

Polimorfismos en el ADN humano

José Fernández Piqueras

Catedrático de Genética, Universidad Autónoma de Madrid

1.-Introducción

- Definición
- Antecedentes
- El advenimiento de la tecnología del ADN recombinante

2.- Tipos de polimorfismos

- Fragmentos de restricción de longitud variable
- Repeticiones en tándem de secuencias cortas
- Otros polimorfismos

3.-Aplicaciones

- Informatividad y análisis de ligamiento
- Elaboración de mapas genéticos
- Búsqueda de genes, diagnóstico anticipado e identificación genética.
- Estudios evolutivos

1.-Introducción

La vigésima primera edición del Diccionario de la Lengua Española (1992) define el término polimorfismo bajo una única acepción (química) como la propiedad de los cuerpos que pueden cambiar de forma sin variar su naturaleza. Para el adjetivo polimorfo (o polimórfico) reconoce como más cercana a nosotros la

acepción bioquímica: "Dícese de las enzimas o proteínas que se presentan bajo varias formas moleculares. Es un fenómeno importante en la genética y en la patología molecular". Así pues, aunque se reconozca su importancia en la genética, a este diccionario se le escapa la posibilidad de que los polimorfismos se puedan reconocer directamente en las propias moléculas de ADN.

Ya en el campo de la genética, Ford había definido en el año 1940 el polimorfismo de una población como la ocurrencia en un mismo ambiente de dos o más formas discontinuas de una especie, en proporciones tales que la más rara de ellas no pueda ser mantenida en la población simplemente por mutación recurrente.

En definitiva un polimorfismo genético aunque sea estimado a través de una pluralidad de formas fenotípicas (morfología externa, variantes enzimáticas etc.) se podría definir como un carácter mendeliano (monogénico) que se presenta al menos bajo dos formas alternativas en una población. Su origen está en la ocurrencia de mutaciones, pero su mantenimiento no puede depender de la recurrencia de una mutación.

Se podría por consiguiente hablar de polimorfismos cromosómicos o de otros fundamentados en alteraciones de la secuencia de ADN, y sólo faltaría el establecimiento de un criterio de valoración de carácter cuantitativo. La mayor parte de los autores coinciden en aceptar que ninguna de las variantes alélicas tenga una frecuencia inferior al 1%.

El primer polimorfismo descrito en la especie humana fue el descubrimiento del grupo sanguíneo ABO (Landsteiner en 1900), y durante los 55 años siguientes todos los polimorfismos descritos se refirieron a diferentes antígenos de superficie de los glóbulos rojos. En 1955 Smithies puso a punto una técnica de electroforesis en geles de almidón que permitía la separación de proteínas no sólo por su carga sino también por su tamaño molecular. Con este nuevo método se detectaron nuevos polimorfismos en proteínas séricas como las haptoglobinas. El procedimiento se enriqueció inmediatamente con variaciones que permitían la catalogación de actividades enzimáticas (polimorfismos enzimáticos).

El advenimiento de la tecnología del ADN recombinante a mediados de los setenta supuso la apertura de una nueva vía para la identificación de marcadores genéticos polimórficos mediante el análisis directo de las propias moléculas de ADN. Nace una nueva generación de polimorfismos: los polimorfismos del ADN.

2.- Tipos de polimorfismos

Los polimorfismos de ADN se pueden ordenar según dos criterios básicos: su propia naturaleza y el sistema de detección. De este modo se podrían distinguir:

–Tipo I: Fragmentos de restricción de longitud variable (RFLPs).– La causa del polimorfismo radica normalmente en la pérdida de una o más dianas de restricción debido a mutaciones puntuales (transiciones o transversiones), y se detectan mediante electroforesis en geles de agarosa.

–Tipo II: Repeticiones en tándem de secuencias cortas (STRs).– En este caso el polimorfismo consiste en la existencia de un número variable de repeticiones en tándem de una secuencia básica que oscila entre un solo nucleótido y una kilobase. Se detectan también mediante electroforesis en el mismo tipo de geles, salvo los microsátélites que, como se verá más adelante, suelen necesitar geles de poliacrilamida.

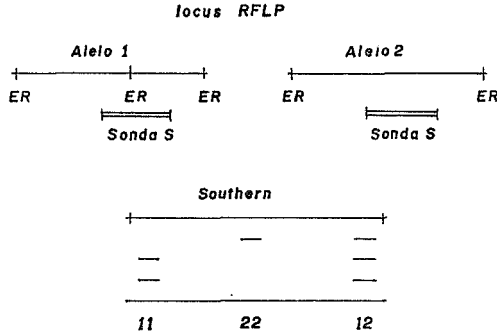
–Tipo III: Polimorfismos debidos a mutaciones puntuales que no afectan a la longitud de los fragmentos de restricción ni alteran la secuencia de ninguna diana conocida.– En estos casos existen métodos indirectos para comprobar la naturaleza polimórfica de un fragmento de ADN (SSCP, DGGE etc.), y lógicamente está el recurso de poder secuenciar directamente el ADN.

