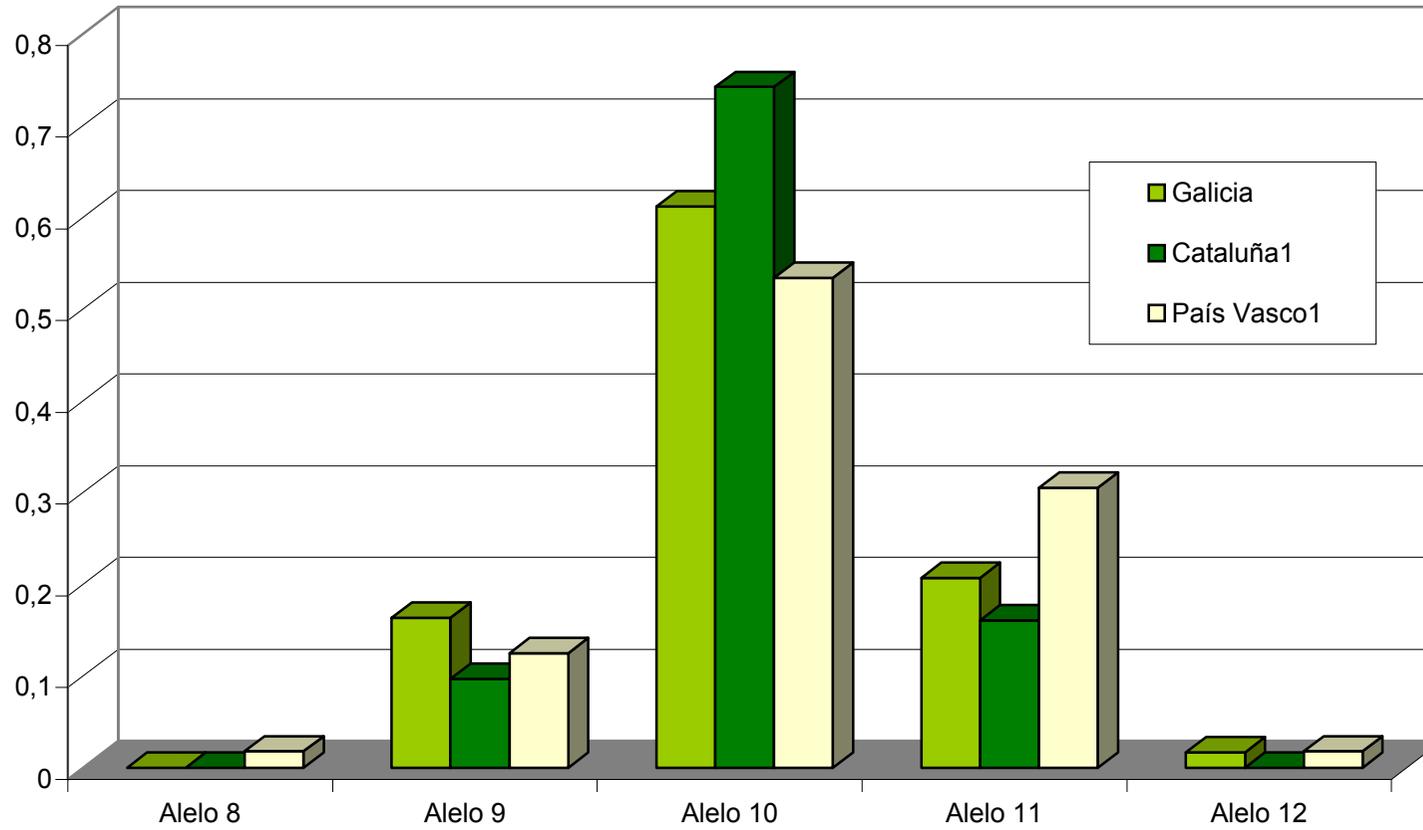


Figura 18. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población Española

Tabla 32. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población Europea.

	n	Alelo 8	Alelo 9	Alelo 10	Alelo 11	Alelo 12
Galicia	116	0.0000	0.1638	0.6121	0.2069	0.0172
Alemania³	250	0.0080	0.1800	0.6880	0.1240	0.0000
Alemania⁵	120	0.0000	0.2500	0.6000	0.1500	0.0000
Alemania⁶	70	0.0000	0.2400	0.5300	0.2300	0.0000
Alemania⁷	70	0.0000	0.2430	0.5290	0.2290	0.0000
Holanda⁶	88	0.0000	0.2700	0.6000	0.1200	0.0000
Holanda⁷	89	0.0000	0.2700	0.6070	0.1240	0.0000
Holanda⁸	110	0.0100	0.2520	0.6120	0.1260	0.0000
Italia⁶	100	0.0000	0.2400	0.5700	0.1800	0.0100
Italia¹²	100	0.0000	0.1600	0.6500	0.1700	0.0200

(La celdillas sombreadas contienen el alelo modal)

•Comparación entre población gallega y población no europea

Los datos para el sistema DYS389I referentes a nuestra población y a otras no pertenecientes al continente europeo se muestran en la Figura 20 y Tabla 33.

Para población caucasoide encontramos de nuevo los alelos 9, 10 y 11 como los más frecuentes. El alelo modal en la población bonaerense e hispana de Nueva York coincide con el de nuestra población (alelo 10), en tanto que para la población de Marruecos está representado por el alelo 11. En ninguna de estas poblaciones está descrito el alelo 8, el alelo 12 no está representado en la población de Buenos Aires y la presencia del alelo 13 se limita al pueblo marroquí.

Sólo hemos encontrado diferencias significativas en este grupo con población marroquí (Tabla 43).

DYS389I

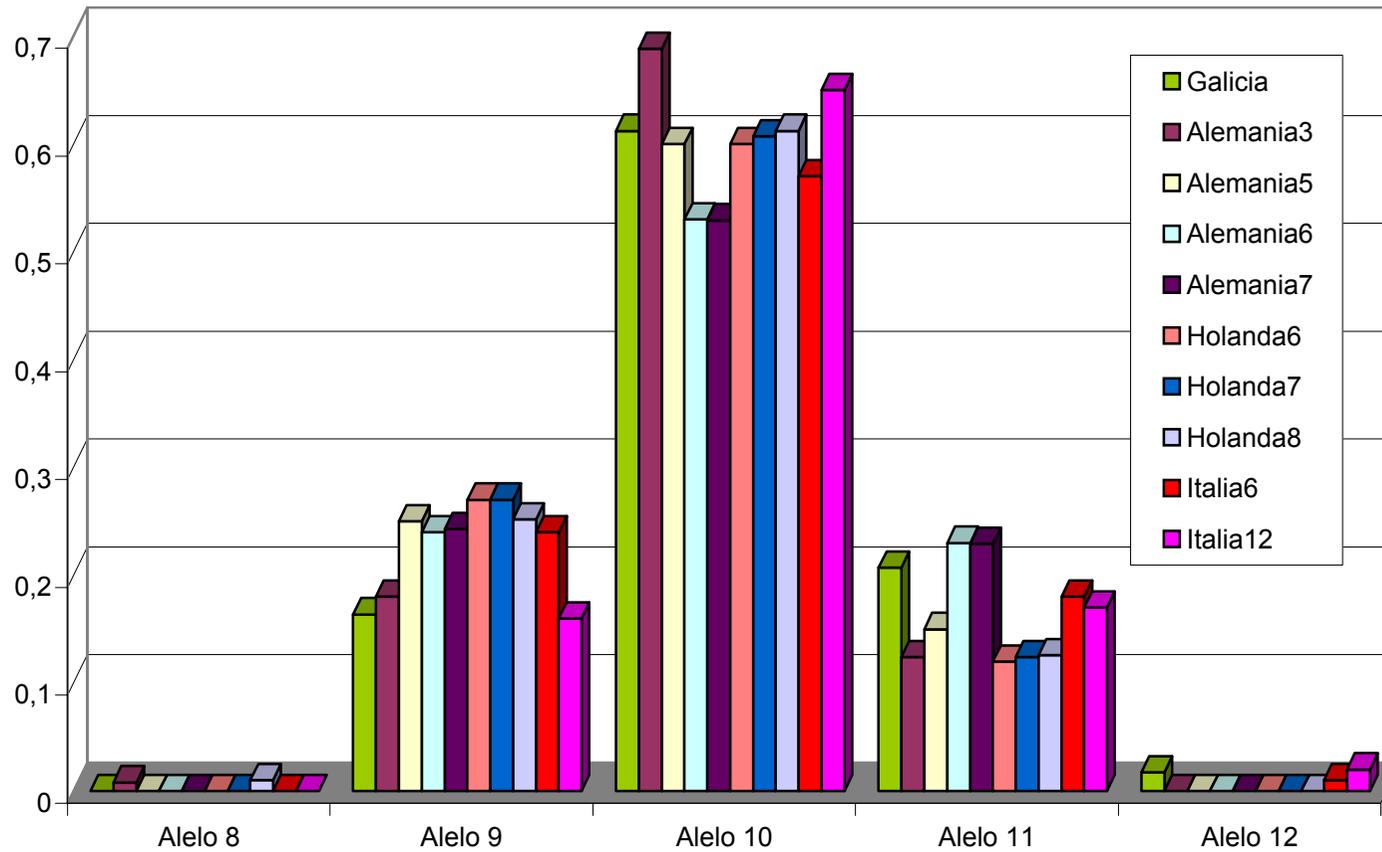


Figura 19. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población Europea

Tabla 33. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población no Europea.

	n	Alelo 8	Alelo 9	Alelo 10	Alelo 11	Alelo 12	Alelo 13
Galicia	116	0.0000	0.1638	0.6121	0.2069	0.0172	0.0000
B.Aires⁶	100	0.0000	0.1100	0.6100	0.2800	0.0000	0.0000
N.York⁶	88	0.0000	0.1400	0.5700	0.2800	0.0100	0.0000
Marruecos⁶	41	0.0000	0.1700	0.3200	0.4400	0.0500	0.0200
China⁶	36	0.0300	0.5100	0.1400	0.3100	0.0000	0.0000
Mongolia⁶	40	0.0200	0.1000	0.7000	0.1700	0.0000	0.0000
Inuit⁶	62	0.0000	0.0800	0.1600	0.6000	0.1400	0.0200
Surinam⁶	54	0.0000	0.1800	0.5400	0.2600	0.0200	0.0000
Yanomami⁶	12	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
Papua⁶	23	0.0000	0.2600	0.5600	0.1700	0.0000	0.0000
Trobriands⁶	63	0.0200	0.3200	0.5100	0.1600	0.0000	0.0000
Pigmea⁶	31	0.1300	0.3200	0.4200	0.0600	0.0600	0.0000

(Las celdillas sombreadas contienen el alelo modal)

En cuanto a población asiática, chinos y mongoles comparten los mismos 4 alelos. Está representado el alelo 8, que no aparece en población gallega, pero no está el 12, que sí está descrito en la nuestra. El alelo modal en la población de Mongolia coincide con el de la gallega, con una frecuencia del 70%, mientras que la población china tiene como alelo más frecuente el 9.

Hemos encontrado que la diferencia es significativa solamente con población china (Tabla 43).

Al comparar nuestra población con varias poblaciones indígenas encontramos que sólo el alelo 10 es común para todas, además es el alelo más frecuente excepto para el pueblo inuit. Trobriands y pigmeos comparten el alelo 8. El alelo 13 únicamente aparece descrito en población inuit. Un caso particular es el pueblo yanomami, en el que todos los integrantes de la exigua muestra están representados por un único alelo.

A pesar de las aparentes diferencias con estas poblaciones indígenas, sólo han resultado ser significativas con los pueblos inuit y pigmeo (Tabla 43).

DYS389I

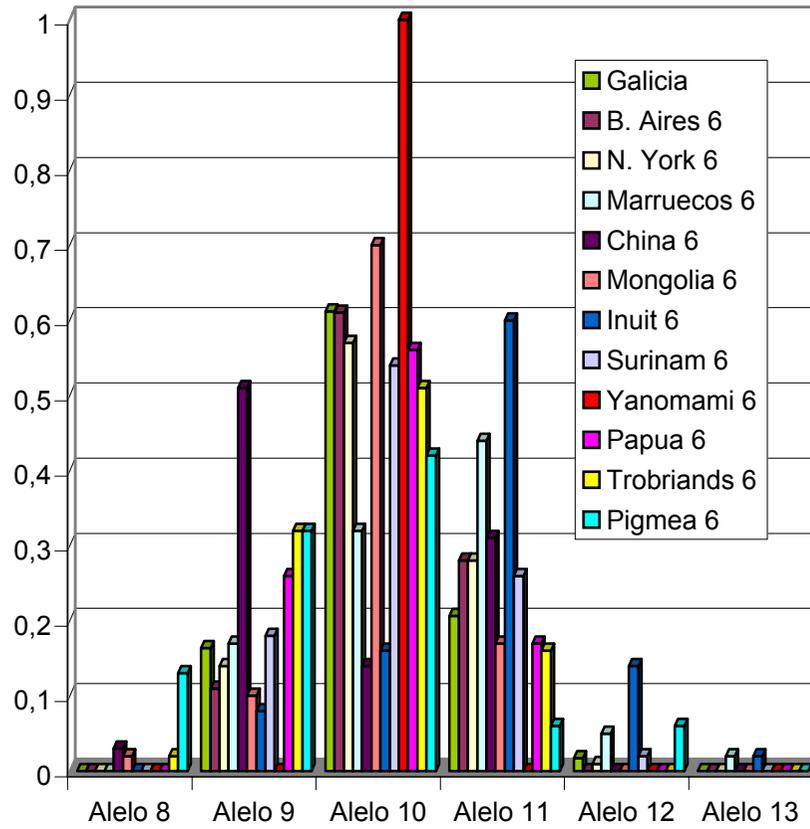


Figura 20. Sistema DYS389I: Comparación de frecuencias alélicas Galicia-Población no Europea

4.3.3.- SISTEMA DYS389II

- **Comparación entre la población gallega y otras poblaciones españolas**

Nuestra casuística está representada por 7 alelos (Figura 21, Tabla 34). En orden de frecuencia, el alelo más común es el 26, al igual que en las otras poblaciones estudiadas. Resulta llamativo el hecho de que en la población vasca el alelo 27 casi iguale en frecuencia al alelo modal y que en la población catalana se hayan descrito tan pocos alelos hasta el momento.

Las diferencias con las poblaciones españolas estudiadas no conllevan diferencias significativas para este sistema (Tabla 43).

Tabla 34. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población Española.

	n	Alelo 24	Alelo 25	Alelo.26	Alelo 27	Alelo 28	Alelo 29	Alelo 30
Galicia	116	0.0172	0.1380	0.5259	0.1983	0.1034	0.0086	0.0086
Cataluña¹	31	0.0000	0.0650	0.5800	0.2260	0.1290	0.0000	0.0000
P. Vasco¹	56	0.0180	0.1960	0.3930	0.3390	0.0180	0.0360	0.0000

(Las celdillas sombreadas contienen el alelo modal)

- **Comparación entre población gallega y población europea**

El número de alelos descritos en población alemana está comprendido entre 5 y 7. Dos de los estudios muestran como alelo modal el 27, aunque su frecuencia está bastante próxima a la del alelo 26 que es el más frecuente en nuestra población y en varias alemanas. En todos los casos aparece el alelo 29 a una frecuencia mucho mayor que en nuestra población.

Los datos para el sistema DYS389II relativos a este grupo se muestran en la Figura 22 y Tabla 35.

DYS389II

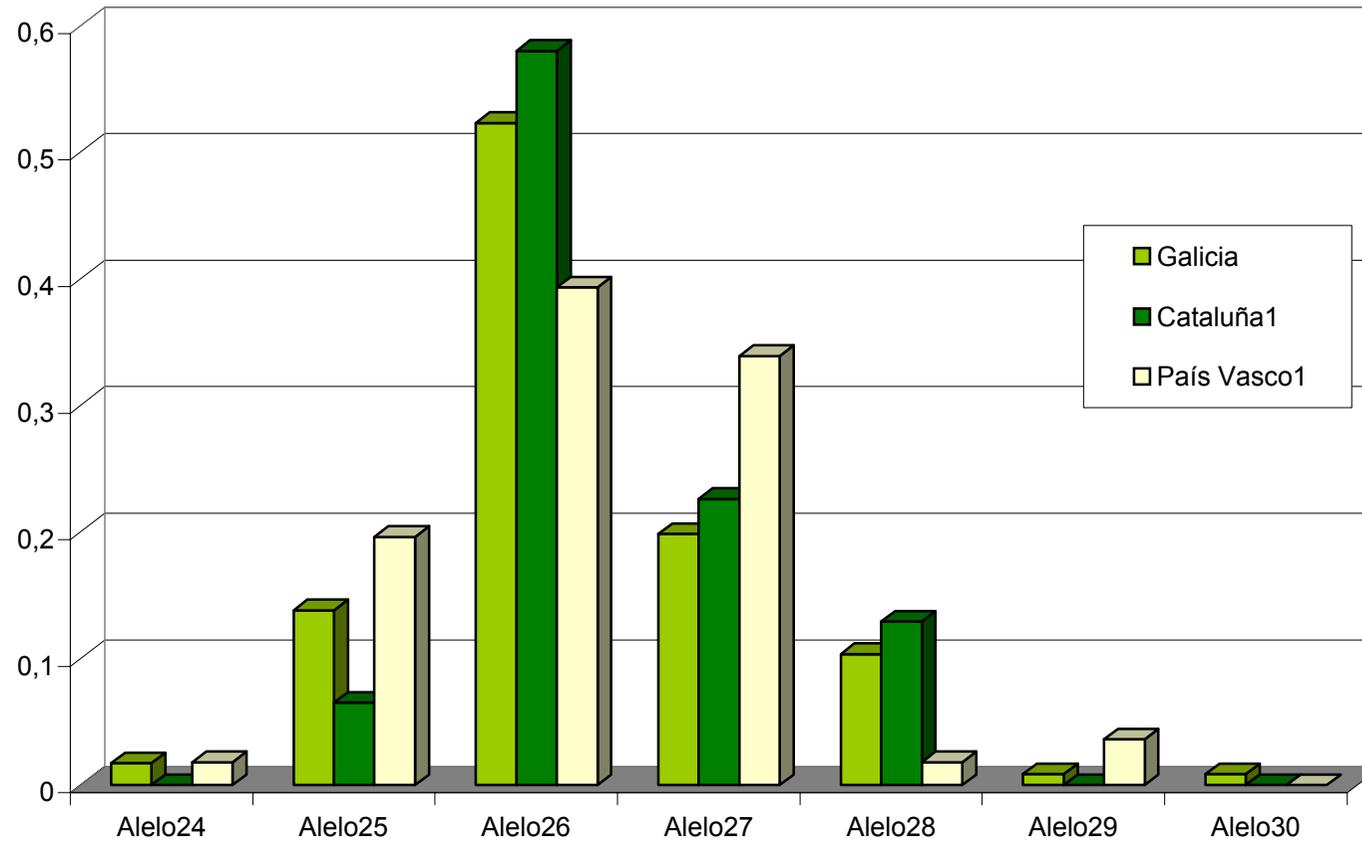


Figura 21. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población española

Dependiendo de los autores, los estudios en población holandesa presentan un número de alelos que oscila entre 4 y 7. Respecto al alelo modal no hay diferencias con nuestra población.

Tabla 35. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población Europea.

	n	Alelo 23	Alelo 24	Alelo 25	Alelo.26	Alelo 27	Alelo 28	Alelo 29	Alelo 30
Galicia	116	0.0000	0.0172	0.1380	0.5259	0.1983	0.1034	0.0086	0.0086
Alemania ³	250	0.0000	0.0160	0.1560	0.3920	0.2800	0.1120	0.0320	0.0120
Alemania ⁵	120	0.0000	0.0200	0.1800	0.3900	0.2800	0.0100	0.0300	0.0080
Alemania ⁶	70	0.0000	0.0000	0.2100	0.2900	0.3300	0.1000	0.0700	0.0000
Alemania ⁷	70	0.0000	0.0000	0.2140	0.2860	0.3290	0.1000	0.0710	0.0000
Holanda ⁶	88	0.0000	0.0000	0.3100	0.4100	0.2500	0.0300	0.0000	0.0000
Holanda ⁷	89	0.0000	0.0110	0.3030	0.4040	0.2470	0.0340	0.0000	0.0000
Holanda ⁸	110	0.0000	0.0190	0.2140	0.4270	0.2520	0.0390	0.0390	0.0100
Italia ⁶	100	0.0100	0.0300	0.1400	0.3600	0.3100	0.1000	0.0400	0.0000
Italia ¹²	100	0.0000	0.0100	0.1400	0.4500	0.3200	0.0600	0.0200	0.0000

(Las celdillas sombreadas contienen el alelo modal)

En población italiana están descritos hasta 7 alelos como en nuestra población, variando el alelo más frecuente dependiendo del estudio al que hagamos referencia. Los alelos 22 y 23 no se han descrito por el momento en población gallega, pero sí el alelo 30 que no aparece representado en población italiana.

Respecto a estas estas poblaciones europeas, no hemos encontrado diferencias significativas con la holandesa o la italiana, pero sí con algunas poblaciones alemanas en función del estudio que hayamos utilizado para la comparación con nuestra población de referencia (Tabla 43).

