

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 023 217**

51 Int. Cl.:

C12Q 1/6818 (2008.01)

C12Q 1/6851 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.09.2020 PCT/EP2020/076410**

87 Fecha y número de publicación internacional: **01.04.2021 WO21058470**

96 Fecha de presentación y número de la solicitud europea: **22.09.2020 E 20772085 (5)**

97 Fecha y número de publicación de la concesión europea: **26.03.2025 EP 4034674**

54 Título: **Sonda y método para el genotipificación de STR**

30 Prioridad:

23.09.2019 EP 19199001

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

30.05.2025

73 Titular/es:

**UNIVERSITEIT GENT (100.00%)
Sint-Pietersnieuwstraat 25
9000 Gent, BE**

72 Inventor/es:

**TYTGAT, OLIVIER;
VAN NIEUWERBURGH, FILIP;
DEFORCE, DIETER y
CORNELIS, SENNE**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 3 023 217 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Sonda y método para el genotipificación de STR

5 **Campo técnico de la invención**

La presente invención se refiere al campo de la toma de huellas de ADN, por ejemplo en un entorno forense. Más específicamente, la presente invención describe una sonda y un método para genotipar repeticiones en tándem cortas polimorfas, basándose en las propiedades de extinción de la fluorescencia de algunos nucleótidos en algunos fluoróforos específicos. Como tal, el grado de complementariedad entre una muestra de ADN amplificado y una sonda específicamente diseñada se puede evaluar midiendo la intensidad de fluorescencia del fluoróforo unido a la sonda con la hibridación o la fusión. Las sondas y el método de la presente invención son adecuados para ser utilizados en un dispositivo de análisis de ADN portátil y menos costoso, y pueden aplicarse en otros campos además del forense, como fraude alimentario, diagnóstico y muchos otros.

15 **Antecedentes de la invención**

Para identificar individuos basándose en pruebas de ADN, un proceso llamado huella genética, se deben analizar polimorfismos de ADN, por ejemplo polimorfismos de un solo nucleótido (SNP) o repeticiones en tándem cortas (STR). Los individuos pueden designar a seres humanos, pero también a animales y otras especies que contienen ADN.

Los polimorfismos en el genoma contienen una enorme cantidad de información. El tipo más común de polimorfismos genéticos se clasifica como SNP. La presencia de un SNP en una determinada posición del genoma indica que, dentro de una determinada población, el mismo nucleótido no se encuentra en esa posición específica en cada individuo de esa población. Los SNP pueden ser bialélicos, lo que indica que 2 nucleótidos posibles ocurren en esta posición dentro de una población determinada. Recientemente, se ha prestado cada vez más atención a la existencia de SNP tri- o tetra-alélicos, que tienen mayor poder discriminatorio que los SNP bialélicos [1]. Una variante causada por la inserción o delección de un nucleótido en una posición determinada del genoma a veces también se denomina SNP, pero también se conoce como indel (abreviatura de 'inserción/delección').

En la actualidad, la genotipificación forense del ADN se realiza casi exclusivamente mediante el examen de las STR. Se trata de regiones de secuencias cortas (unos pocos nucleótidos, normalmente 4) que se repiten varias veces. Las regiones STR son polimorfas en cuanto al número de repeticiones, lo que a su vez define los posibles alelos de una determinada región STR [2].

En el genoma humano se identifican múltiples regiones que contienen este tipo específico de polimorfismo. Se obtiene un perfil estadísticamente único analizando una gran cantidad de loci de STR, ubicados principalmente en regiones no codificantes del genoma humano con fines forenses. En Europa, normalmente se examinaba un panel de 12 STR, llamado Conjunto de Normas Europeas (ESS). Este panel ahora se ha ampliado con cinco loci adicionales [3]. En EE. UU. se utiliza el Sistema de Índice Combinado de ADN (CODIS), que contiene 13 loci principales y 7 loci adicionales [4].

Después de la amplificación mediante la reacción en cadena de la polimerasa (PCR), el número de repeticiones se puede deducir de la longitud del amplicón. Actualmente, esta información se obtiene casi exclusivamente mediante electroforesis capilar, una técnica de separación de ADN muy conocida. La separación por tamaño de los amplicones se logra aplicando un alto potencial eléctrico a través de un capilar lleno de gel a través del cual se mueven los amplicones. Las diferencias en la movilidad electroforética darán como resultado una migración más rápida de amplicones más cortos. Estos fragmentos de ADN marcados con fluorescencia se detectan mediante fluorescencia inducida por láser (LIF) [2].

Aunque se están realizando esfuerzos para crear una variante portátil de las herramientas utilizadas para CE (electroforesis capilar) (por ejemplo, el sistema RapidHIT) [5] o miniaturizando esta técnica en un chip (por ejemplo, vidrio) [6], CE aún requiere un equipo bastante voluminoso. El interés en implementar Lab-on-a-Chip (LoC) en el campo de la genotipificación forense del ADN está creciendo rápidamente, debido a una gran cantidad de beneficios, tales como tiempos de análisis más cortos y menor consumo de reactivos, pero también un alto grado de paralelización y flexibilidad. Además, el coste de producción del equipo disminuiría, y la facilidad de uso aumentaría drásticamente. Otra ventaja importante, especialmente en el campo forense, es el menor riesgo de contaminación. Los LoC muestran un alto nivel de integración, combinando varias funcionalidades en un solo dispositivo y eliminando la necesidad de transportar una muestra de un dispositivo a otro. Al hacer que el análisis de ADN sea portátil, el transporte a un laboratorio acreditado se vuelve redundante, minimizando nuevamente el riesgo de contaminación y permitiendo un tiempo de respuesta más rápido. Los resultados rápidos son cruciales para aquellos encargados de resolver crímenes, ya que las primeras horas de dicha investigación se denominan comúnmente las "horas doradas".

65

Ya se ha realizado un gran esfuerzo para desarrollar nuevos ensayos, principalmente basados en la hibridación. Esto implica que se utilizan oligonucleótidos fabricados sintéticamente, en su mayoría marcados con fluorescencia, para la genotipificación de STR. Estas técnicas muestran algunos beneficios en comparación con la CE, por ejemplo, no se deben aplicar altos voltajes. Se omite otro problema relacionado con la CE, es decir, la detección de artefactos de PCR, ya que estos artefactos no se hibridan con las sondas añadidas al sistema.

Un ejemplo bien conocido de un método basado en hibridación para la genotipificación de STR es el uso de sondas HyBeacon™ para el análisis de la curva de fusión, como se describe en el documento EP3011053A2 [7]. En esta técnica, sólo se utiliza una sonda marcada con fluorescencia por locus, y el número de repeticiones se deduce de la temperatura de fusión de dicha sonda [8]. Las desventajas de este método son múltiples: i) la necesidad de un segundo oligonucleótido que funcione como bloqueador, ii) la presencia de múltiples fluoróforos posicionados internamente, lo que aumenta el costo de las sondas, y iii) las restricciones en el diseño de las sondas, que hacen imposible diseñar un sistema capaz de genotipar todos los loci necesarios para un perfil de ADN completo.

La mayoría de las técnicas alternativas basadas en hibridación se basan en el uso de uno o más fluoróforos combinados con un resto extintor, lo que aumenta el costo de la sonda y la complejidad del sistema. Los ejemplos son las sondas TaqMan (patente de EE.UU. núm. 5.210.015) [9] y balizas moleculares (patente de EE.UU. núm. 6150097A) [10]. Otros sistemas utilizan múltiples fluoróforos, lo que aumenta nuevamente el costo de la sonda y la complejidad del sistema. Ejemplos de estos sistemas son las sondas Scorpion [11] y las sondas ECHO [12]. Cabe señalar que la mayoría de estas sondas no están diseñadas para la genotipificación de STR.

El principio de transferencia de energía resonante fluorescente (FRET) entre un resto donante y un resto aceptor se aprovecha a menudo en métodos basados en sondas de ADN. En la mayoría de los casos, ambos restos están unidos a (uno de) los oligonucleótidos de los que existe el sistema. Halpern et al. desarrollaron un ensayo de genotipificación de curva de fusión basado en FRET, combinando un colorante intercalante y un oligonucleótido marcado con fluorescencia. Al fundirse, la FRET entre el colorante intercalante y el oligonucleótido desaparece, y se puede observar una disminución en la señal fluorescente. La determinación real del alelo de STR en dicho sistema se basa en la detección de T_m , ya que el desemparejamiento da como resultado una T_m más baja en comparación con un emparejamiento perfecto [13], documento US12/276849 ([14]).