El análisis de los polimorfismos de ADN se ha visto enormemente facilitado con la incorporación en el año 1985 de la "técnica de amplificación en cadena de la polimerasa" (PCR) (Saiki, 1990). Con esta nueva metodología, se evitan: la necesidad de extraer y purificar el ADN genómico, el disponer de sondas clonadas, la realización de transferencias, etc. Eso sí, es imprescindible conocer al menos las secuencias de bases adyacentes al fragmento que se pretende amplificar.

Polimorfismos del tipo I (RFLPs).– Desde mediados de los setenta, con la introducción de las endonucleasas de restricción en el análisis genético molecular por el grupo del premio Nobel Nathans (Nathans and Smith, 1975), cualquier molécula de ADN puede ser cortada en sitios específicos produciendo unos fragmentos llamados por ello de restricción. Estas enzimas son un tipo especial de endonucleasas (tipo II) capaces de reconocer secuencias específicas del ADN (dianas de restricción) y catalizar la ocurrencia de cortes diplotómicos

(en ambas cadenas). Los fragmentos liberados se pueden caracterizar mediante electroforesis en geles de agarosa, y si se realiza el mismo tipo de digestiones en el ADN de diferentes individuos, se pueden comparar los fragmentos equivalentes mediante hibridación molecular con sondas específicas utilizando la técnica de Southern (Southern, 1975).

1a



1b

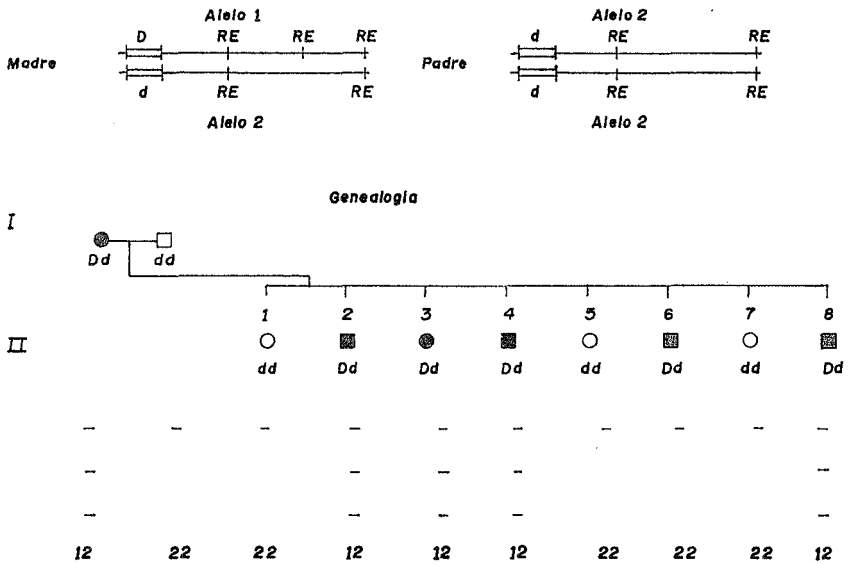


Fig. 1. Fragmentos de restricción polimórficos de longitud variable (a) y su aplicación para la búsqueda de genes (b)

Las sondas de ADN son, normalmente, secuencias génicas o anónimas de ADN, de localización cromosómica conocida, clonadas en plásmidos o cósmidos. Cuando se obtienen a partir de RNAm procesados se habla de sondas ADNc, y cuando derivan directamente de genomas completos sondas genómicas:

Para este tipo de análisis, el ADN humano se suele extraer de linfocitos de sangre periférica. Con 10-20 ml de sangre se puede disponer del ADN suficiente para hacer unas 10 digestiones con endonucleasas e hibridar con otras tantas sondas. Una segunda fuente es la placenta: ahora bastarían unos 10 g. Sin embargo lo ideal sería disponer de la inagotable fuente de líneas celulares (transformadas) procedentes de fibroblastos epiteliales o de las propias células sanguíneas. En todo caso el ADN extraído debe ser siempre de alto peso molecular (no degradado), y el marcaje de las sondas puede hacerse tanto por procedimientos radioactivos como no-isotópicos. Existen en la actualidad excelentes manuales de laboratorio para el desarrollo de estas metodologías (Sambrook et al., 1990; Mathew, 1991).

El polimorfismo del tipo I (RFLPs) se produce cuando la secuencia de reconocimiento de una endonucleasa se ve afectada por una mutación puntual, y se pierde por consiguiente la capacidad de reconocimiento y corte en ese lugar. La consecuencia inmediata es el cambio en los patrones electroforéticos que presentan los fragmentos de restricción. En la figura 1a se esquematiza la identificación de dos variantes alélicas de un locus RFLP con la sonda S.