DYS389II

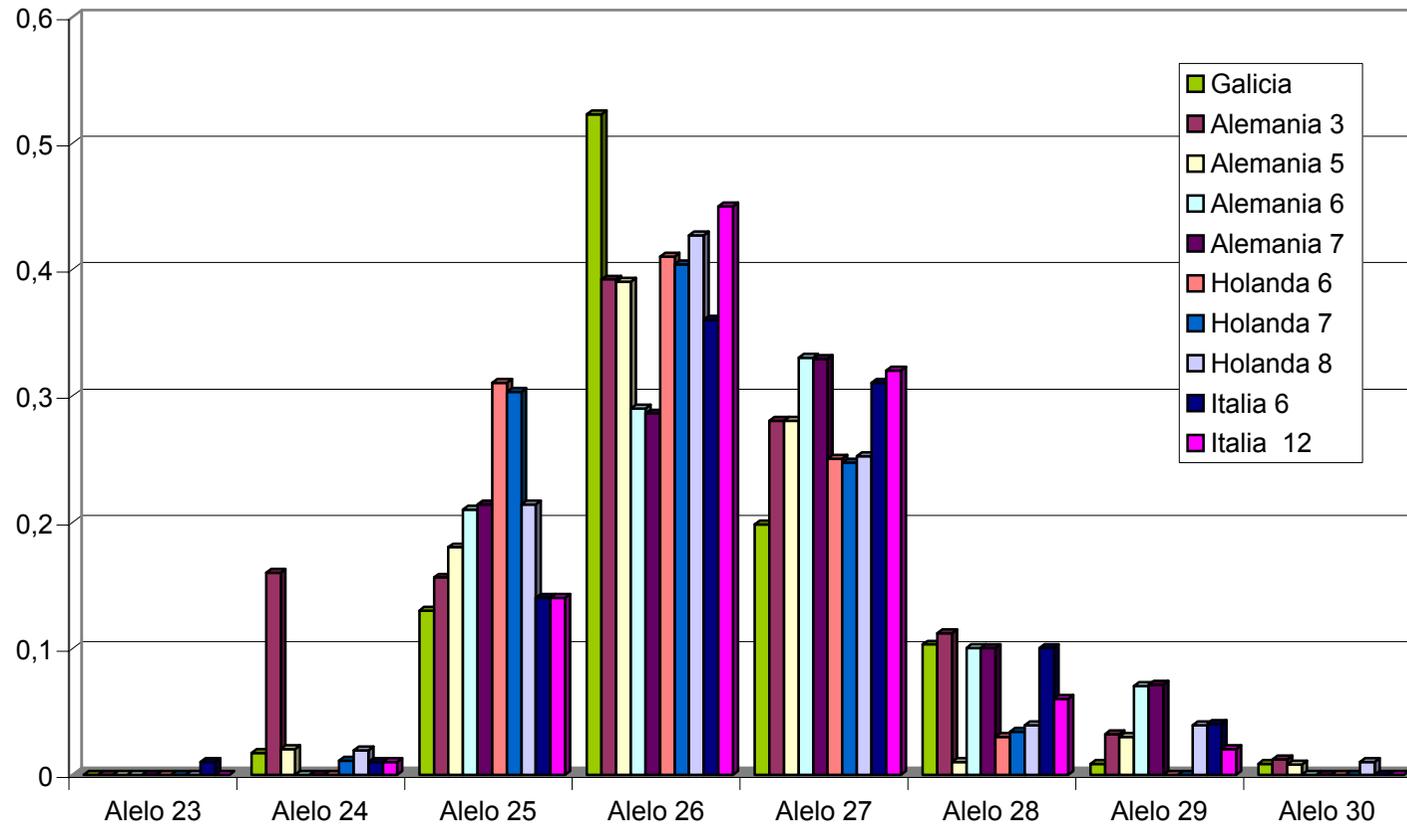


Figura 22. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población Europea

• **Comparación entre población gallega y población no europea**

Los datos relativos a estas poblaciones aparecen representados en la Figura 23 y Tabla 36.

El número de alelos descritos para población caucasoide no perteneciente al continente europeo es menor que en la nuestra. Los alelos 29 y 30 no se han descrito todavía para este grupo (en población marroquí tampoco el 28). El alelo modal coincide con el de la población gallega en la población bonaerense, en la que el alelo 27 casi lo iguala en frecuencia, siendo éste el más frecuente para hispanos de Nueva York y marroquíes.

Tabla 36. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población no Europea

	n	Alelo 23	Alelo 24	Alelo 25	Alelo.26	Alelo 27	Alelo 28	Alelo 29	Alelo 30
Galicia	116	0.0000	0.0172	0.1380	0.5259	0.1983	0.1034	0.0086	0.0086
B.Aires⁶	100	0.0000	0.0000	0.1300	0.3700	0.3200	0.1400	0.0400	0.0000
N.York⁶	88	0.0000	0.0000	0.0700	0.3500	0.4000	0.1500	0.0300	0.0000
Marruecos⁶	41	0.0000	0.0000	0.1900	0.1500	0.5600	0.1000	0.0000	0.0000
China⁶	36	0.0000	0.1100	0.3100	0.1900	0.1900	0.1900	0.0000	0.0000
Mongolia⁶	40	0.0000	0.0500	0.2200	0.4000	0.2000	0.1200	0.0000	0.0000
Inuit⁶	62	0.0000	0.0000	0.0600	0.1400	0.3400	0.2400	0.1900	0.0200
Surinam⁶	54	0.0200	0.0200	0.0600	0.2400	0.3900	0.2000	0.0700	0.0000
Yanomami⁶	12	0.0000	0.0000	0.0000	0.0800	0.5000	0.4200	0.0000	0.0000
Papua⁶	23	0.0000	0.0000	0.0400	0.5200	0.3000	0.0900	0.0000	0.0400
Trobriands⁶	63	0.0000	0.0200	0.1600	0.2700	0.1300	0.3600	0.0600	0.0000
Pigmea⁶	31	0.0000	0.0300	0.3500	0.2600	0.1600	0.0600	0.0600	0.0600

(Las celdillas sombreadas contienen el alelo modal)

Para población caucasoide encontramos diferencias significativas con hispanos de Nueva York y marroquíes. Respecto a población asiática, existen con la población china. Para el restante grupo de poblaciones, encontramos diferencias significativas con todos ellos, excepto con el pueblo papua de Nueva Guinea (Tabla 43).

DYS389II

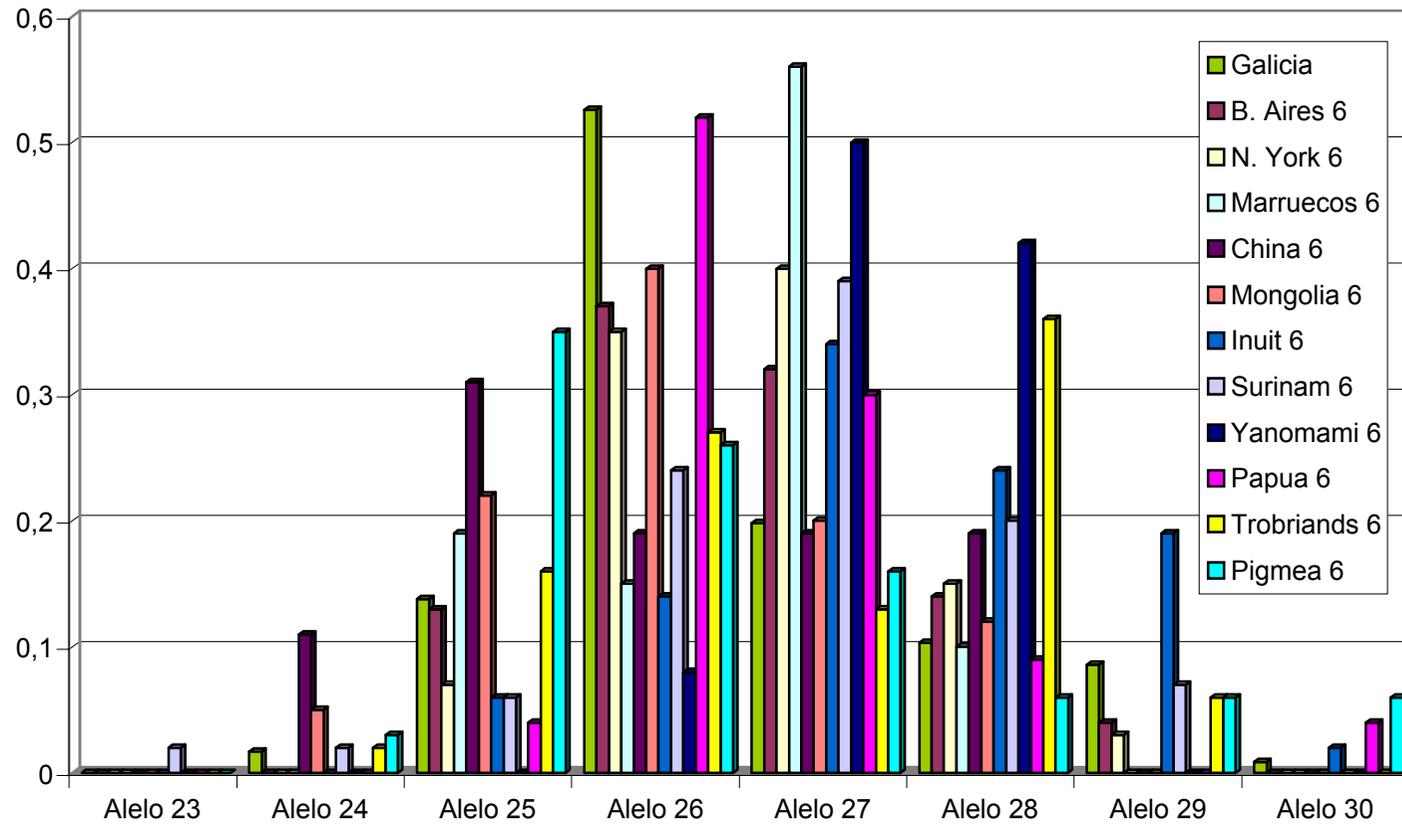


Figura 23. Sistema DYS389II: Comparación de frecuencias alélicas Galicia-Población no Europea

4.3.4.- SISTEMA DYS390

- **Comparación entre población gallega y población española**

Respecto a este sistema (Figura 24, Tabla 37), en nuestra población hemos hallado 5 alelos, número que se reduce para vascos y catalanes. El más frecuente es el 24, dato compartido por los anteriores, aunque con frecuencias superiores.

No hemos encontrado diferencias significativas para este sistema al comparar estas poblaciones con la gallega (Tabla 43).

Tabla 37. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población Española.

	n	Alelo 21	Alelo 22	Alelo 23	Alelo 24	Alelo 25
Galicia	116	0.0259	0.0345	0.2672	0.5345	0.1379
Cataluña¹	29	0.0000	0.0690	0.1720	0.6900	0.0690
P. Vasco¹	53	0.0000	0.0000	0.1700	0.7730	0.0570

(La celdillas sombreada contiene el alelo modal)

- **Comparación entre población gallega y población europea**

Para el sistema DYS390 (Figura 25, Tabla 38) se han descrito, en general, más alelos en población alemana, holandesa e italiana que en la gallega, siendo la holandesa la única europea en la que aparece el alelo 27 (con frecuencia mínima). Los alelos más frecuentes son el 23 (población holandesa) o el 24, coincidiendo en este último caso con el de nuestro medio. En población alemana, puede ser uno u otro en función del estudio. En población italiana puede observarse una distribución bimodal en uno de los estudios.

DYS390

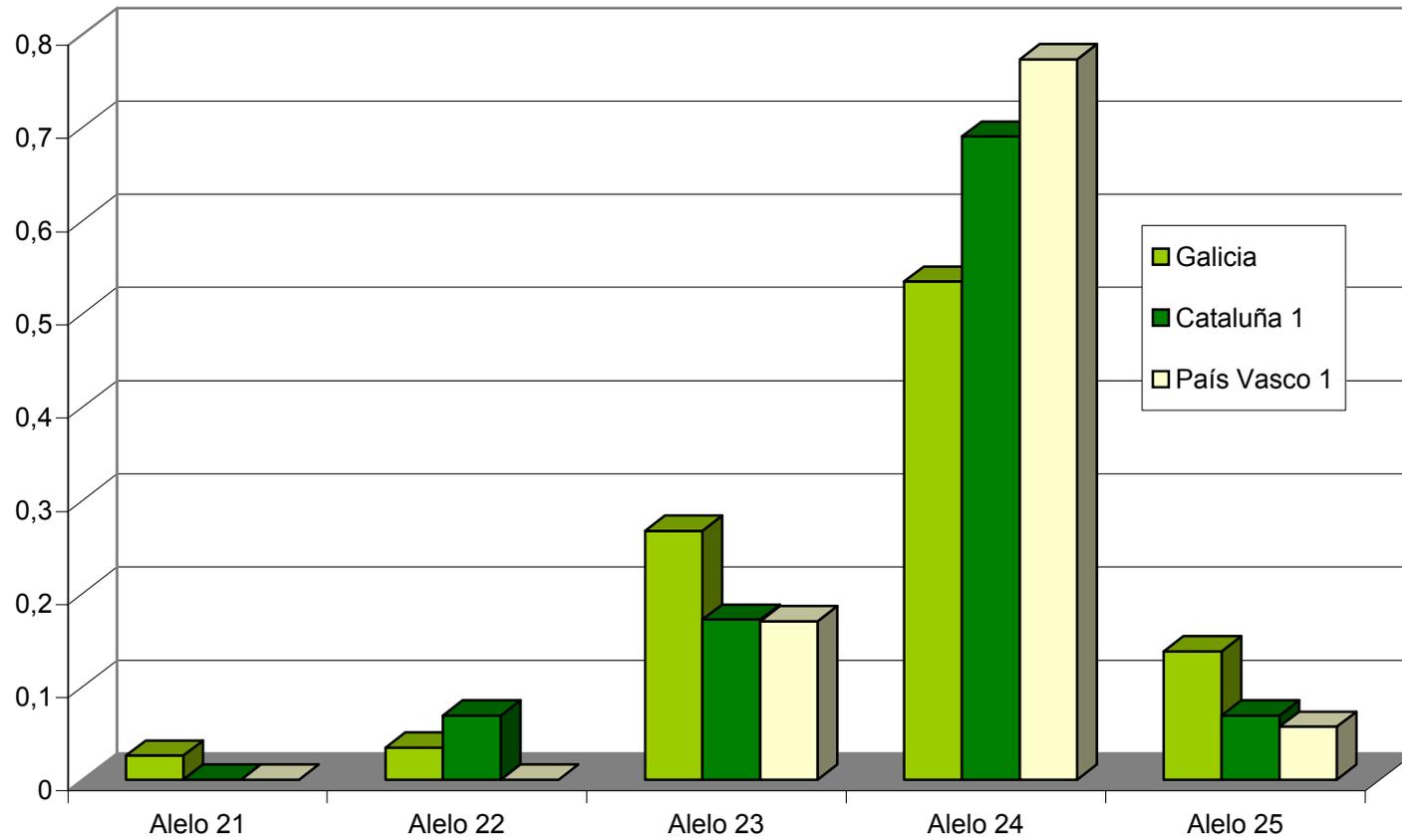


Figura 24. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población Española

Tabla 38. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población Europea.

	n	Alelo 21	Alelo 22	Alelo 23	Alelo 24	Alelo 25	Alelo 26	Alelo 27
Galicia	116	0.0259	0.0345	0.2672	0.5345	0.1379	0.0000	0.0000
Alemania³	250	0.0080	0.0840	0.2560	0.3280	0.3040	0.0200	0.0000
Alemania⁵	119	0.0200	0.2200	0.2000	0.3300	0.2300	0.0080	0.0000
Alemania⁶	70	0.0000	0.0900	0.3600	0.2600	0.2300	0.0700	0.0000
Alemania⁷	70	0.0000	0.0860	0.3570	0.2570	0.2290	0.0710	0.0000
Holanda⁶	88	0.0100	0.1800	0.3900	0.3200	0.0700	0.0200	0.0100
Holanda⁷	89	0.0110	0.1800	0.3820	0.3260	0.0670	0.0220	0.0110
Holanda⁸	110	0.0100	0.1840	0.4080	0.2620	0.1360	0.0000	0.0000
Italia⁶	100	0.0100	0.1500	0.3900	0.3900	0.0500	0.0100	0.0000
Italia¹⁰	114	0.0200	0.1100	0.2600	0.4900	0.1100	0.0100	0.0000
Italia¹²	100	0.0100	0.0600	0.1900	0.6300	0.1100	0.0000	0.0000

(Las celdillas sombreadas contienen el alelo modal)

Comparando nuestra población con las anteriores poblaciones europeas, siempre obtuvimos diferencias significativas con la alemana y la holandesa, mientras que con la población italiana sólo fueron significativas en una ocasión (Tabla 43).

• **Comparación entre población gallega y población no europea**

Cuando comparamos nuestra población con la de otros continentes (Figura 26, Tabla 39), si nos referimos a población caucásica, con la de Buenos Aires y Nueva York compartimos los mismos 5 alelos y con la marroquí solamente 4, al no estar descrito el alelo 21 en esta última. El alelo más frecuente en estas poblaciones es el 24, al igual que en la nuestra.

En población asiática los alelos descritos para la población china coinciden con los de la gallega, siendo el alelo modal diferente. En mongoles se constata la ausencia de los alelos 21 y 26 y la presencia del alelo 27 (este alelo se describió también en población holandesa, como ya hemos visto).

DYS390

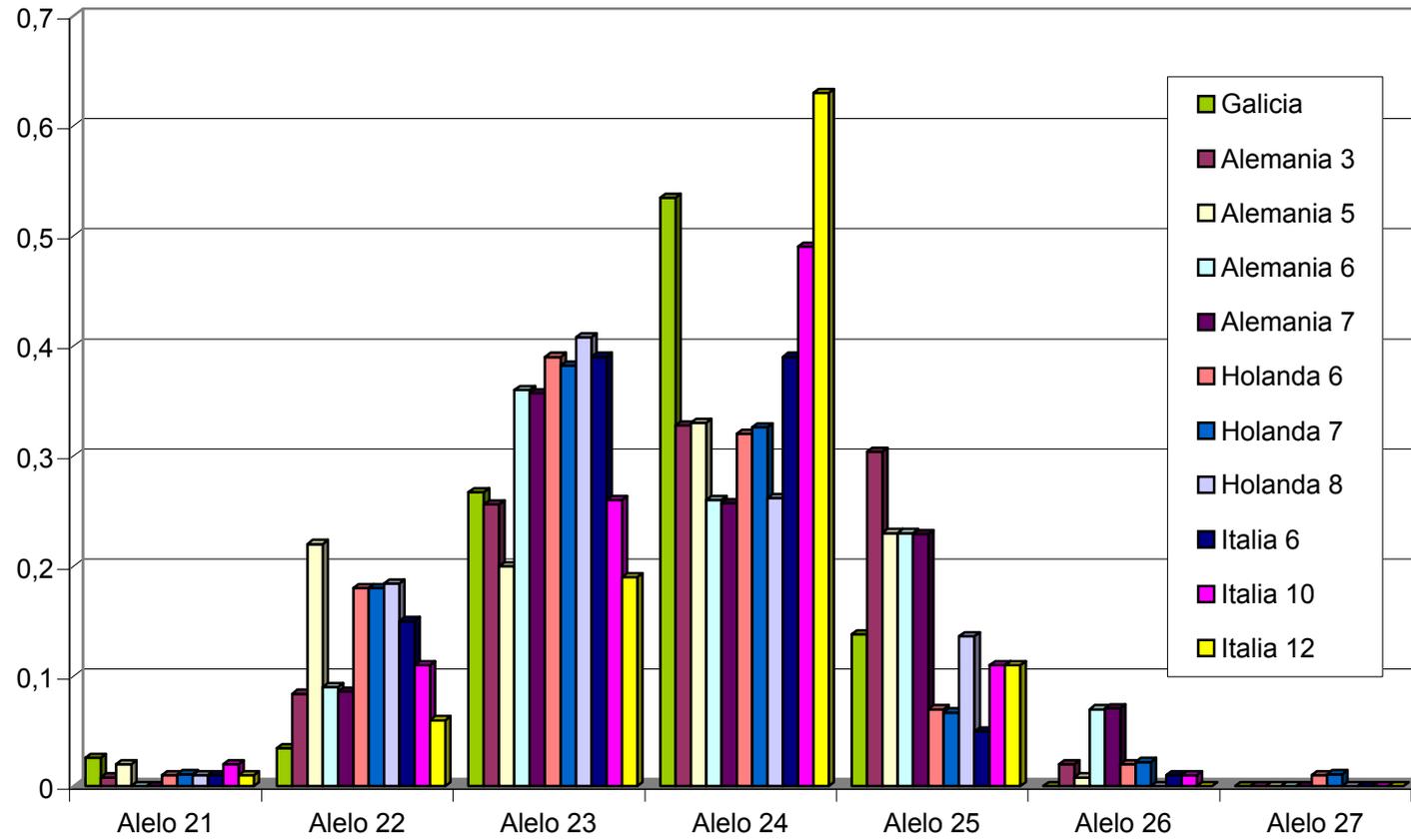


Figura 25. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población Europea

Tabla 39. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población no Europea.

	n	Alelo 20	Alelo 21	Alelo 22	Alelo 23	Alelo 24	Alelo 25	Alelo 26	Alelo 27
Galicia	116	0.0000	0.0259	0.0345	0.2672	0.5345	0.1379	0.0000	0.0000
B Aires⁶	100	0.0000	0.0200	0.1200	0.2800	0.5400	0.0400	0.0000	0.0000
N York⁶	88	0.0000	0.1600	0.0400	0.2500	0.4800	0.0700	0.0000	0.0000
Marruecos⁶	51	0.0000	0.0000	0.1200	0.1800	0.5500	0.1600	0.0000	0.0000
China⁶	36	0.0000	0.0300	0.0500	0.5300	0.2500	0.1300	0.0000	0.0000
Mongolia⁶	40	0.0000	0.0000	0.0700	0.2500	0.3700	0.2700	0.0000	0.0200
Inuit⁶	62	0.0000	0.0000	0.0500	0.0600	0.8600	0.0300	0.0000	0.0000
Surinam⁶	54	0.0000	0.2600	0.1700	0.3000	0.1300	0.1500	0.0000	0.0000
Yanomami⁶	13	0.0000	0.0000	0.8500	0.0800	0.0800	0.0000	0.0000	0.0000
Papua⁶	23	0.0900	0.0000	0.0000	0.0900	0.7000	0.1300	0.0000	0.0000
Trobriands⁶	63	0.0800	0.0200	0.0300	0.3200	0.3300	0.2200	0.0000	0.0000
Pigmea⁶	31	0.0000	0.3200	0.0600	0.2900	0.0600	0.2600	0.0000	0.0000

(Las celdillas sombreadas contienen el alelo modal).