Aunque, en el caso de un desemparejamiento, no se espera que la región reportera de la sonda se hibride con el amplicón, lo que implica la ausencia de colorante intercalante en esta región, sino que más bien se observan curvas de fusión bastante altas y pronunciadas para estas sondas de desemparejamiento, aunque a una temperatura más baja. Probablemente esto se debe a la presencia de colorante intercalante en el resto del dúplex, y a la presencia de colorante intercalante en la proximidad del ADN monocatenario. Sería beneficioso si las curvas de fusión desemparejadas no se produjeran sólo a una temperatura más baja, sino que también tuvieran otra forma (por ejemplo, menor altura de pico, mayor anchura de pico). Esto permitiría una genotipificación correcta de STR a pesar de la presencia de un SNP ubicado en las regiones flanqueantes de la sonda. Un desemparejamiento en la posición del SNP combinado con una coincidencia en la región de STR daría como resultado curvas de fusión con una forma de pico similar a las sondas de emparejamiento normales, pero a una temperatura más baja.

El uso de un colorante intercalante está indisolublemente ligado a algunas desventajas, por ejemplo la inhibición de la PCR. Como las muestras forenses a menudo contienen sólo cantidades muy pequeñas de ADN, la PCR debería realizarse preferiblemente en circunstancias ideales. La inhibición de la PCR es la razón por la que los colorantes intercalantes a menudo se añaden en concentraciones subsaturadas, aumentando de ese modo el riesgo de salto del colorante. Este fenómeno se refiere al colorante intercalante que se libera de un dúplex al fundirse pero que se incorpora a otro dúplex que aún no se ha fundido, lo que da como resultado la ampliación de los picos de fusión. Además, los colorantes intercalantes se unen preferentemente a regiones que contienen gran cantidad de guanina y citosina (las denominadas "regiones ricas en GC"). Esto potencialmente da como resultado una señal más pronunciada para estos loci. Por último, la concentración de colorante intercalante influye de forma bastante pronunciada en la temperatura de fusión de un dúplex, introduciendo así una nueva fuente de variabilidad de la temperatura de fusión, lo que resulta a todos los efectos perjudicial cuando se realiza un ensayo de genotipificación de curva de fusión. [15] Además de estas limitaciones técnicas, se deben tener en cuenta los riesgos para la salud humana relacionados con el uso de colorantes intercalantes.

Todos estos ejemplos indican la necesidad de un ensayo de genotipificación basado en hibridación que sea más simple y robusto por un lado, que permita la integración en un dispositivo portátil con facilidad, y que dé como resultado más información que simplemente una temperatura de fusión por otro lado. En nuestra opinión, la sonda más simple posible contiene sólo un fluoróforo, y genera una señal basada únicamente en la interacción con la muestra, sin el uso de otras modificaciones (por ejemplo, un extintor), otras moléculas (por ejemplo, colorantes intercalantes) o incluso otras sondas (por ejemplo, un bloqueador).

Las sondas que se basan únicamente en las propiedades de extinción de nucleótidos de origen natural ya han demostrado su utilidad en otras aplicaciones, por ejemplo la identificación de especies, qPCR y genotipificación de SNP, como describen Wittwer et al. en el documento EP2927238A1. [16] En dicho documento, se describen las denominadas sondas Q para la genotipificación de SNP. Estas sondas están diseñadas de tal manera que el

nucleótido marcado con fluorescencia siempre se hibrida con la secuencia diana, independientemente de si la sonda se empareja o no con el amplicón en la posición SNP. De esta manera, la extinción aparecerá tras la hibridación, independientemente del genotipo de la muestra examinada. La genotipificación se realiza mediante el análisis de la curva de fusión estándar, ya que una falta de desemparejamiento entre la sonda y la muestra reducirá la temperatura de fusión del dúplex. El efecto de extinción de la muestra sobre el fluoróforo no está asociado al hecho de si la sonda y el amplicón se emparejan perfectamente o no, y la determinación de variantes se basa en un factor: la temperatura de fusión. Esto ha demostrado ser suficiente información para la determinación de SNP, mientras que para la genotipificación de STR es posible una gama bastante amplia de alelos para cada locus investigado, lo que hace que la explotación de sondas como la descrita por Wittwer et al. o las sondas HyBeacon descritas anteriormente no sean útiles. Los loci de STR tienen otras propiedades estructurales en comparación con los loci de SNP, ya que los posibles alelos difieren en longitud en lugar de sólo en secuencia. Por lo tanto, para cada locus, se debe desarrollar una serie de sondas de diferente longitud junto con un método para evaluar si una sonda es o no completamente complementaria a la muestra, lo que contrasta marcadamente con las sondas descritas por Wittwer et al., en las que sólo se diseña una sonda para un locus determinado.

Por lo tanto, todavía es necesario diseñar sondas simples que sean útiles para la genotipificación de STR, ya que algunos elementos estructurales (por ejemplo, una región de anclaje y una región sensora) son indispensables.

Breve descripción de las figuras

Figura 1: Ejemplo de genotipificación de STR para el locus D16S539. Se muestran tres dúplex de sonda (arriba) y amplicón de STR (abajo). Cada amplicón contiene doce repeticiones STR, mientras que las sondas representadas contienen once, doce o trece repeticiones. Las sondas consisten, de 5' a 3', en una región FL1 (negro), un número específico de repeticiones (gris oscuro), y la región FL2 que contiene uno o más fluoróforos (gris claro). Cuando el amplicón contiene exactamente la misma cantidad de repeticiones que la sonda, la región FL2 aparecerá en un estado hibridado, lo que provocará la extinción del fluoróforo o fluoróforos.

Figura 2: Ejemplo típico de un experimento de fusión. Antes de la fusión, la fluorescencia disminuye de forma lineal, y las sondas se hibridan con un amplicón, lo que tiene como consecuencia la extinción del fluoróforo. Al fundirse, se observa un aumento repentino de la fluorescencia debido a la pérdida del efecto extintor de la guanina. Después de la fusión, la fluorescencia disminuye nuevamente de manera lineal, y las sondas pasan a ser monocatenarias.

Figura 3: Curvas de fusión de las sondas de desemparejamiento para el locus D8S1179.

Figura 4: Perfil de temperatura para el experimento D16S539.

Figura 5: Picos de fusión de las sondas diseñadas para el locus D16S539 después de la hibridación con una muestra heterocigótica (9:12).

Figura 6: Picos de fusión de las sondas diseñadas para el locus D16S539 después de la hibridación con una muestra homocigótica (9:9).

Figura 7: Picos de fusión de las sondas diseñadas para el locus D16S539 después de la hibridación con una muestra heterocigótica (11:13).

Figura 8: Picos de fusión de las sondas diseñadas para el locus TH01 después de la hibridación con una muestra heterocigótica (9,3:10).

Figura 9: Picos de fusión de las sondas diseñadas para el locus TH01 después de la hibridación con una muestra homocigótica (9,3:9,3).

Figura 10: Picos de fusión de las sondas diseñadas para el locus D8S1179 después de la hibridación con una muestra heterocigótica (13:13').

Figura 11: Picos de fusión de las sondas diseñadas para el locus D8S1179 después de la hibridación con una muestra heterocigótica (14:15).

Sumario de la invención

La presente invención se refiere a una pluralidad de sondas que representan la variabilidad alélica de un determinado locus de repetición en tándem corta dentro de una población, en la que cada sonda consiste, de 5' a 3' o de 3' a 5', en: 1) una primera región flanqueante que comprende nucleótidos y que se hibrida con una región directamente próxima a la secuencia de ADN específica de interés y que garantiza una hibridación adecuada de la muestra y la sonda y que por lo tanto contiene más nucleótidos que la segunda región flanqueante, 2) una secuencia de ADN específica de interés que comprende al menos una repetición en tándem corta y que se hibrida

con la región de repetición en tándem corta dentro de la muestra, y 3) una segunda región flanqueante que comprende al menos un nucleótido y que contiene al menos un fluoróforo y en la que dicho fluoróforo está unido a un resto de dicha segunda región flanqueante en la posición directamente complementaria a un nucleótido específico capaz de extinguir dicho fluoróforo de manera eficiente de dicha muestra, o unido a un nucleótido adyacente, ya sea en dirección 5' o en dirección 3', a dicha posición, o unido a un nucleótido a 2 posiciones de distancia, ya sea en dirección 5' o en dirección 3', de dicha posición, de modo que se acerca a uno o más nucleótidos específicos capaces de extinguir dicho fluoróforo de una manera eficiente de dicha muestra con la hibridación de dicha segunda región flanqueante con la muestra.