Con esta nueva metodología se empezó a comprobar inmediatamente que los genomas tenían niveles de variación muy superiores a los estimados mediante análisis de proteínas. Jeffreys (1979) estimó que 1 de cada 100 nucleótidos de secuencias no codificantes podía ser polimórfico. Estudios sistemáticos llevados a cabo con secuencias génicas (globinas y albúmina) y no codificantes, apuntaban igualmente una diversidad genética mayor que la demostrada con el análisis de proteínas: 1 de cada 250 o 500 nucleótidos acababa presentando alguna variación (Botstein et al., 1980). La probabilidad de encontrar una variante de un fragmento de restricción definido por una endonucleasa con una diana de 4 pares de bases es del 12,3%. Si la diana es de 6 pares de bases esta probabilidad aumenta hasta el 17.7%.

Polimorfismos tipo II: secuencias cortas repetidas en tándem.—En todos los genomas, y muy especialmente en el humano, existen secuencias cortas que se repiten en tándem y están sometidas a una alta tasa de variabilidad. La longitud

de estas unidades básicas de repetición oscila entre 1 nucleótido y 2 kilobases (Craig et al., 1988). Este tipo de secuencias ha sido bautizado con diferentes nombres, pero parece bastante aceptada la separación en dos categorías según el tamaño de la unidad repetida y el sistema de detección. Cuando la unidad básica es de unas pocas bases (entre una y seis) se suelen denominar microsátélites (Litt and Luty, 1989) o polimorfismos de longitud de secuencia simple (Tautz, 1989). Si la unidad de repetición es mayor (normalmente entre 10 y 15 pares de bases), se habla de minisátélites (Jeffreys et al., 1985) o repeticiones en tándem de número variable (VNTR) (Nakamura et al., 1987). Este tipo de polimorfismos se ven favorecidos por la ocurrencia de "crossing-over" desiguales y otros mecanismos moleculares (como el "slippage" durante la replicación) que contribuyen a su extensión. El resultado son regiones hipervariables y polimorfismos multialélicos, pero hay que tener cuidado a la hora de interpretar sus patrones de segregación porque pueden sufrir modificaciones con una frecuencia superior a la de las mutaciones puntuales, apareciendo en los hijos variantes alélicas que no existían en los padres.

De las dos categorías consideradas, los minisátélites son lógicamente menos polimórficos y, curiosamente, tienden a estar agrupados en las regiones pro-terminales de los cromosomas humanos. Pueden distinguirse fácilmente mediante Southern blot e hibridación, pero los más pequeños necesitan ser amplificados mediante PCR (Armour et al., 1992).

Fue precisamente el desarrollo de esta técnica (PCR), que permitía amplificar segmentos cortos de ADN con gran rapidez y eficacia, lo que hizo posible el descubrimiento de los microsátélites (Tautz, 1989; Litt and Luty, 1989).

Se han descrito un buen número de microsátélites en todos los ADN eucarióticos excepto, curiosamente, en levaduras (Hearne et al., 1992). El comité de la HGM10.5 tenía ya censados 368 microsátélites con unidades de repetición en torno a 4 pares de bases y longitudes totales de unos 20 pares de bases. La mayoría de estos datos están recogidos en la Genome Data Base (GDB) del Howard Hughes Medical Institute y la John Hopkins University (Baltimore, USA).

Según Beckman & Weber (1992) el 76% de las unidades básicas de repetición son A, AC, AAAN y AG (siendo N cualquier nucleótido). Los microsátélites AC son los más frecuentes; aparecen cada 30 kilobases y se distribuyen por igual en las regiones 5' y 3' no traducibles de los genes y en las regiones intrónicas. Como dato anecdótico, el 80% de las repeticiones A, AAAN y AAN, y el

50% de los microsatélites AT, se han encontrado en el extremo 3' de elementos Alu. En el cromosoma X se han descrito microsatélites de 3 y 4 pares de bases cada 300-500 kilobases, aunque esta misma cadencia parece repetirse por todo el genoma humano (Edwards et al., 1992).

En 1990 Weber realizó un exhaustivo estudio sobre el tipo y distribución de los microsatélites (dC-dA)_n (dG-dT)_n en el genoma humano (Weber, 1990). Para ello, rastreó en las secuencias de ADN recogidas en bases de datos, y secuenció numerosos clones M 13 que hibridaban con sondas poly(dC-dA).poly(dG-dT). El resultado fue el hallazgo de tres tipos de microsatélites:

a. Secuencias de repetición perfecta.- Consistentes en repeticiones AC en tándem sin interrupciones ni otro tipo de secuencias repetidas adyacentes.

b. Secuencias de repetición imperfectas.- Consistentes en dos o más series de repeticiones CA interrumpidas por bases nucleotídicas que no se repiten.

c. Secuencias de repetición compuestas.- Pueden ser también perfectas o imperfectas según el status de las series CA.