Refiriéndonos a población caucasoide, sólo hallamos diferencias significativas con la población hispana de Nueva York. Respecto a la asiática, no hemos encontrado que éstas sean significativas. Sí existen al comparar nuestra población con poblaciones indígenas, salvo con la papua y trobriand (Tabla 43).

4.3.5.- SISTEMA DYS393

• Comparación entre población gallega y población española

Para este sistema (Figura 27, Tabla 40), en nuestra población hemos observado mayor número de alelos diferentes que en la de Cataluña y País Vasco. Para estas dos poblaciones no están descritos los alelos 11 y 15. El alelo 13 es más común, pero en población gallega su frecuencia es algo menor que en las otras.

Tabla 40. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población Española.

	n	Alelo 11	Alelo 12	Alelo 13	Alelo 14	Alelo 15
Galicia	116	0.0086	0.1380	0.7327	0.1035	0.0172
Cataluña¹	32	0.0000	0.0630	0.8120	0.1250	0.0000
P. Vasco¹	42	0.0000	0.0710	0.8340	0.0950	0.0000

(Las celdillas sombreadas contienen el alelo modal).

DYS390

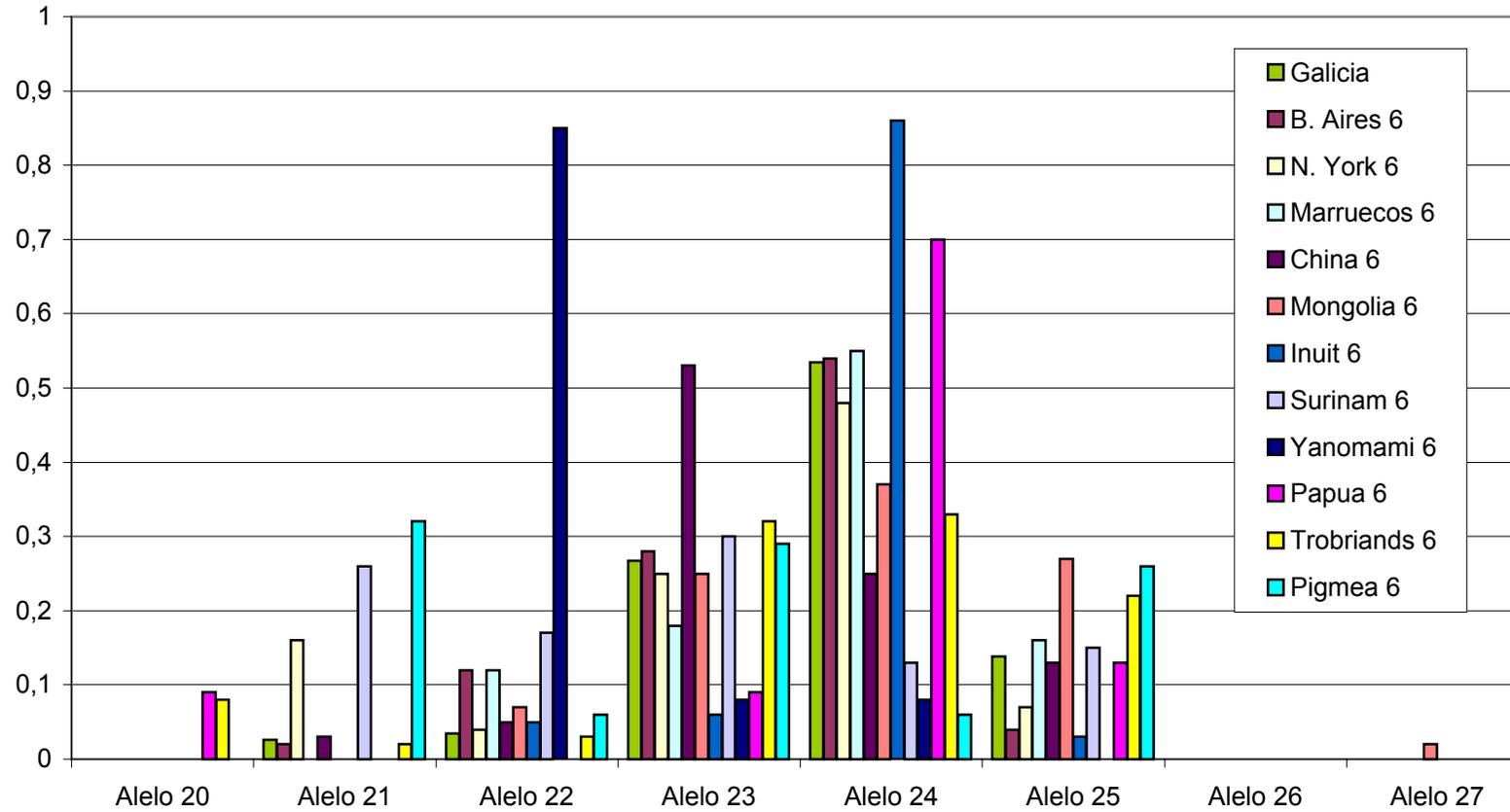


Figura 26. Sistema DYS390: Comparación de frecuencias alélicas Galicia-Población no Europea

Comparando estas poblaciones no hemos encontrado diferencias significativas (Tabla 43).

• **Comparación entre población gallega y población europea**

Como vemos, el alelo más frecuente en todos los estudios para este sistema es el 13 (Figura 28, Tabla 41). La población italiana es la única de las poblaciones europeas estudiadas que comparte el alelo 11 con la nuestra, aunque con una frecuencia mucho mayor

Tabla 41. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población Europea.

	n	Alelo 11	Alelo 12	Alelo 13	Alelo 14	Alelo 15
Galicia	116	0.0086	0.1380	0.7327	0.1035	0.0172
Alemania ²	109	0.0000	0.1190	0.7150	0.1460	0.0180
Alemania ⁶	70	0.0000	0.0400	0.8000	0.1000	0.0600
Alemania ⁷	70	0.0000	0.0430	0.8000	0.1000	0.0570
Holanda ⁶	88	0.0000	0.1100	0.7500	0.1100	0.0200
Holanda ⁷	89	0.0000	0.1120	0.7530	0.1120	0.0230
Holanda ⁸	110	0.0000	0.0290	0.8060	0.1460	0.0190
Italia ⁶	100	0.0300	0.2800	0.5700	0.1100	0.0100
Italia ¹²	100	0.0000	0.1600	0.7300	0.1000	0.0100

(Las celdillas sombreadas contienen el alelo modal).

En el análisis comparativo entre población gallega y europea no hemos encontrado diferencias significativas en ningún caso (Tabla 43).

• **Comparación entre población gallega y población no europea**

Para el sistema DYS393 (Figura 29, Tabla 42), la única referencia al alelo 9 aparece en población caucasoide de Buenos Aires. Respecto al alelo 10, no lo hemos encontrado descrito en ninguna de las poblaciones estudiadas. El alelo 11 está representado en nuestra población exclusivamente. Encontramos solamente un alelo compartido por todas las poblaciones, el 13, que además constituye el alelo modal

DYS393

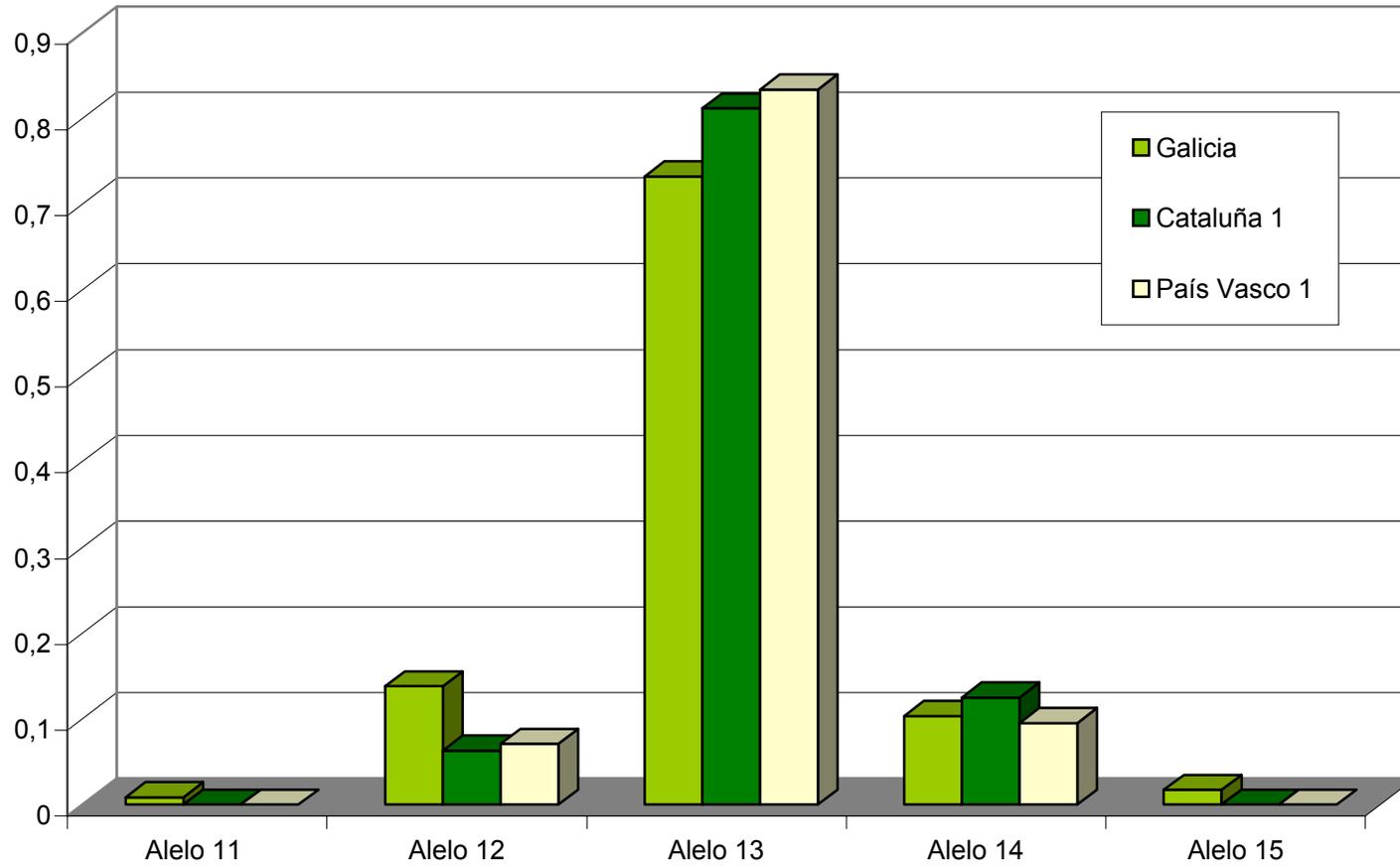


Figura 27. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población Española

para la mayoría. Particularizando, alcanza su frecuencia máxima entre los yanomami, ya que es el único alelo descrito en esta población para este sistema.

Tabla 42. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población no Europea

	n	Alelo 9	Alelo 11	Alelo 12	Alelo 13	Alelo 14	Alelo 15
Galicia	116	0.0000	0.0086	0.1380	0.7327	0.1035	0.0172
B.Aires⁶	100	0.0200	0.0000	0.1600	0.6600	0.1400	0.0200
China⁶	36	0.0000	0.0000	0.5300	0.3100	0.1400	0.0300
Mongolia⁶	40	0.0000	0.0000	0.2200	0.6200	0.1500	0.0000
Inuit⁶	62	0.0000	0.0000	0.0000	0.4800	0.5200	0.0000
Surinam⁶	54	0.0000	0.0000	0.1800	0.4600	0.3100	0.0400
Yanomami⁶	13	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
Papua⁶	23	0.0000	0.0000	0.0000	0.5600	0.3500	0.0900
Trobriands⁶	63	0.0000	0.0000	0.0200	0.8600	0.1300	0.0000
Pigmea⁶	31	0.0000	0.0000	0.0600	0.4200	0.3500	0.1600

(Las celdillas sombreadas contienen el alelo modal).

En la comparación efectuada con estas poblaciones no europeas, existen diferencias significativas respecto a la nuestra, salvo con bonaerenses, mongoles y pueblos yanomami y trobriand (Tabla 43).

A continuación, en la Tabla 43, se muestran los valores que hemos obtenido (significativos o no) en el estudio comparativo de la población gallega con las poblaciones seleccionadas, para los diferentes sistemas analizados.

DYS393

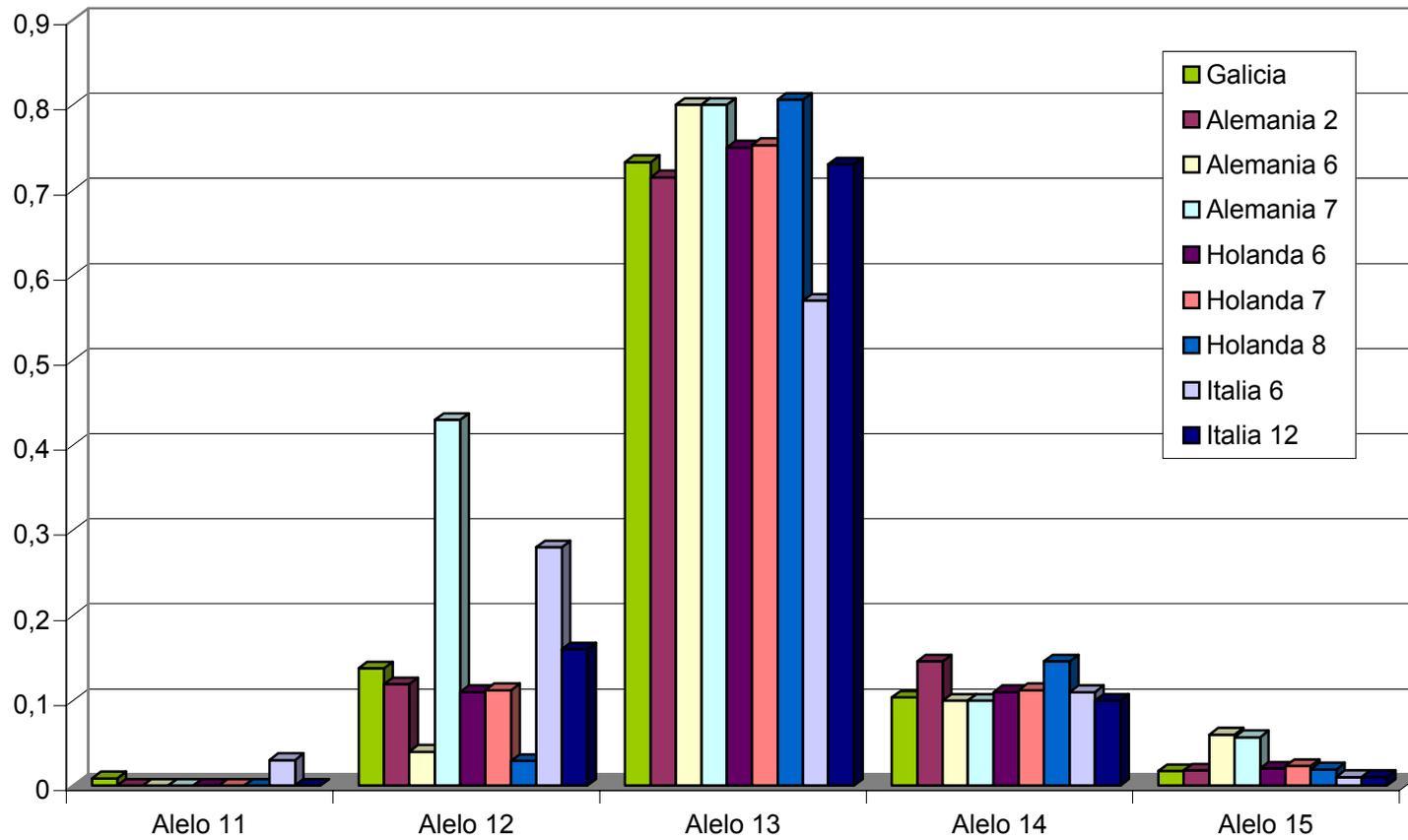


Figura 28. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población Europea

Tabla 43. Comparación Población gallega-Otras poblaciones: Valor de p para diferentes sistemas

Población	Valor de p^*				
	DYS19	DYS389I	DYS389II	DYS390	DYS393
Galicia-Cataluña ¹	0.4042	0.5375	0.8862	0.3882	0.6769
Galicia-P.Vasco ¹	0.6132	0.3664	0.0879	0.0370	0.6177
Galicia-Alemania ²	0.0173	—	—	—	0.7419
Galicia-Alemania ³	0.0000	0.0464	0.2831	0.0002	—
Galicia-Alemania ⁴	0.0170	—	—	—	—
Galicia-Alemania ⁵	0.0038	0.1532	0.0131	0.0001	—
Galicia-Alemania ⁶	0.0000	0.3566	0.0077	0.0003	0.1420
Galicia-Alemania ⁷	0.0002	0.3566	0.0077	0.0003	0.1420
Galicia-Holanda ⁶	0.1469	0.0980	0.0193	0.0003	0.8886
Galicia-Holanda ⁷	0.1444	0.0990	0.0332	0.0004	0.8843
Galicia-Holanda ⁸	0.0442	0.1201	0.1448	0.0000	0.0341
Galicia-Noruega ⁹	0.0006	—	—	—	—
Galicia-Italia ⁶	0.1243	0.5459	0.1539	0.0018	0.0602
Galicia-Italia ¹⁰	0.2578	—	—	0.2528	—
Galicia-Italia ¹¹	0.6541	—	—	—	—
Galicia-Italia ¹²	0.2797	0.5815	0.7839	0.4113	0.8701
Galicia-Austria ¹³	0.0000	—	—	—	—
Galicia-Brasil ¹⁴	0.1122	—	—	—	—
Galicia-Nueva York ⁶	0.0956	0.6217	0.0081	0.0093	—
Galicia-B.Aires ⁶	0.2045	0.2576	0.0743	0.0268	0.4956
Galicia-Marruecos ⁶	0.0000	0.0046	0.0002	0.1508	—
Galicia-China ⁶	0.0000	0.0000	0.0032	0.0296	0.0000
Galicia-Mongolia ⁶	0.0000	0.2950	0.6196	0.0806	0.4794
Galicia-Japon ¹⁵	0.0000	—	—	—	—
Galicia-Japon ¹⁵	0.0000	—	—	—	—
Galicia-Japon ¹⁶	0.0000	—	—	—	—
Galicia-Taiwan ¹⁷	0.0000	—	—	—	—
Galicia-Inuit ⁶	0.0000	0.0000	0.0000	0.0003	0.0000
Galicia-Surinam ⁶	0.0000	0.8238	0.0008	0.0000	0.0039
Galicia-Yanomami ⁶	0.0000	0.0664	0.0044	0.0000	0.3340
Galicia-Papuas ⁶	0.0000	0.6649	0.5869	0.0104	0.0040
Galicia-Trobriands ⁶	0.0000	0.0713	0.0002	0.0101	0.0599
Galicia-Pigmeos ⁶	0.0000	0.0001	0.0055	0.0000	0.0000

*En negrita valores significativos ($p < 0.01$)