La presente invención se refiere además a una pluralidad de sondas como las descritas anteriormente, en la que dichos nucleótidos son análogos de ácidos nucleicos, por ejemplo LNA.

Más específicamente, la presente invención se refiere a una pluralidad de sondas como las descritas anteriormente, en la que dicho fluoróforo se escoge de la lista que comprende fluoresceína (FAM), hexaclorofluoresceína (HEX), tetracloro-6-carboxifluoresceína (TET), 2,7-dimetoxi-4,5-dicloro-6-carboxifluoresceína (JOE) o 6-carboxitetrametilrodamina (TAMRA). El nucleótido específico capaz de extinguir uno de los fluoróforos enumerados anteriormente es la guanina.

Más específicamente, la presente invención se refiere a una pluralidad de sondas como las descritas anteriormente en la que dicho fluoróforo está unido a un resto de citosina de dicha segunda región flanqueante.

La presente invención se refiere también a una pluralidad de sondas como las descritas anteriormente que están inmovilizadas sobre un soporte.

La presente invención también se refiere a un método para genotipar repeticiones en tándem cortas dentro de una muestra, que comprende las etapas de:

- proporcionar una muestra que comprende ADN,
- amplificar el ADN dentro de dicha muestra que comprende una secuencia de ADN específica de interés para obtener secuencias de ADN amplificadas,
- añadir una pluralidad de sondas como se describe anteriormente a dichas secuencias de ADN amplificadas para obtener dúplex de secuencias de ADN monocatenario hibridadas a dicha sonda, y
- desnaturalizar dichos dúplex, seguido del enfriamiento lento de dichos dúplex desnaturalizados mientras se mide continuamente la fluorescencia del fluoróforo de dicha sonda, o calentar lentamente dichos dúplex mientras se mide continuamente la fluorescencia del fluoróforo de dicha sonda, en el que la disminución de la intensidad de fluorescencia o el aumento de la intensidad de fluorescencia, respectivamente, proporciona información sobre si una repetición en tándem corta completamente complementaria específica está presente o no en dicha muestra.

La presente invención se refiere además a un método para genotipar como se describe anteriormente, en el que dicha amplificación dentro de dicha muestra se lleva a cabo mediante una PCR asimétrica para obtener secuencias de ADN monocatenario amplificadas.

La presente invención se refiere además a un método para genotipar como se describe anteriormente, en el que dicha amplificación dentro de dicha muestra se lleva a cabo mediante una PCR simétrica utilizando cebadores marcados con biotina o una digestión posterior con exonucleasa lambda para obtener secuencias de ADN monocatenarias amplificadas.

La presente invención también se refiere a un método como se describe anteriormente en el que dichas sondas se añaden en disolución o se inmovilizan sobre un soporte.

Descripción de la invención

Un aspecto de esta invención se refiere a sondas cuyo funcionamiento depende de las propiedades de extinción de origen natural de nucleótidos específicos en ciertos fluoróforos. El ejemplo más común es el efecto de extinción de la fluoresceína de la guanina [17, 18]. Otro ejemplo es la extinción del pirenobutirato por los nucleótidos de timidina.

Una sonda se define aquí como un oligonucleótido fabricado sintéticamente en el que algunos nucleótidos pueden estar modificados. Ejemplos de modificaciones son, por ejemplo, la presencia de un resto fluorescente, moléculas para fines de fijación, etc. Las sondas generalmente se diseñan de tal manera que interactuarán con la molécula investigada, y la respuesta de la sonda ante esta interacción se observará y se utilizará para obtener información de dicha molécula investigada.

Las sondas de genotipificación de STR descritas en esta invención consisten en tres regiones diferentes, como se muestra en la figura 1: región flanqueante 1 (FL1), la región STR específica, y la región flanqueante 2 (FL2).

- 5 • FL1 es la región que se encuentra directamente al lado de la secuencia de ADN específica y actúa como ancla, asegurando la hibridación adecuada de la muestra y la sonda, evitando deslizamientos. Esto implica que FL1 debería ser significativamente más larga que FL2, un requisito que también se analiza en la bibliografía sobre otras sondas de genotipificación de STR [13]. Si FL1 fuera tan larga o incluso más corta que FL2, FL1 sería monocatenaria en caso de desemparejamiento, y FL2 se hibridaría con la muestra, lo que daría como resultado una señal comparable a la señal generada por un dúplex de emparejamiento. Se puede añadir una funcionalidad adicional al final de FL1 con fines de conexión.
- 10 • La región de STR es la parte polimorfa que difiere entre las sondas para un locus determinado. Una sonda se diseña para cada alelo posible del locus examinado.
- 15 • FL2 es sustancialmente más corta que FL1, y está marcada terminalmente con un fluoróforo, por ejemplo FAM. Este marcaje puede ser de terminal 5' o 3'. FL2 actúa como un sensor, y da una indicación del grado de complementariedad entre la sonda y la muestra.

20 La sonda está diseñada de tal manera que, tras la hibridación con un amplicón complementario, el fluoróforo se acerca a uno o más nucleótidos capaces de extinguir dicho fluoróforo. En una realización más específica de esta invención, dicho fluoróforo es FAM, que se extingue por la presencia de restos de guanina. Estos restos de guanina también ejercen un efecto de extinción sobre otros fluoróforos, como HEX, TET, JOE y TAMRA [19]. Los expertos en la técnica reconocerán que ésta es una lista no limitativa. Cabe señalar que otras combinaciones de fluoróforos y nucleótidos también son aplicables para este propósito. Para lograr una extinción eficiente del fluoróforo FAM, dicho fluoróforo se une a un nucleótido (principalmente citosina) en la posición directamente complementaria al resto de guanina, o se une a un nucleótido adyacente a dicha posición (ya sea en dirección 5' o en dirección 3'), o se une a un nucleótido a 2 posiciones de dicha posición (ya sea en dirección 5' o en dirección 3').

30 En el método descrito aquí, se diseña una matriz de sondas que representan todos los alelos posibles de un determinado locus de STR. La diferencia entre la complementariedad total y la complementariedad parcial de una muestra amplificada y la sonda se puede evaluar midiendo la intensidad de fluorescencia del fluoróforo unido a la sonda tras la hibridación o la fusión. Los gráficos de fluorescencia resultantes en función del tiempo se pueden dividir en 3 partes (véase la figura 2): una parte lineal, durante la cual la fluorescencia disminuye (fenómeno dependiente de la temperatura) y (la mayoría de) las sondas se hibridan con un amplicón, con la extinción del fluoróforo como consecuencia, la parte de fusión durante la cual aumenta la fluorescencia, y una segunda fase lineal en la que las sondas son monocatenarias. El cálculo de la primera derivada de estos gráficos en función de la temperatura proporciona picos de fusión que se utilizan para la interpretación de los datos. Después de la amplificación, la sonda y los amplicones se desnaturalizan mediante calentamiento y después se enfrían lentamente de manera controlada, asegurando una hibridación correcta al evitar el deslizamiento de la sonda. Posteriormente, los dúplex sonda-amplicón se funden, mientras se mide constantemente la fluorescencia. Al fundirse, la distancia entre el fluoróforo y el resto de guanina extintor aumentará, y la intensidad de la fluorescencia aumentará. Este aumento en la intensidad de la fluorescencia ocurre a una temperatura más alta, y es más pronunciado cuando la sonda y el amplicón presentes en el producto de PCR comparten el mismo número de repeticiones, en comparación con una combinación no desemparejada de sonda y amplicón. Aún se puede observar cierta falta de extinción durante la fusión en el caso de una situación de desemparejamiento, debido a la formación de heterodúplex. Estos dúplex contienen un desemparejamiento en la región de repetición, lo que da como resultado la formación de un bucle abultado. Sin embargo, la temperatura de fusión de estos dúplex es menor en comparación con la complementariedad completa, y la eficiencia de hibridación es pronunciadamente menor: para la mayoría de las sondas, la región sensora permanecerá monocatenaria.

Las sondas descritas aquí proporcionan información sobre el grado de complementariedad entre la sonda y la muestra. Además de deducir si la sonda y la muestra tienen o no el mismo número de repeticiones, se puede obtener información sobre el número de repeticiones que difieren entre la muestra y la sonda en el caso de un desemparejamiento. Cuanto mayor sea la diferencia en el número de repeticiones, menor será la señal obtenida. En la figura 4 se ofrece un ejemplo para el locus D8S1179. Se muestran las curvas de fusión de 4 sondas, después de la incubación con la muestra de referencia 2800, que tiene los alelos 14 y 15. Todas las curvas de fusión mostradas se originan a partir de sondas de desemparejamiento. Se puede ver claramente que la sonda 13 muestra la señal más intensa, y la sonda 10 la menos intensa. La intensidad de la señal se puede definir en este ejemplo mediante la altura del pico de fusión y la temperatura de fusión.