En la Tabla 1 quedan recogidos ejemplos de estos tipos de microsatélites con indicación de su frecuencia relativa y la naturaleza de las secuencias adyacentes.

Tabla 1. Ejemplos de diferentes tipos de microsatélites humanos con indicación de su frecuencia (%) y el tipo de secuencias adyacentes (datos de Weber, 1990)

Tipo	Secuencias 5'	Repetición	Secuencias 3'
Perfectos (64%)	TTTAGAAAAA ACAGGCATCA CATGCACGTG	(AC) ₁₉ (AC) ₂₀ (CA) ₂₀	CCCCAAAGCT CACAAAGTGC TACACCAGCT
Imperfectos (26%)	TTGTTGATTT CTTTCTCAGG CAGTCCAAGC ATATAAACAT	(CA) ₁₁ CT(CA) ₄ A(CA) ₇ GAG(AC) ₁₄ A (CA) ₈ G(AC) ₄ AG(AC) ₃ AT(AC) (CA) ₅ G(ACA)G(AC) ₇ A	TACTGATGTG GGCAATGACA ATTTTCATT ACAATTAACA
Compuestos perfectos (4%)	CCTTGTCTCT ATACATACAT ACTCCATCTC	(AC) ₁₆ (TC) ₁₀ (AC) ₂₀ AG(AGAC) ₅ AGA (A) ₁₇ G(CA) ₁₆ C	AGCCAGGCAC ATACTGTCTC TCACTACATT
Compuestos imperfectos (6%)	GGCATGCATG TCTTACACCA	(CA) ₉ A(AC) ₁₉ A(GA) ₇ T(CT) ₁₂ GTT(TC) ₁₁ T(CA) ₁₄ A(AC) ₆	CATGCTGTTC CATGCTGTTC

Como se ha mencionado, la detección de los microsátélites es algo más sofisticada que la de minisátélites. Al tratarse de secuencias muy cortas pueden clonarse y secuenciarse con relativa facilidad a partir de ADN genómico total o de regiones cromosómicas específicas. El procedimiento más utilizado es la elaboración de genotecas en vectores de secuenciación, cortando el ADN genómico con endonucleasas de restricción de corte frecuente (p.e., Sau3AI, o combinaciones de AluI, RsaI y HaeIII). Después se seleccionan los clones que son capaces de hibridar con oligonucleótidos específicos por ejemplo de tipo (AC)_n. Los fragmentos de ADN resultantes son finalmente analizados en geles de poliacrilamida en condiciones desnaturizantes (geles de secuenciación). De esta forma se pueden distinguir, si el tamaño del fragmento no es excesivo, diferencias de tan sólo uno o dos nucleótidos de longitud (Edwards et al., 1991 & 1992).

Polimorfismos del tipo III.– Cuando el origen de los polimorfismos está en la ocurrencia de mutaciones puntuales como las transiciones y transversiones, y estas mutaciones no afectan a ninguna diana de restricción conocida, los fragmentos de ADN variantes pasarían inadvertidos con los sistemas de detección anteriores. Sin embargo, existen métodos adicionales de carácter indirecto (SSCP, DGGE etc.) o directo (secuenciación), que hacen posible también su detección.

Entre los métodos indirectos, uno de los más empleados es el que aprovecha los cambios en la movilidad electroforética que presentan dos cadenas sencillas de ADN cuando se diferencian en una única sustitución de bases. Esta diferente movilidad electroforética se puede poner de manifiesto en geles neutros de poliacrilamida y se basa al parecer en sutiles diferencias conformacionales; por ello se habla de "single strand conformation polymorphism" (SSCP) (Orita et al., 1989). En cuanto al procedimiento directo, es decir la secuenciación, parece cada día más asequible desde la incorporación de los sistemas automatizados, y por la posibilidad de secuenciar directamente los productos amplificados mediante PCR.

3.- Aplicaciones

Informatividad y análisis de ligamiento.– La disponibilidad de marcadores polimórficos de ADN del tipo I (RFLPs) fue aprovechada inmediatamente por Botstein (Botstein et al., 1980) para la elaboración de un mapa genético

más detallado de la especie humana. Sin embargo no todos los marcadores genéticos son igualmente polimórficos, y se impone por tanto la necesidad de adoptar algún criterio que sirva para estimar la valía de un marcador para el análisis de ligamiento, es decir, su informatividad. Los dos procedimientos más utilizados son la determinación del índice de heterocigosidad (H), y el denominado PIC (polymorphic information content). La heterocigosidad de un locus polimórfico viene dada por el porcentaje de individuos que son heterocogóticos, y se calcula mediante la fórmula $H = 1 - (\sum_{i=1}^n p_i^2)$, donde p_i son las frecuencias de los i alelos de ese locus. El PIC fue creado por Botstein para medir la informatividad de un marcador genético polimórfico (marker locus) que se quiere utilizar para mapear un carácter autosómico dominante raro ("index locus"), por ejemplo el alelo responsable de una enfermedad. Para el caso de un locus marcador A_i con n alelos y frecuencias p_i , el valor PIC vendría dado por la fórmula: $PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^n \sum_{j=1}^n 2p_i p_j^2$. Esta fórmula se ha adaptado también para mapear caracteres raros con patrones de herencia diferentes (Chakravarti and Buetow, 1985).

Elaboración de mapas genéticos.— La construcción de un mapa de ligamiento completo de la especie humana (parte esencial del proyecto Genoma Humano) se ha visto facilitada esencialmente por tres factores:

1.— Los avances de la genética molecular que han permitido la identificación de varios miles de polimorfismos de ADN, muchos de ellos multialélicos y altamente informativos (PIC >0.7).

2.— El desarrollo de paquetes estadísticos más adecuados para el establecimiento del orden y distancia entre genes.

3.— La construcción de paneles de familias de referencia como los del "Centre d'Etude du Polymorphisme Humanin" (CEPH) (París) que está compuesto por 60 familias con una media de 8 descendientes cada una. También existe otro excelente panel de familias venezolanas seleccionadas por Gusella (Gusella et al., 1983) para el estudio de la Corea de Huntington.

En la figura 2, se puede observar el número total de marcadores genéticos que se han ido incorporando al mapa genético humano desde el año 1983. En la actualidad hay unos 10.000 loci definidos mediante sondas de ADN, de los cuales la mitad aproximadamente son polimórficos. Esto supondría la disponibilidad de un marcador polimórfico por megabase, pero la distribución no es ho-

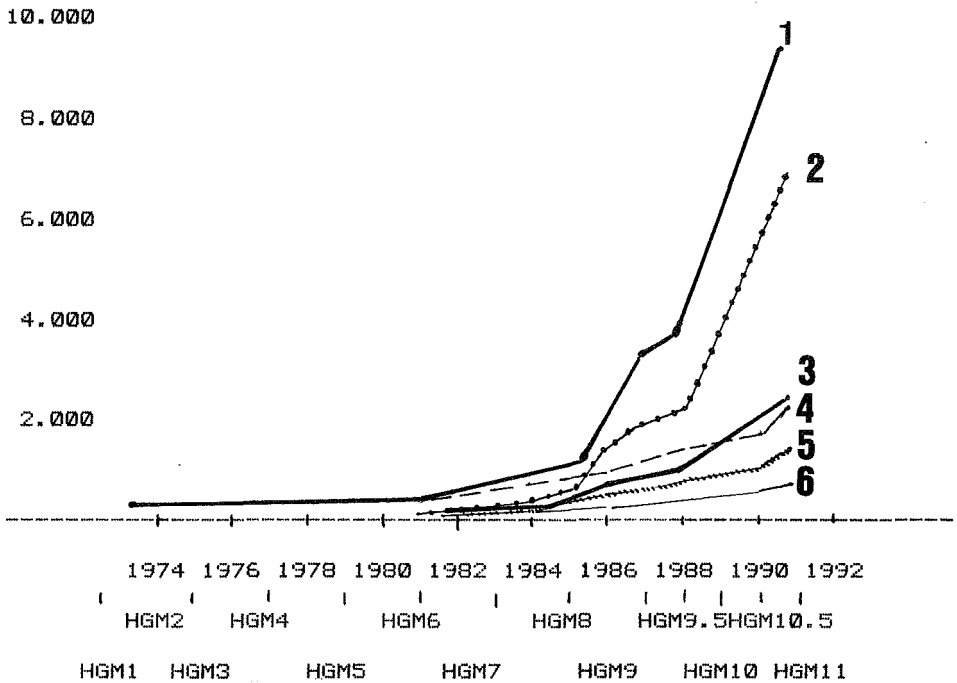


Figura 2.- Crecimiento exponencial del número de loci génicos y marcadores anónimos (D-segments), que se han ido incorporando al mapa genético humano (HGM) desde el año 1974 hasta la conferencia Londres (1991) (Datos tomados de HGM11, 1991). 1, total de loci. 2, Segmentos D. 3, Segmentos D polimórficos. 4, genes. 5, genes clonados. 6, genes polimórficos.

mogénea hay cromosomas como el X o el 21 que tienen más que los demás. El mayor incremento en el número de polimorfismos de ADN se ha producido sin duda en los dos últimos años, como consecuencia de la técnica de amplificación en cadena de la polimerasa (PCR) que permitió la detección de numerosos microsatélites. En la Tabla 2, tomada de Hearne et al. (1992), se refieren los microsatélites humanos más informativos (PICs >0.75) con indicación de su localización cromosómica.

Búsqueda de genes, diagnóstico anticipado e identificación genética.-

Desde el advenimiento de la moderna tecnología del ADN recombinante, hace tan solo unos 15 años, se han clonado y analizado unos pocos cientos de genes responsables de enfermedades congénitas humanas. La gran mayoría de estos genes han podido ser identificados porque se conocía la base bioquímica del

Tabla 2.-Listado de microsátélites humanos con indicación de su localización cromosómica y PIC (datos de Hearne et al., 1992).

LOCUS	MAP	PIC	LOCUS	MAP	PIC	LOCUS	MAP	PIC
1 DIS117 DIS103	q23-q25.1 q32-qter	0.77 0.78	2 GOG	q36-q37	0.78	3 D3S11 GLUT2	p14-p21 q26.1q26.3	0.88 ref 52 0.83
4 D4S230 D4S174 D4S194 D4S175	p16.1-p11 q11-q13 q25-q34 q31	0.84 0.86 0.8 0.82	5 D5S208 D5S39 CRTL1 D5S318 D5S346 IL9 D5S210 CPSIR D5S107	p15.3-15.1 q12-q14 q13-q14.1 q15-q23 q21-q22 q22-232 q31.3-q33.3 q33.3-34	0.76 0.78 0.85 0.78 ref 60 0.81 0.8 0.78 0.85 0.78	6 FTHP1 D6S202 F13A1 D6S105 D6S109 D6S89 Transcription factor IID	p12-p21.3 p22-p21.3 p25-p24 p p p	0.8 0.84 0.76 0.77 0.78 0.88 0.78 ref 50
7 D7S23	q31-q32	0.8	8 LPL D8S133 D8S136 D8S201 MYC D8S161 D8S203 D8S88 D8S164 D8S166 D8S198 D8S199 D8S167	p22 p p p q24 q q q q q q q	0.84 0.77 0.88 0.89 a 0.85 ref 53 0.77 ref 54 0.81 0.82 0.79 a 0.83 a 0.81 a 0.81 a 0.76	9 D9S51 D9S15 GSN ASS D9S105 D9S106 D9S109 D9S116 D9S120 D9S53 D9S55 D9S04	q21-qter q13-q21.1 q33 q34.1 q q q22.1-q32 q31-34 q131-34	0.79 0.75 0.76 0.83 0.8 0.76 0.77 0.81 ref 56 0.78 ref 57 0.83 0.75 0.8
11 TH INT2 D11S527 D11S35 APOC3 D1S809	p15.5 q13 q13.5 q22 q23-qter	0.77 0.78 0.88 0.79 0.91	12 CD4 Mfd109 Mfd129	pter-p12	0.78 0.75 a 0.82 a	14 MYH6 D14S49 D14S51	q11.2-q13	0.79 ref 51 0.8 a 0.8 a
15 ACTC GABR83 D15S87	q11-qter q	0.86 0.83 ref 55 0.85	16 D16S299 D16S298 D16S287 D16S291 D16S289 D16S265 D16S296 D16S285	p11.2-p12.1 p11.2-p12.1 p13.11 p13.3 q22.2-q23.1 q q	0.84 0.78 0.8 0.77 0.75 0.75 0.75 0.82	17 D17S520 D17S250 TCF2 D17S579 D17S254 GH D17S515	p12 q11.2-q12 q11.2-q12 q12-q23 q22-q24 q22-q24	0.85 0.81 0.8 0.79 a 0.91 0.9 ref 61 0.87
18 D18S34	q	0.78	19 APOC2 D19S178 D19S177	q13.2 q	0.79 0.76 0.78	20 ADA D20S41 D20S42 D20S46 D20S48 D20S49 D20S52 D20S55 D20S63 D20S64 D20S66	q12-13.1	0.76 0.84 0.83 0.82 0.82 0.82 0.81 0.77 0.79 0.88 0.77
21 THRA1 D21S167 D21S156 D21S198 D21S212 D21S214 D21S217 D21S11	q11.2-q12 q22.2r q22.3 q22.3 q q q	0.75 ref 58 0.8 0.82 0.81 0.84 0.81 0.82 0.82 ref 59	22 CYP2D IL2RB D22S258	q11.2-qter q13	0.79 0.86 0.8	X SYNARAF MAOB KAL DXS456 DXS424 DMD DXS294 DXS294 DXS425	p11 p11.3 p22.32 q21-q22 q24-q26 q26 q26 q26 q26-q27.1	0.84 0.82 0.75 0.77 0.79 0.78 0.75 0.75 0.79

defecto. En unos pocos casos, como ocurrió con los oncogenes, se pudo aprovechar el efecto "funcional" que se deriva de la presencia del gen anómalo. Sin embargo desafortunadamente en la gran mayoría de los denominados "single gene disorders" recogidos en el catálogo de McKusick (1991) no se dispone de informaciones bioquímicas suficientemente claras que orienten sobre la base genética del defecto. En este caso la disponibilidad de marcadores polimórficos de ADN distribuidos por todo el genoma ha permitido la localización cromosómica de regiones candidatas. El primer ejemplo histórico fue la localización del gen de la Corea de Huntington (Gusella et al., 1983).

El camino seguido para el aislamiento de un gen a partir de la región cromosómica candidata ha sido definido por Francis S. Collins (Collins, 1992) como "positional cloning", y ha dado sus frutos en casos como el de la fibrosis quística, retinoblastoma, síndrome X-frágil, distrofia muscular de Duchenne, distrofia miotónica etc.

Aunque no se haya conseguido el aislamiento del gen "enfermo" es posible realizar no obstante el diagnóstico precoz de la enfermedad, si se puede demostrar la existencia de un ligamiento suficientemente estrecho entre un locus polimórfico y el supuesto gen responsable. Este sería el caso de la Corea de Huntington, donde todavía no se ha aislado el gen, pero el diagnóstico precoz es de la mayor importancia ya que se trata de una degeneración del sistema nervioso central que se suele manifestar en edades avanzadas.

En la Figura 1b se presenta el ejemplo de una familia aquejada de una enfermedad congénita que segrega como un carácter monogénico autosómico y dominante (D,d). Estrechamente ligado al supuesto gen responsable hay un locus marcador (RFLP) con dos variantes alélicas (alelos 1 y 2). El alelo 2 aparece inequívocamente en todos los individuos enfermos, y está ausente en todos los sanos. Por consiguiente este sería un caso claro donde el análisis de un locus RFLP serviría para el diagnóstico precoz de la enfermedad.

Recientemente se ha podido demostrar el hecho singular de que algunas enfermedades congénitas humanas son causadas por la repetición de un número variable de trinucleótidos. Se trata por tanto de microsatélites patógenos (Hearne et al., 1992). Concretamente el síndrome X-frágil se produce cuando el número de trinucleótidos CGG en el exón 1 del gen FMR1 excede las 200 copias. Cuando el número de trinucleótidos es inferior a 60 los individuos son fenotípi-

camente normales, y si el número de copias oscila entre 60 y 200 los individuos son portadores. En el caso de la atrofia muscular, bulbar y espinal el trinucleótido responsable es CAG ; y en la distrofia miotónica los CTGs presentes en la región 3' no traducida del gen candidato parecen los responsables de la enfermedad. Como consecuencia de todo esto se ha creado el concepto de anticipación genética porque la severidad de los síntomas está relacionada con el número de repeticiones de los trinucleótidos, y va aumentando con el transcurso de las generaciones.

Finalmente, el carácter multialélico sobre todo de los microsátélites, permite definir patrones electroforéticos de ADN exclusivos de cada persona (finger printing). Este hecho está siendo aprovechado en medicina forense para resolver casos de paternidad dudosa. Además, dado que se pueden realizar ampliificaciones de ADN mediante PCR a partir de pequeñas muestras de sangre o semen etc. el análisis de estos microsátélites hipervariables está dando excelentes resultados en la identificación de presuntos criminales, violadores etc.(Craig et al., 1988; Sullivan, 1991).

Lejos de estas pretensiones y en tono menos formal, las posibilidades que ofrece el análisis genético molecular para la identificación genética, están causando una auténtica ADN-manía que se quiere explotar comercialmente. Así por ejemplo se ha presentado en el mercado una "variedad" de plumas Cartier que se alimenta con tinta enriquecida con ADN; de este modo la autenticación de una firma se puede hacer mediante "finger printing". También se venden tarjetas que llevan espolvoreado ADN de ídolos del "rock".

Estudios evolutivos.— Los polimorfismos de ADN han sido también aprovechados para la realización de estudios evolutivos de diversa índole, tratando de completar la información que antes habían aportado las proteínas. Pagnier et al (1984) hicieron un estudio de los movimientos migratorios más probables de las poblaciones negras africanas, y determinaron los diferentes orígenes de la población negra americana, mediante el establecimiento de haplotipos con marcadores de ADN de β -globinas. Higgs et al (1986) identificaron variantes RFLPs de β -globinas comunes en todos los grupos raciales, y otras que debieron surgir después del origen de las diferentes razas. Finalmente, en cuanto al establecimiento de filogenias, ha llamado especialmente la atención el análisis comparado del ADN mitocondrial que llevó a plantear

el origen del hombre en una “Eva” africana de hace unos cientos de miles de años (Gibbons, 1992).

Agradecimientos.— La elaboración de este artículo se ha visto facilitada por los conocimientos adquiridos en el desarrollo de los proyectos PM90-0006 (CICYT) y 92/0287 (FIS).

BIBLIOGRAFIA

- Armour J. A. L., Croisier M. and Jeffreys A. J. (1992) Human minisatellite alleles detectable only after PCR amplification. *Genomics* 12: 116-124.
- Botstein D., White R. L., Skolnick H. and Davis R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 32: 314-331.
- Collins F. S. (1992) Positional cloning: let's not call it reverse anymore. *Nature Genetics* 1: 3-6.
- Craig J., Fowler S., Burgoyne L. A., Scaott A. C. and Handing H. V. J. (1988) Repetitive deoxyribonucleic acid (DNA) and human genome variation: A concise review relevant to forensic biology. *Forensic Sci* 33: 1111-1126.
- Chakravarti A and Buetow KH (1985) A strategy for using multiple linked markers for genetic counseling. *Am. J. Hum. Genet.* 37: 984-997.
- Edwards A., Civitello A., Hammond H. A. and Caskey C. T. (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* 49: 746-756.
- Edwards A. L., Hammond H. A., Jin Li, Caskey C. T. and Chakraborty R. (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12: 241-253.
- Gibbons A. (1992) Mitochondrial Eve: wounded but not dead yet. *Science* 257: 873-875.
- Gusella J. F., Wexler N. S., Connealy P. M., Naylor S. L., Anderson M. A. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306: 234-238.
- Hearne C. M., Ghosh S. and Todd J. A. (1992) Microsatellites for linkage analysis of genetic traits. *Trends Genet.* 8: 288-294.
- HGM11 Human Gene Mapping 11. London Conference (1991). Eleventh International Workshop on Human Gene Mapping. *Cytogenet Cell Genet.* 58: 1-2197.
- Higgs D. R., Wainscoat J. S., Flint J., Hill A. V. S., Thein S.L., Nicjhol R.D., Teal H., Ayyub H., Peto T.E.A., Falusi A.G., Jarman A.P., Clegg J.B. and Weatherall D.J. (1986) Analysis of the human β -globin gene cluster reveals

- a highly informative genetic locus. Proc. Natl. Acad. Sci. USA 83: 5165-5169.
- Jeffreys A. J. (1979) DNA sequence variants in the globin-genes of man. Cell 72: 105-117.
- Jeffreys A. J., Wilson V. and Thein S. L. (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314: 67-73.
- Litt M and Luty J. A. (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.
- Mathew C. G. (ed). (1991) Methods in molecular biology. vol 9: Protocols in human molecular genetics. Humana Press. Clifton, New Jersey.
- McKusick V. A. (1991) Mendelian inheritance in man. 9th ed. Johns Hopkins University Press, Baltimore.
- Nakamura Y., Leppert M., O'Connell P., Wolff R., Holm T., Culver M., Martin C., Fujimoto E., Hoff M., Kumlin E. and White R. (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235: 1616-1622.
- Nathans D. and Smith (1975) Restriction endonucleases in the analysis and restructuring of DNA molecules. Annu. Rev. Biochem. 44: 273-293.
- Orita M., Iwahana H., Kanazawa H., Hayashi K. and Sekiya T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Acad. Sci. USA 86: 2766-2770.
- Pagnier J., Mears J.G., Dunda-Beldhodja O., Schaefer-Rego K.E., Beldjord C., Nagel R. L. and Labie D. (1984) Evidence for multicentre origin of the sickle hemoglobin gene in Africa. Proc. Natl. Acad. Sci. USA 81: 1771-1773.
- Saiki R. K. (1990) Amplification of genomic DNA. pp. 13-20. In "PCR protocols", Innis MA et al (eds). Academic Press, New York.
- Sambrook J., Fritsch E. F. and Maniatis T. (1989) Molecular cloning. A laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press.
- Sheffield V. C., Cox D. R. and Myers R. M. (1990) Identifying DNA polymorphisms by denaturing gradient gel electrophoresis. pp. 206-218. In "PCR protocols", Innis MA et al (eds.) Academic Press, New York.

- Southern E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.
- Sullivan K. M. (1991) DNA finger printing and forensic medicine. Chapter 23, pp. 273-286. In "Methods in molecular biology". Vol 9. Protocols in human molecular genetics. Mathew CG ed. Humana Press, Clifton, New Jersey.
- Tautz D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17: 6463-6471.
- Weber J. L. (1990) Informativeness of human (dC-dA)_n(dG-dT)_n polymorphisms. *Genomics* 7: 524-530.