DYS393

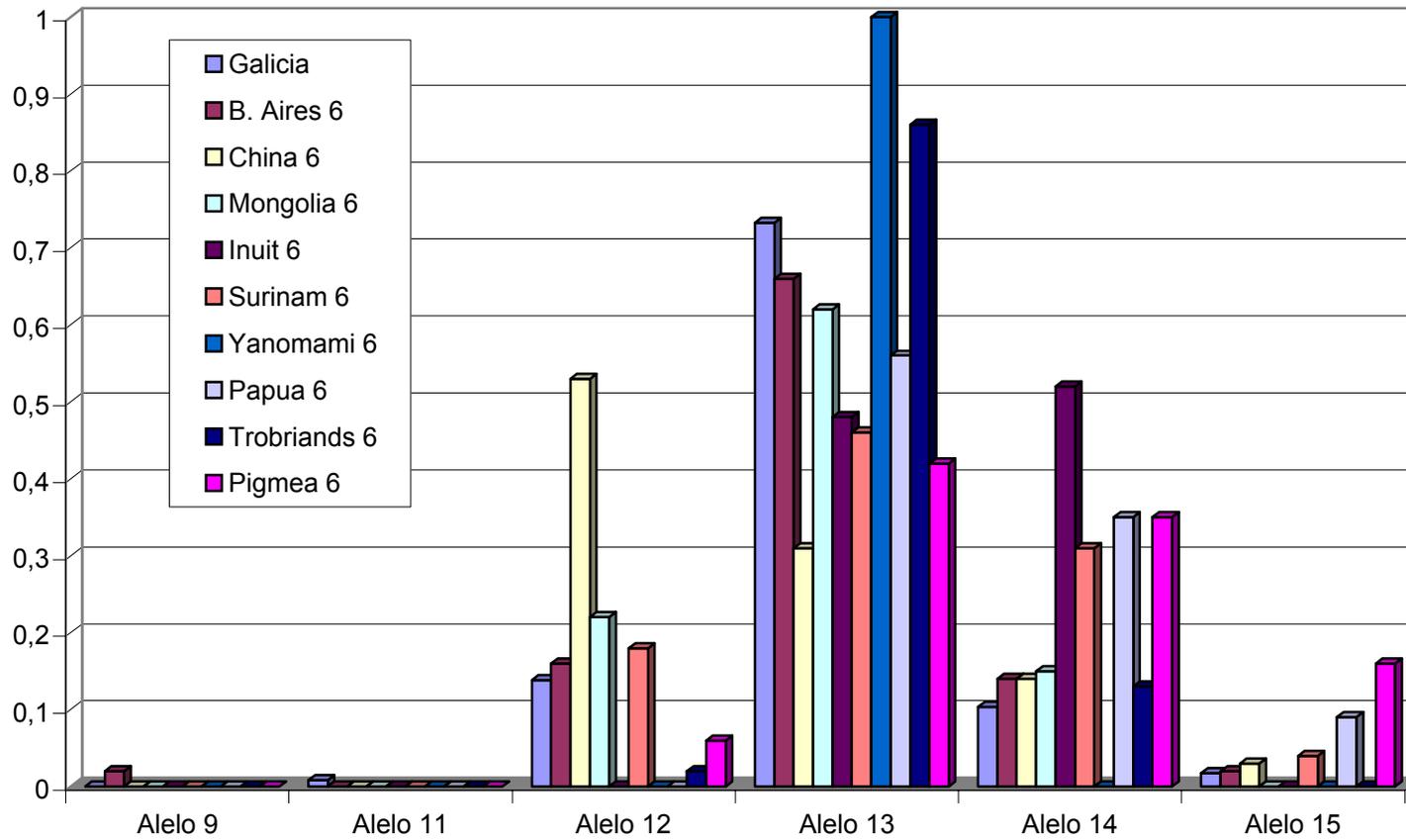


Figura 29. Sistema DYS393: Comparación de frecuencias alélicas Galicia-Población no Europea

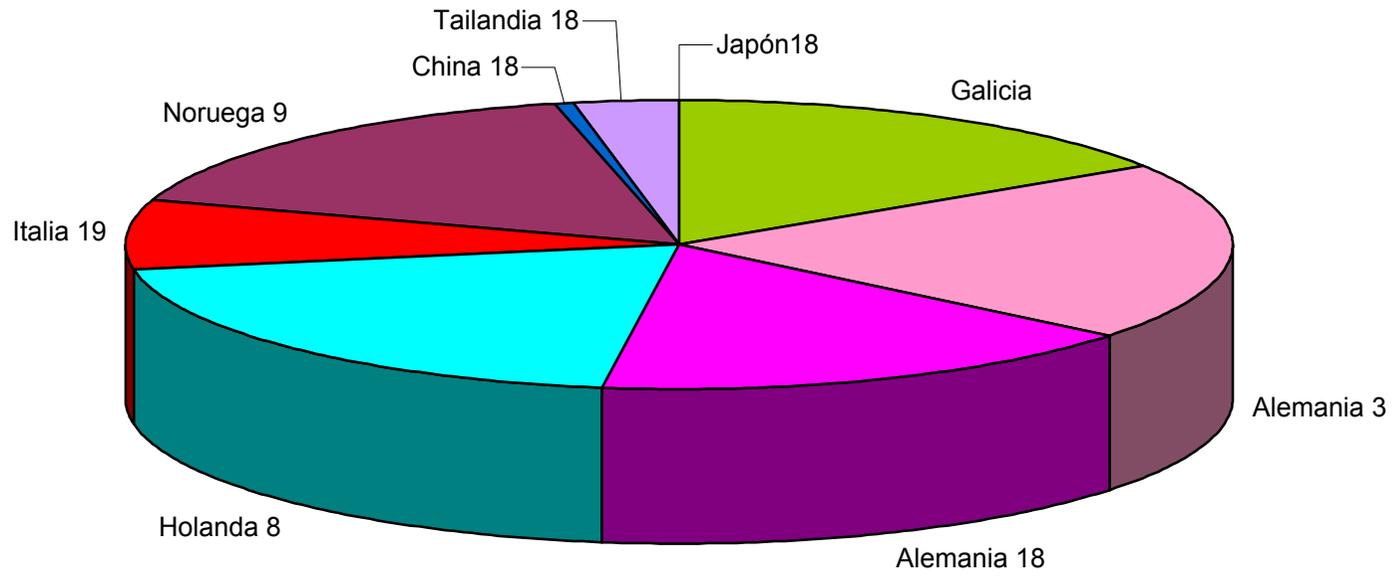
4.3.5.- SISTEMA DYS385

Hemos comparado las frecuencias haplotípicas de nuestra población con las de varias poblaciones europeas y asiáticas (Tabla 44).

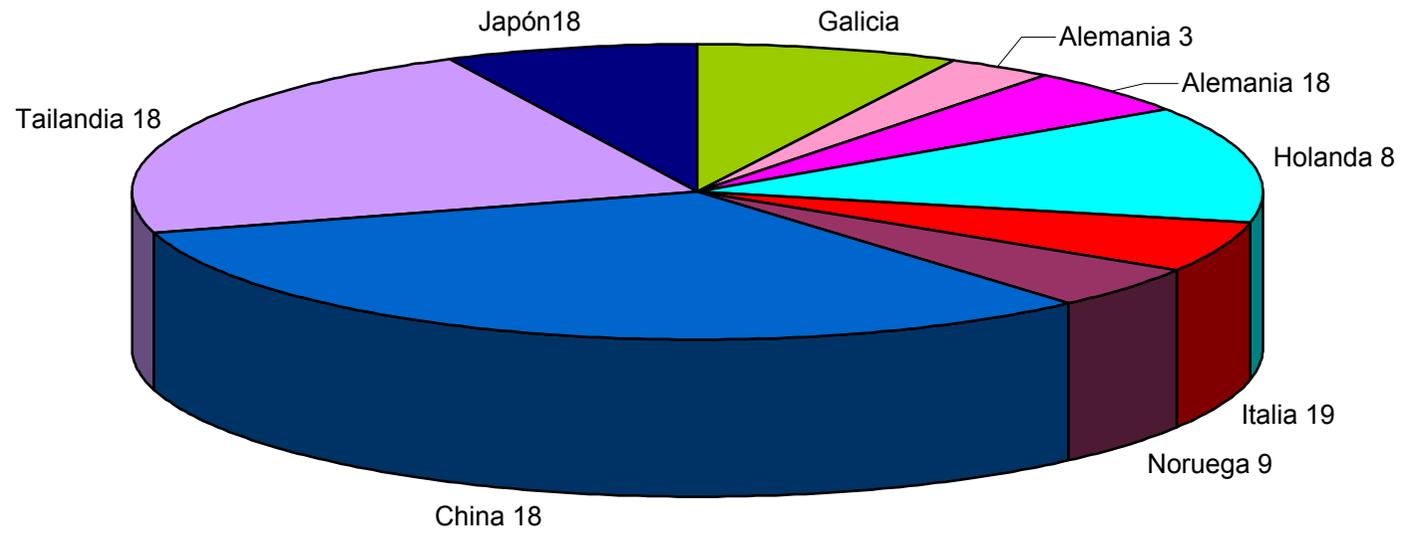
En el sistema DYS385 observamos que las diferencias fenotípicas entre las poblaciones estudiadas (Tabla 45) son notorias. En todas las poblaciones europeas el fenotipo más frecuente es el 11-14 (16-41%), que sólo representa el 1% en la población china y el 6% en la tailandesa, mientras que en la población japonesa ni siquiera está descrito. Para estas últimas poblaciones, los fenotipos más frecuentes son el 13-13 (9% en chinos), 14-18 (7% en tailandeses) y 13-17 (14% en japoneses). En éstos últimos también son notables los fenotipos 10-20 (12%) y 10-19 (10%).

(En la Figura 30 se representa el fenotipo más frecuente en el sistema DYS385 para cada población estudiada y se compara su distribución con la de las restantes poblaciones).

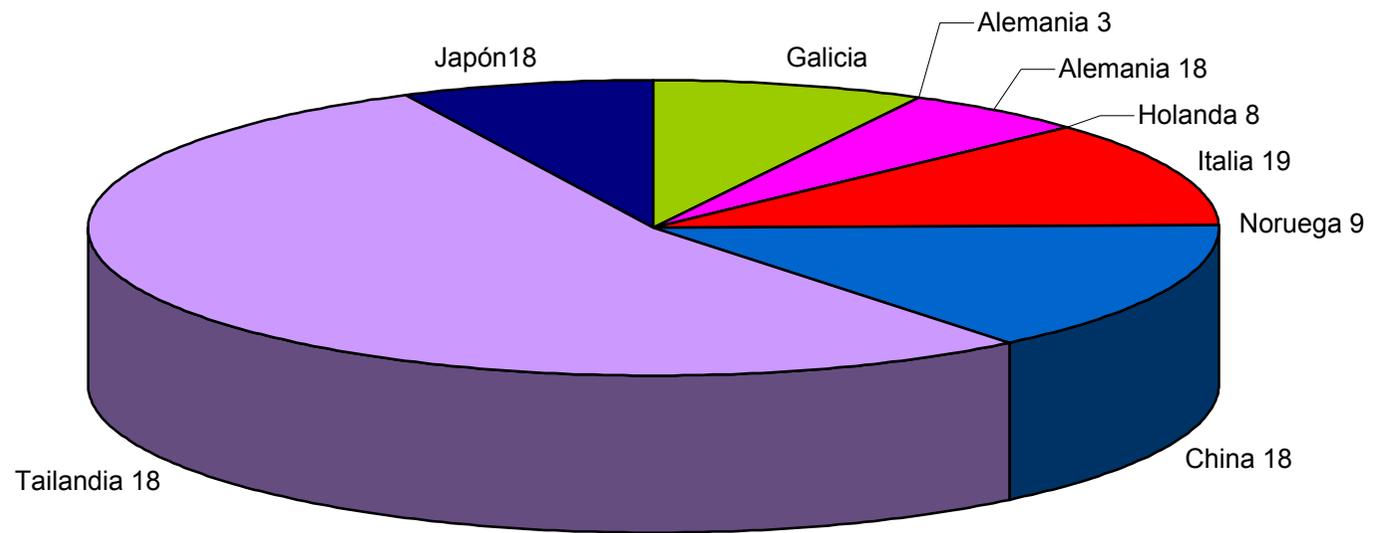
Alelo 11-14



Alelo 13-13



Alelo 14-18



Alelo 13-17

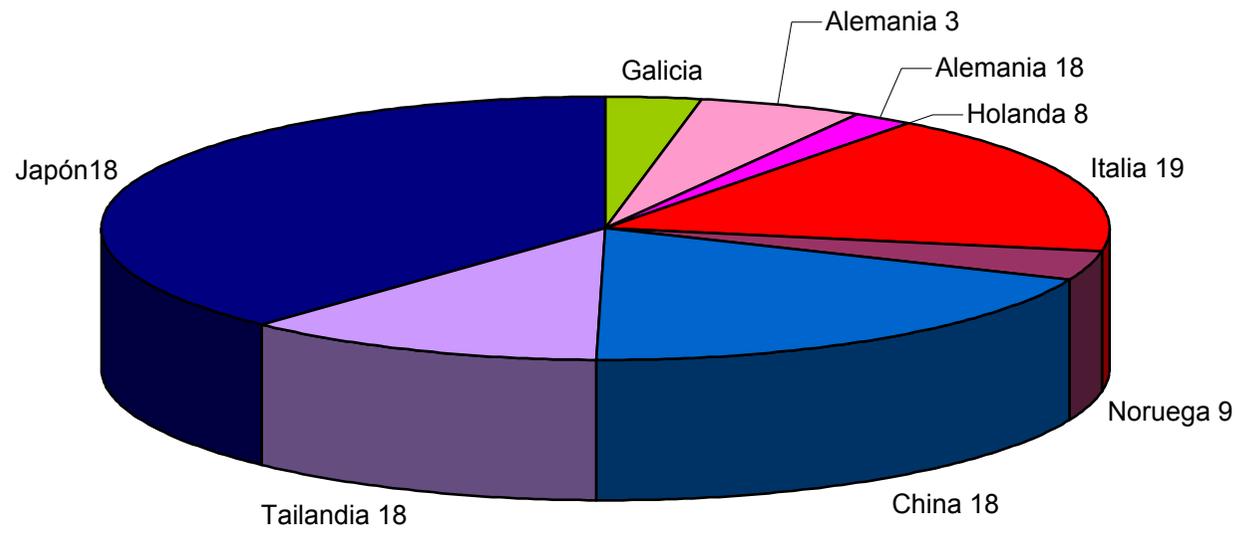


Tabla 44: Sistema DYS385: Frecuencias haplotípicas para Población Gallega, Europea y Asiática

Fenotipo	Galicia	Alemania ³	Alemania ¹⁸	Holanda ⁸	Italia ¹⁹	Noruega ⁹	China ¹⁸	Tailandia ¹⁸	Japón ¹⁸
9-13	0.0108								
9-14		0.0040							
10-14	0.0108	0.0680	0.0490	0.0270	0.0080	0.0130		0.0105	
10-15		0.0200		0.0090		0.0070			0.0200
10-16		0.0080						0.0105	0.0100
10-17									0.0400
10-18					0.0080				0.0700
10-19									0.1000
10-20									0.1200
10-21									0.0100
11-11			0.0070	0.0270	0.0400	0.0070	0.0400	0.0210	0.0100
11-12	0.0108					0.0070	0.0300	0.0105	
11-13	0.0322	0.0360	0.0280	0.0360		0.1070			
11-14	0.3226	0.4120	0.3380	0.4090	0.1600	0.3400	0.0100	0.0631	
11-15	0.0215	0.0680	0.0560	0.0360	0.0400	0.0800		0.0105	
11-16	0.0430	0.0160		0.0090	0.0240	0.0070	0.0200	0.0105	0.0100
11-17					0.0160			0.0105	0.0200
11-18							0.0200		0.0100
11-19							0.0100		0.0300
11-20							0.0100		
11-21							0.0100		
12-12	0.0430		0.0140	0.0090	0.0160		0.0700	0.0210	
12-13	0.0215	0.0040	0.0140	0.0090	0.0160		0.0200	0.0105	0.0100
12-14	0.0645	0.0520	0.0560	0.0450	0.0480	0.0200		0.0210	
12-15	0.0645	0.0080	0.0070	0.0180	0.0240		0.0100	0.0105	
12-16		0.0040	0.0070		0.0080		0.0600	0.0105	0.0500
12-17		0.0040			0.0160		0.0100		0.0800
12-18							0.0400	0.0315	0.0200
12-19			0.0070				0.0400	0.0105	0.0100
12-20							0.0200	0.0421	
12-21								0.0105	
12-22							0.0200	0.0105	
13-13	0.0215	0.0080	0.0142	0.0360	0.0160	0.0130	0.0900	0.0631	0.0200
13-14	0.0322	0.0480	0.0560	0.1000	0.0160	0.0600	0.0100	0.0105	0.0100
13-15	0.0860	0.0360	0.0350	0.0180	0.0400	0.0070	0.0100	0.0105	0.0300
13-16	0.0215	0.0040	0.0070	0.0180	0.0400	0.0130	0.0500	0.0105	0.0100
13-17	0.0108	0.0200	0.0070		0.0640	0.0130	0.0700	0.0421	0.1400
13-18			0.0070		0.0080	0.0070	0.0500	0.0631	0.0300
13-19			0.0140		0.0400	0.0070	0.0800	0.0315	0.0100
13-20					0.0080		0.0600	0.0526	
13-21						0.0130	0.0200	0.0315	
13-22							0.0100		
13-23							0.0100		
13-24							0.0100		
14-14	0.0322	0.0240	0.0490	0.0450	0.0160	0.1130	0.0100	0.0105	0.0100
14-15	0.0215	0.0440	0.0350	0.0270	0.0400	0.1000		0.0105	0.0100
14-16	0.0430	0.0080	0.0070	0.0180	0.0480	0.0130		0.0105	
14-17	0.0215		0.0210		0.0240	0.0070		0.0105	0.0200
14-18	0.0108		0.0070		0.0160		0.0200	0.0736	0.0100
14-19			0.0070				0.0100	0.0105	
14-20							0.0100		
14-21								0.0105	
14-22								0.0105	
14-23							0.0100		
15-15	0.0322	0.0080	0.0140	0.0640	0.0160	0.0330		0.0105	0.0100
15-16		0.0200	0.0210	0.0090	0.0240				
15-17		0.0080	0.0140						0.0300
15-18					0.0080			0.0315	
15-19		0.0040						0.0421	0.0100
15-20			0.0070					0.0631	
15-21								0.0315	
15-22								0.0105	
16-16	0.0108	0.0040	0.0070		0.0240		0.0300		
16-17		0.0080	0.0210		0.0160				
16-18	0.0108	0.0280	0.0140	0.0090	0.0320				
16-19		0.0040		0.0090	0.0400			0.0210	
16-20			0.0070					0.0315	
16-21					0.0080				0.0100
16-22									0.0200
17-17			0.0070	0.0090	0.0080	0.0070		0.0105	
17-18		0.0160	0.0210		0.0160	0.0070			
17-19		0.0040	0.0070		0.0080				
18-19			0.0070				0.0100		
	n=93	n=250	n=146	n=110	n=125	n=150	n=110	n=95n	n=100

(Las celdillas sombreadas contienen el alelo modal)

Tabla 45. Sistema DYS385: Fenotipos en población caucasoide y asiática

Población	n	Fenotipos diferentes	Fenotipo más frecuente	Diversidad Haplótica
Galicia	93	24	11-14 (0.3226)	0.8669
Alemania ³	250	31	11-14 (0.4120)	0.8100
Alemania ¹⁸	146	36	11-14 (0.3380)	0.8700
Holanda ⁸	110	23	11-14 (0.4090)	0.8100
Italia ¹⁹	125	37	11-14 (0.1600)	0.9450
Noruega ⁹	150	24	11-14 (0.3400)	0.8750
China ¹⁸	100	36	13-13 (0.0900)	0.9500
Tailandia ¹⁸	95	44	14-18 (0.0736)	0.9600
Japón ¹⁸	100	33	13-17 (0.1400)	0.9600

Es evidente que la adición del sistema DYS385 incrementa notablemente el poder de discriminación de haplotipos contruidos con STRs de cromosoma Y (Cagliá *et al* 1998).

A falta de más estudios que corroboren estos datos, aparentemente al menos, existe una mayor riqueza haplotípica para este sistema en la población asiática que en la gallega y en las restantes europeas. En la Tabla 46 se muestran los valores obtenidos al comparar nuestra población las con las poblaciones anteriores, siendo las diferencias significativas en algunos casos.

Tabla 46. Comparación Población Gallega-Otras poblaciones: Valor de p para el Sistema DYS385

Valor de p^*	
Población	DYS385
Galicia-Alemania ²	0.0001
Galicia-Alemania ¹⁸	0.1125
Galicia-Holanda ⁸	0.0631
Galicia-Italia ¹⁹	0.0639
Galicia-Noruega ⁹	0.0000
Galicia-China ¹⁸	0.0000
Galicia-Tailandia ¹⁷	0.0000
Galicia-Japón ¹⁸	0.0000

*En negrita valores significativos ($p < 0.01$)

Tabla 47. Relación de autores a los que se hace referencia en el estudio comparativo de poblaciones.

-
- ¹ Pérez-Lezaun *et al* (1997)
 - ² Hidding *et al* (1998)
 - ³ Lessig *et al* (1998)
 - ⁴ Müller *et al* (1994)
 - ⁵ Junge *et al* (1998)
 - ⁶ Kayser *et al* (1997), de Knijff *et al* (1997)
 - ⁷ Roewer *et al* (1996a)
 - ⁸ Kloosterman *et al* (1998)
 - ⁹ Myhre Dupuy *et al* (1998)
 - ¹⁰ Ghio *et al* (1998)
 - ¹¹ Ricci *et al* (1998)
 - ¹² Rossi *et al* (1998)
 - ¹³ Glock *et al* (1998)
 - ¹⁴ Santos *et al* (1993)
 - ¹⁵ Aoyama *et al* (1997)
 - ¹⁶ Wang *et al* (1996)
 - ¹⁷ Lin *et al* (1995)
 - ¹⁸ Schneider *et al* (1998)
 - ¹⁹ Cagliá *et al* (1998)
-

5.- EFICACIA DE LOS SISTEMAS Y SU UTILIDAD MÉDICO-LEGAL

La utilidad de los polimorfismos de cromosoma Y en las ciencias forenses está refrendada por múltiples trabajos (Roewer *et al* 1996, Cooper *et al* 1996, Jobling *et al* 1997, Prinz *et al* 1997, Kayser *et al* 1997, de Knijff *et al* 1997, etc.). La utilización de *ladders* alélicos (Roewer *et al* 1996, Heyer *et al* 1997, Kayser *et al* 1997, de Knijff *et al* 1997) y estudios de colaboración, como los llevados a cabo por Kayser *et al* (1997) y de Knijff (1997), han contribuido en gran medida a potenciar su introducción en el mundo forense.