Una característica única de esta invención es la considerable cantidad de información que se puede recuperar de múltiples parámetros de estas curvas de fusión (T_m , forma del pico, ...). Las curvas de fusión dan como resultado una indicación del grado de complementariedad, mientras que la mayoría de los sistemas sólo dan una respuesta binaria (emparejamiento o desemparejamiento). Los últimos sistemas sólo tienen en cuenta un parámetro, por ejemplo la temperatura de fusión o la intensidad de fluorescencia. La posición única del fluoróforo, en combinación

con los otros elementos estructurales de las sondas, hace que estas sondas de STR sean altamente informativas. El fluoróforo se posiciona en la segunda región flanqueante, actuando así como un sensor: cuando la sonda se empareja con la muestra, FL2 se hibridará con dicha muestra. Por el contrario, cuando la sonda no se empareja con la muestra, FL2 permanece principalmente monocatenaria, o se funde a una temperatura más baja, y la falta de extinción al fundirse se producirá de forma menos repentina. Cuanto mayor sea la distancia entre el fluoróforo y el amplicón con restos extintores, menos intensa será la señal. Por lo tanto, hasta donde sabemos, la sonda de genotipificación de STR analizada aquí es la sonda de genotipificación de STR más elemental e informativa descrita hasta el momento, ya que tanto la temperatura de fusión como la intensidad de fluorescencia son informativas.

Esta información obtenida podrá eventualmente ser analizada de forma automatizada mediante inteligencia artificial. Ya se han descrito algoritmos similares para el análisis de fusión de alta resolución. En el futuro se podría desarrollar un algoritmo personalizado para la determinación de alelos, basándose en una gran cantidad de datos. Para este fin, dicho algoritmo debe ser entrenado para llamar a los alelos correctos, basándose en las curvas de muestras con alelos conocidos.

Un aspecto importante de los métodos de genotipificación basados en hibridación es el requisito de un exceso del amplicón complementario a la sonda. Si no se cumpliera este requisito, ambas cadenas de amplicón se hibridarían preferentemente entre sí, dejando la sonda monocatenaria. Se puede obtener un exceso de una cadena de amplicón adaptando la etapa de amplificación. Después de la preparación de la muestra, se debe realizar una etapa de amplificación para amplificar los loci de STR. Esto normalmente se hace mediante la reacción en cadena de la polimerasa (PCR), una técnica bien conocida por los expertos en la técnica. Las regiones que se amplifican están determinadas por los cebadores utilizados. Se trata de oligonucleótidos cortos complementarios a una secuencia del genoma de la especie examinada. La ADN polimerasa iniciará la amplificación en el extremo 3' del cebador.

En la PCR simétrica, ambos cebadores se añaden en concentraciones iguales, lo que da como resultado amplicones bicatenarios. Cuando se realiza PCR asimétrica, se añade un cebador en exceso. En el primer ciclo, ambos cebadores están presentes, y la PCR se realiza de forma simétrica. En un momento determinado, se agotará un cebador, lo que dará como resultado la amplificación de sólo 1 de las 2 cadenas. A partir de este punto la amplificación no se producirá de forma exponencial sino lineal.

La PCR asimétrica no es la única forma de obtener un exceso de una cadena específica. Tras realizar una PCR simétrica en la que uno de ambos cebadores está marcado con biotina, las hebras en las que se incorpora dicho cebador pueden ser capturadas mediante perlas de estreptavidina. Otra opción es el uso de la enzima lambda exonucleasa, que degrada selectivamente las cadenas de ADN fosforiladas. Esta modificación se puede introducir en 1 de los 2 cebadores. [20]

La sonda descrita anteriormente también puede contener análogos de ácidos nucleicos, por ejemplo LNA. Los primeros son componentes que no se producen de forma natural y que se parecen estructuralmente a los ácidos nucleicos que se producen de forma natural. Entre muchos otros ejemplos están los ácidos nucleicos con una base modificada, o una modificación en el componente de azúcar.

Ejemplos

1. EJEMPLO 1: Hisopos bucales para genotipificación de STR (locus D16S539)

Se sumergieron tres hisopos bucales en un volumen de 200 µl de agua de HPLC estéril. Después de una etapa de vórtice de 30", el hisopo se retiró, y el agua se utilizó como entrada para la PCR. Se realizó PCR asimétrica singleplex con 30 µl de muestra de entrada. Las concentraciones de cebador fueron 0,1 µM de cebador directo y 1,5 µM de cebador inverso. El volumen de la mezcla de PCR fue de 50 µl, que contiene MgCl₂⁺ a una concentración de 0,5 mM, dNTPs a 200 µM cada uno, tampón de PCR Qiagen 1X, y 1,3 U de enzima HotStarTaq. La activación de la polimerasa se realizó calentando la mezcla de PCR a 95 °C durante 15 minutos, seguido de 60 ciclos de 95 °C durante 1 minuto, 59 °C durante 1 minuto, y 72 °C durante 80 segundos. Las secuencias de los cebadores se pueden encontrar en la tabla 1.

Después de la PCR asimétrica, alícuotas de 8,5 µl del producto amplificado se dividieron en una placa de 96 pocillos. A cada pocillo separado, se añadieron 1,5 µl de una sonda particular a una concentración inicial de 1 µM. Estas mezclas se desnaturalizaron durante 10 minutos a 95 °C, seguido de un enfriamiento lento a una velocidad de rampa de 0,04 °C/s mientras se medía continuamente la fluorescencia utilizando un LightCycler (Roche). Lo mismo se hizo calentando lentamente, durante este procedimiento los dúplex se derretirán. Las secuencias de la sonda se pueden encontrar en la tabla 1.

Tabla 1: Secuencia de oligonucleótidos utilizados para el experimento D16S539. 'n' denota el número de repeticiones, y varió entre 9 y 13.

| Nombre | Secuencia | SEQ ID N° |
|-------------------------|--|---|
| D16S539 Cebador directo | 5' GGGGGTCTAAGAGCTTGTA AAAAAG | 1 |
| D16S539 Cebador inverso | 5' GTTTGTGTGTGCATCTGTAAGCATGTATC | 2 |
| D16S539 Sonda | 5' GTTTGTCTTTCAATGA(TATC) _n CAC/36-FAM/ | 3 (n=9), 4 (n=10), 5 (n=11), 6 (n=12), 7 (n=13) |

5 Se calcula la primera derivada de la curva de fusión, lo que da como resultado picos de fusión. De esta manera se pueden examinar las diferencias en la temperatura de fusión debidas a una diferencia en la longitud de la sonda. Todas las muestras examinadas también se genotiparon con análisis de CE convencional como referencia.

10 2. EJEMPLO 2: Hisopos bucales para genotipificación de STR (locus TH01)

Para evaluar la capacidad de este sistema para detectar diferencias bastante sutiles en la longitud del amplicón, causadas por repeticiones parciales, se llevó a cabo un experimento de curva de fusión utilizando las sondas diseñadas para el locus TH01. Un alelo bastante común para este locus es el alelo 9.3, que se caracteriza por la presencia de 10 repeticiones (CATT), de las cuales la cuarta repetición tiene una deleción T. Como consecuencia de ello, los alelos 9.3 y 10 sólo difieren en un nucleótido de longitud, lo que ha resultado ser un desafío incluso para la EC. Se extrajeron dos hisopos bucales, se amplificaron, y analizaron de la misma manera que el experimento para el locus D16S539. A diferencia de la amplificación para este último locus, el cebador directo se añadió en una concentración de 0,1 µM, y el cebador inverso se añadió en una concentración de 1,5 µM. La secuencia de cebadores y sondas utilizados se puede encontrar en la tabla 2. La muestra A tiene los alelos 9.3 y 10; la muestra B es homocigota (9.3:9.3).

