

CAPÍTULO 3

MÉTODOS DE GENOTIPIFICACIÓN

MARCO HISTÓRICO DEL USO DE MARCADORES.
INFORMACIÓN GENÉTICA: CLASIFICACIÓN Y UTILIDADES.
FLUJO DE MUESTRAS. TÉCNICAS DE USO FRECUENTE
EN FORENSE. TÉCNICAS DE NUEVA GENERACIÓN.

*Egle E. Villegas Castagnasso, Elina Francisco, Julián A. Crespi, Silvina Díaz, Claudia
M. Corbi Botto, María V. Ripoli*

3.1. Concepto y marco histórico

En el primer capítulo se definió la genética forense como la especialidad que comprende la aplicación de técnicas de biología molecular para la identificación de individuos/razas/especies con el fin de auxiliar en la resolución de casos judiciales. Para este fin es necesario determinar el genotipo de los individuos o de los restos biológicos que forman parte de la investigación, definiendo como genotipo al conjunto de información genética que posee un organismo. Dentro del genotipo, se analizan determinadas regiones que resultan informativas en la distinción de individuos/razas/especies, ya que las mismas presentan cambios en la secuencia de bases del material genético que generan variabilidad. Estas regiones se denominan polimórficas y cuentan con dos o más variantes en una población/especie, considerando que la variante menos frecuente no puede ser mantenida únicamente por mutación (Ford, 1965).

Inicialmente, las diferencias entre individuos se expresaban mediante la descripción de detalles en su aspecto externo (fenotipo), como por ejemplo el color de capa, la presencia de manchas, remolinos, marcas, tatuajes o alguna otra particularidad destacable. Debido a que las descripciones realizadas eran subjetivas, por depender de cada observador, se incurría en errores que hacían de éste un método escasamente informativo en la resolución de disputas por origen o progenie.

El comienzo del análisis de las regiones polimórficas se produce a principios del siglo XX, cuando el médico austriaco Karl Landsteiner describió diferencias en el sistema sanguíneo ABO (Landsteiner, 1900). Posteriormente, Von Dürgen y Hirsfeld determinaron que este sistema presentaba herencia mendeliana simple, lo que significa que la existencia de cada variante (alelo) en el individuo puede justificarse por la presencia en uno de los progenitores. Esto es así para todos los organismos que tengan reproducción sexual, donde cada progenitor aporta la mitad del material genético para la formación del nuevo individuo. Por ser un sistema inalterable a lo largo de la vida, permite distinguir entre individuos de forma objetiva, analizar vínculos biológicos y excluir a los individuos que no están relacionados. Además, el grupo de trabajo de Landsteiner también describió otros antígenos eritrocitarios y leucocitarios que presentaban polimorfismo y herencia mendeliana como el Rh, MNSs o Duffy. Con estos aportes, hacia mediados del siglo XX los investigadores contaban con un número importante de sistemas informativos (marcadores genéticos) que permitían un análisis objetivo para la resolución de disputas judiciales, pudiendo excluir a los individuos que contaran con una combinación alélica diferente al progenitor putativo o al vestigio biológico hallado.

Las pruebas realizadas basándose en la descripción del aspecto exterior permitían análisis poco confiables. Con la aparición de los marcadores a nivel serológico, las conclusiones aumentaron el poder de discriminación en un 40%, y con la utilización de los polimorfismos del

MHC el poder discriminatorio se elevó al 80%. Así, el desarrollo creciente de diversas técnicas para la identificación del polimorfismo en distintos sistemas hizo que se ampliara el tipo de caracteres analizados, no limitándose sólo a descripciones fenotípicas y grupos sanguíneos sino que también se sumaban los polimorfismos a nivel proteico o bioquímico. De este modo, la certeza con la que se podían dar los resultados de un análisis fue incrementándose, tanto por el aumento del número de sistemas analizados como por el tipo de marcadores utilizados, mejorando considerablemente el poder de discriminación.

En el año 1953, la descripción de la estructura del ADN por James Watson y Francis Crick conformó un avance decisivo para el nacimiento de la biología molecular, siendo la base del desarrollo científico que continúa hasta nuestros días. Este descubrimiento, sumado al desarrollo de nuevas técnicas, permitió identificar regiones del ADN que presentan diferencias entre individuos, pudiendo de este modo distinguirlos de forma fehaciente. Ejemplo de ello son los antígenos del MHC. A partir de 1970, el desarrollo del conocimiento en este campo revolucionó las pruebas biológicas de la paternidad (Mayr, 1970). Sin embargo, estos marcadores genéticos presentaban grandes limitaciones a la hora de analizar muestras degradadas, que se encuentran en escasa cantidad o que no son frescas, como sucede frecuentemente en el ámbito forense. De este modo, poco o nada podía decirse del individuo al que pertenecía un vestigio biológico, principalmente manchas hemáticas, cabello o esperma. Por este motivo, en ese entonces la ayuda brindada a la justicia era realmente limitada.

En el año 1978, los microbiólogos Werner Arber, Daniel Nathans y Hamilton Smith describieron el funcionamiento de las enzimas de restricción, utilizadas por las bacterias como mecanismo de defensa contra virus invasores, que cortan el ADN cuando reconocen una secuencia específica, favoreciendo su degradación. Esta propiedad brindó una nueva herramienta para poder detectar diferencias entre individuos.

A principio de la década siguiente, Alex Jeffreys (1982) describe patrones electroforéticos de bandas de ADN altamente variables y heredables, que permiten la identificación genética de un individuo. El patrón de bandas característico obtenido para cada individuo se conoce como huella genética. Esta técnica fue utilizada por primera vez para resolver un caso judicial en 1985, demostrándose los antecedentes británicos de un niño cuya familia materna era originaria de Ghana. En otro caso forense, la aplicación de esta técnica permitió condenar a un violador y asesino de dos adolescentes en Inglaterra, analizando muestras de semen tomadas de las víctimas. Esta fue la primera condena en el mundo avalada por pruebas de ADN en el campo forense (*para más detalles, ver Cuadro 1.1*). Recién en 1987 los reactivos necesarios para estas pruebas pudieron comercializarse. Desde entonces, el análisis genético ha servido para la administración de justicia. Al tratarse de pruebas vinculantes, ofrecen el medio necesario para que con certeza y rigor científico se determine el perfil genético de un individuo, necesario para auxiliar en un caso judicial. Estos marcadores, que permitían analizar la variabilidad a nivel de ADN, formaron parte del comienzo de las pruebas que se denominan de inclusión, puesto que su poder de discriminación, acompañado por el análisis estadístico necesario, permite determinar si se trata de la combinación alélica que uno espera y que esto no se debe al azar.

En el año 1984, el bioquímico estadounidense Karl Mullis describe la PCR. La técnica amplifica una región específica de ADN, seleccionada a través de cebadores, para obtener un elevado número de copias con gran rapidez y fiabilidad. Este hecho permite que se puedan realizar análisis partiendo de cantidades muy pequeñas de ADN (del orden de los nanogramos), lo que antes no hubiera sido posible. Como esta situación es frecuente en el área forense, la PCR se ha convertido en una invaluable herramienta. Sumado a esto, sus numerosas variantes han permitido poner de manifiesto diferencias tanto a nivel individual como entre razas o especies.

El conocimiento de los genomas comienza en el año 1990 con el desarrollo del proyecto genoma humano, que tenía como finalidad determinar las secuencias de bases nitrogenadas que componen el ADN. Posteriormente, se identificaron y ubicaron en el genoma humano secuencias de interés, principalmente en el área médica. En el caso específico de los genomas de animales domésticos, los primeros en ser descritos fueron los de gallina y perro, seguidos en tercer lugar por el genoma bovino. Hasta la actualidad se ha completado la secuencia de más de 180 genomas (<http://www.genomenewsnetwork.org>). Del mismo modo que en los animales, se ha desarrollado el estudio de genomas vegetales, donde *Arabidopsis* fue el primero en ser completado. Al igual que en el área animal, en los vegetales el desarrollo de los genomas ha estado a cargo de consorcios internacionales con numerosos países implicados. El conocimiento de los diferentes genomas ha permitido la detección de nuevas secuencias polimórficas que se utilizan en la genética forense. En la Figura 1.1 se grafica en una línea temporal los avances tecnológicos descritos anteriormente.

Desde la llegada de las pruebas de ADN, se cuenta con análisis mucho más poderosos en cuanto al poder de discriminación. El análisis simultáneo de numerosas regiones variables del material genético permite obtener conclusiones de hasta un 99,99% de exactitud (ver *Capítulo 5*).

El mejoramiento en los métodos de extracción del ADN, sumado al desarrollo de nuevas técnicas de biología molecular acompañadas del análisis estadístico requerido, ha optimizado los tiempos y las pruebas científicas para ayudar en la resolución de problemáticas judiciales. A continuación se tratará la disposición del material hereditario a nivel celular y las técnicas que permiten poner de manifiesto las regiones que son informativas en el área forense.

3.2. Información genética: localización y utilidades

La información genética de un organismo eucariota cuenta con miles de millones de pares de bases que componen el ADN nuclear, mientras que una porción menor se halla en las organelas citoplasmáticas como las mitocondrias y los cloroplastos (Figura 3.1). El material genético nuclear se hereda en partes iguales tanto vía materna como vía paterna, según patrones mendelianos. En el caso del ADN citoplasmático, la herencia del material genético es únicamente materna, ya que el óvulo es el que aporta las organelas citoplasmáticas del futuro cigoto.

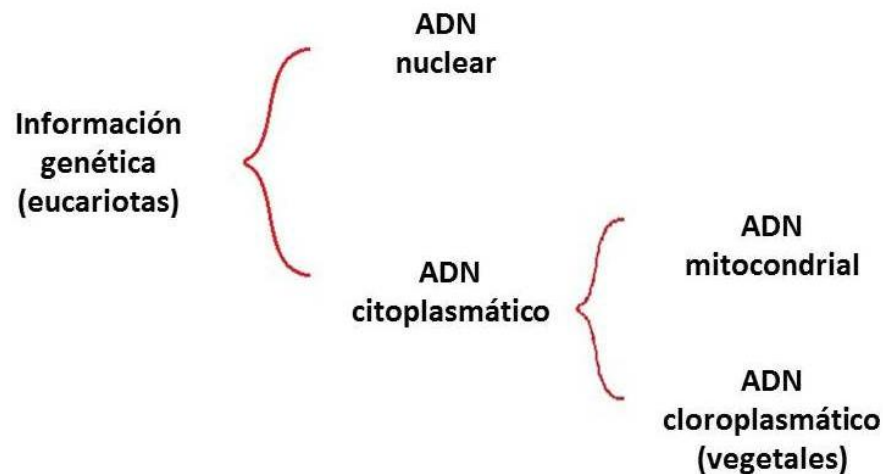


Figura 3.1: Genomas presentes en un organismo eucariota.

3.2.1. ADN nuclear

El ADN nuclear está organizado en secuencias que codifican para productos secundarios (ARN, péptidos y proteínas) que tendrán una función determinada, denominadas *ADN expresivo o codificante*. El resto del ADN lo conforman secuencias que no serán transcritas a proteínas y que se conocen como *ADN no codificante* (Figura 3.2).

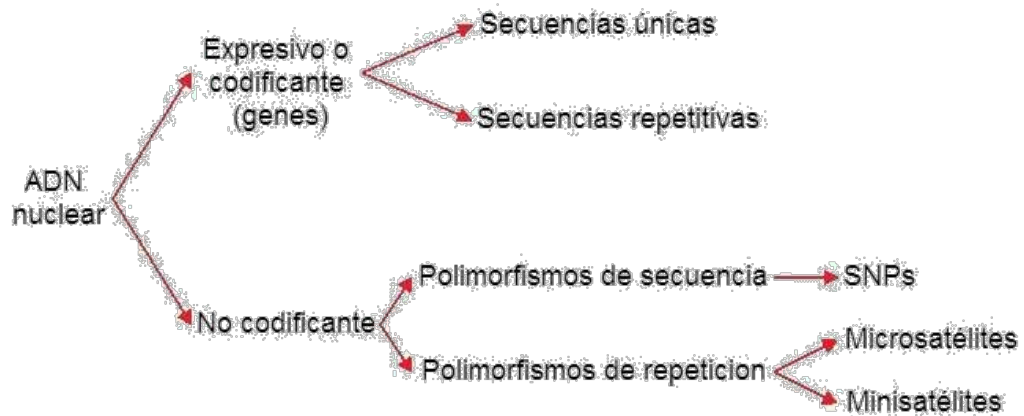


Figura 3.2: Clase de secuencias de ADN presentes en el ADN nuclear de un organismo eucariota.

Del total del ADN nuclear, sólo del 2 al 10% está constituido por secuencias codificantes, siendo éstas genes que constituyen la unidad fundamental y funcional de la herencia. Estos genes codifican para la síntesis de ARNs y proteínas, que son los componentes básicos de las células. En general, estas secuencias son poco variables, presentándose alguna excepción como en el caso de los genes del MHC.

El ADN no codificante, al no estar sujeto a una presión selectiva intensa, admite mayores niveles de variación en comparación con las regiones de ADN codificante. De esta manera, se encuentran regiones altamente variables cuyo análisis evidencia diferencias entre individuos, por lo cual son extensamente empleadas para el análisis forense. Dentro de las secuencias no codificantes, la presencia de las distintas variables puede estar determinada por la composición (secuencia de bases) o la longitud del fragmento de ADN analizado (número de repeticiones).

Los polimorfismos en la secuencia (cambio de uno o más nucleótidos) se producen por mutaciones en las bases nitrogenadas que componen la secuencia de ADN analizada. En el caso de presentarse en un solo sitio, se habla de mutación puntual y el polimorfismo se conoce como

polimorfismo de nucleótido simple (SNP, del inglés *single nucleotide polymorphism*). En este grupo también hallamos aquellos sistemas cuyos alelos se distinguen por el corte diferencial con enzimas de restricción, donde el cambio de secuencia afecta al sitio de corte de la enzima, permitiendo que una variante alélica se corte y la otra no.

Por otro lado, los polimorfismos de longitud son aquellos cuyas variantes puede distinguirse por el cambio en el largo de la secuencia a analizar. Recordemos que estamos hablando siempre de la misma región analizada, ya sea determinada por cebadores específicos en el momento de la PCR o por sondas al momento de la hibridación. Este es el caso de las secuencias microsatélites y minisatélites, donde cada alelo está definido por el número de repeticiones en tándem de la secuencia núcleo (*core*). Así por ejemplo, un alelo estará determinado por tres y otro por cuatro repeticiones. En el caso de los microsatélites, estas secuencias están formadas por dos a seis nucleótidos, mientras que en los minisatélites tienen alrededor de treinta pares de bases. Estas repeticiones se presentan en diferentes lugares (*loci*) del genoma y cada una constituye un marcador genético diferente (Hamata et al, 1982).

En el caso específico de los microsatélites, son los marcadores de elección en el ámbito de la genética forense, por su distribución uniforme a lo largo del genoma, el alto número de alelos que presentan cada sistema, su naturaleza codominante (ambos alelos son observables) y su fácil detección (Chehab et al., 1991; Oudet et al., 1991).

Un caso especial dentro del ADN nuclear es el ADN del cromosoma Y. En los mamíferos sólo lo portan los machos y se transmite de los padres a la descendencia masculina, lo cual permite analizar linajes paternos. Cabe destacar que las regiones analizadas en estos casos se encuentran en la región pseudoautosómica del cromosoma Y, esto es, la región de este cromosoma sin homología con el cromosoma X.

3.2.2. ADN mitocondrial (ADNmt)

El ADNmt presenta herencia materna, es decir que sólo es transmitido por la madre, mientras que el ADN mitocondrial del padre no pasa a la descendencia. Los individuos que provienen del mismo linaje materno poseen el mismo ADNmt. Éste es circular, cerrado y de doble cadena, lo que le confiere mayor estabilidad con respecto al ADN nuclear, frente a fenómenos degradativos como los que se presentan comúnmente en los casos forenses. Sumado a ello, cada célula presenta un alto número de copias de ADNmt (10 a 100 mitocondrias por célula), lo que permite que sea utilizado cuando se presenta una muestra con escaso material periciable.

El análisis del ADNmt se realiza mediante la amplificación de regiones de interés seleccionadas de acuerdo al tipo de análisis requerido, las que luego son secuenciadas permitiendo el análisis comparativo. Este análisis es especialmente útil en el ámbito forense para estudios de identificación, filiación por linajes maternos e identificación de especie de origen (*ver Capítulo 6*).

3.3. Flujo de muestras en forense/procedencias

Debido a que el ADN se encuentra en cualquier resto biológico que contenga células nucleadas, la recolección de muestras forenses tiene múltiples orígenes, como por ejemplo tejido, semen, cabellos, manchas hemáticas, saliva, heces, entre otras. A menudo se trata con vestigios biológicos que están lejos de ser la muestra ideal, tanto en lo que se refiere a su cantidad como a su calidad. Con respecto a la calidad hacemos referencia a muestras degradadas o sometidas a condiciones ambientales adversas y en las que el ADN se encuentra muy fragmentado. También hay que tener en cuenta la presencia de determinadas sustancias que deben ser eliminadas en esta etapa para evitar inhibiciones en la reacción de amplificación como tintes textiles,

altas concentraciones de melanina o hemoglobina. Así, el estudio forense comienza con la toma de la muestra (*ver Capítulo 2*), la cual debe cumplir con ciertos requisitos en los métodos de recolección y debe resguardarse de manera que se conserve la mayor cantidad de células para el análisis.

El segundo eslabón en el estudio forense es la extracción de ADN. Para ello, hay un método específico acorde al origen de la muestra y a su estado de degradación, existiendo numerosas opciones. Algunas de ellas tienen insumos de fácil preparación en el propio laboratorio y otras son kits comerciales que incluyen todos los reactivos necesarios.

Una vez obtenido el material genético, se cuantifica por electroforesis comparativa y/o espectrofotometría, y más recientemente PCR en tiempo real, para saber cuál es la concentración y la calidad de partida. En el caso de la electroforesis comparativa, se realiza la corrida incluyendo una muestra testigo de concentración conocida y que por comparación permite estimar la concentración y el grado de degradación de la muestra problema. La otra alternativa es la utilización del espectrofotómetro, que proporciona con mayor exactitud la concentración y permite estimar la proporción de ciertos contaminantes, como por ejemplo proteínas.

El siguiente paso es la amplificación de la muestra problema. Para llevar a cabo este paso se debe determinar cuál o cuáles son los marcadores adecuados para resolver la problemática planteada y aplicar la técnica necesaria. Posteriormente, se determinarán los alelos en cada sistema elegido, lo que permitirá establecer el perfil genético para realizar el análisis específico. Es importante señalar que la automatización de la mayoría de las técnicas utilizadas en este ámbito permite un mejor sistema de trabajo en los laboratorios, ya que ha permitido aumentar la cantidad de muestras a procesar y ha reducido considerablemente el tiempo de obtención de los perfiles genéticos. Una vez obtenidos, se realizan los análisis necesarios con la evaluación estadística apropiada, lo que dará el marco de certeza de las

conclusiones obtenidas. En la Figura 3.3 se ejemplifica el flujo que siguen las muestras al ingresar al circuito forense para el caso de la tipificación por microsatélites.

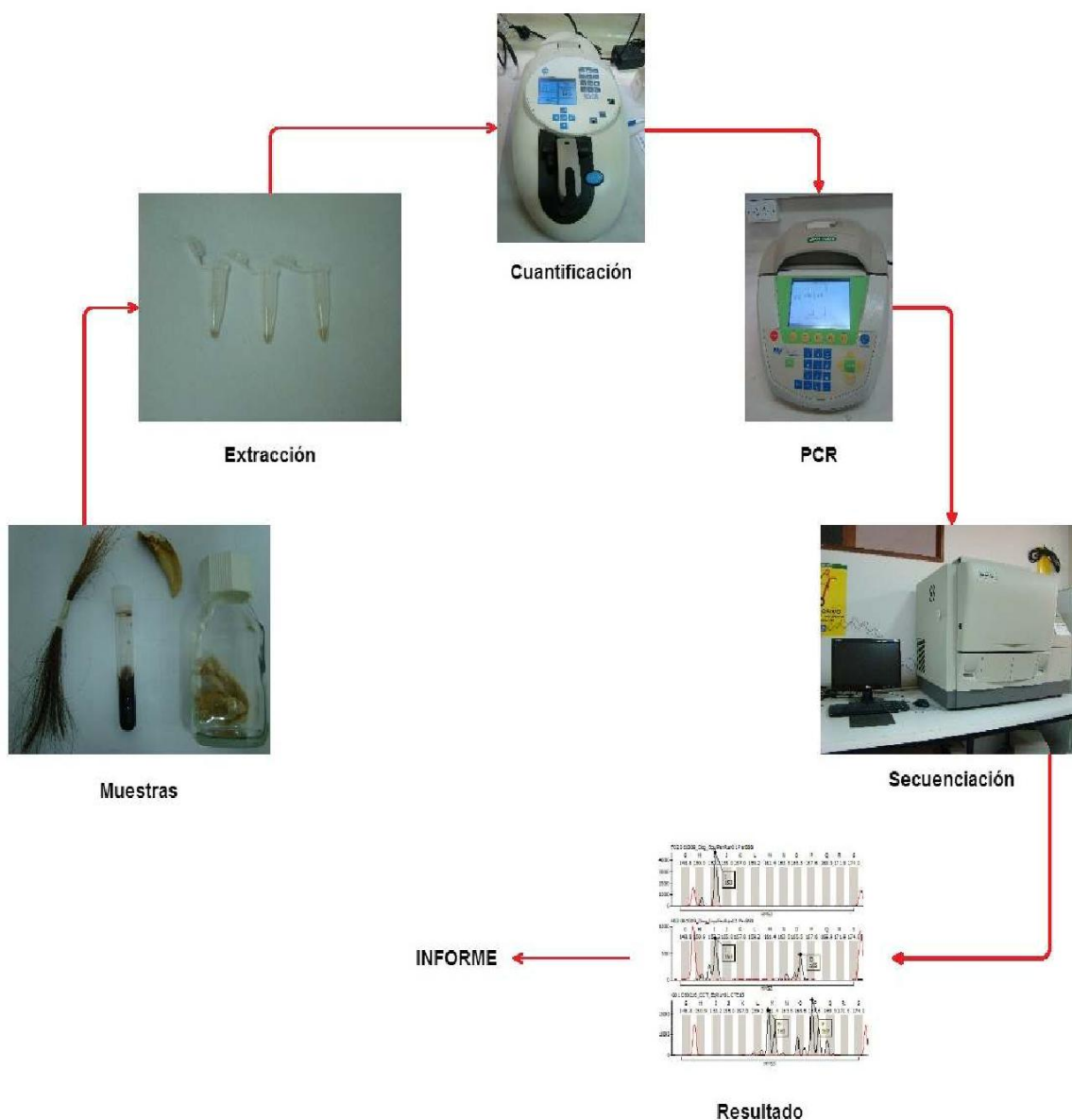


Figura 3.3: Flujo de trabajo para el análisis de una muestra de ADN mediante marcadores de tipo microsatélites.

Por otra parte, las técnicas y metodologías empleadas deben cumplir con normas que aseguren que los distintos procesos, desde la extracción hasta la obtención del perfil genético, se realizan bajo normas estrictas de calidad (*ver Capítulo 10*).

Con respecto a lo edilicio, es necesario delimitar los espacios de trabajo y establecer un flujo unidireccional desde la extracción hasta la obtención de datos, ya que el sentido inverso entre alguna de estas etapas, sobre todo de la extracción a la amplificación, favorecería la posibilidad de contaminación.

3.4. Técnicas de uso frecuente en forense

Como se mencionó anteriormente, en el año 1984 se desarrolló el método de la PCR. Esta técnica representó una revolución tanto en el área de la biología como de la medicina, y dentro de la Genética Forense permitió la resolución de un gran número de casos que no hubieran podido resolverse debido a la cantidad limitante de ADN o a su estado de degradación.

Los polimorfismos del ADN pueden ser de diversos tipos, desde una mutación de una sola base hasta el cambio en el número de unidades repetidas en tándem. La gran versatilidad de la técnica de PCR permitió el desarrollo de diferentes variantes de esta metodología, que se adaptan al tipo de polimorfismo analizado y permiten resolver distintas situaciones. A continuación describiremos en forma breve las principales metodologías utilizadas en genética forense:

3.4.1. PCR-RFLP. PCR de polimorfismo de longitud de los fragmentos de restricción (RFLP, del inglés restriction fragment length polymorphism)

Este método se basa en la detección de diferencias en la longitud de los fragmentos de restricción obtenidos a partir de una secuencia de

ADN determinada. Dicha secuencia o fragmento de ADN debe ser previamente amplificado mediante PCR y sometido luego a la acción de una o varias enzimas de restricción (Damiani *et al.*, 1990; Denicourt *et al.*, 1990; Pinder *et al.*, 1991). Las enzimas de restricción o endonucleasas son proteínas bacterianas que digieren al ADN de doble cadena. Estas enzimas realizan el corte o digestión cuando reconocen en la secuencia del ADN un determinado motivo nucleotídico o diana de restricción (Meselson y Yuan, 1968). Estas secuencias específicas se caracterizan por ser palindrómicas (inversamente repetidas) y tener una longitud de 4 a 10 pares de bases.

La técnica PCR-RFLP consiste en amplificar una región determinada del ADN y luego someterla a la acción de endonucleasas específicas. Las diversas mutaciones que presentan las diferentes moléculas de ADN son las responsables de la aparición o desaparición de los sitios de restricción, y este hecho es a su vez el responsable de generar los diferentes patrones de restricción. Una vez que se realiza la digestión se comparan los patrones de bandas obtenidos a partir de la digestión de moléculas de ADN de diferentes individuos, que han sido sometidas a digestión con las mismas enzimas de restricción. La identificación y comparación de los patrones de bandas se realiza luego de una corrida electroforética en un gel de agarosa o poliacrilamida que permite separar las bandas por tamaño (Figura 3.4).

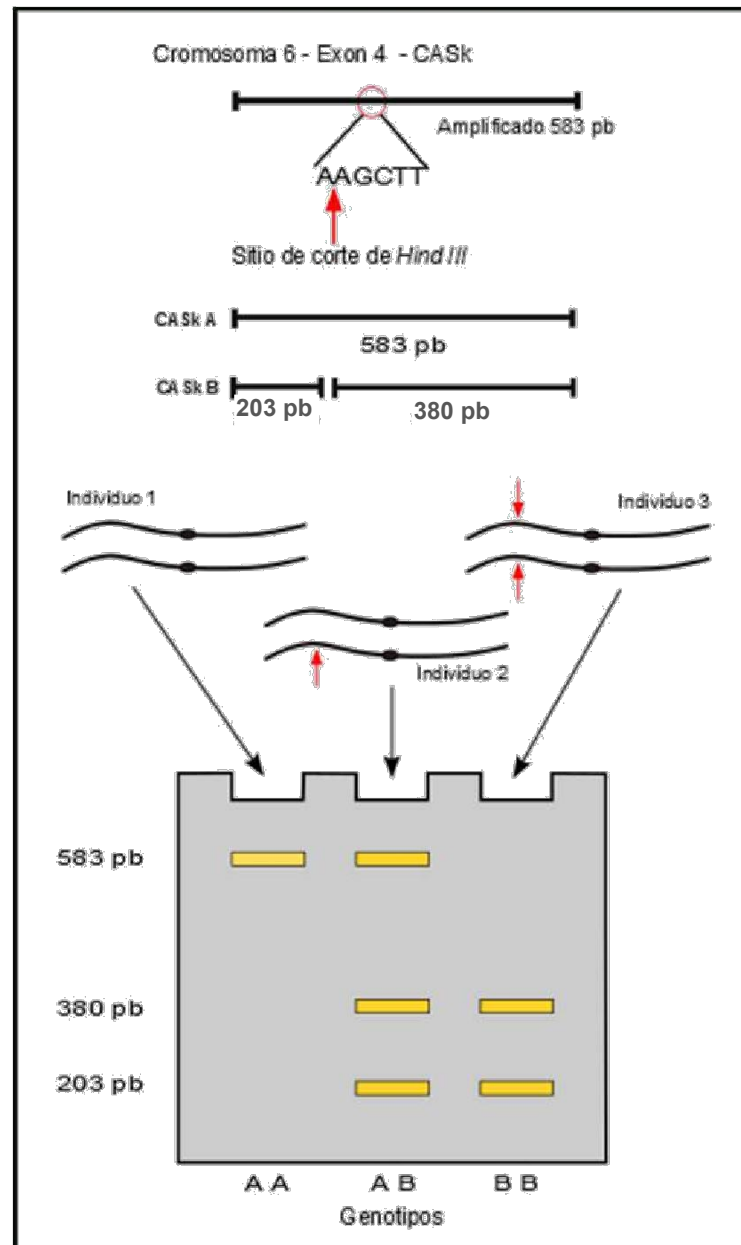


Figura 3.4: Descripción esquemática de la técnica de PCR-RFLP.

3.4.2. PCR-STRs. PCR de repeticiones cortas en tandem (STR, del inglés short tandem repeat)

La técnica de la PCR no sólo permite amplificar y analizar genes estructurales, sino que también permite amplificar y estudiar STR (Litt y Luty, 1989; Tautz, 1989; Weber y May, 1989). Estas secuencias no codificantes se encuentran ampliamente distribuidas por todo el genoma, son altamente polimórficas y además poseen herencia mendeliana. Asimismo, presentan un buen poder de discriminación y facilidades para ser amplificadas de forma simultánea con PCR multiplex, es decir que se pueden amplificar varios sistemas STR simultáneamente a partir de la misma muestra.

Los microsatélites o STRs consisten en repeticiones en tándem o continuas de dos, tres o cuatro nucleótidos (por ejemplo la repetición del dinucleótido AT). En el caso de estos loci, la diferencia entre los alelos está dada por variaciones en el número de repeticiones, así por ejemplo, en un individuo X el dinucleótido (AT) puede estar repetido 7 veces en un lugar específico (locus) de determinado cromosoma, y 9 veces en el mismo locus del cromosoma homólogo (Georges et al., 1991; Kemp et al., 1991; Brezinsky et al., 1992, 1993a, 1993b; Swarbrick et al., 1990; Vaiman et al., 1994; Avraham et al., 1993; Ciampolini et al., 1993, 1995; Ellegren, 1993; Goudarzi et al., 1993; Steffen et al., 1993; Sun et al., 1993a, 1993b, 1994a, 1994b; Sunden et al., 1993). El individuo X presentaría el genotipo 7-9 para el STR analizado.

La técnica de PCR-STR se basa en la amplificación de los microsatélites mediante la utilización de cebadores adyacentes a dicha secuencia. Los productos de la PCR pueden someterse a electroforesis en geles de poliacrilamida o bien pueden detectarse las distintas variantes mediante la utilización de secuenciadores automáticos. En el primer caso se realiza la amplificación de los STRs de interés seguida de una corrida electroforética en geles desnaturizantes. Las diferentes bandas, correspondientes a los fragmentos con distinto número de repeticiones, se visualizan posteriormente mediante tinción del gel con

nitrate de plata. En el segundo caso, para la amplificación de los STRs deben utilizarse cebadores marcados con fluorescencia. El análisis de los STRs se efectúa en un secuenciador de ADN en el que se realiza una electroforesis capilar. Posteriormente se realiza la interpretación de los resultados con un software específico (Figura 3.5).

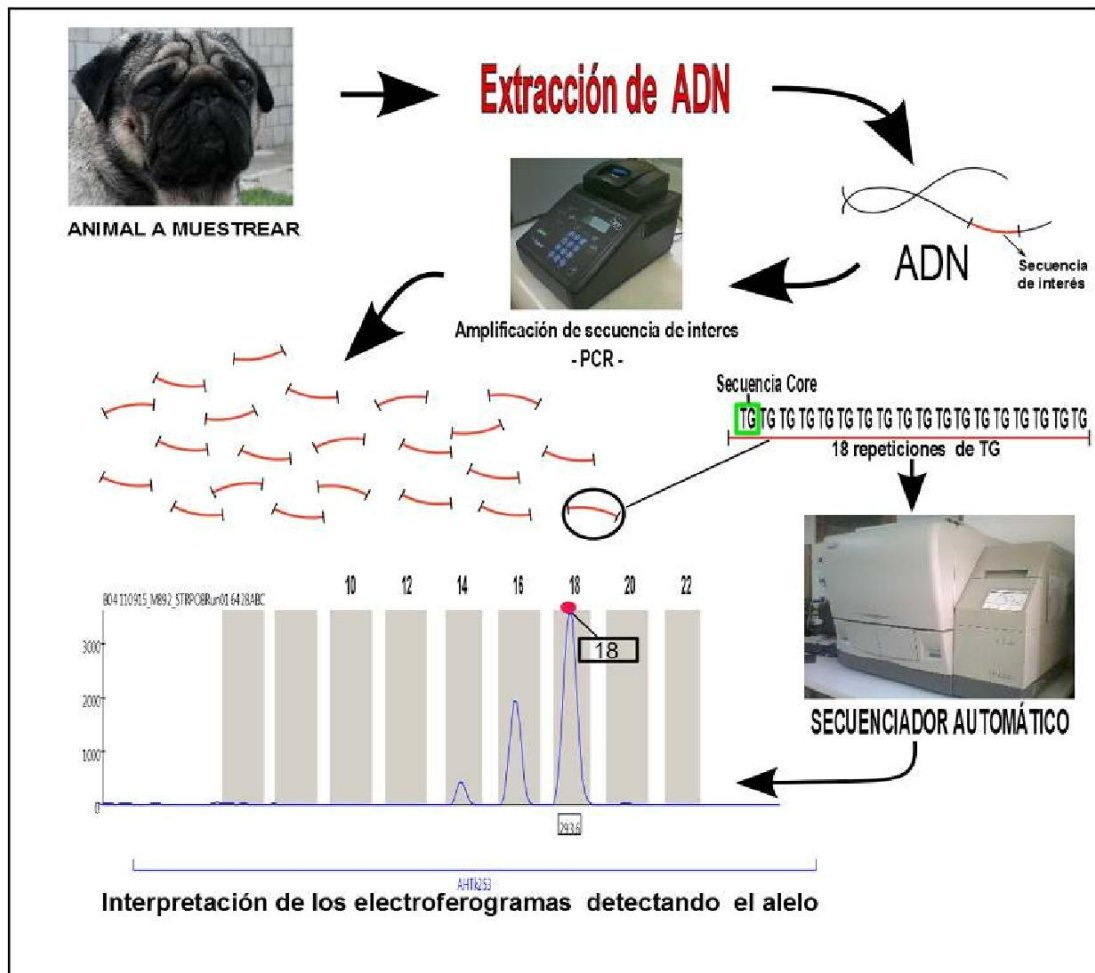


Figura 3.5: Descripción esquemática de la técnica de PCR-STRs (análisis de fragmentos).

3.4.3. PCR-SNPs. PCR de polimorfismos de nucleótido simple (SNP, del inglés *single nucleotide polymorphisms*)

Dentro de las metodologías que se están utilizando actualmente, deben mencionarse las técnicas de análisis de los polimorfismos de nucleótido simple (SNPs). Este tipo de polimorfismo consiste en el cambio de un nucleótido por otro (transición o transversión) dentro de una secuencia determinada de ADN, es decir que cada SNP posee usualmente dos nucleótidos alternativos en una posición determinada. El hecho que los SNPs se encuentren distribuidos a lo largo de todo el genoma, sean marcadores bialélicos de herencia codominante y tiendan a ser relativamente estables, los hace marcadores biológicos de excelentes características. Otra particularidad de interés de los SNPs es su baja tasa mutacional, lo que los hace marcadores ideales para estudios de paternidad. Si bien los marcadores más ampliamente utilizados dentro de la genética forense son los STRs, en los casos de muestras de ADN muy degradadas o con escasa cantidad, los marcadores más apropiados son los SNPs, ya que como el polimorfismo sólo afecta una base, la posibilidad de que la mutación esté intacta en material muy degradado es alta.

Por otra parte, la simplicidad de los SNPs los hace susceptibles de análisis a gran escala y particularmente de análisis con chips o microarreglos de ADN. Actualmente, los SNPs se están utilizando para la resolución de casos complejos de parentesco o de identificación dado que cuando las muestras se encuentran muy degradadas los STRs no funcionan. Otra de las grandes utilidades de los SNPs es la determinación del origen geográfico de una muestra biológica y también la determinación de ciertas características físicas. Sin embargo, los SNPs tienen también limitaciones, por ejemplo el número de SNPs que se necesitan para realizar un estudio es alrededor de cuatro veces el número de STRs que serían necesarios (Sobrinho et al., 2005).

En animales domésticos aún no se han llevado a cabo estudios tan extensivos como en humanos; sin embargo, los datos disponibles

indican que existiría una alta densidad de SNPs en las regiones que han sido estudiadas. Los estudios realizados en animales domésticos para la detección de SNPs se enfocan principalmente hacia el análisis de marcadores genéticos potencialmente asociados a caracteres de producción (genes candidatos). Sin embargo, la información que se está obteniendo a partir de estos polimorfismos resulta de utilidad tanto para ser aplicada en programas de selección como así también en identificación genética (trazabilidad, abigeato) y en farmacogenética.

No existe un método ideal para la tipificación de los SNPs, por lo tanto, la selección de una técnica apropiada depende de los requerimientos del investigador. Para comprender las técnicas de tipificación de SNPs se debe tener en cuenta que existen distintas reacciones para discriminar alelos, diversos métodos de detección de los alelos discriminados y diferentes formatos de ensayo. A continuación se describen algunas de las técnicas más utilizadas para el estudio de los SNPs:

3.4.3.i. AS-PCR. PCR alelo-específica (AS, del inglés *allele specific*)

Esta metodología permite discriminar las variantes alélicas de un marcador basándose en la utilización de dos cebadores, cada uno específico para un alelo determinado y un cebador común (David y Deuhtch, 1992). Los cebadores alelo-específicos se diferencian en una base para que sólo se unan a la mutación específica. Para cada individuo analizado se realizan dos PCR simultáneas en las que se incluyen el mismo ADN molde, pero mientras en un tubo se agrega un cebador específico para uno de los alelos más el cebador común, en el otro tubo se añade el cebador complementario al otro alelo y además el oligonucleótido común. Los productos de PCR se someten a electroforesis y se determina el genotipo del individuo en estudio de acuerdo a la presencia o ausencia de las bandas correspondientes a cada reacción (Figura 3.6).

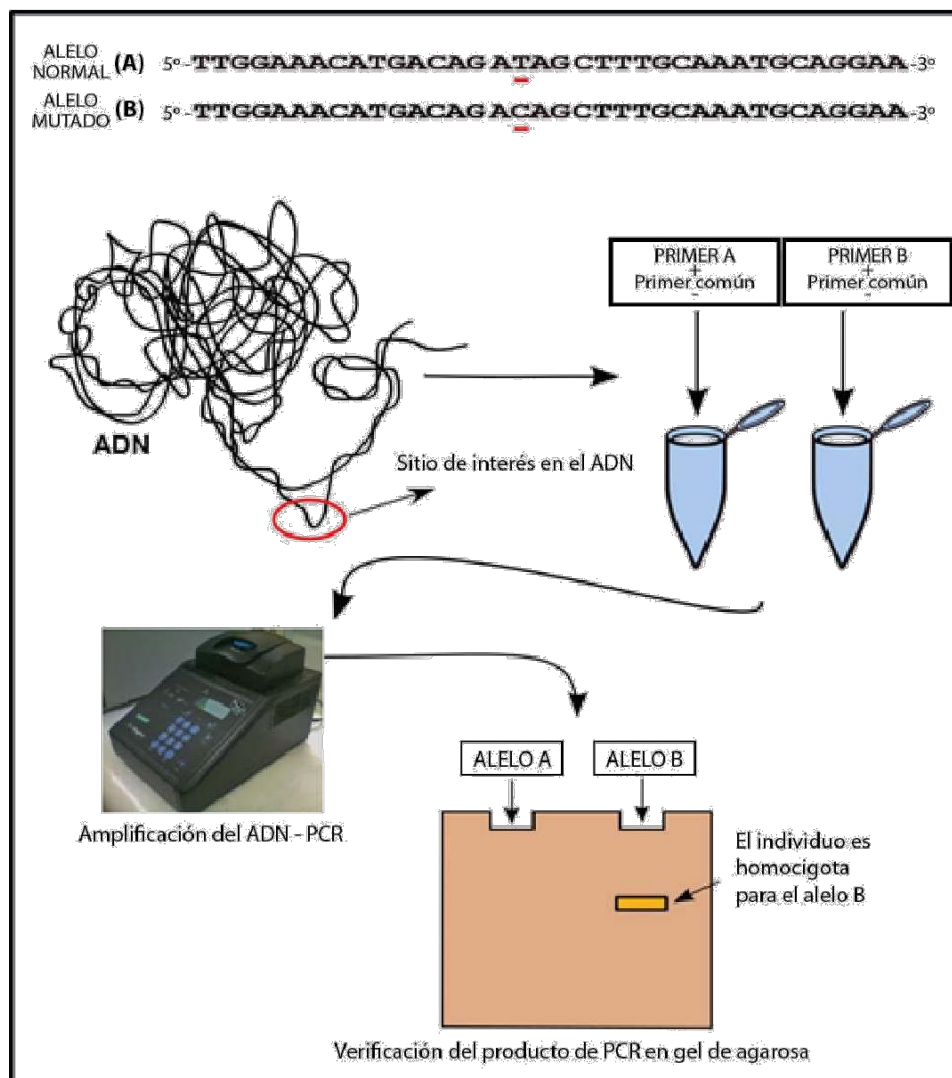


Figura 3.6: Descripción esquemática de la técnica de AS-PCR.

3.4.3.ii. ASO-PCR. PCR de oligonucleótidos específicos de alelos (ASO, del inglés allele specific oligonucleotide)

La hibridación con sondas u oligonucleótidos alelo-específicos es una técnica que permite discriminar las variantes de un marcador determinado mediante la utilización de sondas alelo-específicas marcadas con fluorescencia. Esta técnica permite distinguir fragmentos de ADN que sólo difieren en una base. Para poder realizar este tipo de

reacciones es necesario diseñar dos sondas, una específica para cada alelo. Los oligonucleótidos tienen un tamaño de 15 a 20 pb y generalmente la base polimórfica se ubica en el centro de cada sonda. En condiciones óptimas de hibridación, los híbridos ADN - sonda complementaria serán estables, mientras que los híbridos que difieran en una base serán inestables. Los genotipos pueden inferirse a partir de las señales de hibridación que sólo serán emitidas por los híbridos estables.

El principio empleado para la detección de los híbridos estables es el de transferencia de energía fluorescente mediante resonancia (FRET, del inglés *fluorescence resonance energy transfer*). En este caso, la fluorescencia se emite como resultado de un cambio en la distancia física entre un colorante fluorescente llamado donante o *reporter* y una molécula atenuadora de la fluorescencia llamada atenuador o *quencher*. Estas dos moléculas se encuentran en la sonda. Si la sonda no se une al ADN molde por falta de complementariedad, no se produce emisión de fluorescencia, ya que el colorante y el atenuador se encuentran en posiciones cercanas. En el caso que la sonda sea complementaria al ADN, se une al mismo y de esta manera se incrementa la distancia entre el *reporter* y el *quencher*. De este modo, la molécula atenuadora de la fluorescencia no puede actuar sobre la molécula fluorescente y esto permite la emisión y consiguiente detección de la fluorescencia.

Entre las técnicas desarrolladas para tipificar SNPs utilizando este tipo de reacciones, podemos mencionar la detección de actividad exonucleásica (TaqMan) y las guías moleculares:

1. Detección de actividad exonucleásica (TaqMan): Esta técnica aprovecha la actividad exonucleásica 5' de la Taq polimerasa que elimina los nucleótidos desde el extremo 5' de una sonda unida a un fragmento de ADN a medida que avanza la polimerización desde un cebador, generando de esta manera una señal fluorescente (Holland et al., 1991; Livak et al., 1995; Sobrino et al., 2005). Su mecanismo de acción es similar al videojuego PacMan, de ahí su nombre. En este caso

se utiliza un cebador común y dos sondas que difieren en sólo una base, cada una complementaria a un alelo. En los extremos de las sondas se encuentran los colorantes fluorescentes, el *reporter* en un extremo y el *quencher* en el otro. La metodología consiste en permitir que se unan al fragmento de ADN molde un cebador y la sonda complementaria con el colorante fluorescente, con un espacio entre ambos de manera que pueda actuar la polimerasa. Se realiza entonces una PCR y a medida que avanza la polimerización a partir del cebador hacia la sonda, la Taq polimerasa también realiza una actividad exonucleásica $5' \rightarrow 3'$ que degrada a la sonda. De esta manera, el colorante fluorescente se separa de la molécula atenuadora y comienza a emitir fluorescencia, ya que a medida que aumenta la distancia entre las dos moléculas el atenuador deja de actuar y permite que el colorante emita fluorescencia. La emisión de fluorescencia sólo ocurre si la sonda es específica para uno de los alelos; de lo contrario no se detecta ningún tipo de emisión debido a que la sonda, al no ser complementaria, no se une al ADN en estudio (Figura 3.7).

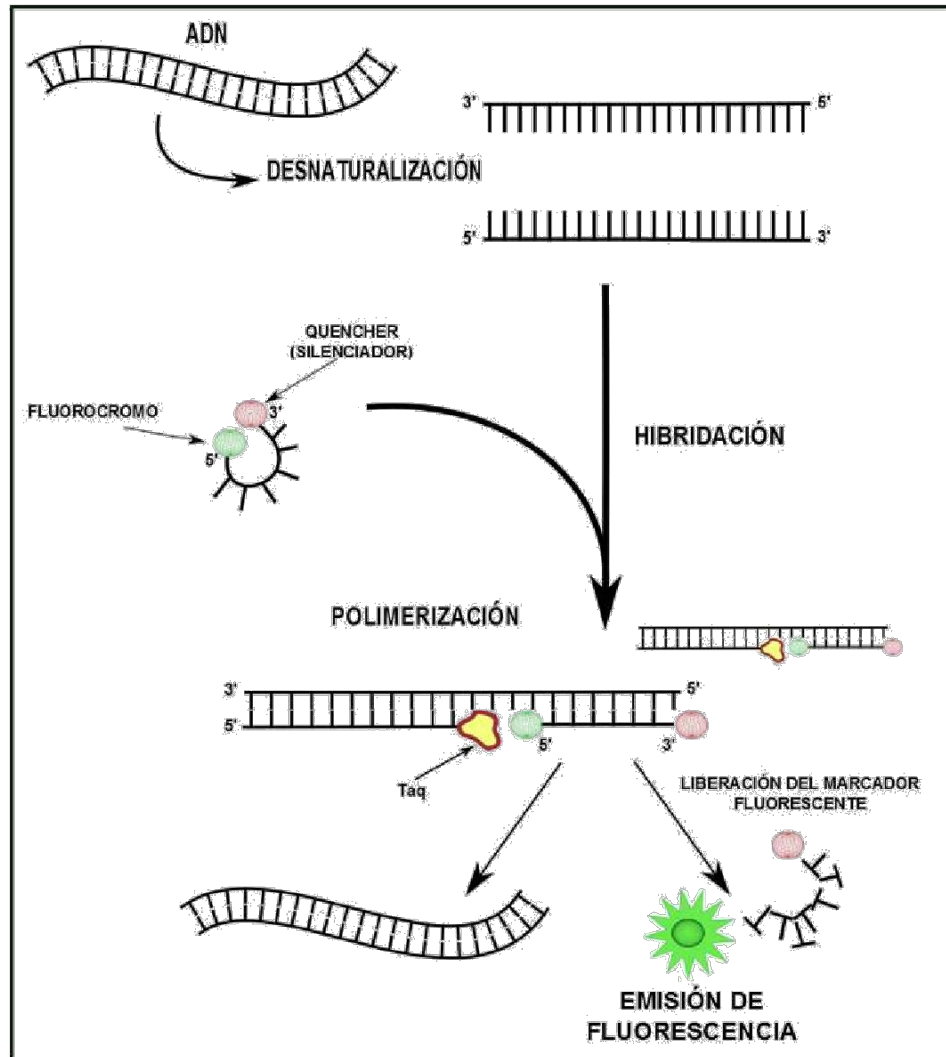


Figura 3.7: Descripción esquemática de la técnica de detección de actividad exonucleásica (TaqMan).

2. Guías moleculares (*Molecular beacons*): Para esta técnica se utilizan sondas que son complementarias al fragmento de ADN en el que se encuentra la mutación de interés. Estas sondas poseen en ambos extremos secuencias que son complementarias entre sí, y que a su vez llevan un marcador fluorescente en el extremo 5' y un atenuador en el 3'. Debido a la presencia de esas secuencias complementarias en los extremos, la sonda adquiere una configuración de bucle cuando no hibrida con el ADN blanco, razón por la cual la molécula fluorescente está inhibida por el atenuador y no se produce emisión de fluorescencia.

Cuando la sonda se aparea correctamente con el ADN analizado, el marcador y el atenuador se separan produciéndose la emisión de fluorescencia (Tyagi y Kramer, 1996; Sobrino et al., 2005). Para la tipificación de los SNPs deben utilizarse sondas específicas para cada alelo, marcadas con colorantes fluorescentes diferentes (Figura 3.8).

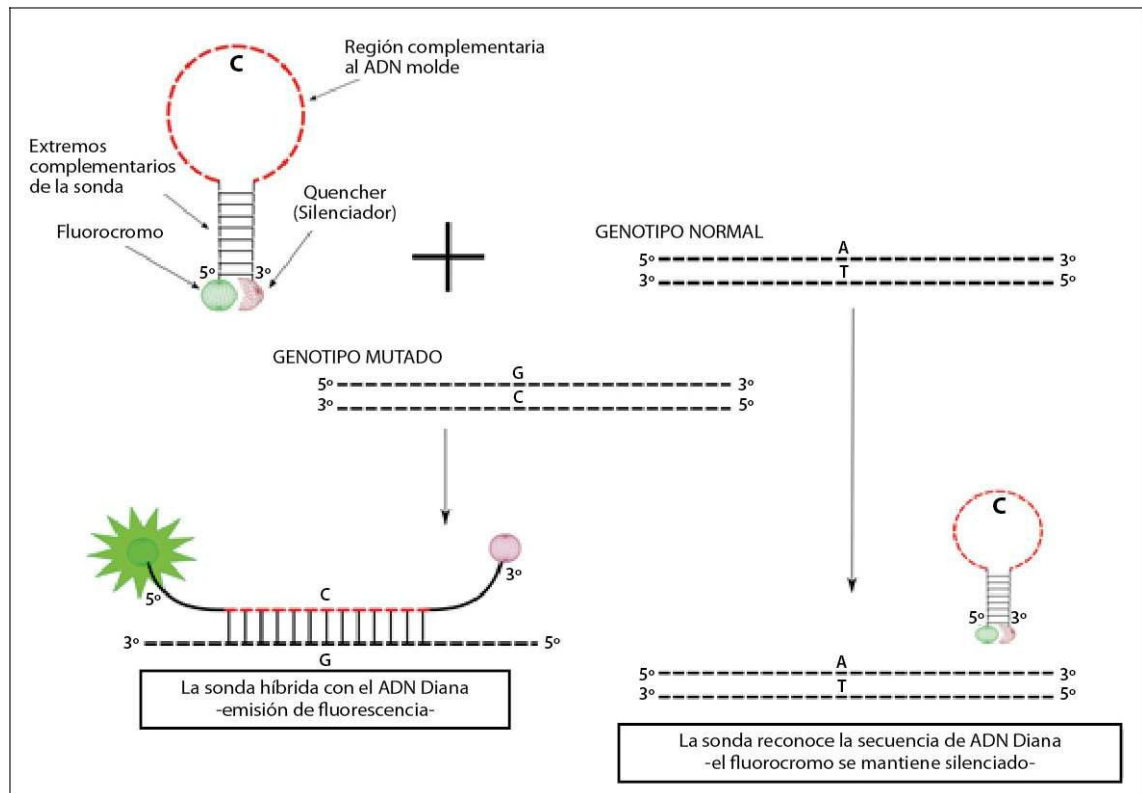


Figura 3.8: Descripción de la técnica de guías moleculares (*Molecular beacons*).

3.4.3.iii. Microarreglos de ADN

Los microarreglos de ADN, también llamados microarrays de ADN o chips de ADN, representan una tecnología de alto rendimiento que ha revolucionado en los últimos años distintas áreas de la investigación, incluida la genética forense (Kobilinsky, 2011). Los microchips surgen como consecuencia de una combinación entre técnicas

microelectrónicas empleadas para la fabricación de microprocesadores informáticos y materiales biológicos. La gran ventaja de esta metodología es que permite el análisis simultáneo de un gran número de moléculas de ADN en muy poco espacio.

Un microarreglo consiste básicamente en una superficie sólida del tamaño de un portaobjetos, generalmente de vidrio, sílice o plástico, a la cual se fijan pequeñas gotas que contienen fragmentos cortos de ADN o sondas (20 a 50 pb). El tamaño de las gotas es del orden de los micrómetros. Las sondas tienen una secuencia conocida y quedan inmovilizadas sobre la superficie sólida en posiciones conocidas. Sobre estos arreglos de sondas se hibrida el ADN que necesita ser evaluado. Estos chips permiten detectar desde cientos a cientos de miles de marcadores (Figura 3.9).

Dentro de los chips de ADN, encontramos los microarreglos de SNPs y de STRs. Estos resultan de utilidad dentro del área de la genética forense ya que permiten realizar perfiles de la muestra en evaluación. Ambos tipos de microarreglos han sido utilizados tanto en humanos como en animales para establecer relaciones de parentesco entre individuos y para efectuar estudios de identificación, con resultados exitosos (Divne et al., 2005; Li et al., 2006; Teletchea et al., 2008; Rönn et al., 2009; Park et al., 2010; Kling et al., 2012; Ziętkiewicz et al., 2012).

El funcionamiento de los microarreglos se basa en la capacidad de las moléculas complementarias de ADN de hibridar entre sí. Básicamente, se utilizan dos métodos de tipificación, un método se basa en la hibridación específica de alelos y el otro en la extensión de la sonda inmovilizada que actúa como un cebador, con dNTPs marcados con fluorescencia. En el primer caso las sondas fijadas en el chip representan todas las posibles variantes de un determinado sitio polimórfico; de este modo el ADN en estudio sólo se unirá a la sonda con la que la coincidencia sea exacta. Luego el microarreglo es escaneado para poder medir la fluorescencia y la mutación se identifica

de acuerdo a la señal fluorescente emitida. En este caso, el principio empleado es el de las sondas ASO descritas anteriormente.

En el segundo caso, la tipificación se basa en la extensión de la sonda usando como molde al ADN en estudio. Los cuatro nucleótidos, que contienen diferentes marcas fluorescentes, se añaden junto con la ADN polimerasa. La base incorporada, que es complementaria a la base ubicada en el sitio polimórfico del ADN molde, se identifica por la naturaleza de su señal fluorescente. En este caso el microarreglo también es escaneado para detectar la fluorescencia.

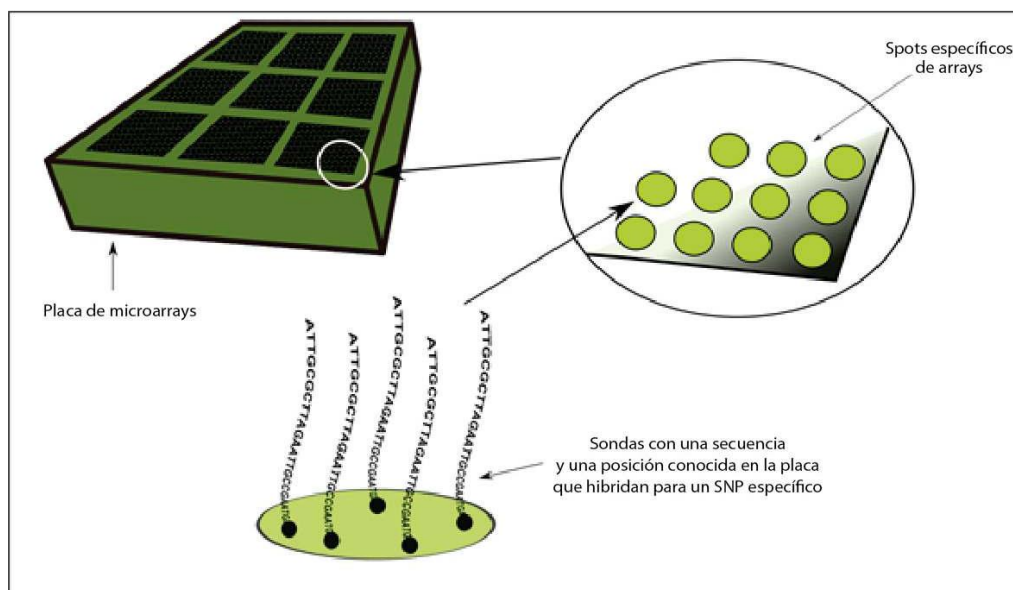


Figura 3.9: Descripción esquemática de la técnica de microarreglos.

3.4.3.iv. Minisecuenciación (*Minisequencing primer extension*)

Esta metodología se basa en la utilización de un cebador cuyo extremo 3' se ubica junto a la base que precede al SNP o mutación que define una variante alélica. La ADN polimerasa sólo extiende al cebador en una base (*single-base extension primer*), es decir que sólo se incorpora un nucleótido que es complementario a la mutación (Syvanen

et al., 1990; Sokolov 1990). Mediante esta técnica pueden detectarse varios SNPs simultáneamente (*multiplex reaction*). En este caso es necesario diseñar cebadores específicos para cada SNP analizado. La minisequenciación es utilizada en genética forense para la identificación de haplogrupos del ADN mitocondrial (La Neve et al., 2007; Daga et al., 2008).

Existen distintas técnicas que permiten analizar los productos de extensión del cebador, entre ellas pueden mencionarse la electroforesis y detección de fluorescencia, la espectrometría de masas (MALDI-TOF MS) y microarreglos.

1. Electroforesis y detección de fluorescencia (SNaPshotTM): Para lograr la extensión del cebador ubicado junta a la mutación se agrega una ADN polimerasa y ddNTPs marcados con cuatro colorantes fluorescentes. Luego los productos de la reacción se separan mediante electroforesis capilar en un secuenciador de ADN (Figura 3.10).

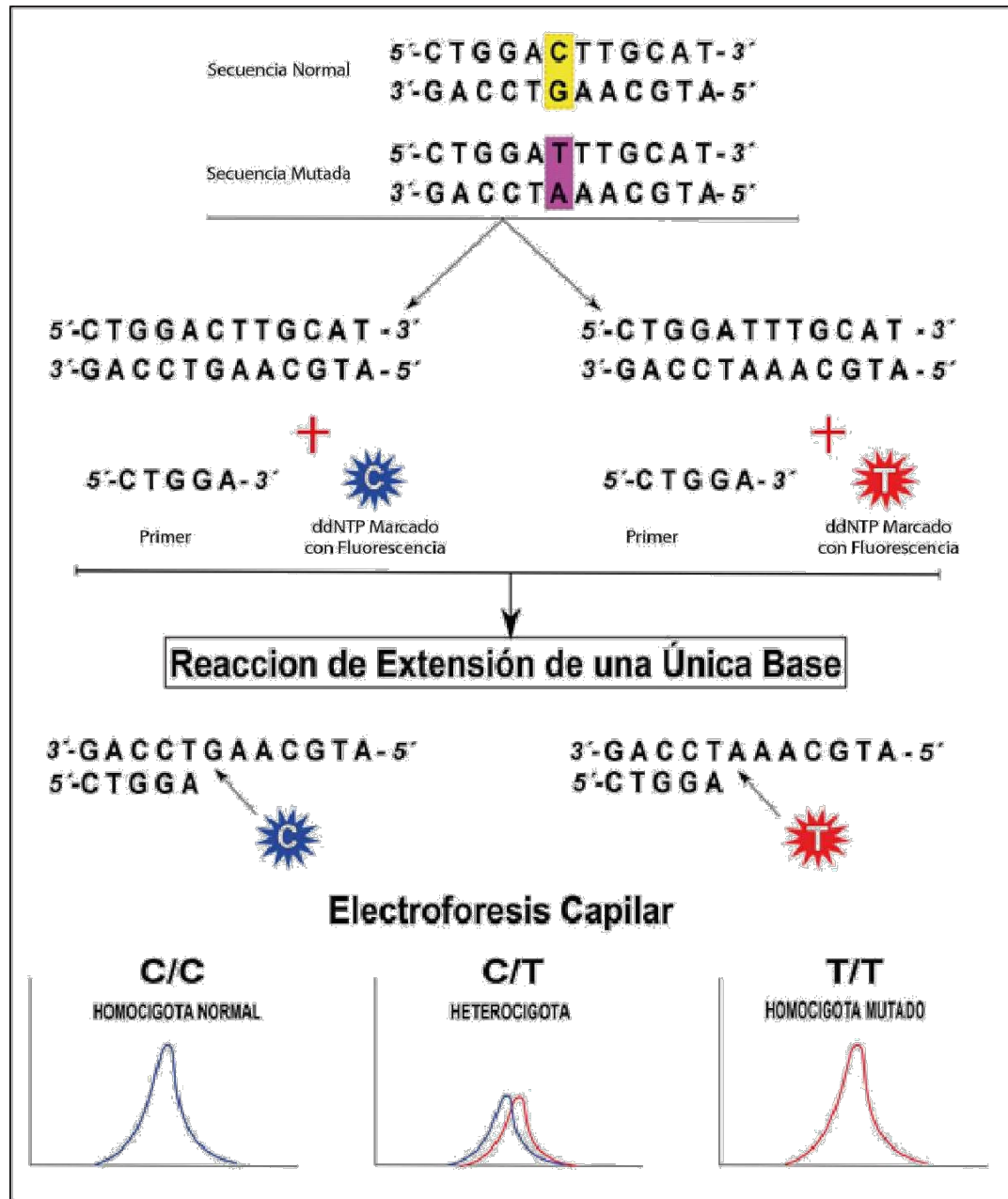


Figura 3.10: Descripción esquemática de la técnica de minisequenciación (SNaPshot™).

2. Espectrometría de masas (MALDI-TOF MS): esta técnica mide el peso molecular de los productos obtenidos a partir de una minisequenciación. La masa precisa del producto de la reacción varía de acuerdo al tipo de dNTP incorporado, ya que cada nucleótido presenta

un peso molecular diferente. En este caso no es necesario marcar los nucleótidos, pero sí es preciso disponer de equipos automatizados y además debe asegurarse la pureza de la muestra sometida al ensayo.

3. Extensión de cebadores fijados a un chip (APEX, del inglés *arrayed primer extension*): Otra forma de detectar el nucleótido incorporado durante una minisequenciación es mediante el uso de microarreglos. En este caso los cebadores se fijan a la matriz y son extendidos mediante una polimerasa que incorpora ddNTPs marcados. Como molde se utiliza el ADN bajo estudio. Posteriormente el microarreglo se escanea para medir la emisión de fluorescencia.

3.4.3.v. Pirosecuenciación

Esta tecnología también se utiliza para el estudio de SNPs y se basa en la ocurrencia de una cascada de reacciones enzimáticas que producen la emisión de luz cada vez que se incorpora un nucleótido al ADN molde. Mediante la actividad de una polimerasa se van incorporando nucleótidos a partir de un cebador unido a un ADN molde. Cada vez que se incorpora un nucleótido complementario a la cadena molde se libera un pirofosfato (PPi) que se transforma en ATP por acción de una ATPsulfurilasa. El ATP es entonces utilizado por la enzima luciferasa para producir un pico de luz que es captado por una lente. Los nucleótidos no incorporados son degradados por una enzima llamada apirasa para que no afecten las reacciones de incorporación sucesivas. La incorporación de cada nucleótido se registra en un pirograma en el cual se puede leer la secuencia del ADN molde. La gran ventaja de este método es que se puede detectar cualquier nuevo polimorfismo y analizar un gran número de individuos simultáneamente. La desventaja es que se requiere de un equipamiento automatizado específico (Figura 3.11).

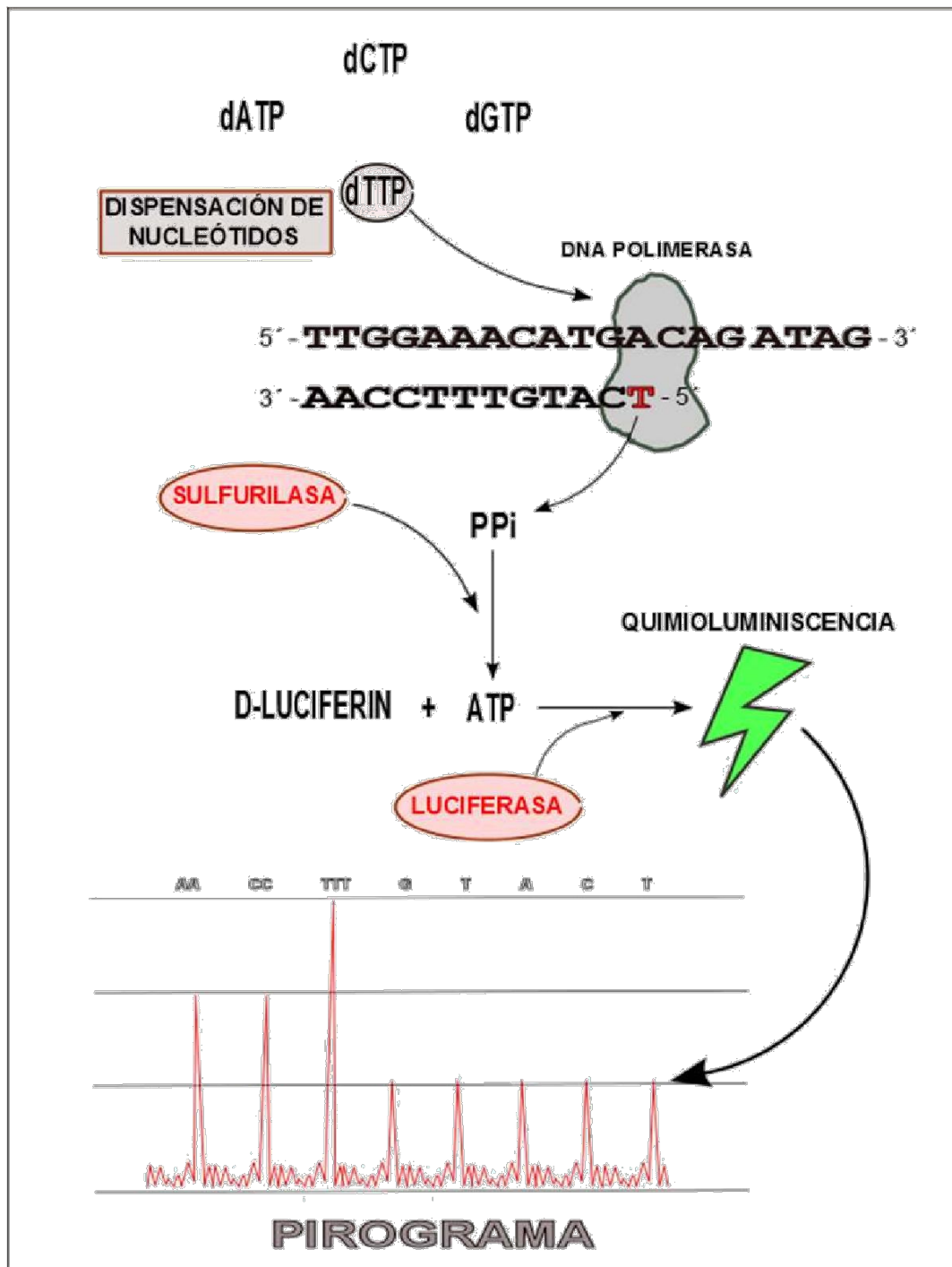


Figura 3.11: Descripción esquemática de la técnica de pirosecuenciación.

3.4.4. PCR-Secuenciación

En el año 1977, Maxam y Gilbert (1977) y Sanger et. al. (1977) desarrollaron dos métodos para poder determinar la secuencia de un fragmento de ADN. Si bien ambos métodos son muy similares, la metodología más comúnmente utilizada es el método enzimático de Sanger. Esta técnica se basaba originalmente en la extensión específica de la cadena de ADN a partir de un cebador, mediante la acción de la ADN polimerasa que incorpora dNTPs en la cadena en formación. Además, a la mezcla de reacción se le agrega ddNTPs que carecen de un residuo -OH en la posición 3' de la desoxirribosa. Estos nucleótidos se incorporan a la cadena en crecimiento por medio del trifosfato presente en posición 5', pero al carecer del hidroxilo en la posición 3' no pueden formar enlaces fosfodiéster con nucleótidos sucesivos y la cadena queda interrumpida. La técnica original consiste en realizar cuatro reacciones de polimerización diferentes, en las cuales debe incluirse el fragmento de ADN que se desea secuenciar, dNTPs convencionales que sí llevan el residuo -OH en posición 3' y los ddNTPs. La diferencia entre las cuatro reacciones es que una lleva ddATPs, otra ddTTPs, la tercera ddCTPs y la cuarta ddGTPs. De este modo, los dNTPs y los ddNTPs compiten entre ellos generando poblaciones de oligonucleótidos terminados en A, C, G o T (Sambrook et al., 1989). Los fragmentos obtenidos, que difieren en la última base, se separan mediante electroforesis en geles. Posteriormente la secuencia se determina leyendo el gel de abajo hacia arriba. De esta manera se logra conocer la secuencia del fragmento original (Figura 3.12).

Actualmente, la secuenciación se realiza en equipos automatizados que utilizan los mismos principios básicos establecidos por Sanger et al. (1977) pero empleando los cuatro ddNTPs en una misma reacción, los cuales se diferencian porque son marcados con cuatro fluorocromos distintos que luego son excitados por un láser leídos para que puedan ser detectados por el equipo. De esta manera los secuenciadores

automáticos generan un electroferograma de cuatro colores en el que cada pico representa a cada una de las cuatro bases del ADN.

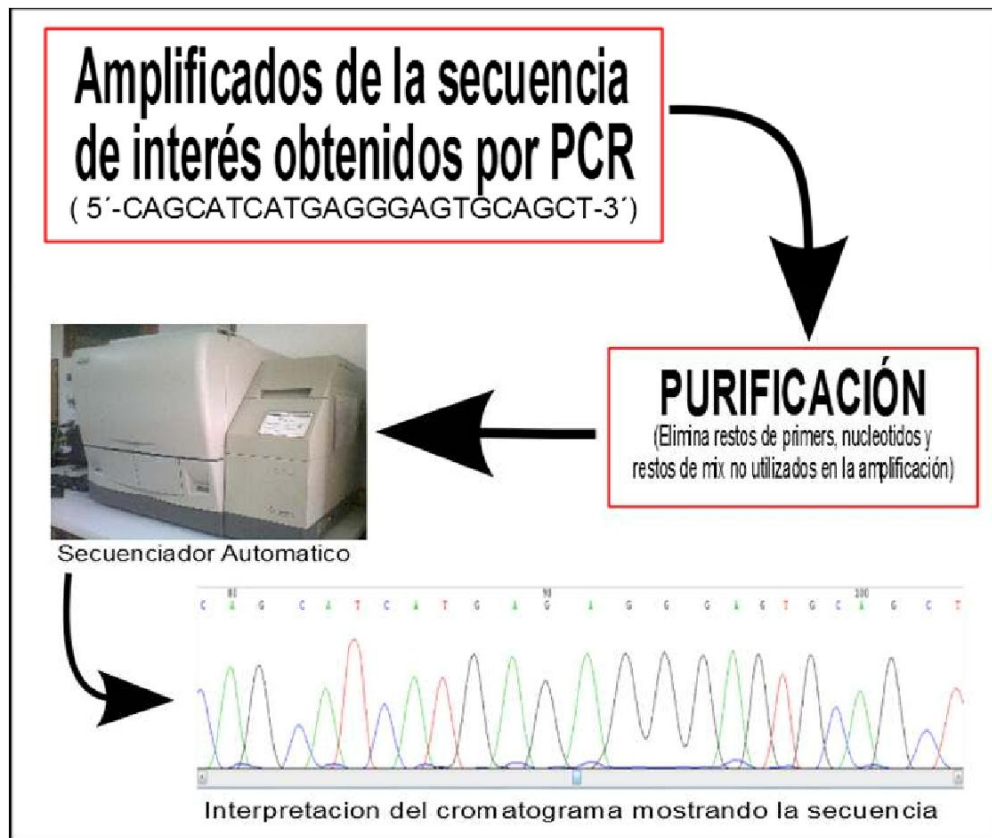


Figura 3.12: Descripción esquemática de la técnica de PCR-secuenciación.

3.5. PCR cuantitativa y Genética Forense

En 1993, Higuchi perfeccionó la metodología convencional de PCR adicionando una videocámara al termociclador que actuaba en el mismo momento en que se estaba produciendo la amplificación. La cámara captaba la fluorescencia emitida por el bromuro de etidio al intercalarse con las hebras recién sintetizadas. La cantidad de fluorescencia emitida era directamente proporcional a la cantidad de ADN producido por la PCR dando la posibilidad de medir en tiempo real dicha amplificación. Esta metodología es conocida como PCR cuantitativa (qPCR) o PCR en

tiempo real (RT-PCR), aunque se ha ido modificando la forma en que se emite la fluorescencia.

Dada las bajas cantidades de ADN usualmente encontradas en las evidencias biológicas forenses, determinar la cantidad de ADN humano, animal o vegetal extraído de una muestra de la escena del crimen es un paso crucial en el análisis. Este dato no sólo es necesario para optimizar las condiciones de reacción sino también para minimizar el consumo de la muestra. Actualmente, muchos laboratorios de genética forense cuentan con esta tecnología debido a la importancia de conocer la cantidad de ADN con la que se cuenta para realizar la amplificación. Además, en muchos casos, es necesario conocer si se está en presencia o no de inhibidores de la amplificación y esto también puede determinarse por esta técnica.

Si bien actualmente existen kits comerciales basados en qPCR para uso en genética forense humana (Nicklas et al., 2003), hasta el momento no existen kits comerciales disponibles para cuantificar ADN animal o vegetal. Sin embargo, se han publicado metodologías desarrolladas mediante sondas TaqMan para cuantificar el ADN en caninos, felinos, equinos y bovinos (Evans et al., 2007; Lindquist et al., 2011; Kanthaswamy y Premasuthan, 2012; Kanthaswamy et al., 2012) y productos vegetales (Johnson et al., 2013). Estas técnicas resultan de gran utilidad ya que muchas veces es necesario cuantificar con precisión en casos de forense animal o vegetal, como por ejemplo, en ataques de una especie a otra, o para evaluar la contribución de una especie en la manufactura de alimentos.

3.5.1. Fundamentos de la PCR cuantitativa

Durante la amplificación de ADN por PCR, la cantidad de templado se duplica con cada ciclo, por lo tanto la cantidad de producto (**P**), luego de un ciclo será:

$$P = (2)^n T$$

siendo **T** el tiempo transcurrido. En consecuencia, existe una relación exponencial entre la cantidad de producto amplificado y la cantidad de templado original. Esto se cumple sólo durante la fase exponencial de la reacción, cuando los reactivos no son limitantes y la formación de producto es proporcional a la cantidad de templado original. La relación exponencial existente se va a mantener duplicando el ADN en cada ciclo, y **P** va a depender tanto del ADN original a cuantificar (**X**) como de la eficiencia de la reacción (**Er**) será:

$$P = X(1+Er)^n$$

Para poder realizar la medición del producto se deben utilizar estándares de ADN con cantidades conocidas, y así poder extrapolar los datos desconocidos. El umbral (*threshold*) es un punto dentro de las curvas de calibración que se toma para determinar cuántos ciclos tardará una determinada muestra en alcanzar una determinada cantidad de producto **Y**. El número de ciclos que una muestra tarda en alcanzar una determinada cantidad de producto se denomina **Ct** (Figura 3.13).

A mayor concentración de ADN en la muestra, más rápidamente se alcanzara el umbral y el **Ct** será menor. Por lo tanto, existe una relación exponencial entre la cantidad de producto amplificado y la cantidad de ADN del templado original:

$$Y = X(1+Er)^n$$

siendo **Y** el producto amplificado en cada ciclo; durante la fase exponencial **Er** se mantiene relativamente constante y el producto (**Y**) es función de la cantidad de ADN molde (**X**).

Cabe destacar que se puede realizar el análisis de varios *targets* o secuencias blanco a la vez utilizando sondas marcadas con diferentes fluoróforos. Esta técnica se conoce como Multiplex qPCR.

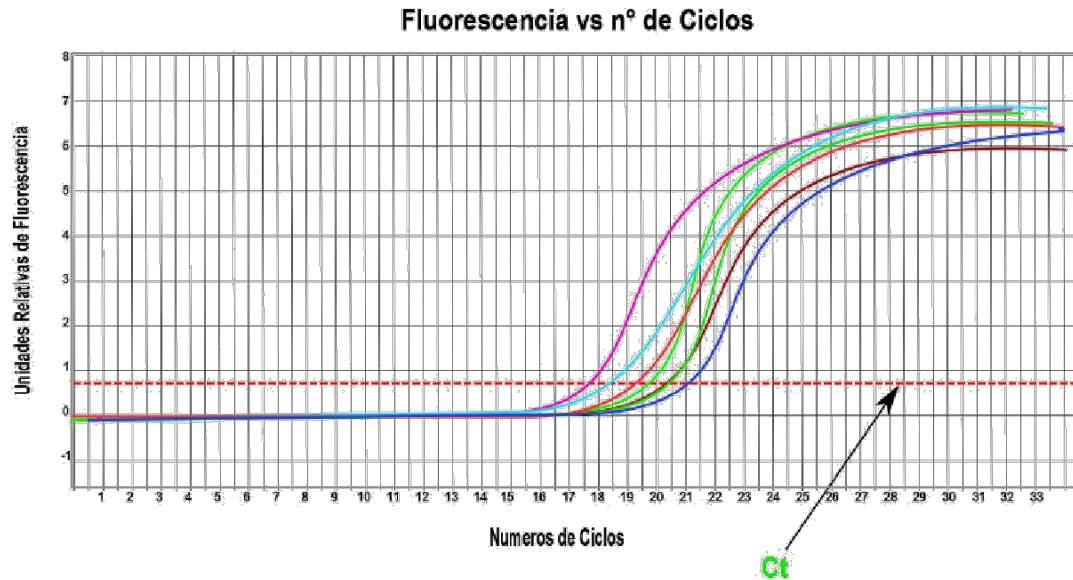


Figura 3.13: Curvas de amplificación obtenidas mediante la técnica de PCR cuantitativa.

3.5.2. Métodos de detección

Los métodos de detección son variados, pero todos ellos se basan en la transferencia de energía fluorescente mediante resonancia (FRET) entre dos moléculas. De este modo, el incremento del ADN en cada ciclo es proporcional a la cantidad de la fluorescencia emitida. En genética forense las sondas TaqMan son la metodología más utilizada para la cuantificación del ADN, dada su gran especificidad. Por otro lado, los agentes intercalares como el *SYBR Green I* son también muy utilizados dada la facilidad en optimizar las condiciones de reacción.

3.5.2.i Agentes intercalares

Los agentes intercalares son fluorocromos que aumentan notablemente la emisión de fluorescencia cuando se unen a ADN de doble hélice. El más empleado en RT-PCR es el *SYBR Green I*, cuya ventaja radica en que la optimización de las condiciones de reacción es

fácil y económica. El inconveniente de su utilización reside en que los agentes intercalares tienen baja especificidad, debido a que se unen de manera indistinta tanto a los productos específicos como a los generados inespecíficamente o a los dímeros de cebadores, que son muy frecuentes en la PCR.

3.5.2.ii Sondas de hidrólisis: TaqMan

Como se mencionó anteriormente, el principio de las sondas TaqMan se basa en la actividad de exonucleasa 5'-3' de la enzima Taq polimerasa. El fundamento de esta metodología radica en que al escindir una sonda marcada con un fluorocromo donante en el extremo 5' (que emite fluorescencia al ser excitado) y un aceptor en el extremo 3' (que absorbe la fluorescencia liberada por el donador al estar ambas moléculas espacialmente próximas) se produce la emisión de fluorescencia.

Durante la amplificación, la sonda se hibrida con su cadena complementaria. La actividad 5' exonucleasa de la Taq polimerasa hidroliza el extremo libre 5' de la sonda, produciéndose la liberación del fluorocromo donante. Este último, al encontrarse espacialmente alejado del aceptor, puede emitir fluorescencia que es captada por el lector (Figura 3.7). Al igual que en otros métodos de RT-PCR, la señal de fluorescencia resultante permite mediciones cuantitativas de la acumulación del producto durante las etapas exponencial de la PCR. La ventaja de las sondas TaqMan es que aumentan significativamente la especificidad de la detección, dado que están diseñadas para hibridar dentro de una región de ADN amplificada por un conjunto específico de cebadores. Por lo tanto, la utilización de estas sondas es ventajosa ya que resulta una metodología sensible, precisa y reproducible.

3.6. Técnicas de secuenciación masiva (NGS, del inglés *Next generation sequencing*)

La secuenciación a gran escala ha sido revolucionada en los últimos años por el desarrollo de varias tecnologías de secuenciación masiva (NGS). Cuando hablamos de técnicas de secuenciación masiva nos referimos a la secuenciación simultánea de millones de fragmentos pequeños de ADN, lo que permite crear una colección masiva de datos. Cada colección de datos puede alcanzar tamaños del orden de los gigabytes, lo que es equivalente a 1 billón (1.000.000.000) de pares de bases de ADN. Comparando estas nuevas tecnologías con la secuenciación de primera generación (secuenciación automática de Sanger), esta última podía secuenciar un fragmento de ADN por vez, generando quizás entre 500 a 1000 pares de bases de ADN en una única reacción (Metzker, 2010). Por lo tanto, este desarrollo ha incrementado drásticamente el número de bases obtenidas en cada secuenciación, y al mismo tiempo ha disminuido el costo de cada base secuenciada.

Las técnicas de secuenciación masiva son ventajosas en ciertos aspectos, entre los que podemos destacar que proveen alternativas de alto rendimiento mucho más económicas que la secuenciación de Sanger. Por lo tanto, se puede secuenciar un genoma completo en un solo día. Como consecuencia de estas características favorables, la secuenciación de alto rendimiento de genomas facilitó, entre otros, el descubrimiento de genes y elementos regulatorios asociados con enfermedades, la identificación de mutaciones causantes de patologías, así como información del transcriptoma completo (secuenciación del ARNm) de una muestra en un único análisis, sin requerir el conocimiento previo de la secuencia genética de un organismo.

Como desventajas, en vías de solución, podemos mencionar que para algunos laboratorios las plataformas para NGS aún pueden ser muy costosas. En forma adicional, el análisis de los datos puede llevar mucho

tiempo y puede requerir conocimientos especiales de bioinformática para obtener información adecuada de los datos de las secuencias. Quizás el mayor desafío al que se enfrentan los laboratorios de genética es el manejo y almacenamiento de las grandes cantidades de datos de secuencia que se generan por NGS, que pueden alcanzar a 600 gigabytes en una sola corrida (equivalente a alrededor de 90.000 canciones en un reproductor mp3).

Este desafío del conocimiento y aplicación de la bioinformática ha dado lugar al desarrollo de plataformas y programas que brindan soluciones crecientes para el tratamiento del formato de los datos, el ensamble *de novo* de las secuencias, el mapeo de lecturas en referencia a un genoma, la cuantificación y detección de variantes de secuencia y, fundamentalmente, el almacenamiento de los datos (Grada y Weinbrecht, 2013).

La primer tecnología NGS que estuvo disponible comercialmente fue el secuenciador genómico Roche/454 (454 Life Sciences, Branford, CT; Margulies et al., 2005). Mientras que los primeros equipos podían entregar longitudes de lectura de alrededor de 100 bases, la generación actual de instrumentos alcanza lecturas de más de 400 bases.

Todas las tecnologías NGS disponibles difieren de la secuenciación automática de Sanger en que no requieren el clonado del molde de ADN en vectores bacterianos. En la mayoría de las estrategias de NGS el molde de ADN se fragmenta, se une a un sustrato y se amplifica por PCR para generar representaciones clonales de los fragmentos originales que son separados espacialmente para la reacción de secuenciación subsecuente (Nowrousian, 2010). La secuenciación en sí misma se logra mediante un cierto número de métodos que hacen uso de diferentes enzimas (ligasas o polimerasas) y de diversas propiedades químicas para generar señales lumínicas que son registradas por métodos de detección altamente sensibles. Un aspecto común de todas las tecnologías NGS es el alto grado de simultaneidad, dado que tienen

lugar millones de reacciones de secuenciación al mismo tiempo y en pequeños volúmenes de reacción (Figura 3.14).

En la actualidad existen tres tecnologías NGS principales: el Analizador Genómico de Illumina (Solexa), el Ion Torrent y el Pirosecuenciador Roche/454 FLX. Si bien la forma en que cada plataforma NGS ejecuta la secuenciación de ADN es única, los distintos sistemas poseen aspectos comunes (Metzker, 2010). Generalmente inician el proceso fragmentando el ADN genómico al que se unen oligonucleótidos adaptadores; estas moléculas son amplificadas clonalmente para la secuenciación en paralelo de millones o billones de lecturas. Para dimensionar comparativamente, con la secuenciación de Sanger se necesitaron billones de dólares y 15 años de trabajo para completar un único genoma en el proyecto Genoma Humano (HGP, del inglés *Human Genome Project*). No obstante, esta metodología aún se utiliza en la mayoría de los laboratorios de investigación y permanece como una herramienta clave para la verificación de clones y la secuenciación basada en PCR (Mardis, 2008; Ansorge y Wilhelm, 2009).

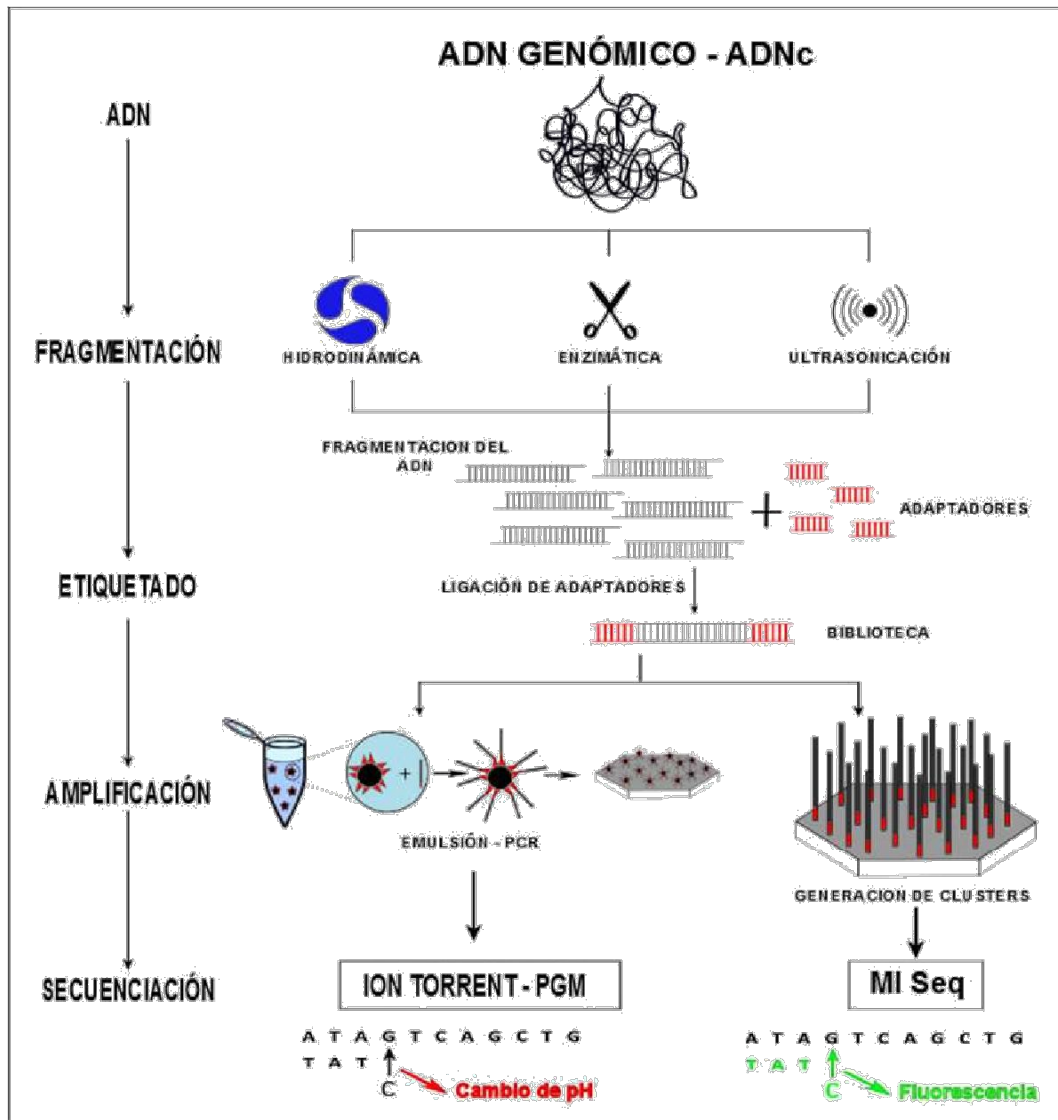


Figura 3.14: Descripción esquemática de la técnica de secuenciación masiva (NGS).

3.6.1. Aplicaciones de las tecnologías NGS

La tipificación genética tradicional implica el análisis paso a paso de un número de genes, para tratar de identificar dónde se encuentra la alteración genética. Un ejemplo que ilustra esta situación es la tipificación genética en casos de cáncer hereditario donde usualmente se secuencian dos genes, BRCA1 y BRCA2, en búsqueda de

alteraciones en los pacientes afectados. Esto puede requerir la secuenciación de más de 60 fragmentos de ADN para identificar dónde puede estar presente un único cambio de bases en el ADN. En contraste, usando NGS estos genes pueden ser secuenciados juntos en un solo test, mostrando claramente el potencial de la tecnología NGS de brindar información en numerosos genes de una sola vez. Otro ejemplo interesante es el caso de los numerosos genes involucrados en el desarrollo de tumores, tales como el cáncer de pulmón. El análisis del perfil molecular de una muestra del tumor utilizando NGS puede llevar a identificar una alteración que podría utilizarse para definir la mejor opción del tratamiento con drogas para un paciente.

Las aplicaciones de las tecnologías NGS son múltiples, entre ellas pueden mencionarse la secuenciación de genomas, el análisis de la organización genómica, el análisis de expresión génica, el estudio de microorganismos eucariotas, así como también para estudios de metagenómica y para la detección de variaciones de secuencia dentro de genomas individuales. En este último caso pueden mencionarse a modo de ejemplo, los SNPs, indels, variantes estructurales, análisis de variación genética individual, dinámica genómica y estudios de expresión. Además, estas tecnologías han sido adoptadas rápidamente para estudios de alto rendimiento que previamente se realizaban por métodos basados en hibridación, tales como los microarreglos. Esto incluye el uso de NGS para transcriptómica (RNA-seq) o análisis de genomas completos de ADN/interacciones proteicas (ChIP-seq).

Una de las aplicaciones más generalizadas de las técnicas de secuenciación con alto rendimiento y a bajo costo es la secuenciación de genomas. Podemos distinguir entre la resecuenciación, es decir la secuenciación de genomas de una especie cuando ya hay disponible un genoma de referencia, y la secuenciación *de novo*. La resecuenciación es actualmente una de las mayores áreas de aplicación de NGS. Las plataformas con lecturas cortas pero con alto rendimiento tienen un gran potencial porque con un genoma de referencia disponible, aún las

lecturas relativamente cortas pueden ser mapeadas con un alto nivel de confianza en la secuencia de referencia. Para genomas grandes como los de mamíferos, esta estrategia es muy útil; actualmente ya se han secuenciado genomas individuales de humanos y bovinos con la tecnología NGS (Xia et al., 2012; Derek et al., 2012). Además, las secuencias que son mapeadas con un genoma de referencia pueden emplearse para identificar SNPs, Indels, variaciones en el número de copias (CNV) y variantes estructurales, colaborando de esta manera en la comprensión de las bases genéticas de las diferencias fenotípicas.

3.7. Referencias Bibliográficas

- Ansorge, W.J. (2009) *Next-generation DNA sequencing techniques*. New Biotechnology.
- Avraham, A., Yoffre, M.B., Shani, M., Ron, M. (1993) "Bovine dinucleotide repeat polymorphism at the aro26 locus", en *Animal Genetics*. Número 24(2), pp. 147.
- Bickhart, D.M., Hou, Y., Schroeder, S.G., Alkan, C., Cardone, M.F., et al., (2012) "Copy number variation of individual cattle genomes using next-generation sequencing" en *Genome Res.*, Número 22(4) pp. 778–790.
- Brezinsky, L., Kemp, S.J., Teale, A. (1993a) "Five polymorphic bovine microsatellites (ILSTS010-014)" en *Animal Genetics*, Número 24 (1) pp. 74.
- Brezinsky, L., Kemp, S.J., Teale, A. (1992) "ILSTS001: A polymorphic bovine microsatellite" en *Animal Genetics*, Número 23 (1) pp. 81.
- Brezinsky, L., Kemp, S.J., Teale, A. (1993b) "ILSTS005: A polymorphic bovine microsatellite" en *Animal Genetics*, Número 24(1) pp. 73.
- Chehab, E.F., Johnson, J., Louie, E., Goossens, M., Kawasaki, E., Erlich, H. (1991) "A dimorphic 4-bp repeat in the cystic fibrosis gene is in absolute linkage disequilibrium with the $\Delta F508$ mutation: implications for

prenatal diagnosis and mutation origin” en *Am J Hum Genet.*, Número48 pp. 223–226.

Ciampolini, R., Goudarzi, K., Vaiman, D., Leveizel, H. (1993)“A new bovine dinucleotide repeat microsatellite: microsatellite INRA 30” en *Animal Genetics*, Número24(3) pp. 221.

Ciampolini, R., Moazami-Goudarzi, K., Vaiman, D., Dillmann, C., Mazzanti E., Foulley, J.L., Leveziel, H., Cianci, D. (1995) “Individual multilocus genotypes using microsatellite polymorphisms to permit the analysis of the genetic variability within and between Italian beef cattle breeds” en *Journal of Animal Science*, Número 73 pp. 3259–3268.

Corach, D. (2010)“Manual 14vo Curso de Actualización sobre Técnicas Moleculares de Identificación Humana Mediante Análisis de ADN”.SHDG.

Daca, P., Mielnik, M., Rogalla, U., Skonieczna, K., Linkowska, K., Grzybowski, T. (2008) “The application of minisequencing reactions for haplogroup assignment of mitochondrial DNA” en *Arch Med Sadowej Kryminol*, Número 58(4) pp. 212-217.

Damiani, G., Ferretti, L., Rognoni, G., Sgaramelli, V. (1990) “Restriction fragment length polymorphism analysis of the kappa-casein locus in cattle” en *Animal Genetics*, Número21 pp. 107-114.

David, V., Deuhtch, A. (1992) “Detection of α S1-casein genomic variants using the allele-specific polymerase chain reaction” en *Animal Genetics*, Número 23 pp. 425-429.

Denicourt, D., Sabour, M.P., Mcallister, A.J. (1990) “Detection of bovine -casein genomic variants by the polymerase chain reaction method” en *Animal Genetics*, Número 21 pp. 215-216.

Divne, A.M., Allen, M. (2005) “A DNA microarray system for forensic SNP analysis” en *Forensic Sci Int.*, Número 154(2-3) pp. 111-121.

Ellegren, H. (1993) “Genome analysis with microsatellite markers. Department of animal breeding and genetics” Swedish University of Agricultural Sciences. Uppsala.

- Evans, J.J., Wictum, E.J., Penedo, M.C., Kanthaswamy, S. (2007) "Real-time polymerase chain reaction quantification of canine DNA" en *J Forensic Sci.*, Número 52(1) pp. 93-96.
- Ford, E.B. (1965) "Polymorphism" en *Biology Reviews*, Numero20 pp. 73 - 88.
- Georges, M., Gunawardana, A., Threagil, ID.W., Lathrop, M., Olsaker, I., et al. "Characterization of a set variable number of tandem repeat markers conserved in bovidae" en *Genomics*, Número11 pp. 24-32.
- Goudarzy, K., Ciampolini, R., Vainman, D., Leveziel, H. (1993) "A new bovine dinucleotide repeat microsatellite: microsatellite INRA 18" en *Animal Genetics*, Número 24(3) pp. 221.
- Grada, A., Weinbrecht, K., (2013) "Next-Generation Sequencing: Methodology and Application" en *Journal of Investigative Dermatology*, Número133 pp.11.
- Hamada, H., Petrino, M.G., Kakunaga, T. (1982) "Molecular structure and evolutionary origin of human cardiac muscle actin gene" en *Proc Natl Acad Sci USA*, Número 79(19) pp. 5901–5905.
- Higuchi, R., Fockler, C., Dollinger, G., Watson, R. (1993) "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions" en *Biotechnology*, Número 11(9) pp. 1026-1030.
- Holland, P.M., R.D. Abramson, R. Watson, D.H. Gelfand (1991) "Detection of specific polymerase chain reaction product by utilizing the 50–30 exonuclease activity of *Thermus aquaticus* DNA polymerase" en *Proc. Natl. Acad. Sci. USA.*, Número 88 pp. 7276–7280.
- Johnson, C.E., Premasuthan, A., Satkoski Trask, J., Kanthaswamy S. (2013) "Species Identification of *Cannabis sativa* Using Real-Time Quantitative PCR (qPCR)" en *J Forensic Sci.*, Número 58(2) pp. 486-490.
- Kanthaswamy, S., Premasuthan, (2012) A. "Quadriplex real-time PCR (qPCR) assay for human-canine-feline species identification and nuclear DNA quantification" en *Forensic Sci Int Genet.*, Número 6(3) pp. e97-8.

- Kanthaswamy, S., Premasathan, A., Ng, J., Satkoski, J., Goyal, V. (2012) "Quantitative real-time PCR (qPCR) assay for human-dog-cat species identification and nuclear DNA quantification" en *Forensic Sci Int Genet.*, Número 6(2) pp. 290-295.
- Kemp, S.J., Teale, A.J. (1991) "Dinucleotide repeat polymorphism at the bovine locus FSHB" en *Animal Genetics*, Número 22(5) pp. 435.
- Kling, D., Welander, J., Tillmar, A., Skare, O., Egeland, T., Holmlund, G. (2012) "DNA microarray as a tool in establishing genetic relatedness. Current status and future prospects" en *Forensic Sci Int Genet.* Número 6(3) pp. :322-329.
- Kobilinsky, L. (2011) "Forensic Chemistry Handbook" Ed John Wiley & Sons. 528 pgs.
- La Neve, F., Civera, T., Mucci, N., Bottero, M.T. (2008) "Authentication of meat from game and domestic species by SNaP short minisequencing analysis" en *Meat Science*, Número 80 pp. 216–224.
- Landsteiner, K. (1900) "Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe" en *Zentralblatt Bakteriologie*, Número 27 pp. 357-362.
- Mayr, W.R. (1970) "Die genetik des HLA systems" en *Hum Genet* Número 12 pp. 195-199.
- Li, L., Li, C. T., Li, R. Y., Liu, Y., Lin, Y., Que, T. Z., Sun M.Q., Li, Y. (2006) "SNP genotyping by multiplex amplification and microarrays assay for forensic application" en *Forensic Science International*, Número 162(1) pp. 74-79.
- Lindquist, C.D., Evans, J.J., Wictum, E.J. (2011) "Developmental validation of feline, bovine, equine, and cervid quantitative PCR assays" en *J Forensic Sci.*, Número 56 (1) pp. 29-35.
- Litt, M., Luty, J.A. (1989) "A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene" en *American Journal of Human Genetics*, Número 44 pp. 397-401.
- Livak, K.J., S.J. Flood, J. Marmaro, W. Giusti, K. (1995) "Deetz, Oligonucleotides with fluorescent dyes at opposite ends provide a

quenched probe system useful for detecting PCR product and nucleic acid hybridization” en *PCR Methods Appl.*, Número 4 pp. 357–362.

Mardis, E.R. (2008) "Next-Generation DNA Sequencing Methods" en *Annual Reviews Genomics and Human Genetics*.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., et al. (2005) "Genome sequencing in microfabricated high-density picolitre reactors" en *Nature*, Número 437 pp. 376–380.

Martinez, D.A., Nelson, M.A. (2010) "The next generation becomes the new generation" en *PLoS Genet.* 6:e1000906.

Maxam, A.M., Gilbert, W. (1977) "A new method for sequencing DNA" en *Proc. Natl. Acad. Sci. USA*, Número 74 pp. 560-564.

Meselson, M, Yuan, R. (1968) "DNA restriction enzyme from E. coli" en *Nature*, Número 217 pp. 1110-1114.

Metzker, M.L. (2010) "Sequencing technologies - the next generation" en *Nat Rev Genet.*, Número 11(1) pp. 31-46.

Nicklas, J.A., Buel, E. (2003) "Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples" en *J Forensic Sci.*, Número 48(5) pp. 936-944.

Nowrousian, M. (2010) "Next-Generation Sequencing Techniques for Eukaryotic Microorganisms: Sequencing-Based Solutions to Biological Problems" en *Eukaryot Cell*, Número 9(9) pp. 1300–1310.

Oudet, C., Heilig, R., Hanauer, A., Mandel, J.L. (1991) "Nonradioactive assay for new microsatellite polymorphisms at the 5' end of the dystrophin gene, and estimation of intragenic recombination" en *Am J Hum Genet.*, Número 49(2) pp. 311-319.

Ozsolak, F., Platt, A.R., Jones, D.R., Reifengerger, J G., Sass, L.E., et al. (2009) "Direct RNA sequencing" en *Nature*, Número 461 pp. 814–818.

Park, J.Y., Kim J.H., An Y.R., Kim M.J., Lee W.S., et al. (2010) "A DNA microarray for species identification of cetacean animals in Korean water" en *BioChip Journal*, Número 4(3) pp. 197-203.

- Pinder, S.J., Perry, B.N., Skidmore, C.J., Savva, D. (1991) "Analysis of polymorphism in the bovine casein genes by use of the polymerase chain reaction" en *Animal Genetics*, Número 22 pp. 11-20.
- Rönn, A.C., Andrés O., López-Giráldez, F., Johnsson-Glans, C., Verschoor, E.J., et al. (2009) "First generation microarray-system for identification of primate species subject to bushmeat trade" en *Endang. Species Res.* Published online June 19, 2009.
- Sambrook, J., Fritsch, E.F., Maniatis, ST. (1989) "Molecular Cloning: A laboratory manual" Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1626 pp.
- Sanger, F., Nicklen, S., Coulson, A.R. (1977) "DNA sequencing with chain – terminating inhibitors" en *Proc. Natl. Acad. Sci. USA*. Número 74 pp. 5463 – 5468.
- Sobrino, B., Brión, M., Carracedo, A. "SNPs in forensic genetics: a review on SNP typing methodologies" en *Forensic Sci Int.*, Número 154(2-3) pp. 181-94.
- Sokolov, B.P. (1990) "Primer extension technique for the detection of single nucleotide in genomic DNA" en *Nucleic Acids Res.*, Número 18 pp. 3671.
- Steffen, P., Eggen, A., Dietz, A., Womack, J.E., Stranzinger, G., Fries, R. (1993) "Isolation and mapping of polymorphic microsatellites in cattle" en *Animal Genetics*, Número 24 pp. 1221-1224.
- Sun, H., Dentine, M.R., Barensse, W., Kirkpatrick, B.W. (1994) "UWCA19 and UWCA20: Polymorphic bovine microsatellite" en *Animal Genetics*, Número 24(2) pp. 142.
- Sun, H., Dentine, M.R., Kirkpatrick, B.W. (1994) "A polymorphic bovine microsatellite" en *Animal Genetics*, Número 24(2) pp. 149.
- Sun, H., Dentine, M.R., Kirkpatrick, B.W. (1993) "A polymorphic microsatellite (UWCA4) detected on bovine chromosome 23" en *Animal Genetics*, Número 24(2) pp. 149.

- Sun, H., Hart, G.L., Kirkpatrick, B.W. (1993) "A polymorphic microsatellite (UWCA1) detected on bovine chromosome 23" en *Animal Genetics*, Número 24(2) pp. 142.
- Sunden, S.L.F., Stone, R.T., Bishop, M.D., Kappes, S.M., Keele, J.W., Beattie, C.W. (1993) "A highly polymorphic bovine microsatellite locus: BM2113" en *Animal Genetics*, Número 24(1) pp. 69.
- Swarbrick, P.A., Dietz, A.B., Womack, J.E., Aston, W.P. (1990) "Ovine and bovine dinucleotide repeat polymorphism at the MAF46 locus" en *Animal Genetics*, Número 23 pp. 182.
- Syvanen, A.C., Aalto-Setälä, K., Harju, L., Kontula, K., Soderlund, H. (1990) "A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E" en *Genomics*, Número 8 pp. 684–692.
- Tautz, D. (1989) "Hypervariability of simple sequences as a general source for polymorphic DNA markers" en *Nucleic Acids Research*, Número 17 (16) pp. 6463-6471.
- Teletchea, F., Bernillon, J., Duffraisse, M., Laudet, V. and Hänni, C. (2008) "Molecular identification of vertebrate species by oligonucleotide microarray in food and forensic samples" en *Journal of Applied Ecology*, Número 45 pp. 967–975.
- Tyagi, S., Kramer, F.R. (1996) "Molecular beacons: probes that fluoresce upon hybridization" en *Nat. Biotechnol.*, Número 14 pp. 303–308.
- Vaiman, D., Mercier, D., Moazami-Goudarzi, K. "A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism" en *Mammalian Genome*, Número 5 pp. 288-297.
- Wang, Z., Zhang, J., Luo, H., Ye, Y., Yan, J., Hou, Y. (2012) "Screening and confirmation of microRNA markers for forensic body fluid identification" en *Cytotherapy*, Número 14(6) pp. 752-66
- Watson, J., Crick, F. (1953) "Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid" en *Nature*, Número 171 pp. 737-738.

Weber, J.L., May, P.E. (1989) "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction" en *American Journal of Human Genetics*, Número 44 pp. 388-396.

Xia, J., Wang, Q., Jia, P., Wang, B., Pao, W., Zhao, Z. (2012) "NGS catalog: A database of next generation sequencing studies in humans" en *Hum Mutat.*, Número 33(6) pp. 2341-2355.

Ziętkiewicz, E., Witt, M., Dąca, P., Żebracka-Gala, J., Goniewicz, M., Jarzab, B., Witt, M. (2012) "Current genetic methodologies in the identification of disaster victims and in forensic analysis" en *J. Appl. Genetics*, Número 53 pp. 41-60.