Las características de los sistemas que hemos estudiado permiten simplificar procesos como la puesta a punto de multiplexes o la secuenciación, ya que únicamente precisamos analizar un alelo (exceptuando los sistemas DYS-389I/II y DYS385) y no dos, como en el caso de los STRs autosómicos. El haber empleado *ladders* alélicos secuenciados, además de una nomenclatura estandarizada basada en el número total de repeticiones, va a facilitar enormemente la reproducibilidad y la comparación de resultados entre laboratorios.

Los siete STRs descritos en nuestro estudio conforman haplotipos de cromosoma Y informativos desde el punto de vista forense, con valores de CE y PD de 0.9794 (Tabla 27). Además, los sistemas que hemos puesto a punto podemos considerarlos específicamente masculinos, como ya hemos comentado, ya que tras amplificar ADN proveniente de mujeres (bajo condiciones muy rigurosas) no se ha detectado producto amplificado.

Como pequeña contribución al estudio de mutaciones, hemos analizado 35 parejas padre/hijo con paternidad probada y no hemos encontrado mutaciones para ninguno de los sistemas. En otros estudios sobre este tema (Roewer *et al* 1996, Kayser *et al* 1997, Lessig y Edelman 1998), se han documentado esporádicamente en sistemas como DYS389II, DYS390 y DYS385, aunque no en varios sistemas para la misma serie padre/hijo.

En nuestra experiencia hemos obtenido excelentes resultados en casos forenses críticos, especialmente para la identificación de ADN masculino en mezclas varón/mujer muy degradadas en casos de violación, incluso múltiple, y en análisis de paternidad, tanto en casos de deficiencia como en estadios tempranos de embarazos consecutivos a una agresión sexual (en este último caso se realiza la determinación en líquido amniótico). Cuando se parte de mezclas de material biológico masculino y femenino la detección de microsatélites de cromosoma Y proporciona mayor sensibilidad que la de microsatélites autosómicos, al aumentar la capacidad de identificación del componente masculino aunque exista una gran desproporción entre ambos (Prinz *et al* 1997).

Los STRs de cromosoma Y que hemos estudiado pueden constituir un método de screening rápido y fiable, ya que permiten la exclusión de sospechosos, pero no pueden ser utilizados como criterio de inclusión en casos de criminalística: siempre hay que tener presente que todos los familiares varones por línea paterna comparten el mismo haplotipo. Esto constituye una importante limitación, tanto en casos de paternidad o violación con implicados pertenecientes a la misma familia como cuando se trata, por ejemplo, de identificar a familiares varones víctimas de catástrofes.

De lo expuesto anteriormente podemos inferir que, polimorfismos informativos de cromosoma Y, conformados como haplotipos, pueden ser perfectamente válidos en algunos casos de rutina forense (obviamente en aquellos casos con implicación de uno o más varones), pero no podemos hacer una identificación individualizada basada exclusivamente en ellos, por lo que de momento sigue siendo conveniente combinar STRs de cromosoma Y con marcadores autosómicos para proporcionar a ambos más opciones útiles en estudios forenses y pruebas de paternidad.



IV. CONCLUSIONES

PRIMERA.- El sistema DYS19 es un STR simple con una parte variable $(TAGA)_n$. Su nomenclatura se definió según el número de repeticiones de la parte variable.

SEGUNDA.- El sistema DYS389I posee dos motivos repetitivos $(TCTA)$ y $(TCTG)$ pero sólo el primero es variable en Caucasoides. El sistema DYS389II incluye la variación del sistema DYS389I y posee en su región variable específica tres trectos variables $(TGTG)_n(TCTA)_m(TCTA)_p$.

La nomenclatura de estos sistemas es difícil según se tenga en cuenta la inclusión o no de la región 389I y si se usan partes variables o constantes, aunque hemos optado por incluir sólo las primeras.

TERCERA.- El sistema DYS390 es un STR compuesto con dos unidades variables $(TCTG)$ y $(TCTA)$. En su nomenclatura hemos sumado cuatro repeticiones correspondientes a un tracto $(TCTA)_4$ constante para coincidir con otra nomenclatura previamente utilizada.

CUARTA.- El sistema DYS393 posee dos motivos de repetición $AGAT$ y $ATGT$ siendo variable solamente el primero. Para su nomenclatura hemos considerado solamente la parte variable.

QUINTA.- Se han fabricado escaleras alélicas (*ladders*) secuenciados que incluyen todos los alelos observados pertenecientes a los sistemas DYS19, DYS389I, DYS389II, DYS390, DYS393.

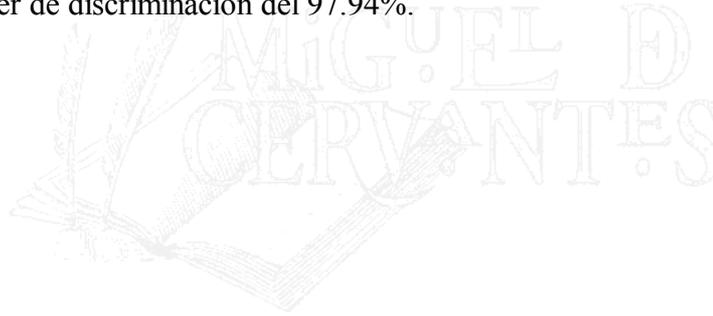
SEXTA.- Los cinco sistemas secuenciados se pueden amplificar simultáneamente (PCR multiplex) y detectar en secuenciadores automáticos monocromáticos (ALF-APB). Las ventajas de los secuenciadores que admiten varios flurocromos (ABI377, Perkin Elmer) son evidentes para la secuenciación.

SÉPTIMA.- Se analizaron los seis sistemas en una muestra de población gallega, y los valores observados están en el rango esperado de las poblaciones europeas

Conclusiones

diferenciándose notablemente de otras poblaciones no europeas, negroides o mongoloides. Las poblaciones ibéricas analizadas hasta ahora son particularmente semejantes.

OCTAVA.- Los STR analizados conforman haplotiupos de cromosoma Y muy informativos desde el punto de vista forense, con una probabilidad de exclusión a priori y un poder de discriminación del 97.94%.





V. BIBLIOGRAFÍA

- Alford RL, Hammond HA, Coto I, Carkey CT** (1994). Rapid and efficient resolution of parentage by amplification and short tandem repeats. *Am J Hum Genet* **55**: 190-195.
- Alvarez-García A, Muñoz I, Pestoni C, Lareu MV, Rodríguez-Calvo MS, Carracedo A** (1996). Effect of environmental factors on PCR-DNA analysis from dental pulp. *Int J Legal Med* **109**: 125-129.
- Andersen J, Martin P, Carracedo A, Dobosz M, Eriksen B, Johnsson V, Kimpton C, Kloosterman A, Kornialis C, Kratzer A, Phillips P, Mevag B, Pfitzinger H, Rand S, Rosén B, Schmitter H, Schneider P, Vide M** (1996). Report on the third EDNAP collaborative STR exercise. *Forensic Sci Int* **78**: 83-93.
- Anderson S, Bankier AT, Barrell BG, deBruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AHJ, Staden R, Young IG** (1981). Sequence and organisation of the human mitochondrial genome. *Nature* **290**: 457-465.
- Aoyama M, Matsuki T, Kishi K, Iida R, Tsubota E** (1997). DYS19 marker in Japanese populations. *Int J Legal Med* **110**: 112-113.
- Armour JAL, Wong Z, Wilson V, Royle NJ, Jeffreys AJ** (1989). Sequences flanking the repeat arrays of human minisatellites: association with tandem and dispersed repeat elements. *Nucleic Acids Res.* **17**: 4925-4935.
- Armour JAL, Povey S, Jeremiah S, Jeffreys AJ** (1990). Systematic cloning of human minisatellites from ordered array Charomid libraries. *Genomics* **8**: 501-512.
- Armour JAL** (1996). Tandemly repeated minisatellites: generating human genetic diversity via recombinational mechanisms. In: *Human Genome Evolution*. Ed. by M Jackson, T Strachan, G Dover. Bios Scientific publishers, pp. 172-190.
- Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B** (1997). DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeats systems. *Int J Leg Med* **110**: 175-176.
- Barros F, Carracedo A, Lareu MV, Rodríguez-Calvo MS** (1991). Electrophoretic human leukocyte antigen HLA-DQA1 DNA typing after polymerase chain reaction amplification. *Electrophoresis* **12**: 1041-1045.
- Barros F, Lareu MV, Carracedo A** (1992). Detection of polymorphisms of human DNA after polymerase chain reaction by minaturized SDS-PAGE. *Forensic Sci Int* **55**: 27-36.

- Barros F, Muñoz I, Larev MV, Rodríguez-Calvo MS, Carracedo A** (1994). Double -and single- strand conformation polymorphism analysis of point mutations and short tandem repeats. *Electrophoresis* **15**: 566-571.
- Beckman JS, Weber JL** (1992). Survey of human and rat microsatellites. *Genomics* **12**: 627-631.
- Boerwinkle E, Xiong W, Fourest E, Chan L** (1989). Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. *Proc Natl Acad Sci USA* **86**: 212-216.
- Botstein D, White RL, Skolnick M, Davis RW** (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* **32**: 314-331.
- Bouzekri N, Dover GA, Jeffreys AJ, Jobling MA** (1995). Digital coding and "null" repeats at the Y-chromosome specific minisatellite MSY1 (DYF155S1). Euroconference, Barcelona.
- Brinkmann B, Butler R, Lincoln P, Mayr W, Rossi U** (1992). 1991 report concerning recommendations of the DNA commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms. *Forensic Sci Int* **52**: 125-130.
- Brinkmann B, Möller A, Wiegand P** (1995). Structure of new mutations in 2 STR systems. *Int J Legal Med.* **107**: 201-203.
- Brinkmann B** (1996). The STR approach. In: *Advances in Forensic Haemogenetics* 6. Ed. by A Carracedo, B Brinkmann, W Bär. Springer-Verlag, Berlin, Heidelberg, pp. 41-51.
- Brinkmann B, Klitschar M, Neuhuber F, Huhn J, Rolf B** (1998). Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet* **62**: 1408-1415.
- Budowle B, Chakraborty R, Giusti AM, Einsenberg J, Allen RC** (1990). Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am J Hum Genet* **38**: 137-144.
- Cagliá A, Dobosz M, Boschi I, d'Áloja E, Pascali VL** (1998). Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus. *Int J Legal Med* **111**: 142-146.
- Casanova M, Leroy P, Boucekkine C, Weissenbach J, Bishop C, Fellous M, Purrello M, Fiori G, Siniscalco M** (1985). A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* **230**: 1403-1406.

- Chen H, Lowther W, Avramopoulos D, Antonarakis SE** (1994). Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)_n and map to human X and Y chromosomes. *Hum Mut* **4**: 208-211.
- Cooper G, Amos W, Hoffman D, Rubinsztein D** (1996). Network analysis of human Y microsatellite Haplotypes. *Hum Mol Genet* **5**: 1759-1766.
- Corach D, Sala A, Penacino G, Sotelo A** (1995). Mass disasters: rapid molecular screening of human remains by means of short tandem repeats typing. *Electrophoresis* **16**: 1617-1623.
- Corach D, Sala A, Penacino G, Sotelo A** (1996). Combined STR, MVR and MVR typing and mtDNA sequencing, led to the identification of human remains emergent from the AMIA explosion in Buenos Aires. In: *Advances in Forensic Haemogenetics* **6**. Ed. by A Carracedo, B Brinkmann and W Bär. Springer-Verlag, Berlin, Heidelberg, pp. 337-339.
- Choo KH, Vissel B, Nagy A, Earle E, Kalitsis P** (1991). A survey of the genomic distribution of alpha satellite DNA and all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res* **19**: 1179-1182.
- D'Esposito MA, Ciccodicola A, Gianfrancesco F, Esposito T, Flagiello L, Mazzarella R, Schlessinger D, D'Urso M** (1996). A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X-inactivation. *Nature Genet* **13**: 227-229.
- de Knijff P, Kayser M, Cagliá A, Corach D, Fretwell, Gehrig C, Graziosi G, Heidorn F, Herrman S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L** (1997). Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* **110**: 134-140.
- Disteche CM, Casanova M, Saal H, Friedman C, Sybert V, Graham J, Thuline H, Page DC, Fellous M** (1986). Small deletions of the short arm of the Y chromosome in 46,XY females. *Proc Natl Acad Sci USA* **84**: 7841-7844.
- Dorit RL, Akashi H, Gilbert W** (1995). Absence de polymorphism at the ZFY locus on the human Y chromosome. *Science* **268**: 1183-1185.
- Edwards A, Civitello A, Hammond HA, Caskey CT** (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* **49**: 746-756.
- Ellis N, Goodfellow PN** (1989a). The mammalian pseudoautosomal region. *Trends genet* **5**: 406-410.

- Ellis N, Goodfellow PJ, Pym B, Smith M, Palmer M, Frischauf A-M, Goodfellow PN** (1989b). The pseudosomal boundary in man is defined by an Alu repeat sequence inserted on the Y chromosome. *Nature* **337**: 81-84.
- Ellison J, Passage M, Yu LC, Yen P, Mohandas TK, Shapiro L** (1992). Directed isolation of human genes that escape X inactivation. *Somat Cell Mol Genet* **18**: 256-268.
- Erlich HA, Sheldon EL, Horn G** (1986). HLA typing using DNA probes. *Bio Technol* **4**: 975-981.
- Erlich HA, Bugawan TL** (1989). HLA class II gene polymorphism: DNA typing, evolution, and relationship to disease susceptibility. In: *PCR Technology: Principles and Applications for DNA Amplification*. Ed. by HA Erlich. Stockton Press, Nueva York, USA, pp. 193-208.
- Evett IW, Buffery C, Willot G, Stoney D** (1991a). A guide to interpreting single locus profiles of DNA mixture in forensic cases. *J Forensic Sci Soc* **31**: 41-47.
- Evett IW, Gill P** (1991b). A discussion of the robustness of methods for assessing the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis* **12**: 226-230.
- Fisher EMC, Beer-Romero P, Brown LG, Ridley A, McNeil JA, Lawrence JB, Willard HF, Bieber FR, Page DC** (1990). Homologous ribosomal protein genes on the human X and Y chromosomes: escape from inactivation and possible implications for Turner syndrome. *Cell* **63**: 1205-1218.
- Fisher RA** (1951). Standard calculations for evaluating a blood group system. *Heredity* **5**: 95-102.
- Ford EB** (1940). Polymorphism and taxonomy. In: *New systematics*. Ed by JS Huxley. Clarendon Press, Oxford, pp. 493.
- Forensic Science International** (1997). Special Issue: Proceedings of the European Symposium: Ethical and Legal Issues of DNA Typing in Forensic Medicine. *Forensic Sci Int* **88**: 1-110.
- Fowler JCS, Borgoyne LA, Scott AC, Harding HWJ** (1988). Repetitive deoxyribonucleic acid (DNA) and human genome variation. A concise review relevant to forensic biology. *J Forensic Sci* **33**: 1111-1126.
- Freije D, Helms C, Watson MS, Donis-Keller** (1992). Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science* **258**: 1784-1787.
- Frommer M, Prosser J, Tkachuk D, Reisner AH, Vincent PC** (1982). Simple repeated sequences in human satellite DNA. *Nucl Acids Res* **10**: 547-563.

- Ghio F, Betti F, Piccinini A** (1998). North Italian population genetic data for the Y STR systems DYS19, DYS390, DYS392 and their possible application to forensic routine casework. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp 524-526.
- Gill P, Kimpton CP, Sullivan KM** (1992a). A rapid polymerase chain reaction methods for identifying fixed specimens. *Electrophoresis* **13**: 173-175.
- Gill P, Woodroffe S, Bär W, Brinkmann B, Carracedo A, Eriksen B, Jones S, Kloosterman AD, Ludes B, Mevag B, Pascali VL, Rudler M, Schmitter H, Schneider PM, Thomson JA** (1992b). A report of an international collaborative experiment to demonstrate the uniformity obtainable using DNA profiling techniques. *Forensic Sci Int* **47**: 515-523.
- Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K** (1994a). Identification of the remains of the Romanov family by DNA analysis. *Nature Genetics* **6**: 130-135.
- Gill P, Kimpton C, D'Aloja E, Andersen JF, Bar W, Brinkmann B, Holgersson S, Johnsson V, Kloosterman AD, Lareu MV, Nellesmann L, Pfitzinger H, Phillips CP, Schmitter H, Schneider PM, Stenersen M** (1994b). Report of the European DNA profiling group (EDNAP) towards standardisation of short tandem repeat (STR) loci. *Forensic Sci Int* **65**: 51-59.
- Gill P, Kimpton CP, Urquhart A, Oldroyd N, Millicam E, Natson SK, Downes TJ** (1995). Automated STR analysis in forensic casework-a strategy for the future. *Electrophoresis* **16**: 1543-1552.
- Gill P, Kimpton C, D'Aloja E, Andersen JF, Bar W, Carracedo A, Dupuy B, Eriksen B, Jangbland M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider PM, Vide MC** (1997). Considerations from the European DNA Profiling group (EDNAP) concerning STR nomenclature. *Forensic Sci Int* **87**: 185-192.
- Glock B, Schwartz DWM, Dauber EM, Schwartz-Jungl EM, Mayr WR** (1998). Austrian population data on the polymorphic STR loci DYS19 and HUMF13B. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 291-293.
- Gomolka M, Hundrieser J, Nurnberg P, Roewer L, Epplen JT, Epplen C** (1994). Selected di- and tetranucleotide microsatellites from chromosomes 7, 12, 14, and Y in various Eurasian populations. *Hum Genet* **93**: 592-596.
- Goodfellow P, Banting G, Sheer D, Ropers HH, Caine A, Ferguson-Smith MA, Povey S, Voss R** (1983). Genetic-evidence that a Y-linked gene in man is homologous to a gene on the X-chromosome. *Nature* **302**: 346-349.

- Gough NM, Gearing DP, Nicola NA, Baker E, Pritchard M, Callen DF, Sutherland GR** (1990). Localization of the human GM-CSF receptor gene to the X-Y pseudoautosomal region. *Nature* **345**: 734-736.
- Hagelberg E, Gray IC, Jeffreys AJ** (1991). Identification of the skeletal remains of a murder victim by DNA analysis. *Nature* **352**: 427-429.
- Hammer MF** (1994). A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol* **11**: 749-761.
- Hammer MF** (1995a). A recent common ancestry for human Y chromosomes. *Nature* **378**: 376-378.
- Hammer MF, Horai S** (1995b). Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* **56**: 951-962.
- Hassold TJ, Sherman SL, Pettay D, Page DC, Jacobs PA** (1991). XY-chromosome non disjunction in man associated with diminished recombination in the pseudoautosomal region. *Am J Hum Genet* **49**: 253-260.
- Heukeshoven J, Dernick R** (1985). Simplified methods for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining. *Electrophoresis* **6**: 103-112.
- Heyer E, Puymirat j, Dieltjes P, Bakker e, de Knijff P** (1997). Estimating Y-chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* **6**: 799-803.
- Hidding M, Staak M, Schmitt C** (1998). Y-chromosomal STR systems: applications of a triplex-PCR in forensic stain analysis. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 515-517.
- Higuchi R, von Beroldingen CH, Sensabaugh GF, Erlich HA** (1988). DNA typing from single hairs. *Nature* **322**: 543-546.
- Holmlund G, Engdahl M, Marttila M, Lindblom** (1998). Allelic diversity in the variable number of tandem repeat marker D2S44 (YNH24) shown by multi variant repeat mapping. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and P Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 234-236.
- Hollis M, Hindley J** (1988). Satellite II of human lymphocytes: Tandem repeats of a single sequence element. *Nucl Acids Res* **16**: 363.

- Horn GT, Richards B, Klinger KW** (1989). Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. *Nucleic Acids Res* **17**: 2140.
- ISFH** (1989) -Editorial-Recommendations of the International Society for Forensic Haemogenetics concerning DNA polymorphisms. *Forensic Sci Int* **43**: 109-111.
- ISFH** (1992a) -Editorial-Recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. *Forensic Sci Int* **43**: 109-111.
- ISFH** (1992b) -Editorial-1991 Report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms. *Forensic Sci Int* **52**: 125-130.
- ISFH** (1995) -Recommendations-1994 report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. *Forensic Sci Int* **69**: 103-104.
- ISFH** (1997) -Editorial-DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *Forensic Sci Int* **87**: 103-104.
- Ishiyama I, Yoshii T, Honma M, Mukaida M, Yamaguchi T** (1988). DNA fingerprints: the importance in forensic medicine, II. The significance of examining DNA polymorphisms of placental tissues for the purpose of paternity determination during the early stages within the first trimester. *Z Rechtsmed* **99**: 241-248.
- Jakubiczka S, Arnemann J, Cooke HJ, Krawczak M, Schmidtke J** (1989). A search for restriction fragment length polymorphism on the human Y chromosome. *Hum Genet* **84**: 86-88.
- Jeffreys AJ, Wilson V, Thein SL** (1985a). Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**: 67-73.
- Jeffreys AJ, Wilson V, Thein SL** (1985b). Individual-specific 'fingerprints' of human DNA. *Nature* **316**: 76-79.
- Jeffreys AJ, Brookfield JFY, Semeonoff R** (1985c). Positive identification of an immigration test-case using human DNA fingerprints. *Nature* **317**: 818-819.
- Jeffreys AJ, Wilson V, Neumann R, Keyte J** (1988). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **16**: 10953-10971.
- Jeffreys AJ, MacLeod A, Tamaki K, Neil DL, Monckton DG** (1991). Minisatellite repeat coding as a digital approach to DNA typing. *Nature* **354**: 204-209.