Tabla 2: Secuencia de oligonucleótidos utilizados para el experimento TH01. 'n' denota el número de repeticiones, y varió entre 6 y 10.

| Nombre | Secuencia | SEQ ID N° |
|----------------------|---|---|
| TH01 Cebador directo | 5'GTGATTCCCATTGGCCTGTTC | 8 |
| TH01 Cebador inverso | 5'GTTTGTGTGTGCATCTGTAAGCATGTATC | 9 |
| TH01 Sonda | 5'/56-FAM/CCT(CATT) _n CACCATGGAGTCTGTGTTCCCTGTG | 10 (n=6), 11 (n=7), 12 (n=8), 13 (n=9), 14 (n=10) |
| TH01 Sonda 9,3 | 5'/56-FAM/CCT(CATT)3CAT(CATT)6CACCATGGAGTCTGTGTTCCCTGTG | 15 |

25 3. EJEMPLO 3: Muestras de referencia de genotipificación de STR (locus D8S1179)

Para evaluar la capacidad de este sistema para detectar isoalelos, causados por un SNP en una repetición, se llevó a cabo un experimento de curva de fusión utilizando las sondas diseñadas para el locus D8S1179. La muestra de referencia 9947a es homocigota para el locus D8S1179 (13:13), pero está genotipada mediante secuenciación paralela masiva como 13:13'. Las secuencias correspondientes a los alelos 13 y 13' se pueden encontrar en la tabla 3. La muestra de referencia 2800 es heterocigota para el locus D8S1179 (14:15). Ambas muestras de referencia se amplifican y analizan de la misma manera que el experimento para el locus D16S539. A diferencia de la amplificación para este último locus, el cebador directo se añadió en una concentración de 0,1 µM, y el cebador inverso se añadió en una concentración de 1,5 µM. La secuencia de cebadores y sondas utilizados se puede encontrar en la tabla 3.

ES 3 023 217 T3

Tabla 3: Secuencia de oligonucleótidos utilizados para el experimento D8S1179.

| Nombre | Secuencia | SEQ ID N° |
|----------------------------|---|-----------|
| D8S1179 Cebador directo | 5' ATTGCAACTTATATGTATTTTTGTATTTTCATG | 16 |
| D8S1179 Inverso cebador | 5' ACCAAATTGTGTTTCATGAGTATAGTTTC | 17 |
| D8S1179 Sonda 10 | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA) ₁₀ TTCCC/3 6-FAM/ | 18 |
| D8S1179 Sonda 11 | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA) ₁₁ TTCCC/36 -FAM/ | 19 |
| D8S1179 Sonda 12 | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA) ₁₂ TTCCC/36 -FAM/ | 20 |
| D8S1179 Sonda 13 | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA)(TCTG)(TCT A) ₁₁ TTCCC/36-FAM/ | 21 |
| D8S1179 Sonda 13' | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA) ₁₃ TTCCC/36 -FAM/ | 22 |
| D8S1179 Sonda 14 | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA)(TCTG)(TCT A) ₁₂ TTCCC/36-FAM/ | 23 |
| D8S1179 Sonda 14' | 5' TTGTATTTTCATGTGTACATTCGTA(TCTA) ₁₄ TTCCC/36 -FAM/ | 24 |

| Nombre | Secuencia | SEQ ID N° |
|------------------|---|-----------|
| D8S1179 Sonda 15 | 5' TTGTATTCATGTGTACATTCGTA(TCTA)(TCTG) (TCTA) ₁₃ TTCCC/36-FAM/ | 25 |
| D8S1179 Sonda 16 | 5' TTGTATTCATGTGTACATTCGTA(TCTA)(TCTG) (TCTA) ₁₄ TTCCC/36-FAM/ | 26 |

Resultados

1. EJEMPLO 1: Hisopos bucales para genotipificación de STR (locus D16S5339)

Se calculó la primera derivada de las curvas de fusión obtenidas, los picos de fusión resultantes se muestran en las figuras 5-7. La figura 5 muestra los picos de fusión obtenidos de la muestra 7 con los alelos 9 y 12, en la que la sonda del alelo 12 (P12) se funde a una temperatura más alta en comparación con la sonda del alelo 9 (P9).

Como se puede ver en la figura 5, cada sonda muestra algún tipo de pico de fusión. Sin embargo, P9 y P12 muestran una altura de pico mucho mayor y una anchura de pico más estrecha. P11 muestra el pico de fusión más intenso de las sondas de desemparejamiento, lo cual es significativo ya que se trata de un alelo vecino de la sonda de emparejamiento 12. Sin embargo, la diferencia de T_m entre P11 y P12 es demasiado grande, lo que indica una hibridación no específica de P11.

La figura 6 muestra las curvas de fusión de una muestra homocigótica (alelos 9:9). Los picos de fusión de las sondas de emparejamiento son más pronunciados en comparación con las muestras heterocigotas.

La figura 7 muestra las curvas de fusión de una muestra heterocigótica (alelos 11 y 13). La sonda con 12 repeticiones es una sonda vecina de ambas sondas de emparejamiento, pero aún así se puede hacer una distinción clara entre emparejamiento y desemparejamiento.

En resumen, se puede deducir información suficiente para la genotipificación a partir de experimentos de hibridación o fusión. Se debe tener en cuenta que, cuando se realiza un experimento de fusión, debe preceder un procedimiento de hibridación lento para garantizar la hibridación específica de las sondas.

2. EJEMPLO 2: Hisopos bucales para genotipificación de STR (locus TH01)

Cabe señalar que, para la mayoría de los loci examinados, las sondas de emparejamiento muestran 2 picos, mientras que las sondas de desemparejamiento sólo 1. Lo más probable es que esto se deba a la presencia de otro alelo (muestras heterocigotas), picos entrecortados, y productos de PCR no específicos. La evaluación de estas curvas de fusión es, en consecuencia, menos complicada. Sin embargo, para el locus D16S539, los alelos coincidentes muestran sólo un pico, que muy probablemente esté relacionado con la FL1 más corta de estas sondas.

Para la muestra A, 2 sondas muestran picos de fusión a una temperatura más alta, y además estos picos se caracterizan por un denominado "hombro", que en realidad es un segundo pico como se mencionó anteriormente. Las 2 sondas corresponden a los alelos correctos. Para la muestra homocigótica B, sólo una sonda muestra un pico de fusión a una temperatura más alta, correspondiente al alelo 9.3. Aunque la sonda 10 muestra un pico de fusión más alto, ocurre a una temperatura más baja, y no hay pico de hombro. Por tanto, se puede concluir que el alelo 10 no está presente en la muestra examinada. Además de esto, se puede concluir que las sondas y el sistema descritos son capaces de distinguir entre los alelos 9.3 y 10.

3. EJEMPLO 3: Muestras de referencia de genotipificación de STR (locus D8S1179)

Para la muestra 9947a, tanto la sonda 13 como la 13' muestran picos de fusión a una temperatura más alta. Además de eso, ambos picos de fusión muestran los denominados hombros, similares a las sondas TH01. Por lo tanto, se puede concluir que los amplicones complementarios para ambas sondas 13 y 13' están presentes en la

muestra 9947a. Para la muestra 2800, las sondas 14 y 15 muestran picos de fusión a una temperatura más alta, con hombros. La sonda 14', sin embargo, no muestra un hombro, y se produce a una temperatura más baja. La elevada altura del pico se puede explicar por la presencia de ambos alelos, el 14 (que tiene la misma longitud) y el 15 (que es un vecino cercano). Se puede concluir que este método es capaz de distinguir isoalelos, siendo por tanto más informativo que la electroforesis capilar.