- Jeffreys AJ, Allen MJ, Hagelberg E, Sonnberg A** (1992). Identification of the skeletal remains of Josef Mengele by DNA analysis. *Forensic Sci Int* **56**: 65-76.
- Jeffreys AJ, Neumann R** (1997). Somatic mutation processes at a human minisatellite. *Hum Mol Genet* **1**: 129-136.
- Jobling MA** (1994a). A survey long-range DNA polymorphism on the human Y chromosome. *Hum Mol Genet* **3**: 107-114.
- Jobling MA, Fretwell N, Dover GA, Jeffreys AJ** (1994b). Digital coding of human Y chromosomes: MVR-PCR at Y-specific minisatellites. *Cytogenet Cell Genet* **67**: 390.
- Jobling MA and Tyler-Smith C** (1995). Fathers and sons: the Y chromosome and human evolution. *Trends Genet* **11**: 449-456.
- Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L Crawford MH, Tyler-Smith C** (1996). Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet* **5**: 1767-1775.
- Jobling MA, Pandya A, Tyler-Smith C** (1997). The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* **110**: 118-124.
- Jones DA** (1972). Blood samples: probability of discriminations. *J Forensic Sci Soc* **12**: 355-359.
- Junge A, Madea B** (1998). Population studies of the Y-Chromosome specific polymorphisms DYS19, DYS 389 I+II and DYS390 in a Western German population (Rhine area). In: *Progress in Forensic Genetics 7*. Ed by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 518-520.
- Kasai K, Nakamura Y, White R** (1990). Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. *J Forensic Sci* **5**: 1196-1200.
- Kayser M, Cagliá A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P, Roewer L** (1997a). Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* **110**:125-133.
- Kermouni A, Vanroost E, Arden KC, Vermeesch JR, Weiss S, Godelaine D, Flint J, Lurquin C, Szikora JP, Higgs DR, Marynen P, Renaud JC** (1995). The IL-9 receptor gene (IL-9R)-genomic structure, chromosomal localization in

the pseudoautosomal region of the long arm of the sex-chromosomes, and identification of IL-9R pseudogenes at 9qter, 10pter, 16pter, and 18pter. *Genomics* **29**: 371-382.

Kimpton CP, Walton A, Gill P (1992). A further tetranucleotide repeat polymorphism in the vWF gene. *Hum Mol Genet* **1**: 287.

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993). Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods and Applic* **3**: 13-22.

Kimpton C, Gill P, D'Aloja E, Andersen JF, Bar W, Holgersson S, Jacobsen S, Johnsson V, Kloosterman AD, Lareu MV, Nellemann L, Pfitzinger H, Phillips CP, Rand S, Schmitter H, Schneider PM, Sternersen M, Vide MC (1995). Report on the second EDNAP collaborative STR exercise. *Forensic Sci Int* **71**: 137-152.

Kimura M (1983). The neutral theory of molecular evolution. Cambridge, Cambridge University Press.

Kloosterman A, Janssen H (1997). Forensic DNA testing and its legislation in the Netherlands. *Forensic Sci Int* **88**: 55-58.

Kloosterman AD, Pouwels M, Daselaar P, Janssen HJT (1998). Population genetic study of Y-Chromosome specific STR-loci in Dutch caucasians. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 491-493.

Koopman P, Gubbay J, Vivian N, Goodfellow P, LovellbadgeR (1991). Male development of chromosomally female mice transgenic for SRY. *Nature* **351**: 117-121.

Korenberg JR, Rykowski MC (1988). Human genome organization: Alu, Lines, and molecular structure of metaphase chromosome bands. *Cell* **53**: 391-400.

Kremer E, Baker E, Dandrea RJ, Slim R, Phillips H, Moretti P, López AF, Petit C, Vadas MA, Sutherland GR, Goodall GJ (1993). A cytokine receptor gene-cluster in the X-Y pseudoautosomal region. *Blood* **82**: 22-28.

Kwok C, Tyler-Smith C, Mendoza BB, Hughes I, Berkovitz GD, Goodfellow PN, Hawkins JR (1996). Mutation analysis of the 2 Kb 5' to SRY in XY females and XY intersex subjects. *J Med Genet* **33**: 465-468.

Lahn B, Page D (1997). Functional coherence of the human Y chromosome. *Science* **278**: 675-680.

- Lareu MV, Pestoni C, Barros F, Salas A, Carracedo A** (1996). Sequence variation of a hypervariable short tandem repeat at the D12S391 locus. *Gene* **182**: 151-153.
- Lareu MV, Barral S, Salas A, Pestoni C, Carracedo A** (1998). Sequence variation of a hypervariable short tandem repeat at the D1S1656 locus. *Int J Leg Med* **111**: 244-247.
- Lessig R, Edelmann J** (1998). Y-Chromosome polymorphisms and haplotypes in West Saxony (Germany). *Int J Legal Med* **111**: 215-218.
- Litt M, Luty JA.** (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* **44**: 397-401.
- Lin SJ, Ko YH, Wang WP, Lai ML** (1995). Population genetic study of selected tetranucleotide repeat DNA polymorphisms on chromosomes Y and 12. *J Formos Med Assoc* **94**: 318-321.
- Lucotte G, Ngo NY** (1985). p49f, a highly polymorphic probe, that detects Taq I RFLPs on the human Y chromosome. *Nucleic Acids Res* **13**: 8285.
- Malaspina P, Persichetti F, Novelletto A, Iodice C, Terrenato L, Wolfe J, Ferraro M, Prantera G** (1990). The human Y chromosome shows a low level of DNA polymorphism. *Ann Hum Genet* **54**: 297-305.
- Mathias N, Bayés M, Tyler-Smith C** (1994). Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* **3**:115-124.
- Mills KA, Even D, Murray JC** (1992). Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). *Hum Mol Genet* **1**: 779.
- Moos M, Gallwitz D** (1983). Structure of two human beta-actin-related processed genes one of which is located next to a simple repetitive sequence. *EMBO J* **2**: 757-761.
- Morton NE** (1991). Parameters of the human genome. *Proc Natl Acad Sci USA* **88**: 7474-7476.
- Müller S, Gomolka M, Walter H** (1994). The Y-specific SSCP of the locus DYS19 in four different European samples. *Hum Heredity* **44**: 298-300.
- Mullis KB, Faloona FA** (1987). Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Meth Enzymol* **155**: 335-350.
- Myhre Dupuy B, Gjesti A, Olaisen B** (1998). Population database of DYS19 and DYS385 in 150 Norwegians. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 536-538.

- Nakahori Y, Tamura T, Yamada M, Nakagome Y** (1989). Two 47z (DXYS5) RFLPs on the X and Y chromosome. *Nucleic Acids Res* **17**: 2152.
- Nakahori Y, Takenaka O, Nakagome Y** (1991). A human X-Y homologous region encodes Amelogenin. *Genomics* **9**: 264-269.
- Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R** (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**: 1616-1622.
- National Research Council [NRC]** (1992). DNA Technology in Forensic Science. National Academy Press, Washington D.C.
- National Research Council [NRC]** (1997). DNA Technology in Forensic Science. National Academy Press, Washington D.C.
- Ngo KY, Vergnaud G, Johnsson C, Lucotte G, Weissenbach J** (1986). A DNA probe detecting multiple haplotypes of the human Y chromosome. *Amer J Hum Genet* **38**: 407-418.
- Oakey R and Tyler-Smith C** (1990). Y chromosome DNA halotyping suggest that most European and Asian men are descended from one of two males. *Genomics* **7**: 325-330.
- Odelberg SJ, Demers DB, Westin EH, Hossaini AA** (1988). Establishing paternity using minisatellite DNA probes when the putative father is unavailable for testing. *J Forensic Sci* **33**: 921-928.
- Ohno Y, Sebetan IM, Akaishi S** (1982). A simple method for calculating the probability of excluding paternity with any number of codominant alleles. *Forensic Sci Int* **19**: 93-98.
- Olaisen B, Stenersen M, Mevag B** (1997). Identification by DNA analysis of the victims of the August 1996 Spitsbergen civil aircraft disaster. *Nature Genetics* **15**: 402-405.
- Page DC, Mosher R, Simpson EM, Fisher C, Mardon G, Pollack J, McGillivray B, de la Chapelle A, Brown LG** (1987). The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* **51**: 1091-1104.
- Pena SDJ, Chakraborty R** (1994). Paternity testing in the DNA era. *Trends genet* **10**: 204-209.
- Pérez-Lezaun A, Calafell F, Sielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J** (1997). Population genetics of Y Chromosome Short Tandem Repeats in humans. *J Mol Evol* **45**: 265-270.

- Polymeropoulos MH, Rath DS, Xiao H, Merrill CR (1991a).** Tetranucleotide repeat polymorphism at the human coagulation factor XIII A subunit gene (F13A1). *Nucleic Acids Res* **19**: 4306.
- Polymeropoulos MH, Rath DS, Xiao H, Merrill CR (1991b).** Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene (FES). *Nucleic Acids Res* **19**: 4018.
- Polymeropoulos MH, Rath DS, Xiao H, Merrill CR (1992a).** Tetranucleotide repeats polymorphism at the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2). *Nucleic Acids Res.* **20**: 1432.
- Polymeropoulos MH, Xiao H, Merrill CR (1992b).** Tetranucleotide repeat polymorphism at the human myelin basic protein gene (MBP). *Hum Mol Genet* **1**: 658.
- Prinz M, Boll K, Baum H, Shaler B (1997).** Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Sci Int* **85**: 209-218.
- Prosser J, Frommer M, Paul C, Vincent PC (1986).** Sequence relationships of three human satellite DNA's. *J Mol Biol* **187**: 145-155.
- Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B (1992).** Population genetics and forensic efficiency data of 4 AMPFLPs. *Int J Legal Med* **104**: 329-333.
- Ricci U, Vieri F, Scarselli B, Knusel P, Giovannucci Uzielli ML (1998).** Analysis of the Y-Chromosome: the Y-27H39 polymorphism in a population sample of Tuscany region (Central Italy). In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 531-532.
- Rodríguez-Calvo MS, Bellas S, Soto JL, Barros F, Carracedo A (1996).** A comparison of different electrophoretic methods for PCR-MVR typing. *Electrophoresis* **17**: 1294-1299.
- Roewer L (1998).** The Y chromosome: Forensic application and evolutionary aspects. In: *Progress in Forensic Genetics 7*. Ed. by B Olaisen, B Brinkmann and PJ Lincoln. Elsevier Science BV, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, pp. 407-412.
- Roewer L and Epplen JT (1992a).** Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. *Forensic Sci Int* **53**: 163-171.
- Roewer L, Arnemann J, Spurr NK, Grzeschik K-H, Epplen JT (1992b).** Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum. Genet* **89**: 389-394.

- Roewer L, Kayser M, Dieltjes M, Nagy M, Bakker E, Krawczak M, de Knijff P** (1996). Analysis of molecular variance (AMOVA) of Y-chromosome specific microsatellites in two closely related human populations. *Hum Mol Genet* **5**: 1029-1033.
- Rossi E, Rolf B, Schürenkamp M, Brinkmann B** (1998). Y-Chromosome STR haplotypes in an Italian population sample. *Int J Legal Med* **112**: 78-81.
- Royle NJ, Clarkson RE, Wong Z, Jeffreys AJ** (1988). Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. *Genomics* **3**: 352-360.
- Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Arnheim N** (1985). Enzymatic amplification of beta-globin genomic sequences and restriction analysis of sickle cell anemia. *Science* **230**: 1350.
- Saiki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA** (1986). Analysis of enzymatically amplified beta-globin and HLA-DQ α DNA with allele-specific oligonucleotide probes. *Nature* **324**: 163.
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT** (1988). Analysis of DNA with a thermostable DNA polymerase. *Science* **139**: 487-491.
- Santos FR, Pena SDJ, Epplen JT** (1993). Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. *Hum Genet* **90**: 655-656.
- Santos FR, Pena SDJ, Tyler-Smith C** (1995). PCR haplotypes for the human Y chromosome based on alphoid satellite DNA variants and heteroduplex analysis. *Gene* **165**: 191-198.
- Santos FR, Bianchi NO, Pena SDJ** (1996). Worldwide distribution of human Y-chromosome haplotypes. *Genome Research* **6**: 601-611.
- Schmitt C, Schmutzler A, Prinz M, Staak M** (1994). High sensitive DNA typing approaches for the analysis of forensic evidence: comparison of nested variable number of tandem repeats (VNTR) amplification and short tandem repeats (STR) polymorphism. *Forensic Sci Int* **66**: 129-141.
- Schneider PM, Fimmers R, Woodroffe S, Werret DJ, Bär W, Brinkmann B, Eriksen B, Jones S, Kloosterman AD, Mevag B, Pascali VL, Rittner C, Schmitter H, Thomson JA, Gill P** (1991). Report of a European collaborative exercise comparing DNA typing results using a single locus VNTR probe. *Forensic Sci Int* **49**: 113-125.
- Schneider PM, Meuser S, Waiyawuth W, Seo Y, Rittner C** (1998a). Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and

- frequency studies in the German and three Asian populations. *Forensic Sci Int* **97**: 61-70.
- Sharma V, Litt M** (1992). Tetranucleotide repeat polymorphism at the D21S11 locus. *Hum Mol Genet* **1**: 67.
- Seielstad MT, Herbert JM, Lin AA, Underhill PA, Ibrahim M, Vollath D, Cavalli-Sforza LL** (1994). Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum Mol Genet* **3**: 2159-2161.
- Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths B, Smith MJ, Foster JW, Frischauf A-M, Lovell-Badge R, Goodfellow PN** (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**: 240-244.
- Singer MF** (1982a). Highly repeated sequences in mammalian genomes. *Int Rev Cytol* **76**: 67-112.
- Singer MF** (1982b). SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* **28**: 433-434.
- Slagel V, Flemington E, Traina-Dorge V, Bradshaw H, Deininger P** (1987). Clustering and subfamily relationships of the Alu family in the human genome. *Mol Biol Evol* **4**: 19-29.
- Spurdle AB and Jenkins T** (1992a). The search of Y chromosome polymorphism is extended to negroids. *Hum Mol Genet* **1**: 169-170.
- Spurdle AB, Jenkins T** (1992b). Y-chromosome probe p49a detects complex PvuII haplotypes and many new TaqI haplotypes in Southern African populations. *Am J Hum Genet* **50**: 107-125.
- Sullivan KM, Piercy R, Benson N, Ivanov P, Mannucci AK, Gill P** (1994). Solid phase sequencing of mitochondrial DNA: towards a fully automated forensic DNA test. In: *Advances in Biomagnetic Separation*. Ed. by M Uhlen, E Hornes, O Olsvic. Eaton Publ. Natick MA, USA, pp. 135-148.
- Tamaki K, Monckton DG, MacLeod A, Neil DL, Allen M, Jeffreys AJ** (1992). Minisatellite variant repeat (MVR) mapping: analysis of 'null' repeat units at D1S8. *Hum Mol Genet* **1**: 401-406.
- Tautz D** (1993). Notes on the definition and nomenclature of tandemly repeated DNA sequences. In: *DNA fingerprinting: state of the science*. Ed. by SDJ Pena, R Chakraborty, JT Epplen, AJ Jeffreys. Birkhauser Verlag, Basel, pp. 21-28.
- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1989). Guidelines for a Quality Assurance Program for DNA Restriction Fragment Length Polymorphism Analysis. *Crime Lab Digest* **15**: 109-113.

- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1990). Guidelines for a Proficiency Testing program for DNA Restriction Fragment Length Polymorphism Analysis. *Crime Lab Digest* **17**: 50-60.
- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1991). Guidelines for a Quality Assurance Program for DNA Analysis. *Crime Lab Digest* **18**: 44-75.
- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1993). A Guide for Conducting a DNA Quality Assurance Audit. *Crime Lab Digest* **20**: 8-18.
- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1994). Guidelines for DNA Proficiency Test Manufacturing and Reporting. *Crime Lab Digest* **21**: 27-32.
- Technical Working Group on DNA Analysis Methods [TWGDAM]** (1995). Guidelines for a Quality Assurance Program for DNA Analysis. *Crime Lab Digest* **22**: 21-43.
- Trabetti E, Casartelli A, Pignatti PF** (1994). Application of a Y-linked tetranucleotide repeat polymorphism to paternity cases. In: *Advances in Forensic Haemogenetics* **5**. Ed. by W Bär, A Fiori, and U Rossi. Springer Verlag, Berlin, Heidelberg, pp. 187-189.
- Ugozzoli L, Wallace RB** (1992). Applications of an allele-specific polymerase chain reaction to the direct determination of ABO blood group genotypes. *Genomic* **12**: 670.
- Underhill PA, Jin L, Zemmans R, Oefner PJ, Cavalli-Sforza LL** (1996). A pre-columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* **93**: 196-200.
- Underhill PA, Jin L, Lin AA, Medhdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner P** (1997). Detection of numerous Y chromosome biallelic polymorphisms by denaturing high performance liquid chromatography (DHPLC). *Genome Research* **7**: 996-1005.
- Urquhart A, Kimpton CP, Downes TJ, Gill P** (1994). Variation in Short Tandem Repeat sequences-a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Legal Med* **107**: 13-20.
- Valverde E, Cabrero C, Cao R, Rodríguez-Calvo MS, Díez A, Barros F, Alemany J, Carracedo A** (1993). Population genetics of three VNTR polymorphisms in two different Spanish populations. *Int J Legal Med* **105**: 251-256.

- Veitia R, Ion A, Barbaux S, Jobling MA, Souleyreau N, Ennis K, Ostrer H, Tosi M, Meo T, Chibain J, Fellous M, Mc Elreavey K** (1997). Mutations and sequence variants in the testis-determining region of the Y chromosome in individuals with a 46,XY female phenotype. *Hum Genet* **99**: 648-652.
- Vogt P** (1990). Potential genetic functions of tandem repeated DNA sequence blocks in the human genome are based on highly conserved chromatin folding code. *Hum Genet* **84**: 301-336.
- Wang W, Kishida T, Fukuda M, Tamaki Y** (1996). The Y-27H39 polymorphism in a Japanese population. *Int J Legal Med* **109**: 157-158.
- Watchel SS, Ohno S, Koo GC, Boyse EA** (1975). Possible role of H-Y antigen in the primary determination of sex. *Nature* **257**: 235-236.
- Waye JS, Willard HF** (1987). Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: A survey of aliphoid sequences from different chromosomes. *Nucl Acids Res* **15**: 7549-7580.
- Weber JL, May PE** (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* **44**: 388-396.
- Weber JL, Wong C** (1993). Mutation of human short tandem repeats. *Hum Mol Genet* **2**: 1123-1128.
- Weichhold GM, Keil W, Bayer B** (1996). The Y-linked locus Y27H39(DYS19), frequency distribution in South Bavarian and application to paternity testing. In: *Advances in Forensic Haemogenetics* **6**. Ed. by A Carracedo, B Brinkmann and W Bär. Springer-Verlag, Berlin, Heidelberg, pp. 641-643.
- Werret DJ** (1997). The National DNA Database. *Forensic Sci Int* **88**: 33-42.
- Whitfield LS, Sulston JE, Goodfellow PN** (1995). Sequence variation of the human Y Chromosome. *Nature* **378**: 379-380.
- Wiegand P, Schuerenkamp M, Schuette U** (1992). DNA extraction from mixtures of body fluid using mild preferential lysis. *Int J Legal Med* **104**: 359-360.
- Wolf U, Schempp W, Scherer G** (1992). Molecular biology of the human Y chromosome. *Reviews in Physiology, Biochemistry and Pharmacology* **121**: 147-213.
- Wolff RK, Nakamura Y, White R** (1988). Molecular characterization of a spontaneously generated new allele at a VNTR locus: no exchange of flanking DNA sequence. *Genomics* **3**: 247-351.

- Wong Z, Wilson V, Patel I, Povey S, Jeffreys AJ** (1987). Characterization of a panel of highly variable minisatellites cloned from human DNA. *Ann Hum Genet* **51**: 269-288.
- Wu S, Seino S, Bell GI** (1990). Human collagen, type II, alpha 21, (Col2A1) gene: VNTR polymorphism detected by gene amplification. *Nucleic Acids Res* **18**: 3102.
- Wyman AR, White R** (1980). A highly polymorphic locus in human DNA. *Proc Natl Acad Sci USA* **77**: 6754-6758.
- Yi HF, Donohue SJ, Klein DC, McBride OW** (1993). Localization of the hydroxyindole-o-methyltransferase gene to the pseudoautosomal region-implications for mapping of the psychiatric-disorders. *Hum Mol Genet* **2**: 127-131.
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Evgrafov O, Tyler-Smith C** (1997). Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* **60**: 1174-1183.
- Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Diehl SR** (1992). Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics* **14**: 1026-1031.



VI. ANEXO

Y-CHROMOSOME STR HAPLOTYPES: CONSTRUCTION OF ALLELIC LADDERS, SEQUENCING DATA AND POPULATION DATA FROM GALICIA (NW SPAIN)

C. Pestoni, M.L. Cal, S. Bellas, S. Barrai, M.V. Lareu, M.S. Rodriguez-Calvo and A. Carracedo.

Institute of Legal Medicine, University of Santiago de Compostela, 15705 Santiago de Compostela, Galicia, Spain

Y-chromosome polymorphisms are increasingly being used for anthropological and forensic purposes [1]. The non-pseudoautosomal part of the Y chromosome is particularly interesting due to its peculiarities such as its lack of recombination and mode of inheritance. Forensic usefulness of recently described Y-polymorphisms (DYS19, DYS385, DYS389-I, DYS389-II, DYS390 and DYS393) was explored in order to introduce these systems in forensic routine. Tetrameric STRs polymorphisms can be used to construct highly discriminative Y haplotypes useful for some deficient forensic cases. Our study includes population data from Galicia (NW Spain), sequencing data and construction of sequenced allelic ladders. This is particularly important for anthropological and forensic application of Y haplotypes, since the availability of sequenced allelic ladders is crucial for intercomparison purposes.

MATERIAL AND METHODS

116 male blood samples were obtained from healthy individuals from Galicia. In addition 35 father-son combinations were analyzed for studying mutations.

PCR reactions:

DYS393, DYS390: 1-25 ng DNA, 0.2 μ M each primer, and 0.65 U Taq DNA polymerase in 25 μ L

DYS19, DYS385, DYS389 I and II: 1-25 ng DNA, 0.5 μ M each primer, and 1 U Taq DNA polymerase (2 U for DYS19 DYS385) in 25 μ L.

Primer sequences were obtained from the GDB: DYS19 (GDB locus:121409); DYS385 (GDB locus:316257); DYS389 I and II (GDB locus:366108); DYS390 (GDB locus:366115); DYS393 (GDB locus:456649). Forward primer was labeled with fluorescein. PCR cycling conditions: slightly modified from the GDB.

Each locus was amplified individually, except DYS389 I/II and DYS385 (two loci each with the same set of primers). All of them contain tetranucleotides as repeat units.

Detection systems: Detection of the amplified products was carried out using the Automatic Laser Fluorescent (ALF) DNA sequencer (Pharmacia) [2]

Sequence reactions: The DNA Sequencing Kit, Dye Terminator Cycle Sequencing Ready Reaction (with AmpliTaq DNA Polymerase, FS) (PE Applied Biosystems), the ABI Prism™ 377 DNA Sequencer and the DNA sequencing analysis software (Perkin-Elmer) were used for sequencing analysis.

RESULTS

SEQUENCE DATA, ALLELIC DESIGNATION AND ALLELIC LADDERS

The consensus sequences of the different systems are shown in table 1. According with Urquhart A.J. et al (1994) [3] and Gill P. et al (1997) [4] these STRs can be classified as simple or compound STRs. Allelic designation of the systems was made according to the International Society of Forensic Haemogenetics (1997) and the EDNAP group [4]. Variant repeat regions defined allele nomenclature (Pestoni et al., sent for publication). Allelic ladders were constructed with a mixture of the sequenced common alleles found in our population. For the DYS385 system an allelic ladder of five alleles (11, 13, 15, 17, 19) was kindly supplied by P. Schneider.

POPULATION DATA

Allele frequencies of the systems and gene diversity values are shown in tables 2 and 3. DYS385 system show two alleles of two loci that cannot be unambiguously assign to each loci because of overlapping sizes. So they

Table 3. DYS385 phenotype frequencies (two loci amplified together) (n: 91) and DYS389-II frequencies. DYS385 haplotype diversity value : 0.8617

Phenotype DYS385	Frequency	Phenotype DYS385	Frequency	Locus	Allele (bp)	Frequency	Gene Diversity
9-13	0.0110	13-14	0.0330	DYS389-II (n: 116)	25 (359)	0.0172	0.6539
10-14	0.0110	13-15	0.0879		26 (363)	0.1380	
11-12	0.0110	13-16	0.0549		27 (367)	0.5259	
11-13	0.0330	13-17	0.0110		28 (371)	0.1983	
11-14	0.3297	14-14	0.0330		29 (375)	0.1034	
11-15	0.0220	14-15	0.0110		30 (379)	0.0086	
11-16	0.0220	14-16	0.0330		31 (383)	0.0086	
12-12	0.0439	14-17	0.0220				
12-13	0.0220	14-18	0.0110				
12-14	0.0658	15-15	0.0220				
12-15	0.0658	16-16	0.0110				
13-13	0.0220	16-18	0.0110				

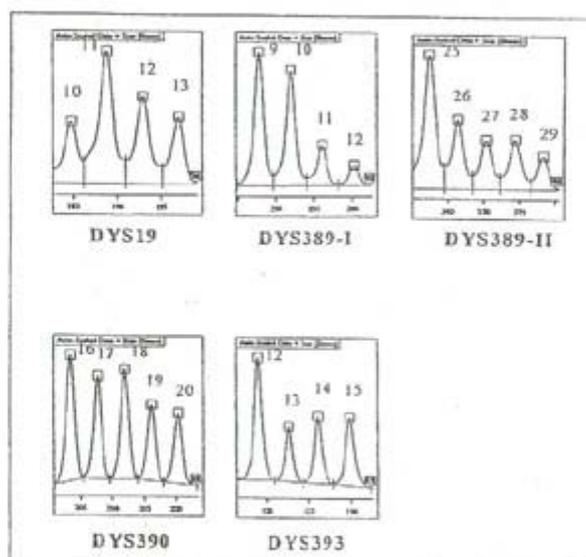


Figure 1. Allelic ladders of the studied STRs and nomenclature of alleles.

REFERENCES

- 1- Jobling MA, Tyler-Smith C (1995) Fathers and Sons: the Y chromosome and human evolution. *Trends Genet* 11:449-456
- 2- Pestoni C, Lareu MV, Rodriguez MS, Muñoz I, Barros F, Carracedo A (1996). The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data from Galicia (NW Spain) (1995) *Int J Legal Med* 107:283-290
- 3- Urquhart AJ, Kimpton CP, Downes TJ, Gill P (1994) Variation in short tandem repeat sequences-a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Leg Med* 107:13-20
- 4- Gill P, Brinkmann B, d'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Ertiksen B, Jangblad M, Johansson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Int J Leg Med* 87:185-192
- 5- Roewer L, Kayser M, Deltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet* 7 1029-1033

were analyzed as an haplotype. Significant differences were found between some European populations in some of the STRs ($p < 0.01$) [5] Pestoni et al (sent for publication).

FORENSIC USEFULNESS

The analysis of these polymorphisms could be useful in certain cases of forensic routine. The main usefulness of these systems concern to the cases in which degraded similar cells of male-female are mixed. In paternity cases the finding of different Y chromosome haplotype will exclude the suspect, always considering the mutation rate, which appears as a change in a single repeat unit of one of the systems. No mutations were found in a number of 35 father-son combinations.

Y-haplotypes were construct with the seven loci together and 80 different haplotypes were found in 93 samples (haplotype diversity value= $PD=CE= 0.9778$). The typing of base substitutions may be also useful to differentiate identical haplotypes of Y-STRs polymorphisms and will add important information for anthropological purposes [1]. Y polymorphisms are male inherited and one can only expect to analyze male samples. But it is technically simple and these markers seem to be very useful in forensic studies, particularly in mixed DNA samples. Sequenced allelic ladders and standardization of nomenclature will facilitate intercomparisons with forensic and anthropological purposes.

Table 1. Consensus sequences of the systems DYS19, DYS389-I and II, DYS390 and DYS393.

DYS19: alleles (10-14)	
Consensus structure:	
P1(21bp)- 8bp (at) ₃ agtatt (at) ₄ agtggt (at) ₅ agtggtt (TAGA) ₃ tagg (TAGA) ₁₀₋₁₄ tata 31bp -P2(19bp)	
DYS389-I: alleles (9-12)	
Consensus structure:	
P1(25bp) TATC (TGTC) ₃ (TATC) ₉₋₁₂ cctccctcTATCaatcTATCtattTATCtagc 71bp TGTC 43bp -P2(20bp)	
DYS389-II: alleles (25-30)	
Consensus structure:	
P1(25bp)-4bp (TGTC) ₅₋₆ (TATC) ₁₀₋₁₄ 44bp TATC (TGTC) ₃ (TATC) ₉₋₁₂ 150bp -P2 (20bp)	
DYS390: alleles(16-20):	
Consensus structure:	
P1(23bp)-26bp (TCTA) ₂ (TCTG) ₈₋₉ (TCTA) ₈₋₁₂ TCTG (TCTA) ₄ TCA (TCTA) ₂ 29bp-P2(23bp)	
DYS393: alleles (11-15)	
Consensus structure:	
P1(24bp)- (AGAT) ₁₁₋₁₅ (ATGT) ₂ 17bp -P2(22bp)	

Table 2. Allele frequencies of the systems. *Gene Diversity* value computed as $1-\sum p_i^2$; p_i allele frequencies

Locus	Allele (bp)	Frequency	Gene Diversity	Locus	Allele (bp)	Frequency	Gene Diversity
DYS19 (n: 92)	10 (186)	0.1413	0.5084	DYS390 (n: 116)	16 (204)	0.0259	0.6221
	11 (190)	0.6740			17 (208)	0.0345	
	12 (194)	0.1195			18 (212)	0.2672	
	13 (198)	0.0543			19 (216)	0.5345	
	14 (202)	0.0108			20 (220)	0.1379	
DYS389-I (n: 116)	9 (247)	0.1638	0.5553	DYS393 (n: 116)	11 (115)	0.0086	0.4331
	10 (251)	0.6121			12 (119)	0.1380	
	11 (255)	0.2069			13 (123)	0.7327	
	12 (259)	0.0172			14 (127)	0.1035	
			15 (131)		0.0172		

ORIGINAL ARTICLE

C. Pestoni · M. L. Cal · M. V. Lareu
M. S. Rodríguez-Calvo · A. Carracedo

Y chromosome STR haplotypes: genetic and sequencing data of the Galician population (NW Spain)

Received: 6 October 1997 / Received in revised form: 24 March 1998

Abstract Recently described Y-STR polymorphisms can be analysed as informative haplotypes which are useful in the forensic field. In order to include these systems in our forensic routine, we have carried out a population study in Galicia (NW Spain) analysing seven Y-STR polymorphisms (DYS19, DYS389-I, DYS389-II, DYS390, DYS393 and DYS385: two loci). The results were compared with other population studies. In addition various alleles for each system (except DYS385) were sequenced and the corresponding allelic ladders constructed.

Key words Y chromosome haplotypes · STRs · Sequencing data · Galician population

Introduction

In recent years a number of Y-chromosome polymorphisms have been described although the Y chromosome appears to be the least polymorphic human chromosome. Several Y polymorphisms have been studied such as some microsatellites (Roewer et al. 1992, 1996; Mathias et al. 1994; Kayser et al. 1997; de Knijff et al. 1997), base substitutions (Hammer and Horai 1995; Whitfield et al. 1995; Underhill et al. 1995) and other polymorphisms (Santos et al. 1995). Some of them such as base substitutions are quite useful to distinguish between main population groups. However, STRs are more informative to study variation within a particular population due to their polymorphism. There are several important applications of these Y-STRs in human evolution, genealogical and forensic studies (Jobling and Tyler-Smith 1995). The non-

pseudoautosomal part of the Y chromosome and mitochondrial DNA have some similarities because of the lack of recombination and the mode of inheritance. The Y chromosome is paternally inherited and passed down from father to son unchanged except by the accumulation of mutations.

Tetrameric STR polymorphisms can be used to construct highly discriminative Y haplotypes which are useful in forensic science (Roewer et al. 1996; Cooper et al. 1996; Jobling et al. 1997; Prinz et al. 1997; Kayser et al. 1997; de Knijff et al. 1997). The availability of allelic ladders (Roewer et al. 1996; Heyer et al. 1997; Kayser et al. 1997; de Knijff et al. 1997) and the collaborative studies carried out by Kayser et al. (1997) and de Knijff et al. (1997) have greatly contributed to the introduction of Y-STRs in the forensic field. Here we present sequencing and population data from Galicia (NW Spain) on the Y chromosome STRs DYS19, DYS385, DYS389-I, DYS389-II, DYS390 and DYS393 and a comparison of the results with other population studies.

Material and methods

Samples

Blood samples from 116 random males were obtained from healthy individuals from Galicia (NW Spain) and from routine paternity cases (fathers). Additionally 35 father/son combinations were included for studying mutations. DNA was extracted from EDTA blood using a phenol-chloroform procedure (Valverde et al. 1993) and quantified using the DyNA Quant 200 fluorometer (Hofer Pharmacia Biotech, San Francisco, Calif.). Oligonucleotides used as primers were obtained from Pharmacia Biotech (Sweden) and selected primers were 5'end labeled with fluorescein.

PCR conditions

Standard PCR reactions were performed in a total of 25 µl for all the systems. DYS393, DYS390: 1–20 ng DNA, 0.2 µM each primer and 0.65 U Taq DNA polymerase. DYS19, DYS385, DYS389I and II: 1–20 ng DNA, 0.5 µM each primer and 1 U Taq DNA polymerase (2 U for DYS19 and DYS385). PCR cycling conditions are shown in Table 1.

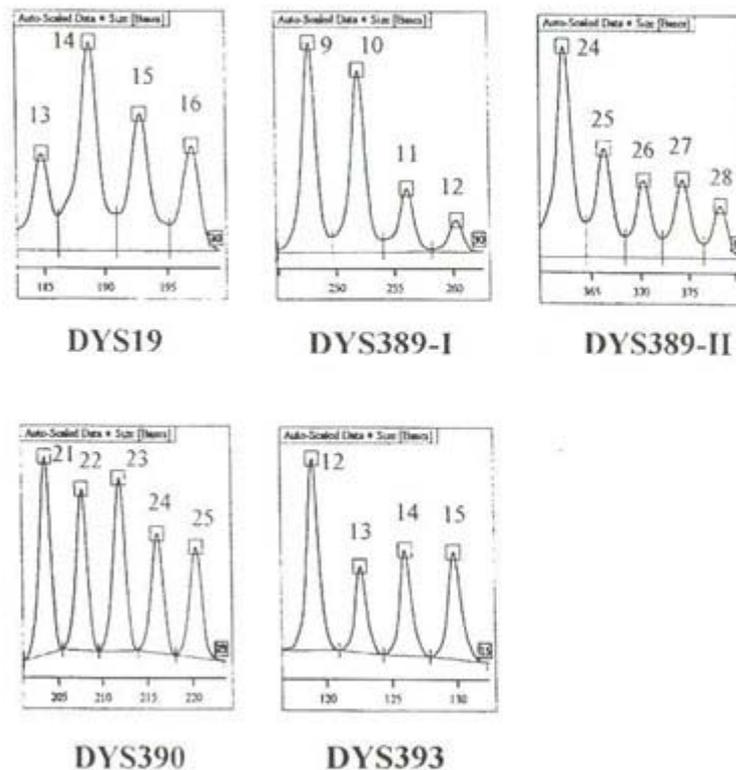
C. Pestoni · M. L. Cal · M. V. Lareu · M. S. Rodríguez-Calvo
A. Carracedo (✉)
Institute of Legal Medicine,
University of Santiago de Compostela,
E-15705 Santiago de Compostela, Galicia, Spain
e-mail: apimlang@uscmail.usc.es
Tel: +34 81 582327; Fax: +34 81 580336

Table 1 PCR cycling conditions of the systems included in this study

	Prdenatura-tion	Denaturation	Annealing	Extension	Cycles	Final extension
DYS393						
DYS389-I	94°C-2 min	94°C-15 s	58°C-15 s	72°C-20 s	30	72°C-10 min
DYS389-II						
DYS390	94°C-2 min	94°C-15 s	58°C-15 s	72°C-20 s	5	
		94°C-15 s	54°C-15 s	72°C-20 s	30	72°C-10 min
DYS19		94°C-1 min	58°C-1 min	72°C-2 min	30	
DYS385	94°C-3 min	94°C-30 s	59°C-30 s	72°C-1 min	3	
		94°C-30 s	57°C-30 s	72°C-1 min	2	
		94°C-30 s	56°C-30 s	72°C-1 min	29	

Fig. 1 Sequenced allelic ladders of the STRs studied. Nomenclature according to Kayser et al. (1997) and de Knijff et al. (1997)

Ladders



Primer sequences

DYS19 (locus GDB:121409):

Primer1: *5'CTACTGAGTTTCTGTTATAGT3'

Primer2: 5'ATGGCATGTAGTGAGGACA3'

DYS385 (locus GDB:316257):

Primer1: *5'AGCATGGGTGACAGAGCTA3'

Primer2: 5'GGGATGCTAGGTAAGCTG3'

DYS389 I and II (locus GDB:366108):

Primer 1: *5'CCAACTCTCATCTGTATTATCTATG3'

Primer 2: 5'TCTTATCTCCACCCACCAGA3'

DYS390 (locus: GDB:366115):

Primer 1: *5'TATATTTTACACATTTTGGGCC3'

Primer 2: 5'TGACAGTAAAAATGAACACATTGC3'

DYS393 (locus GDB:456649):

Primer 1: *5'GTGGTCTTCTACTTGTGTCAATAC3'

Primer 2: 5'AACTCAAGTCCAAAAAATGAGG3'

* Forward primer was labeled with fluorescein.

Table 3. DYS385 phenotype frequencies (two loci amplified together) (n: 91) and DYS389-II frequencies. *DYS385 haplotype diversity value : 0.8617*

Phenotype DYS385	Frequency	Phenotype DYS385	Frequency	Locus	Allele (bp)	Frequency	Gene Diversity
9-13	0.0110	13-14	0.0330	DYS389-II (n: 116)	25 (359)	0.0172	0.6539
10-14	0.0110	13-15	0.0879		26 (363)	0.1380	
11-12	0.0110	13-16	0.0549		27 (367)	0.5259	
11-13	0.0330	13-17	0.0110		28 (371)	0.1983	
11-14	0.3297	14-14	0.0330		29 (375)	0.1034	
11-15	0.0220	14-15	0.0110		30 (379)	0.0086	
11-16	0.0220	14-16	0.0330		31 (383)	0.0086	
12-12	0.0439	14-17	0.0220				
12-13	0.0220	14-18	0.0110				
12-14	0.0658	15-15	0.0220				
12-15	0.0658	16-16	0.0110				
13-13	0.0220	16-18	0.0110				

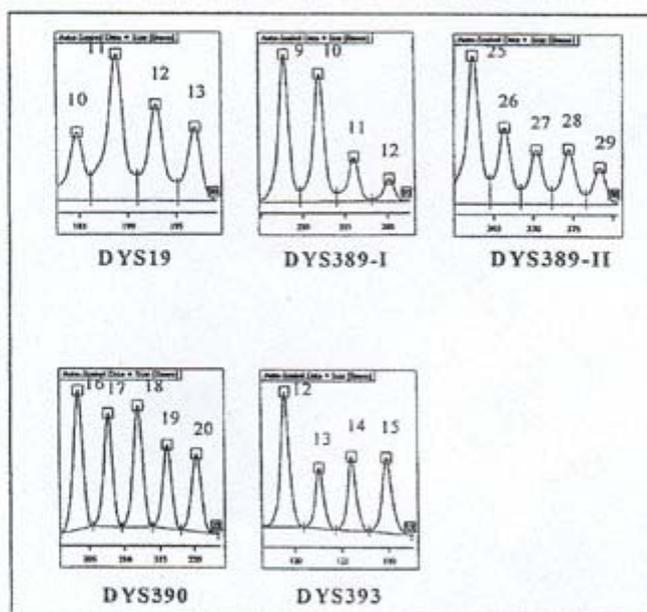


Figure 1. Allelic ladders of the studied STRs and nomenclature of alleles.

REFERENCES

- 1- Jobling MA, Tyler-Smith C (1995) Fathers and Sons: the Y chromosome and human evolution. *Trends Genet* 11:449-456
- 2- Pestoni C, Lareu MV, Rodriguez MS, Muñoz I, Barros F, Carracedo A (1996). The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data from Galicia (NW Spain) (1995) *Int J Legal Med* 107:283-290
- 3- Urquhart AJ, Kimpton CP, Downes TJ, Gill P (1994) Variation in short tandem repeat sequences-a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Leg Med* 107:13-20
- 4- Gill P, Brinkmann B, d'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Ertiksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Int J Leg Med* 87:185-192
- 5- Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawzak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet* 7 1029-1033

Table 2 Sequencing data of the systems DYS19, DYS389-I and II, DYS390 and DYS393

DYS19	
Consensus structure: P1(21bp)-8bp (at) ₃ agtatt (at) ₄ agtatt (at) ₅ agtgatt (TAGA) ₃ tagg (TAGA) _n tata 31bp -P2(19bp)	
Allele (bp)	Sequence
13 (186): P1(21bp)-8bp (at) ₃ agtatt (at) ₄ agtatt (at) ₅ agtgatt (TAGA) ₃ tagg (TAGA) ₁₀ tata 31bp -P2(19bp)	
14 (190): P1(21bp)-.....(TAGA) ₁₁-P2(19bp)	
15 (194): P1(21bp)-.....(TAGA) ₁₂-P2(19bp)	
16 (198): P1(21bp)-.....(TAGA) ₁₃-P2(19bp)	
17 (202): P1(21bp)-.....(TAGA) ₁₄-P2(19bp)	
DYS389-I	
Consensus structure: P1(25bp) TATC (TGTC) ₃ (TATC) _n cctcccctcTATCaateTATCtattTATCtagc 71bp TGTC 43bp-P2(20bp)	
Allele (bp)	Sequence
9 (247): P1(25bp) TATC (TGTC) ₃ (TATC) ₉ cctcccctcTATCaateTATCtattTATCtagc 71bp TGTC 43bp -P2(20bp)	
10 (251): P1(25bp)-.....(TATC) ₁₀-P2(20bp)	
11 (255): P1(25bp)-.....(TATC) ₁₁-P2(20bp)	
12 (259): P1(25bp)-.....(TATC) ₁₂-P2(20bp)	
DYS389-II	
Consensus structure: P1(25bp)-6bp (TCTG) _n (TCTA) _m 48bp (TCTG) ₃ (TCTA) _p 152bp-P2 (20 bp)	
Allele (bp)	Sequence
24 (359): P1(25bp)-6bp (TCTG) ₅ (TCTA) ₁₀ 48bp (TCTG) ₃ (TCTA) ₈ 152bp -P2(20bp)	
25 (363): P1(25bp)-6bp (TCTG) ₄ (TCTA) ₁₂ 48bp(TCTA) ₉ 152bp -P2(20bp)	
26 (367): P1(25bp)-6bp (TCTG) ₅ (TCTA) ₁₁ 48bp(TCTA) ₁₀ 152bp -P2(20bp)	
27 (371): P1(25bp)-6bp (TCTG) ₅ (TCTA) ₁₁ 48bp(TCTA) ₁₁ 152bp -P2(20bp)	
28 (375): P1(25bp)-6bp (TCTG) ₅ (TCTA) ₁₁ 48bp(TCTA) ₁₂ 152bp -P2(20bp)	
29 (379): P1(25bp)-6bp (TCTG) ₅ (TCTA) ₁₄ 48bp(TCTA) ₁₀ 152bp -P2(20bp)	
DYS390	
Consensus structure: P1(23bp)-26bp (TCTA) ₂ (TCTG) _n (TCTA) _m TCTG (TCTA) ₄ TCA (TCTA) ₂ 29bp-P2(23bp)	
Allele (bp)	Sequence
21 (204): P1(23bp)-26bp (TCTA) ₂ (TCTG) ₈ (TCTA) ₈ TCTG (TCTA) ₄ TCA (TCTA) ₂ 29bp -P2(23bp)	
22 (208): P1(23bp)-26bp(TCTG) ₈ (TCTA) ₉ 29bp -P2(23bp)	
23 (212): P1(23bp)-26bp(TCTG) ₈ (TCTA) ₁₀ 29bp -P2(23bp)	
P1(23bp)-26bp(TCTG) ₉ (TCTA) ₉ 29bp -P2(23bp)	
24 (216): P1(23bp)-26bp(TCTG) ₈ (TCTA) ₁₁ 29bp -P2(23bp)	
25 (220): P1(23bp)-26bp(TCTG) ₈ (TCTA) ₁₂ 29bp -P2(23bp)	
DYS393	
Consensus structure: P1(24bp)- (AGAT) _n (ATGT) ₂ 17bp-P2(22bp)	
Allele (bp)	Sequence
11 (115): P1(24bp)- (AGAT) ₁₁ (ATGT) ₂ 17bp -P2(22bp)	
12 (119): P1(24bp)- (AGAT) ₁₂ 17bp -P2(22bp)	
13 (123): P1(24bp)- (AGAT) ₁₃ 17bp -P2(22bp)	
14 (127): P1(24bp)- (AGAT) ₁₄ 17bp -P2(22bp)	
15 (131): P1(24bp)- (AGAT) ₁₅ 17bp -P2(22bp)	

Each locus was amplified individually except for DYS389 I/II and DYS385 (two loci each with the same set of primers). All loci contain tetranucleotides as repeat units.

Detection systems

Detection of the amplified products was carried out using the Automatic Laser Fluorescent (ALF) DNA sequencer (Pharmacia) following the conditions of Pestoni et al. (1995). Fragment sizes were

determined automatically using the Fragment Manager 1.2 software and typed by comparison with the sequenced allelic ladders.

Sequence reactions

PCR products were purified through MicroSpin S-300 HR columns (Pharmacia Biotech) before sequencing. Alleles from the systems DYS389-I and II were separated in a 3% agarose gel and eluted with the Sephaglass BandPrep Kit (Pharmacia Biotech).

The DNA Sequencing Kit, Dye Terminator Cycle Sequencing Ready Reaction (with AmpliTaq DNA Polymerase, FS) (PE Applied Biosystems), the ABI Prism 377 DNA Sequencer and the DNA sequencing analysis software (Perkin-Elmer) were used for sequencing analysis.

Allelic ladders and allelic designation

Sequenced allelic ladders containing all the common alleles found in our population study were constructed (Fig. 1). For the DYS385 system an allelic ladder with five alleles (11, 13, 15, 17, 19) was kindly supplied by P. Schneider (Institute of Legal Medicine, Mainz, Germany). Allelic designation of the systems was made according to Kayser et al. (1997) and de Knijff et al. (1997).

Results and discussion

Sequencing data

Sequencing data of the alleles of the different systems and their consensus structure are shown in Table 2. According to Urquhart et al. (1994) and Gill et al. (1997) these STRs can be classified as simple or compound STRs. Most of these systems have constant and variable repeat sequences and the nomenclature was made according to Kayser et al. (1997) and de Knijff et al. (1997) which is widely accepted in a large number of forensic laboratories.

The sequence of the system DYS389-I is included in the DYS389-II sequence (Cooper et al. 1996) and one set of primers amplifies both loci together because the forward primer anneals in two different places on the same strand. Unambiguous assignment of each allele to each locus is possible because the alleles of each locus differ by at least 100 bp.

Population data

Allele frequencies of the systems and gene diversity values are shown in Tables 3 and 4. The DYS385 system showed two alleles from two loci that cannot be unambiguously assigned to each loci because of overlapping sizes, so they were analysed as a haplotype. A comparative study was carried out between allele frequencies of the Galician sample and two Spanish population samples (Catalans and Basques) (Pérez-Lezaun et al. 1997) by means of χ^2 contingency tests and no significant differences were found in any of the systems ($p > 0.01$). Y-haplotypes were constructed with the seven loci together (Table 5) and the haplotype diversity value is shown in Table 6. The typing of base substitutions may also be useful to differentiate between identical haplotypes of Y-STR polymorphisms and will provide important additional information for anthropological purposes.

Forensic usefulness

The gene and haplotype diversity values have the same value as the power of discrimination (PD) and chance of

Table 3 Allele frequencies of the systems

Locus	Allele (bp)	Frequency	Gene diversity value
DYS19 (<i>n</i> = 93)	13 (186)	0.1613	0.5613
	14 (190)	0.6236	
	15 (194)	0.1398	
	16 (198)	0.0645	
	17 (202)	0.0107	
DYS389-I (<i>n</i> = 116)	9 (247)	0.1638	0.5553
	10 (251)	0.6121	
	11 (255)	0.2069	
	12 (259)	0.0172	
DYS389-II (<i>n</i> = 116)	24 (359)	0.0172	0.6539
	25 (363)	0.1380	
	26 (367)	0.5259	
	27 (371)	0.1983	
	28 (375)	0.1034	
	29 (379)	0.0086	
	30 (383)	0.0086	
DYS390 (<i>n</i> = 116)	21 (204)	0.0259	0.6221
	22 (208)	0.0345	
	23 (212)	0.2672	
	24 (216)	0.5345	
	25 (220)	0.1379	
DYS393 (<i>n</i> = 116)	11 (115)	0.0086	0.4331
	12 (119)	0.1380	
	13 (123)	0.7327	
	14 (127)	0.1035	
	15 (131)	0.0172	

Gene diversity value computed as $1 - \sum p_i^2$; p_i : allele frequencies

Table 4 DYS385 phenotype frequencies (two loci amplified together) (*n*:93)

Phenotype	Frequency	Phenotype	Frequency
9-13	0.0108	13-14	0.0322
10-14	0.0108	13-15	0.0860
11-12	0.0108	13-16	0.0215
11-13	0.0322	13-17	0.0108
11-14	0.3226	14-14	0.0322
11-15	0.0215	14-15	0.0215
11-16	0.0430	14-16	0.0430
12-12	0.0430	14-17	0.0215
12-13	0.0215	14-18	0.0108
12-14	0.0645	15-15	0.0322
12-15	0.0645	16-16	0.0108
13-13	0.0215	16-18	0.0108

Haplotype diversity value: 0.8669

exclusion (CE) i.e. $1 - (\text{matching probability}) = 1 - \sum p_i^2$; (p_i : allele or haplotype frequencies). The seven STRs described in this study result in informative Y-haplotypes with CE and PD values of 0.9794 (Table 6). A total of 35 male offspring were analysed where paternity had been positively confirmed in order to increase the number of

Table 5 Y-STR haplotypes of the Galician population sample ($n = 93$). The systems conforming the haplotypes are shown in increasing order of the PCR fragment sizes (from left to right) in order to facilitate the analysis of the data

Haplotype	<i>n</i>	DYS393	DYS19	DYS390	DYS389-I	DYS389-II	DYS385
1	1	11	13	22	9	25	11-14
2	1	12	13	24	10	26	11-14
3	1	12	13	24	10	26	16-18
4	1	12	14	22	10	26	11-14
5	1	12	14	23	10	26	14-17
6	1	12	14	23	10	27	13-15
7	1	12	14	23	10	27	13-16
8	1	12	14	23	10	29	14-14
9	1	12	14	23	11	28	12-15
10	1	12	14	24	10	25	11-14
11	1	12	14	24	10	27	12-15
12	1	12	15	23	10	26	13-15
13	1	12	15	24	10	26	12-14
14	1	12	15	24	10	28	13-15
15	1	13	13	23	10	27	11-14
16	1	13	13	24	10	28	16-16
17	1	13	13	24	11	27	13-13
18	2	13	13	24	11	27	13-14
19	1	13	13	24	11	27	13-15
20	1	13	13	24	11	27	14-14
21	1	13	13	24	12	28	11-16
22	1	13	13	25	10	26	12-14
23	1	13	13	25	11	28	13-15
24	1	13	13	25	12	28	11-16
25	1	13	14	22	9	26	14-15
26	1	13	14	23	9	25	13-15
27	2	13	14	23	10	26	11-14
28	1	13	14	23	10	26	12-12
29	1	13	14	23	10	26	12-14
30	1	13	14	23	10	26	13-17
31	1	13	14	23	11	27	10-14
32	1	13	14	23	11	27	12-13
33	1	13	14	23	11	27	12-14
34	1	13	14	24	9	25	11-14
35	1	13	14	24	9	26	11-14
36	1	13	14	24	10	24	11-14
37	1	13	14	24	10	26	11-12
38	2	13	14	24	10	26	11-13
39	9	13	14	24	10	26	11-14
40	1	13	14	24	10	26	11-16
41	1	13	14	24	10	26	12-13
42	1	13	14	24	10	26	12-14
43	1	13	14	24	10	26	13-15
44	1	13	14	24	10	26	14-15
45	1	13	14	24	10	26	14-16
46	1	13	14	24	10	27	11-14
47	1	13	14	24	11	26	11-14
48	2	13	14	24	11	27	11-14
49	1	13	14	24	11	27	11-15
50	1	13	14	24	11	28	11-14
51	2	13	14	25	9	25	11-14
52	1	13	14	25	9	25	12-14
53	1	13	14	25	10	26	9-13
54	1	13	14	25	10	26	11-14
55	1	13	14	25	10	26	11-16
56	1	13	14	25	10	26	12-15
57	1	13	14	25	10	26	13-15
58	1	13	15	22	10	25	12-12

Table 5 (continued)

Haplotype	n	DYS393	DYS19	DYS390	DYS389-I	DYS389-II	DYS385
59	1	13	15	23	9	26	13-13
60	1	13	15	23	9	26	14-18
61	1	13	15	23	10	26	13-16
62	1	13	15	24	10	26	11-13
63	1	13	15	24	10	26	11-14
64	1	13	15	24	11	26	11-15
65	1	13	16	21	9	26	14-16
66	1	13	16	23	10	25	12-12
67	1	13	16	25	9	25	14-16
68	1	13	16	25	10	28	11-14
69	1	13	17	24	10	25	12-12
70	1	14	13	24	10	28	13-14
71	1	14	14	23	9	26	14-14
72	1	14	14	24	10	26	11-14
73	1	14	14	24	10	27	12-15
74	1	14	15	23	10	26	14-16
75	1	14	15	23	11	28	15-15
76	1	14	15	23	11	30	15-15
77	1	14	16	21	9	26	14-17
78	1	14	16	23	10	26	15-15
79	2	15	14	24	11	27	12-15
Total:	93						

Table 6 Haplotype diversity values

	Haplotype 1 (DYS389-I, II, DYS390, DYS393)	Haplotype 1+DYS385, DYS19
Number of individuals	116	93
Different haplotypes	50	79
Haplotype diversity value	0.9214	0.9794

samples studied for calculating the mutation rate of these systems. No mutations were found. In addition DNA from 20 females was amplified in order to demonstrate that they are exclusively male polymorphic systems and no amplification products were obtained.

The study was carried out by amplifying all the systems individually. Multiplexing of the systems has also been performed with excellent results. The analysis of these polymorphisms could be useful in certain forensic cases (Prinz et al. 1997; Jobling et al. 1997; Kayser et al. 1997).

In our experience we have also obtained excellent results in some critical forensic cases. We have obtained such results especially in degraded male-female mixtures in rape cases and in paternity testing of male offspring using amniotic fluid taken in the early stages of a pregnancy which has been the result of a rape.

Acknowledgements This work was supported by a grant from The Xunta de Galicia (XUGA 20806B97). We also appreciate the technical assistance of Amelia Rodríguez.

References

- Cooper G, Amos W, Hoffman D, Rubinsztein DC (1996) Network analysis of human Y microsatellite haplotypes. *Hum Mol Genet* 11: 1759-1766
- De Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* 110: 134-149
- Gill P, Brinkmann B, d'Aloja E, Andersen J, Bär W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Int J Legal Med* 87: 185-192
- Hammer MF, Horai S (1995) Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* 56: 951-962
- Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* 6: 799-803
- Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. *Trends Genet* 11: 449-456
- Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* 110: 118-124
- Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* 110: 125-133; appendix 141-149

- Mathias N, Bayés M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3:115-123
- Pérez-Lezaun A, Calafell F, Scielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y-chromosome short tandem repeats in humans. *J Mol Evol* 45:265-270
- Pestoni C, Lareu MV, Rodríguez MS, Muñoz I, Barros F, Carracedo A (1995) The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data from Galicia (NW Spain). *Int J Legal Med* 107:283-290
- Prinz M, Boll K, Baum H, Shaler B (1997) Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Sci Int* 85:209-218
- Roewer L, Arnemann J, Spurr NK, Grzeschik KH, Epplen JT (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet* 89:389-394
- Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawzak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet* 7:1029-1033
- Santos FR, Pena SDJ, Tyler-Smith C (1995) PCR haplotypes for the human Y chromosome based on alphoid satellite DNA variants and heteroduplex analysis. *Gene* 165:191-198
- Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1995) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93:196-200
- Urquhart AJ, Kimpton CP, Downes TJ, Gill P (1994) Variation in short tandem repeat sequences—a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Legal Med* 107:13-20
- Valverde E, Cabrero C, Cao R, Rodríguez-Calvo MS, Díez A, Barros F, Alemany J, Carracedo A (1993) Population genetics of three VNTR polymorphisms in two different Spanish populations. *Int J Legal Med* 105:251-256
- Whitfield LS, Suiston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. *Nature* 378:379-380

Y-CHROMOSOME STR HAPLOTYPES: CONSTRUCTION OF ALLELIC LADDERS, SEQUENCING DATA AND POPULATION DATA FROM GALICIA (NW SPAIN)

C. Pestoni, M.L. Cal, S. Bellas, S. Barral, M.V. Lareu, M.S. Rodríguez-Calvo and A. Carracedo.

Institute of Legal Medicine, University of Santiago de Compostela, 15705 Santiago de Compostela, Galicia, Spain

Y-chromosome polymorphisms are increasingly being used for anthropological and forensic purposes [1]. The non-pseudoautosomal part of the Y chromosome is particularly interesting due to its peculiarities such as its lack of recombination and mode of inheritance. Forensic usefulness of recently described Y-polymorphisms (DYS19, DYS385, DYS389-I, DYS389-II, DYS390 and DYS393) was explored in order to introduce these systems in forensic routine. Tetrameric STRs polymorphisms can be used to construct highly discriminative Y haplotypes useful for some deficient forensic cases. Our study includes population data from Galicia (NW Spain), sequencing data and construction of sequenced allelic ladders. This is particularly important for anthropological and forensic application of Y haplotypes, since the availability of sequenced allelic ladders is crucial for intercomparison purposes.

MATERIAL AND METHODS

116 male blood samples were obtained from healthy individuals from Galicia. In addition 35 father-son combinations were analyzed for studying mutations.

PCR reactions:

DYS393, DYS390: 1-25 ng DNA, 0.2 µM each primer, and 0.65 U Taq DNA polymerase in 25 µl.

DYS19, DYS385, DYS389I and II: 1-25 ng DNA, 0.5 µM each primer, and 1 U Taq DNA polymerase (2 U for DYS19 DYS385) in 25 µl.

Primer sequences were obtained from the GDB: DYS19 (GDB locus:121409); DYS385 (GDB locus:316257); DYS389 I and II (GDB locus:366108); DYS390 (GDB locus:366115); DYS393 (GDB locus:456649). Forward primer was labeled with fluorescein. PCR cycling conditions: slightly modified from the GDB.

Each locus was amplified individually, except DYS389 I/II and DYS385 (two loci each with the same set of primers). All of them contain tetranucleotides as repeat units.

Detection systems: Detection of the amplified products was carried out using the Automatic Laser Fluorescent (ALF) DNA sequencer (Pharmacia) [2]

Sequence reactions: The DNA Sequencing Kit, Dye Terminator Cycle Sequencing Ready Reaction (with AmpliTaq DNA Polymerase, FS) (PE Applied Biosystems), the ABI Prism™ 377 DNA Sequencer and the DNA sequencing analysis software (Perkin-Elmer) were used for sequencing analysis.

RESULTS

SEQUENCE DATA, ALLELIC DESIGNATION AND ALLELIC LADDERS

The consensus sequences of the different systems are shown in table 1. According with Urquhart A.J. et al (1994) [3] and Gill P. et al (1997) [4] these STRs can be classified as simple or compound STRs. Allelic designation of the systems was made according to the International Society of Forensic Haemogenetics (1997) and the EDNAP group [4]. Variant repeat regions defined allele nomenclature (Pestoni et al., sent for publication). Allelic ladders were constructed with a mixture of the sequenced common alleles found in our population. For the DYS385 system an allelic ladder of five alleles (11, 13, 15, 17, 19) was kindly supplied by P. Schneider.

were analyzed as an haplotype. Significant differences were found between some European populations in some of the STRs ($p < 0.01$) [5] Pestoni et al (sent for publication).

FORENSIC USEFULNESS

The analysis of these polymorphisms could be useful in certain cases of forensic routine. The main usefulness of these systems concern to the cases in which degraded similar cells of male-female are mixed. In paternity cases the finding of different Y chromosome haplotype will exclude the suspect, always considering the mutation rate, which appears as a change in a single repeat unit of one of the systems. No mutations were found in a number of 35 father-son combinations.

Y-haplotypes were construct with the seven loci together and 80 different haplotypes were found in 93 samples (haplotype diversity value= $PD=CE= 0.9778$). The typing of base substitutions may be also useful to differentiate identical haplotypes of Y-STRs polymorphisms and will add important information for anthropological purposes [1]. Y polymorphisms are male inherited and one can only expect to analyze male samples. But it is technically simple and these markers seem to be very useful in forensic studies, particularly in mixed DNA samples. Sequenced allelic ladders and standardization of nomenclature will facilitate intercomparisons with forensic and anthropological purposes.

Table 1. Consensus sequences of the systems DYS19, DYS389-I and II, DYS390 and DYS393.

DYS19: alleles (10-14)	
Consensus structure:	
P1(21bp)- 8bp (at) ₃ agtatt (at) ₄ agtggt (at) ₅ agtggtt (TAGA) ₃ tagg (TAGA) ₁₀₋₁₄ tata 31bp -P2(19bp)	
DYS389-I: alleles (9-12)	
Consensus structure:	
P1(25bp) TATC (TGTC) ₃ (TATC) ₉₋₁₂ cctccctcTATCaatcTATCtattTATCtagc 71bp TGTC 43bp -P2(20bp)	
DYS389-II: alleles (25-30)	
Consensus structure:	
P1(25bp)-4bp (TGTC) ₅₋₆ (TATC) ₁₀₋₁₄ 44bp TATC (TGTC) ₃ (TATC) ₉₋₁₂ 150bp -P2 (20bp)	
DYS390: alleles(16-20):	
Consensus structure:	
P1(23bp)-26bp (TCTA) ₂ (TCTG) ₈₋₉ (TCTA) ₈₋₁₂ TCTG (TCTA) ₄ TCA (TCTA) ₂ 29bp-P2(23bp)	
DYS393: alleles (11-15)	
Consensus structure:	
P1(24bp)- (AGAT) ₁₁₋₁₅ (ATGT) ₂ 17bp -P2(22bp)	

Table 2. Allele frequencies of the systems. *Gene Diversity value* computed as $1 - \sum p_i^2$; p_i : allele frequencies

Locus	Allele (bp)	Frequency	Gene Diversity	Locus	Allele (bp)	Frequency	Gene Diversity
DYS19 (n: 92)	10 (186)	0.1413	0.5084	DYS390 (n: 116)	16 (204)	0.0259	0.6221
	11 (190)	0.6740			17 (208)	0.0345	
	12 (194)	0.1195			18 (212)	0.2672	
	13 (198)	0.0543			19 (216)	0.5345	
	14 (202)	0.0108			20 (220)	0.1379	
DYS389-I (n: 116)	9 (247)	0.1638	0.5553	DYS393 (n: 116)	11 (115)	0.0086	0.4331
	10 (251)	0.6121			12 (119)	0.1380	
	11 (255)	0.2069			13 (123)	0.7327	
	12 (259)	0.0172			14 (127)	0.1035	
			15 (131)		0.0172		