Referencias

1. Westen, A.A., et al., Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples. *Forensic Science International: Genetics*, 2009. 3(4): p. 233-241.
2. Butler, J.M., *Forensic DNA typing: biology, technology, and genetics of STR markers*. 2005: Elsevier.
3. Schneider, P.M., Expansion of the European standard set of DNA database loci-the current situation. *Profiles in DNA*, 2009. 12: p. 6-7.
4. Hares, D.R., Selection and implementation of expanded CODIS core loci in the United States. *Forensic Science International: Genetics*, 2015. 17: p. 33-34.
5. Hennessy, L.K., et al., Developmental validation studies on the RapidHIT™ human DNA identification system. *Forensic Science International: Genetics Supplement Series*, 2013. 4(1): p. e7-e8.
6. Bruijns, B., et al., Microfluidic devices for forensic DNA analysis: A review. *Biosensors*, 2016. 6(3): p. 41.
7. French, D., et al., Fluorophore-based oligonucleotide probes with a universal element. 2016, documento EP3011053A2.
8. Gale, N., et al., Rapid typing of STRs in the human genome by HyBeacon® melting. *Organic & biomolecular chemistry*, 2008. 6(24): p. 4553-4559.
9. Gelfand, D.H., et al., Homogeneous assay system using the nuclease activity of a nucleic acid polymerase. 1993, patente de EE. UU. núm. 5.210.015.
10. Tyagi, S. y F.R. Kramer, Nucleic acid detection probes having non-FRET fluorescence quenching and kits and assays including such probes. 2000, US6150097.
11. Thelwell, N., et al., Mode of action and application of Scorpion primers to mutation detection. *Nucleic acids research*, 2000. 28(19): p. 3752-3761.
12. Okamoto, A., ECHO probes: a concept of fluorescence control for practical nucleic acid sensing. *Chemical Society Reviews*, 2011. 40(12): p. 5815-5828.
13. Halpern, M.D. y J. Ballantyne, An STR melt curve genotyping assay for forensic analysis employing an intercalating dye probe FRET. *Journal of forensic sciences*, 2011. 56(1): p. 36-45.
14. Halpern, M. y P.M. Ellis, Dye probe fluorescence resonance energy transfer genotyping. 2010, documento US12/276849.
15. Reed, G.H., J.O. Kent, y C.T. Wittwer, High-resolution DNA melting analysis for simple and efficient molecular diagnostics. 2007.
16. Wittwer, C.T., et al., Single-labeled oligonucleotide probes for homogeneous nucleic acid sequence analysis. 2003, patentes de Google.
17. Cooper, J.P. y P.J. Hagerman, Analysis of fluorescence energy transfer in duplex and branched DNA molecules. *Biochemistry*, 1990. 29(39): p. 9261-9268.
18. Lee, S.P., et al., A fluorometric assay for DNA cleavage reactions characterized with BamHI restriction endonuclease. *Analytical biochemistry*, 1994. 220(2): p. 377-383.
19. Mao, H., et al., The mechanism and regularity of quenching the effect of bases on fluorophores: the base-quenched probe method. 2018. 143(14): p. 3292-3301.
20. Marimuthu, C., et al., Single-stranded DNA (ssDNA) production in DNA aptamer generation. *Analyst*, 2012. 137(6): p. 1307-1315.

REIVINDICACIONES

- 5 1. Una pluralidad de sondas que representan la variabilidad alélica de un determinado locus de repetición en tándem corta dentro de una población que consiste, en la que cada sonda consiste, de 5' a 3' o de 3' a 5', en: 1) una primera región flanqueante que comprende nucleótidos y que se hibrida con una región directamente próxima a la secuencia de ADN específica de interés y que contiene más nucleótidos que la segunda región flanqueante, 2) una secuencia de ADN específica de interés que comprende al menos una repetición en tándem corta y que se hibrida con la región de repetición en tándem corta dentro de la muestra, y 3) una segunda región flanqueante que comprende al menos un nucleótido y que contiene al menos un fluoróforo y en la que dicho fluoróforo está unido a un resto de dicha segunda región flanqueante en la posición directamente complementaria a un nucleótido específico capaz de extinguir dicho fluoróforo de manera eficiente de dicha muestra, tras la hibridación de dicha segunda región flanqueante dentro de la muestra, y en la que dicho nucleótido específico capaz de extinguir dicho fluoróforo de manera eficiente es guanosina, y en la que dicho fluoróforo se escoge de la lista que comprende fluoresceína (FAM), hexaclorofluoresceína (HEX), tetracloro-6-carboxifluoresceína (TET), 2,7-dimetoxi-4,5-dicloro-6-carboxifluoresceína (JOE) o 6-carboxitetrametilrodamina (TAMRA).
- 10 2. Una pluralidad de sondas según la reivindicación 1, en la que dicho fluoróforo está unido a un resto de citosina de dicha segunda región flanqueante.
- 15 3. Una pluralidad de sondas según las reivindicaciones 1 o 2, en la que dichos nucleótidos son análogos de ácidos nucleicos.
- 20 4. Una pluralidad de sondas según cualquiera de las reivindicaciones 1-3 que están inmovilizadas sobre un soporte.
- 25 5. Un método para genotipar repeticiones en tándem cortas dentro de una muestra, que comprende las etapas de:
 - proporcionar una muestra que comprende ADN,
 - amplificar el ADN dentro de dicha muestra que comprende una secuencia de ADN específica de interés para obtener secuencias de ADN amplificadas,
 - añadir una pluralidad de sondas según las reivindicaciones 1-4 a dichas secuencias de ADN amplificadas para obtener dúplex de secuencias de ADN monocatenario hibridadas a dicha sonda, y
 - desnaturalizar dichos dúplex, seguido del enfriamiento lento de dichos dúplex desnaturalizados mientras se mide continuamente la fluorescencia del fluoróforo de dicha sonda, o calentar lentamente dichos dúplex mientras se mide continuamente la fluorescencia del fluoróforo de dicha sonda, en el que la disminución de la intensidad de fluorescencia o el aumento de la intensidad de fluorescencia, respectivamente, proporciona información sobre si una repetición en tándem corta completamente complementaria específica está presente o no en dicha muestra.
- 30 6. Un método para genotipar según la reivindicación 5, en el que dicha amplificación de ADN dentro de dicha muestra se realiza mediante una PCR asimétrica para obtener secuencias de ADN monocatenario amplificadas.
- 35 7. Un método para genotipar según la reivindicación 5, en el que dicha amplificación de ADN dentro de dicha muestra se lleva a cabo mediante una PCR simétrica utilizando cebadores marcados con biotina o una digestión posterior con exonucleasa lambda para obtener secuencias de ADN monocatenario amplificadas.
- 40 8. Un método según las reivindicaciones 5-7, en el que dichas sondas se añaden en disolución o se inmovilizan sobre un soporte.
- 45

Figura 2

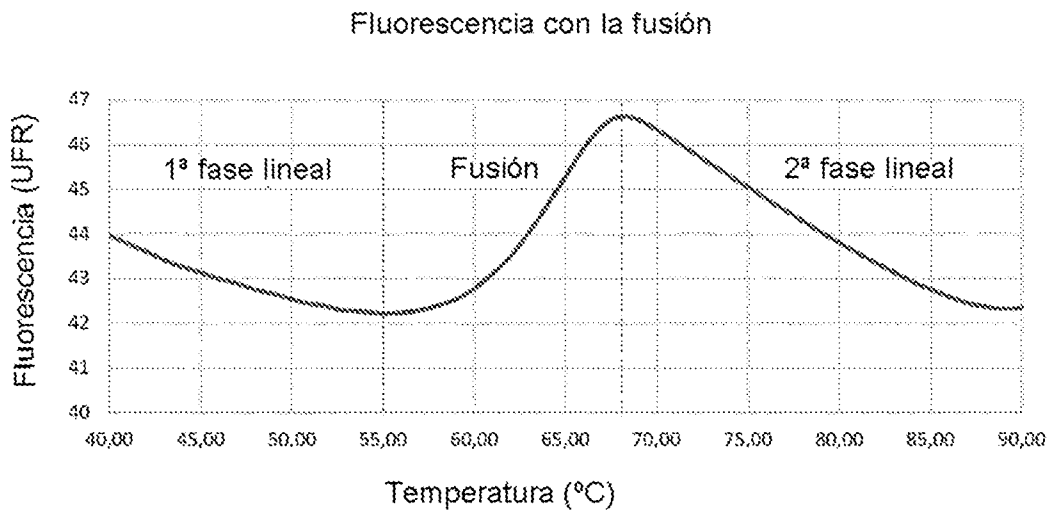


Figura 3

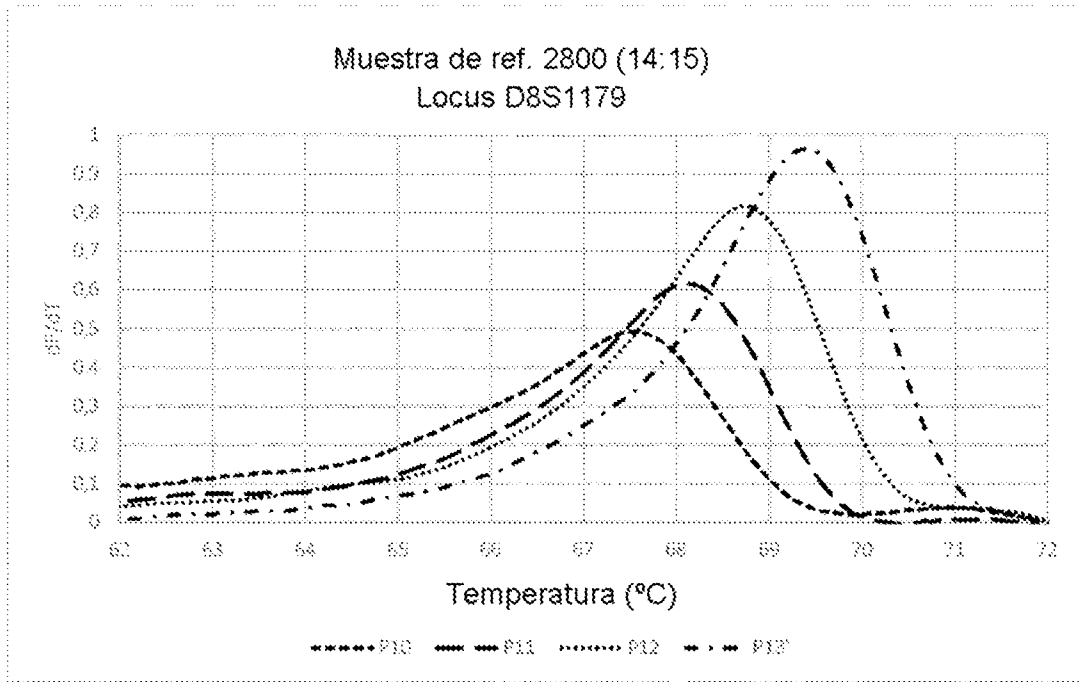


Figura 4:

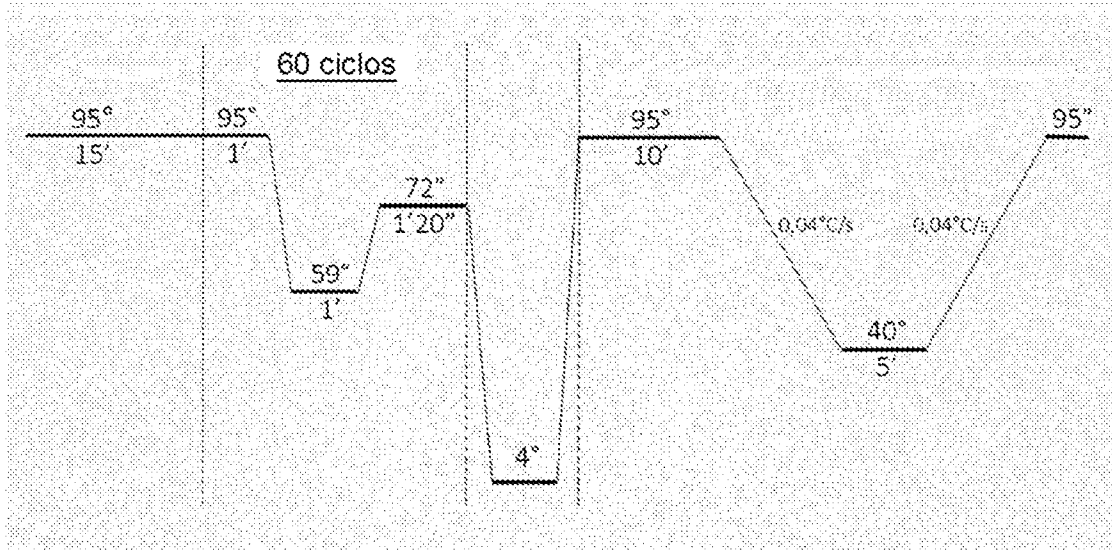


Figura 5:

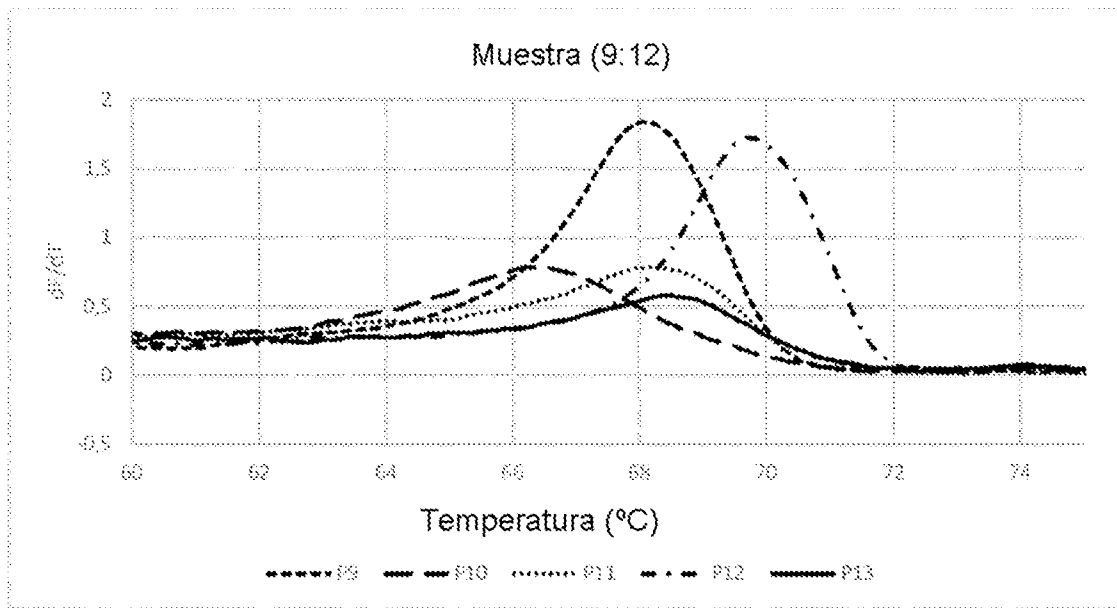


Figura 6:

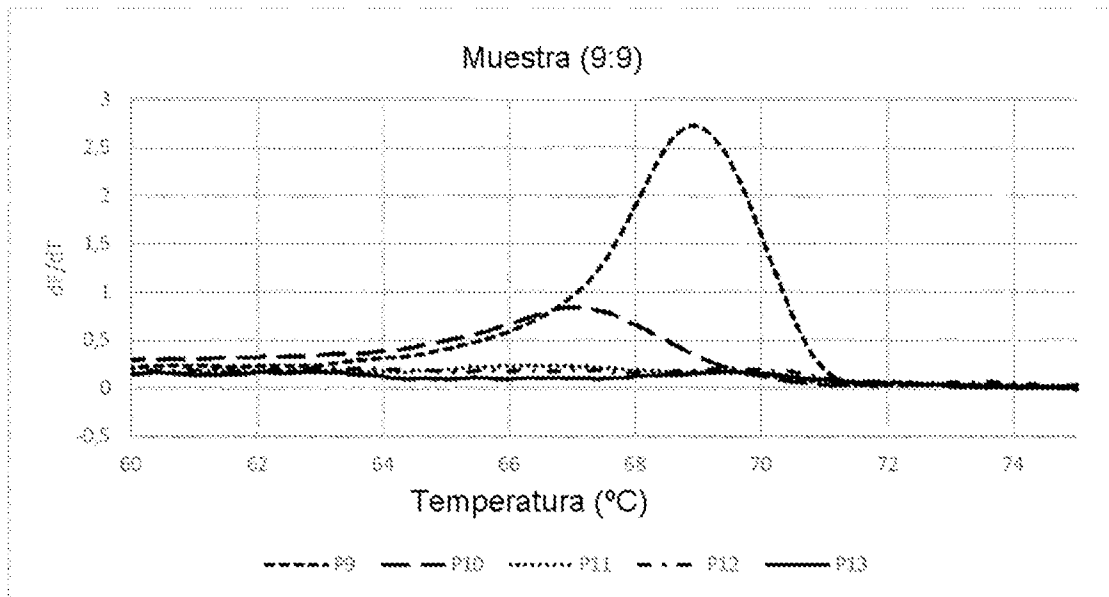


Figura 7:

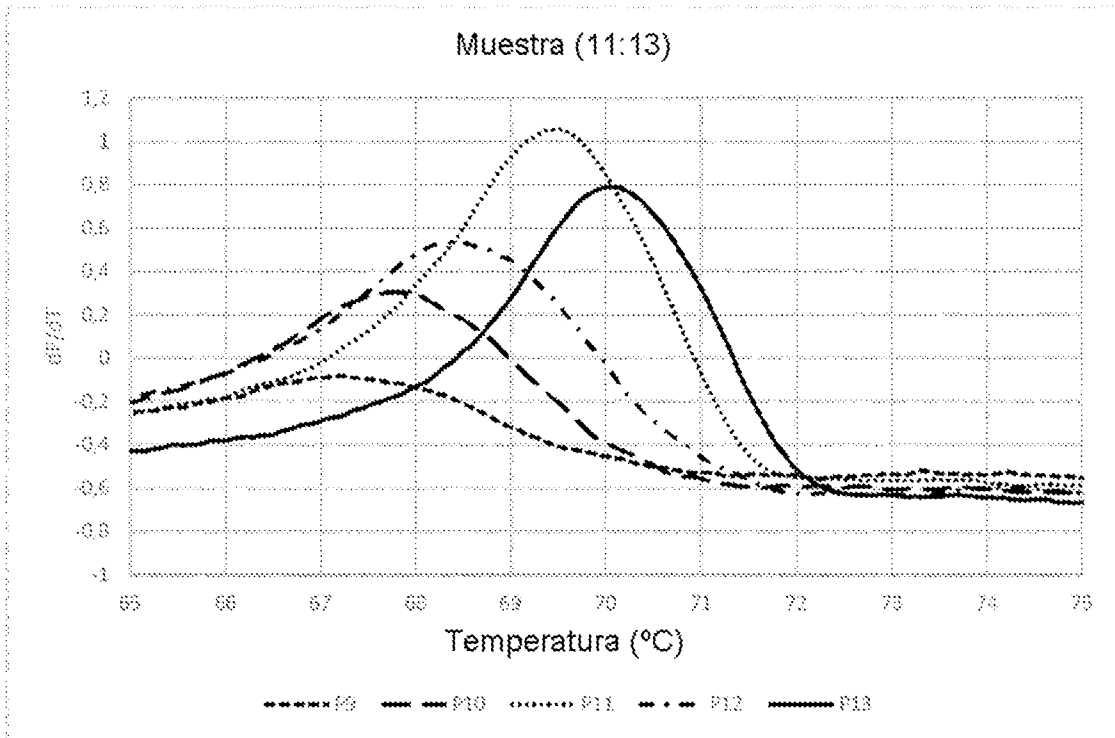


Figura 8:

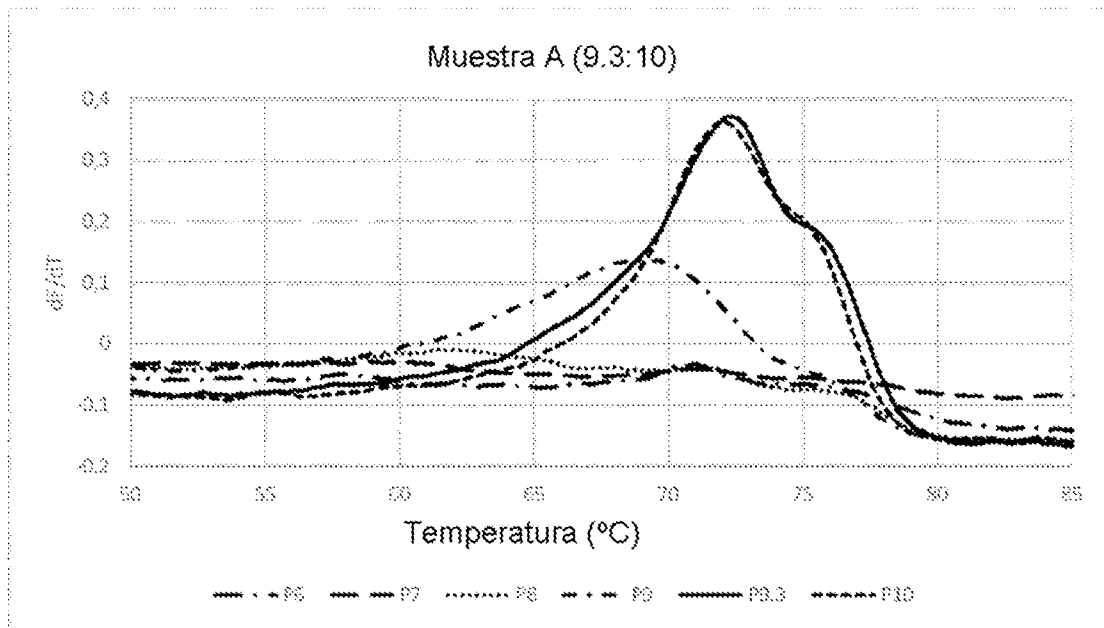


Figura 9:

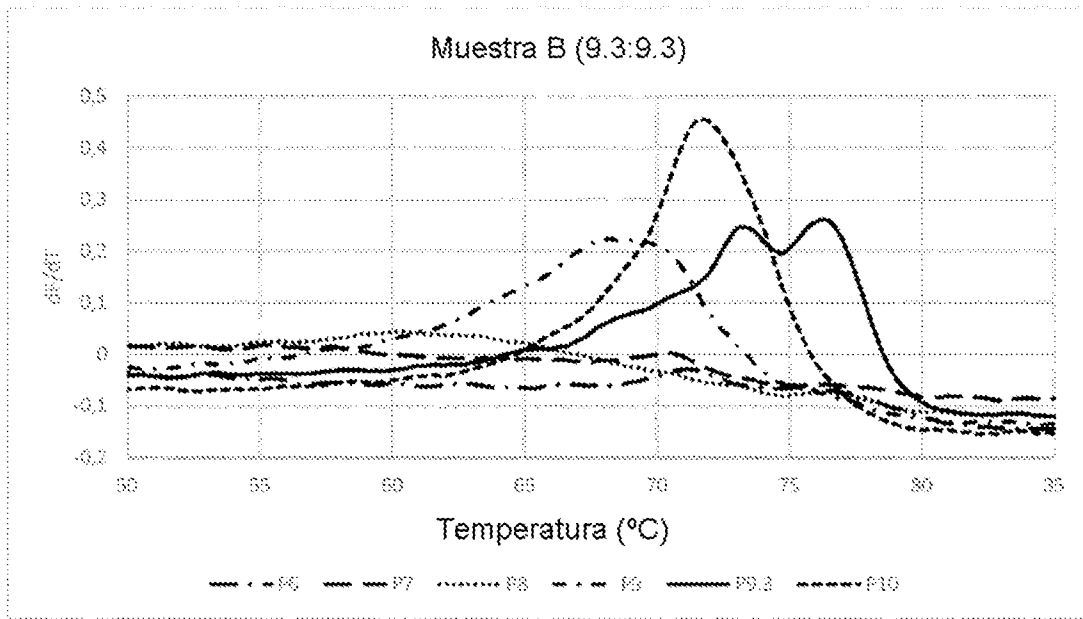


Figura 10:

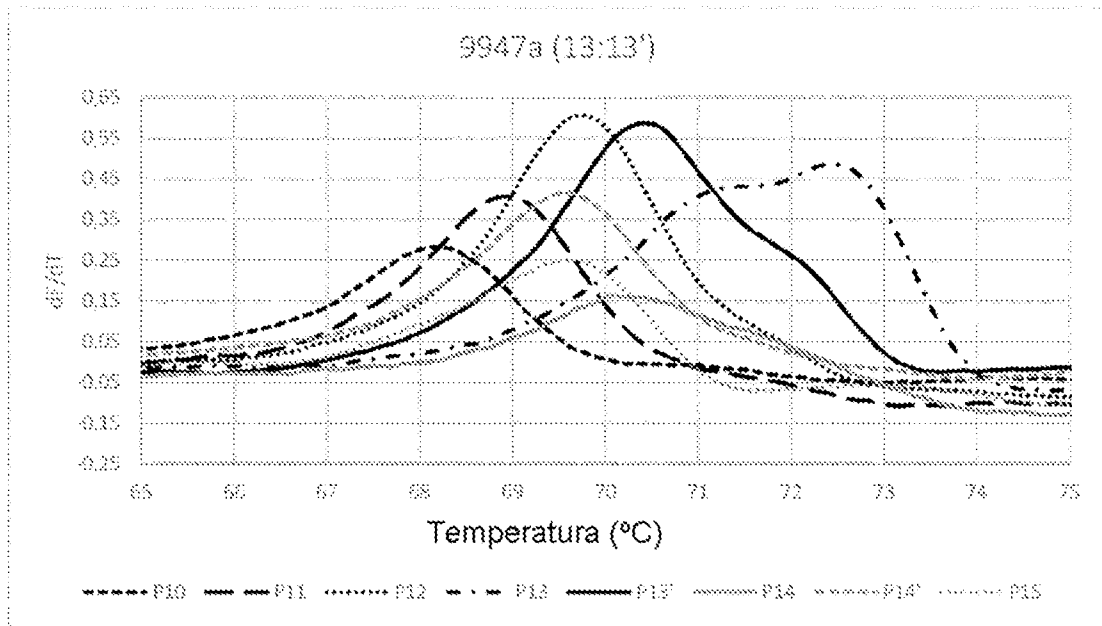


Figura 11:

