

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 705 423**

51 Int. Cl.:

**C12Q 1/68**

(2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.06.2015 PCT/GB2015/051605**

87 Fecha y número de publicación internacional: **10.12.2015 WO15185916**

96 Fecha de presentación y número de la solicitud europea: **02.06.2015 E 15728090 (0)**

97 Fecha y número de publicación de la concesión europea: **24.10.2018 EP 3149198**

54 Título: **Métodos para reducir el sesgo de GC dependiente de la densidad en la amplificación**

30 Prioridad:

**02.06.2014 GB 201409777**

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

**25.03.2019**

73 Titular/es:

**ILLUMINA CAMBRIDGE LIMITED (100.0%)  
Chesterford Research Park, Little Chesterford,  
Nr. Saffron Walden  
Essex CB10 1XL, GB**

72 Inventor/es:

**BOUTELL, JONATHAN MARK;  
SHANAHAN, SUSAN y  
RIGATTI, ROBERTO**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 705 423 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

Métodos para reducir el sesgo de GC dependiente de la densidad en la amplificación

**Campo de la invención**

5 La invención se refiere a métodos de amplificación de secuencias de polinucleótidos y, en particular, se refiere a métodos de amplificación de secuencias de polinucleótidos para minimizar sesgos específicos de secuencia. Los métodos de acuerdo con la presente invención son adecuados para la amplificación en fase sólida, por ejemplo, usando celdas de flujo.

**Antecedentes**

10 La reacción en cadena de la polimerasa o PCR (Saiki et al. (1985) *Science* 230:1350) se ha convertido en una técnica estándar de la biología molecular que contempla la amplificación de moléculas de ácidos nucleicos. Este método *in vitro* es una herramienta poderosa para la detección y el análisis de pequeñas cantidades de ácidos nucleicos y otras tecnologías de ácidos nucleicos recombinantes.

15 En resumen, la PCR usa típicamente una serie de componentes: una molécula de ácido nucleico diana, un exceso molar de un cebador directo e inverso que se une a la molécula de ácido nucleico diana, trifosfatos de desoxirribonucleósidos (dATP, dTTP, dCTP y dGTP) y una enzima polimerasa.

20 La reacción PCR es una reacción de síntesis de ADN que depende de la extensión de los cebadores directo e inverso hibridados a las cadenas opuestas de un molde de ADN bicatenario que se ha desnaturizado (las dos cadenas se han separado aplicando calor) a alta temperatura (de 90 °C a 100 °C). Usando repetidas etapas de desnaturización por calor, hibridación y extensión que generalmente se llevan a cabo a diferentes temperaturas, se generan copias del molde original de ADN.

25 La amplificación de secuencias molde por PCR típicamente se basa en el conocimiento de la secuencia molde que se va a amplificar, de modo que los cebadores pueden hibridarse específicamente al molde. El uso de múltiples pares de cebadores diferentes para amplificar simultáneamente diferentes regiones de la muestra se conoce como "PCR múltiple", y tiene numerosas limitaciones, que incluyen altos niveles de dimerización del cebador y la pérdida de representación de la muestra debido a las diferentes eficiencias de amplificación de las diversas regiones.

30 Para el análisis múltiple de grandes cantidades de fragmentos diana, a menudo es deseable llevar a cabo una reacción de amplificación simultánea para todas las dianas de la mezcla, usando un solo par de cebadores para todas las dianas. En ciertas realizaciones, se puede inmovilizar uno o más de los cebadores sobre un soporte sólido. Estas reacciones de amplificación universal se describen más detalladamente en la publicación de la solicitud de patente de EE. UU. N.º 2005/0100900. Los métodos de amplificación isotérmica para la amplificación de ácidos nucleicos se describen en la publicación de la solicitud de patente de EE. UU. N.º 2008/0009420. Los métodos involucrados pueden depender de la unión de regiones adaptadoras universales, que permiten la amplificación de todos los moldes de ácido nucleico a partir de un solo par de cebadores.

35 La reacción de amplificación universal todavía puede sufrir limitaciones en la eficiencia de amplificación relacionada con las secuencias de los moldes. Una manifestación de esta limitación es que la masa o el tamaño de los diferentes agrupamientos de ácidos nucleicos varía de una manera dependiente de la secuencia. Por ejemplo, los agrupamientos ricos en AT pueden adquirir más masa o volverse más grandes que los agrupamientos ricos en GC. Como resultado, el análisis de diferentes agrupamientos puede llevar a sesgos. Por ejemplo, en aplicaciones en las que los agrupamientos se analizan usando la secuenciación mediante técnicas de síntesis, los agrupamientos ricos en GC pueden parecer más pequeños o más tenues, de manera que los agrupamientos se detectan de forma menos eficiente. Esto da como resultado una menor representación de los datos de secuencia para los agrupamientos ricos en GC que los agrupamientos ricos en AT más brillantes (más intensos) y más grandes. Esto puede tener como resultado una representación más baja y una determinación de secuencia menos precisa para los moldes ricos en GC, un efecto que puede denominarse "sesgo de GC".

50 La presencia de sesgos específicos de secuencia durante la amplificación da lugar a dificultades para determinar la secuencia de ciertas regiones del genoma, por ejemplo, regiones ricas en GC tales como islas de CpG en regiones de promotores. La falta de representación de secuencia resultante en los datos de agrupamientos de diferentes composiciones de GC se traduce en problemas de análisis de datos tales como aumentos en el número de brechas en la secuencia analizada; un rendimiento de contigios más cortos, dando lugar a un ensamblaje *de novo* de menor calidad; identificación de los SNP con menos precisión en las regiones ricas en GC debido a la baja cobertura de estas regiones; y una necesidad de una mayor cobertura para secuenciar un genoma, lo que aumenta el coste de secuenciación de genomas.

55 El problema del sesgo puede agudizarse cuando la densidad de los agrupamientos en el soporte sólido es alta. En ciertas situaciones, a medida que los agrupamientos crecen, los cebadores de amplificación en el soporte sólido se extienden por completo, y por lo tanto, los agrupamientos adyacentes no pueden expandirse uno

encima del otro debido a la falta de cebadores de amplificación disponibles. La sobre-amplificación de secuencias ricas en AT causa un rápido consumo de los cebadores en la superficie, y por lo tanto, se reduce la capacidad de amplificación de las secuencias ricas en GC.

5 Los siguientes documentos pueden ser útiles para comprender la invención. El documento WO2011/106368 describe métodos para amplificar ácidos nucleicos. Los métodos se pueden usar para minimizar el sesgo específico de secuencia causado por la amplificación preferencial de ciertas secuencias del ácido nucleico. Se describen métodos que pueden disminuir la eficiencia de los moldes ricos en AT en relación con los moldes ricos en GC, minimizando así el sesgo de GC durante las reacciones de amplificación con múltiples moldes de diferentes secuencias. Los métodos son adecuados para la amplificación en fase sólida, por ejemplo, usando  
10 celdas de flujo.

15 El documento WO2012/106072 describe métodos de amplificación de ácidos nucleicos. Un método ilustrativo incluye las etapas de: (a) proporcionar una superficie que comprende una pluralidad de parches de cebadores, (b) proporcionar una pluralidad de diferentes moléculas de ácidos nucleicos, (c) poner en contacto la pluralidad de diferentes moléculas de ácidos nucleicos con la superficie en condiciones en las que las moléculas de ácido nucleico se unen a cebadores solo en un subconjunto de los parches, (d) amplificar las moléculas de ácidos nucleicos en condiciones de saturación del subconjunto de parches con copias de las moléculas de ácidos nucleicos, y repetir las etapas (c) y (d), aumentando así el número de parches que se saturan con copias de las moléculas de ácido nucleico.

20 El documento WO2007/107710 se refiere a un método para la amplificación isotérmica de una pluralidad de diferentes ácidos nucleicos diana, en la que los diferentes ácidos nucleicos diana se amplifican usando cebadores universales y se pueden distinguir entre sí las colonias producidas de este modo. Por lo tanto, el método genera colonias distintas de secuencias de ácidos nucleicos amplificadas que pueden analizarse por diversos medios para proporcionar información particular a cada colonia distinta.

25 El documento WO2014/043143 describe métodos para la amplificación de ácidos nucleicos, que comprenden formar una mezcla de reacción y someter la mezcla de reacción a condiciones adecuadas para la amplificación de los ácidos nucleicos.

30 Los métodos y composiciones presentados en la presente memoria tienen como objetivo reducir el sesgo de GC dependiente de la densidad en las reacciones de amplificación isotérmica de ácidos. La presente invención se refiere a un método y a un kit para reducir el sesgo de GC dependiente de la densidad y/o el daño del ácido nucleico en la amplificación en puente de un molde de ácido nucleico bicatenario en una superficie y también, a un sistema para una amplificación en puente como se reivindica con posterioridad. Se exponen realizaciones preferidas de la invención en las reivindicaciones dependientes.

También se describen métodos asociados en la presente memoria para ayudar a comprender la invención, pero estos no forman parte de la invención reivindicada.

35 Los ejemplos o realizaciones descritos en la presente memoria que no caen dentro de la definición de las reivindicaciones no forman parte de la presente invención.

Un aspecto de la invención proporciona un método para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en la amplificación en puente de un molde de ácido nucleico bicatenario en una superficie que comprende:

40 a. desnaturalizar el molde de ácido nucleico con una primera solución que comprende formamida y un solo tipo de dNTP y/o un solo tipo de rNTP, o una mezcla de diferentes dNTP y/o rNTP para producir cadenas de molde de ácido nucleico monocatenario;

b. opcionalmente reemplazar la primera solución por una segunda solución;

45 c. hibridar las cadenas del molde de ácido nucleico monocatenario con cebadores de oligonucleótidos unidos a la superficie; y

d. reemplazar la primera o la segunda solución con una solución que comprende una polimerasa, por lo que los cebadores de oligonucleótidos unidos a la superficie se extienden completamente;

en donde se reduce el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en la amplificación isotérmica del molde de ácido nucleico bicatenario.

50 Preferiblemente, al menos una de la primera y segunda solución comprende al menos un aditivo, opcionalmente en donde:

i) el aditivo comprende un agente quelante;

ii) el aditivo comprende al menos un citrato, opcionalmente en donde el aditivo comprende al menos un citrato seleccionado de: citrato monosódico, citrato disódico, citrato trisódico, citrato de potasio;

- iii) el aditivo comprende EDTA;
- iv) el aditivo comprende betaína;
- v) el aditivo comprende DMSO.

Preferiblemente, se genera un agrupamiento de cadenas de ácido nucleico idénticas.

5 El aditivo puede comprender una mezcla de diferentes dNTP y/o rNTP.

En una realización, la segunda solución comprende agua.

Preferiblemente, la segunda solución comprende menos de aproximadamente 100 mM de sal.

En una realización, la segunda solución está sustancialmente libre de cualquier sal.

En una realización, la segunda solución comprende una solución de premezcla.

10 Preferiblemente, la solución que comprende la polimerasa comprende menos de aproximadamente 100 mM de cualquier sal.

En una realización, la solución que comprende la polimerasa está sustancialmente libre de cualquier sal.

15 Otro aspecto de la invención proporciona un sistema para una amplificación en puente que comprende un aparato que tiene al menos una entrada y al menos una salida; y medios para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en una amplificación en puente de un molde de ácido nucleico bicatenario en una superficie en donde el aparato está programado para coordinar las etapas del método.

El aparato puede comprender medios para inmovilizar cebadores sobre una superficie.

20 Opcionalmente, el aparato comprende una celda de flujo y las soluciones se aplican a través de la entrada y se sacan a través de la salida mediante un procedimiento de intercambio de solución.

Opcionalmente, el sistema comprende medios de detección para detectar una señal fluorescente.

Otro aspecto más de la invención proporciona un kit para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en una amplificación en puente de un molde de ácido nucleico bicatenario en una superficie que comprende:

25 una primera solución desnaturalizante para producir cadenas de moldes de ácido nucleico monocatenario que comprenden formamida y al menos uno de un grupo que comprende: i) un solo tipo de dNTP, ii) un solo tipo de rNTP; iii) una mezcla de dNTP, y iv) una mezcla de rNTP; opcionalmente al menos un aditivo; y una polimerasa.

30 Preferiblemente, el al menos un aditivo comprende: un agente quelante, EDTA, un citrato, una sal de citrato, citrato monosódico, citrato disódico, citrato trisódico, citrato de potasio, betaína y/o DMSO.

El kit puede comprender cebadores y/o instrucciones para poner en práctica el método.

### Breve descripción

35 Se proporciona un método para reducir el sesgo de GC dependiente de la densidad en una amplificación isotérmica en puente de un molde de ADN bicatenario (dsDNA, por sus siglas en inglés) en una superficie de celda de flujo, que comprende: a) desnaturalizar el molde de ADN bicatenario con una primera solución que comprende un primer aditivo para producir cadenas del molde de ADN monocatenario (ssDNA, por sus siglas en inglés); b) reemplazar la primera solución por una segunda solución que puede comprender un segundo aditivo, por lo que las cadenas del molde de ADN monocatenario se hibridan a cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; y c) reemplazar la segunda solución por una tercera solución que

40 comprende una ADN polimerasa, por lo que los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo se extienden completamente; en donde se reduce el sesgo de GC dependiente de la densidad en la amplificación isotérmica del molde de ADN bicatenario. En una realización, se genera un agrupamiento de cadenas de ADN idénticas. En otra realización, la primera solución comprende formamida. En una realización, la primera solución comprende, además, EDTA. En una realización, la primera solución comprende formamida

45 junto con una alta concentración de betaína. En algunas realizaciones, el aditivo es una mezcla de diferentes dNTP o es un tipo único de dNTP. En una realización adicional, el aditivo es citrato trisódico. En otras realizaciones adicionales, el aditivo es una mezcla de diferentes dNTP. En algunas realizaciones, la segunda solución comprende, además, una solución de premezcla, en particular en donde la solución de premezcla

50 comprende una o más sales y uno o más tampones, más particularmente en donde la solución de premezcla comprende DMSO y/o betaína. En otra realización, la segunda solución comprende agua. Incluso en otra realización, la ADN polimerasa es una Bst ADN polimerasa. En otros ejemplos, una tercera solución

comprende, además, una mezcla de diferentes dNTP. En un ejemplo adicional más, una tercera solución comprende una o más sales y uno o más tampones, particularmente en donde la solución de premezcla comprende DMSO y/o betaína. En otro ejemplo, una tercera solución no comprende sulfato de amonio. En ejemplos adicionales, las etapas (a) a (c) se realizan dos o más veces, particularmente en donde las etapas (a) a (c) se realizan alrededor de 35 veces, y más particularmente en donde las etapas (a) a (c) se realizan a alrededor de 60 °C.

En algunos ejemplos, la temperatura a la que se realizan las etapas (a) a (c) puede ajustarse de acuerdo con la temperatura óptima para el tipo de ADN polimerasa usado.

También se describe un método para evaluar el efecto de los reactivos y/o aditivos de amplificación de agrupamientos sobre el daño en el ADN durante la amplificación isotérmica en puente de las cadenas del molde de ADN monocatenario en una superficie de la celda de flujo, que comprende: a) sembrar la superficie de la celda de flujo con las cadenas del molde de ADN monocatenario, por lo que las cadenas del molde de ADN monocatenario se hibridan a cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; b) extender los cebadores de oligonucleótidos usando una ADN polimerasa; c) desnaturalizar las cadenas del molde de ADN bicatenario, por lo que se producen moléculas de ADN monocatenario unidas de forma covalente a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; d) realizar un método de amplificación isotérmica simulada sobre las moléculas de ADN monocatenario unidas covalentemente a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo, que comprende ciclar reactivos y/o aditivos de amplificación de agrupamientos; e) realizar un método de amplificación isotérmica sobre las moléculas de ADN monocatenario covalentemente unidas a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; y f) visualizar agrupamientos de cadenas de ADN idénticas, en las que los agrupamientos de cadenas de ADN idénticas se tiñen con un colorante de unión a ADN y se toman imágenes; en donde el número de agrupamientos de cadenas de ADN idénticas es inversamente proporcional a la cantidad de daño en el ADN causado por los reactivos de amplificación de agrupamientos y/o aditivos. En un ejemplo, el método de amplificación isotérmica simulada comprende: i) agregar una primera solución que comprende un primer aditivo a las moléculas de ADN monocatenario unidas covalentemente a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; ii) reemplazar la primera solución por una segunda solución que comprende un segundo aditivo; y iii) reemplazar la segunda solución por una tercera solución que no comprende una ADN polimerasa. En un ejemplo alternativo, la etapa (d) puede reemplazarse por una etapa (d) alternativa de bombeo de reactivos de amplificación y/o aditivos sustancialmente de manera continua a las moléculas de ADN monocatenario unidas covalentemente a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo.

En un ejemplo, la primera solución, la segunda solución y/o la tercera solución usadas en la etapa (d) no comprenden aditivos.

En otro ejemplo, el método de amplificación isotérmica comprende: i) agregar una primera solución que comprende un primer aditivo a las moléculas de ADN monocatenario unidas covalentemente a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; ii) reemplazar la primera solución por una segunda solución que comprende un segundo aditivo; y iii) reemplazar la segunda solución por una tercera solución que comprende una ADN polimerasa.

En un ejemplo, la primera solución, la segunda solución y/o la tercera solución usadas en la etapa (e) de amplificación isotérmica no comprenden aditivos.

En algunos ejemplos, la primera solución comprende formamida. El primer aditivo puede comprender uno o más de los siguientes: un agente quelante, EDTA, una mezcla de dNTP, un solo tipo de dNTP, una mezcla de rNTP, un solo tipo de rNTP, un citrato, una sal de citrato, citrato de sodio, citrato disódico, citrato trisódico, citrato de potasio, betaína, DMSO.

En un ejemplo, el primer aditivo comprende EDTA. En algún ejemplo, el primer aditivo es una mezcla de diferentes dNTP o el primer aditivo es un tipo único de dNTP. En un ejemplo adicional, el primer aditivo es citrato trisódico. En otros ejemplos adicionales, el segundo aditivo es una mezcla de diferentes dNTP. En algunos ejemplos, la segunda solución comprende una solución de premezcla, en particular en donde la solución de premezcla comprende una o más sales y uno o más tampones, más particularmente en donde la solución de premezcla comprende DMSO y/o betaína. En otro ejemplo, la segunda solución comprende agua.

El segundo aditivo puede comprender uno o más de los siguientes: un agente quelante, EDTA, una mezcla de dNTP, un solo tipo de dNTP, un citrato, una sal de citrato, citrato monosódico, citrato disódico, citrato trisódico, citrato de potasio, betaína, DMSO.

Incluso en otro ejemplo, la ADN polimerasa es una Bst ADN polimerasa. En otros ejemplos, la tercera solución comprende, además, una mezcla de diferentes dNTP. En un ejemplo adicional más, la tercera solución comprende una o más sales y uno o más tampones, particularmente en donde la solución de premezcla comprende DMSO y/o betaína. En otro ejemplo, la tercera solución no comprende sulfato de amonio ni otra sal. En ejemplos adicionales, las etapas (a) a (d) se realizan dos o más veces, particularmente en donde las etapas (a) a (d) se realizan alrededor de 26 veces, más particularmente en donde las etapas (a) a (d) se realizan

alrededor de 35 veces o menos, y aún más particularmente en donde las etapas (a) a (d) se realizan a alrededor de 60 °C.

En algunos ejemplos, la temperatura a la que se realizan las etapas (a) a (d) puede ajustarse de acuerdo con la temperatura óptima para el tipo de ADN polimerasa usado en el procedimiento de amplificación.

5 También se describe un método para reducir el daño del ADN en una amplificación isotérmica en puente de un molde de ADN bicatenario en una superficie de la celda de flujo, que comprende: a) desnaturalizar el molde de ADN bicatenario con una primera solución que comprende formamida para producir cadenas del molde de ADN monocatenario (ssDNA); b) reemplazar la primera solución por una segunda solución que comprende los dNTP, por lo que las cadenas del molde del ADN monocatenario se hibridan a cebadores de oligonucleótidos unidos a la superficie de la celda de flujo; y c) reemplazar la segunda solución por una tercera solución que comprende una ADN polimerasa, por lo que los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo se extienden completamente; en donde se reduce el daño del ADN en la amplificación isotérmica del molde de ADN bicatenario.

15 La Figura 1 ilustra un diagrama de flujo de un ejemplo de un protocolo de amplificación isotérmica en puente para la amplificación de agrupamientos sobre una superficie de celda de flujo;

Las Figuras 2A y 2B muestran un gráfico del tamaño de agrupamiento y un gráfico de la intensidad de agrupamiento, respectivamente, de los agrupamientos generados a partir de monomoldes con un inserto de 400 pb con diferente composición de GC;

20 La Figura 3 muestra una captura de pantalla de una imagen de agrupamientos que muestra agrupamientos ricos en GC y agrupamientos ricos en AT, en donde se sembraron cuatro monomoldes a una densidad relativamente baja;

La Figura 4 ilustra un diagrama de flujo de un ejemplo de un ensayo de daños en el ADN para evaluar el efecto de los reactivos de amplificación de agrupamientos y/o la adición de aditivos en la generación de agrupamientos durante la amplificación isotérmica en puente;

25 La Figura 5 muestra gráficamente las etapas del ensayo de daño en el ADN de la Figura 4;

La Figura 6 muestra un gráfico de barras del porcentaje de moléculas de ADN no dañadas en muestras tratadas con formamida, premezcla y/o premezcla más dNTP, en comparación con un carril tratado solo con tampón de lavado y se evaluó usando el ensayo de daños en el ADN de la Figura 4;

30 La Figura 7 muestra un gráfico de barras del porcentaje de moléculas de ADN no dañadas en muestras tratadas con formamida ± dNTP, premezcla ± dNTP o agua ± dNTP y se evaluó usando el ensayo de daño en el ADN de la Figura 4;

La Figura 8 muestra una curva de respuesta a la dosis del efecto de la concentración de dNTP en formamida en el número de agrupamientos durante la amplificación isotérmica simulada;

35 La Figura 9 muestra un gráfico de barras de la efectividad de cada dNTP para evitar el daño en el ADN cuando se usa como aditivo formamida;

La Figura 10 muestra paneles de agrupamientos teñidos con SYBR® Green generados por amplificación isotérmica usando formamida con y sin dNTP agregado;

La Figura 11 muestra una tabla de datos de métrica de secuenciación para una ejecución de secuenciación que evalúa el "método de los dNTP" y el "método del agua";

40 Las Figuras 12A y 12B muestran curvas de sesgo de GC y curvas de sesgo de GC normalizadas al carril de baja densidad 1, respectivamente, de la ejecución de secuencia descrita con referencia a la Figura 11;

La Figura 13A muestra un diagrama esquemático resumido de soluciones de amplificación en puente de acuerdo con los métodos de la invención; y

45 La Figura 13B muestra un diagrama esquemático resumido de las soluciones de amplificación en puente según un método estándar.

50 La Figura 14 muestra el efecto de varios aditivos de formamida sobre el daño en el ADN ocasionado por los reactivos de amplificación isotérmica de agrupamientos (FC61WEBAAXX, exp. 12374). Se sembró CT180 a 1 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Las moléculas individuales se trataron luego con 26 ciclos de amplificación isotérmica simulada (28 µl de formamida, 28 µl de H<sub>2</sub>O, 36 µl de premezcla con dNTP). El carril de control se trató con tampón de lavado para toda la duración. TL = tampón de lavado. dNTP = 200 µM de cada nucleótido. AMP = monofosfato de adenosina a una concentración final de 800 µM. Betaína 2 M en formamida = formamida al 60 % que contiene betaína a una concentración final de 2 M y agua. Betaína 1 M en formamida = formamida al 80 %

que contiene betaina a una concentración final de 1 M y agua. Betaina 0,5 M en formamida = formamida al 90 % que contiene betaina a una concentración final de 0,5 M y agua.

Figura 15. Efecto protector contra el daño en el ADN de cada componente individual del tampón de lavado (FC62263AAXX exp 12402). Se sembró CT180 a 0,6 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Las moléculas individuales se trataron luego con 26 ciclos de amplificación isotérmica simulada (28 µl de formamida con varios aditivos, 28 µl de H<sub>2</sub>O, 36 µl de premezcla con dNTP). El carril 1 actuó como control y se trató con tampón de lavado durante todo el tratamiento. La segunda solución y la tercera solución fueron 28 µl de H<sub>2</sub>O y 36 µl de premezcla/dNTP, respectivamente, para todos los carriles, excepto para el carril 1 en el que se bombeó el tampón de lavado. TL = tampón de lavado. dNTP = 200 µM de cada nucleótido. Se agregó cada aditivo en una concentración que es idéntica a la que se alcanzaría al agregar 10 % de tampón de lavado (NaCl 4,5 mM, citrato de Na 0,45 mM, Tween 20 al 0,01 %).

Figura 16. Efecto de la concentración del citrato de sodio en formamida sobre el daño en el ADN (FC62272AAXX, exp 12411). Se sembró CT180 a 0,6 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Las moléculas individuales se trataron luego con 26 ciclos de amplificación isotérmica simulada (28 µl de formamida, 28 µl de H<sub>2</sub>O y 36 µl de premezcla con dNTP). Se usó el carril 1 como control y se lo trató bombeando el tampón de lavado durante todo el tratamiento. Al final del tratamiento, se llevó a cabo la amplificación del agrupamiento usando 26 ciclos de amplificación isotérmica (28 µl de formamida + 64 µl de mezcla Bst). Luego, los agrupamientos se tiñeron con SYBR® Green y se tomaron imágenes de tres placas por carril usando un microscopio y una cámara. El número de agrupamientos se determinó usando un programa informático llamado Firecrest.

Figura 17. Efecto de la presencia en formamida de la molécula quelante de EDTA y del ión magnesio sobre el daño en el ADN (FC6254LAAXX, exp 12473). El diseño del experimento se muestra en (A). Se sembró CT180 a 1 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Las moléculas individuales se trataron luego con 26 ciclos de amplificación isotérmica simulada (28 µl de formamida, 28 µl de H<sub>2</sub>O, 36 µl de premezcla con dNTP excepto el carril 1, que se trató con tampón de lavado para toda la duración). La premezcla fue ya sea estándar, que contenía sulfato de amonio 10 mM, o una premezcla que no contenía sulfato de amonio. Se añadió EDTA (0,5 M, pH 8,0) en formamida hasta una concentración final de 1 mM. Los dNTP se agregaron a la formamida a una concentración final de 200 µM de cada nucleótido (concentración total de 800 µM). En el carril 7, se añadió sulfato de Mg a una concentración final de 3 mM a los dNTP que contenían solución de formamida. Después de la amplificación en puente, los agrupamientos se tiñeron con SYBR® Green y se tomaron imágenes de tres placas por carril usando un microscopio y una cámara. El número de agrupamientos se determinó usando un programa informático llamado Firecrest. El número de agrupamientos se representa como un porcentaje del carril 1 (carril de control tratado con tampón de lavado).

Figura 18. Efecto del aumento de la concentración de sulfato de amonio en la mezcla Bst sobre el sesgo de GC dependiente de la densidad. Resumen de ejecución (A) y curvas de sesgo de GC (B y C) a partir del experimento 100526\_EAS20\_0170\_FC61LRUAAXX. Se sembró CT3576 (genoteca humana estándar, tamaño promedio del inserto de 300 pb) a baja densidad (1 pM) o bien, a alta densidad (5,5 pM). Después de la primera extensión de la cadena estándar con Phusion, se amplificó el agrupamiento con 26 ciclos de amplificación isotérmica usando 28 µl de formamida seguido de 28 µl de agua y 36 µl de la mezcla Bst con diferentes concentraciones de sulfato de amonio (que varía entre la concentración estándar de 10 mM y hasta 80 mM). Las curvas de sesgo en C se obtuvieron a partir de las curvas en (B) dividiendo cada curva por la curva de carril de baja densidad (carril 1) como una forma de enfocarse exclusivamente en el componente dependiente de la densidad del sesgo de GC creado durante la amplificación de grupo.

Figura 19. Efecto combinado de los dNTP en formamida y menores concentraciones de sulfato de amonio en la mezcla Bst sobre el sesgo de GC dependiente de la densidad. El diseño del experimento se muestra en (A). Después de sembrar una genoteca humana estándar (CT3576) a alta y baja densidad y la primera extensión de la cadena con Phusion, se amplificaron los agrupamientos usando 26 ciclos de amplificación isotérmica con las condiciones que se muestran en la tabla. Después de la preparación de la lectura 1, se secuenció la celda de flujo con un instrumento GAIIX. Esta ejecución en particular fue un experimento final emparejado usando la secuenciación química 95G de 36 bases en cada lectura. El análisis se llevó a cabo con el programa informático Pipeline 1.6. (B) muestra el resumen de la lectura 1, mientras que en C se muestran las curvas de sesgo de GC normalizadas (obtenidas al dividir las curvas de sesgo de GC originales por la curva de baja densidad (carril 1)).

Figura 20. Combinación de citrato de sodio con una mezcla Bst sin amonio y ciclos para completar. El diseño del experimento se muestra en (A). Después de sembrar la genoteca humana estándar (CT3576) a alta y baja densidad y la primera extensión de la cadena con Phusion, se amplificaron los agrupamientos usando 26 o 32 ciclos de amplificación isotérmica con las condiciones que se muestran en la tabla. Después de la preparación de la lectura 1, se secuenció la celda de flujo con un instrumento GAIIX. Este experimento en particular fue una ejecución final pareada usando la secuenciación química 95G de 36 bases en cada lectura. El análisis se llevó a cabo con el programa informático Pipeline 1.8. En (B), se muestra el resumen de la

lectura 1, y las curvas de sesgo de GC normalizadas (obtenidas al dividir las curvas de sesgo de GC originales por la curva en el carril 1 de baja densidad) se muestran en C. Una comparación directa en términos de sesgo de GC dependiente de la densidad de los dNTP en formamida y agua y citrato de sodio en formamida, ambos con una mezcla Bst sin amonio, se muestra en D.

5 Figura 21. Efecto de la concentración de citrato de sodio y betaína en formamida. El diseño del experimento se muestra en (A). Después de la siembra de una genoteca humana estándar (CT3576) a alta y baja densidad y la primera extensión de la cadena con Phusion, se amplificaron los agrupamientos usando 35 ciclos (con reactivos estándar, carriles 1 y 2) o 32 ciclos de isotérmicos (con una mezcla Bst sin amonio, carriles 3 a 8).  
 10 Después de la preparación de la lectura 1, se secuenció la celda de flujo con un instrumento GAIIX. Este experimento en particular fue una ejecución de lectura única usando la secuenciación química 95G (36 ciclos de SBS). El análisis se llevó a cabo con el programa informático Pipeline 1.8. (B) muestra el resumen de la lectura 1, mientras que las curvas de sesgo de GC normalizadas (obtenidas al dividir las curvas de sesgo de GC originales por la curva en el carril 1 de baja densidad) se muestran en C.

### Definiciones

15 Como se usa en la presente memoria, los siguientes términos tienen los significados indicados.

Las expresiones "normalizar" o "reducir el sesgo específico de secuencia" (por ejemplo, reducir el sesgo de GC dependiente de la densidad) cuando se usan en referencia a la amplificación de los moldes de ácido nucleico, significan alterar la proporción de moléculas de diferente tipo obtenidas durante un procedimiento de amplificación de manera que el número de moléculas de un tipo particular en la población se asemeja más al número de moléculas de otro tipo en la población. Por lo tanto, para una reacción de amplificación llevada a  
 20 cabo en una población de moldes de ácido nucleico de secuencia diferente, normalizar la amplificación puede significar disminuir cualquier sesgo específico de secuencia que de otro modo, daría lugar a que ciertos miembros de la población aumenten en número más que otros miembros de la población. El procedimiento de normalización se puede usar para producir proporciones relativas de los fragmentos en la población que son las mismas después de la amplificación que en la población antes de la amplificación. Así, por ejemplo, una población que comprende 1 millón de moléculas de secuencia diferente contendrá, después de la amplificación, en promedio, el mismo número de copias de cada uno del millón de fragmentos sin ningún sesgo específico para ciertas secuencias. Se entenderá que esta es una medida estadística y que la ausencia de sesgo puede estar dentro de una variación aceptable tal como dentro de 0,5, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45 o 50  
 25 % de variación en el número de copias para cada fragmento. Cuando se lleva a cabo en un soporte sólido para crear agrupamientos de ácido nucleico, la normalización de la amplificación da lugar a una serie de agrupamientos con un número similar de moléculas en cada agrupamiento, y por lo tanto, tamaños e intensidades de señal similares.

El término "diferente" cuando se usa en referencia a dos o más ácidos nucleicos significa que los dos o más ácidos nucleicos tienen secuencias de nucleótidos que no son las mismas. Por ejemplo, dos ácidos nucleicos pueden diferir debido a que una secuencia es más larga que la otra y, a la inversa, una secuencia es más corta que la otra. Dos ácidos nucleicos pueden diferir en el contenido y en el orden de los nucleótidos en la secuencia de un ácido nucleico en comparación con el otro ácido nucleico, independientemente de cualquier diferencia en la longitud de la secuencia entre los dos ácidos nucleicos. El término se puede usar para describir ácidos nucleicos, ya sea que se denominen copias, amplicones, moldes, dianas, cebadores, oligonucleótidos, polinucleótidos o similares.  
 35

Como se describe en la presente memoria, los moldes de ácidos nucleicos que contienen un alto nivel de bases A y T típicamente se amplifican con mayor eficiencia que los moldes de ácido nucleico con un alto nivel de bases G y C. Los moldes de ácido nucleico con secuencias que contienen un alto nivel de bases A o T en comparación con el nivel de las bases G o C se conocen como moldes ricos en AT o moldes con alto contenido de AT. Por consiguiente, los moldes ricos en AT pueden tener niveles relativamente altos de bases A, bases T o bases A y T. De manera similar, los moldes de ácido nucleico con secuencias que contienen un alto nivel de bases G o C en comparación con el nivel de bases A o T se denominan en toda la memoria como moldes ricos en GC o moldes con un alto contenido de GC. Por consiguiente, los moldes ricos en GC pueden tener niveles relativamente altos de bases G, bases C o bases tanto G como C. Las expresiones "rico en GC" y "con alto contenido de GC" se usan indistintamente. De manera similar, las expresiones "rico en AT" y "con alto contenido de AT" se usan indistintamente. Las frases "rico en GC" y "rico en AT", como se usan en la presente memoria, se refieren a una secuencia de ácido nucleico que tiene un número relativamente alto de bases G y/o C o bases A y/o T, respectivamente, en su secuencia o en una parte o región de su secuencia, en relación con el contenido de la secuencia dentro de un control. En este caso, el control puede ser secuencias de ácidos nucleicos, genes o genomas similares a partir de los cuales se originan las secuencias de ácidos nucleicos. En general, las secuencias de ácido nucleico que tienen un contenido de GC o AT superior a aproximadamente el 52 % se consideran secuencias ricas en GC o ricas en AT. Opcionalmente, el contenido de GC o AT es mayor que 55, 60, 65, 70, 75, 80, 85, 90, 95 o 99 %. Las secuencias de ácido nucleico que contienen regiones discretas de alto contenido de GC a AT también pueden considerarse ricas en GC o ricas en AT, respectivamente. Los métodos proporcionados en la presente memoria normalizan las eficiencias o los niveles de amplificación de moldes con diferentes secuencias, por ejemplo, con alto contenido de AT y/o GC.  
 45  
 50  
 55  
 60

La expresión "ciclo de amplificación" se refiere a una o más etapas de un procedimiento de amplificación que son suficientes para producir una o más copias de un molde de ácido nucleico. A modo de ejemplo, un ciclo de amplificación incluye proporcionar uno o más moldes de ácido nucleico, desnaturalizar los moldes de ácido nucleico para producir moldes de ácido nucleico monocatenario, hibridar uno o más cebadores a los moldes de ácido nucleico monocatenario y extender los cebadores para producir copias de los moldes de ácido nucleico. Como se describe en la presente memoria, dichos ciclos pueden repetirse una o más veces en condiciones que favorecen a los moldes ricos en AT o GC. Por lo tanto, un ciclo de amplificación puede incluir una unidad de una o más etapas que se repiten en una ronda de amplificación.

Tal como se usa en toda la memoria, la frase "que favorece los moldes ricos en AT" significa que la eficiencia de la amplificación de los moldes ricos en AT no se reduce ni se inhibe en relación con los moldes que no son ricos en AT. A modo de ejemplo, bajo condiciones de amplificación estándar, los moldes ricos en AT se amplifican con una mayor eficiencia que los moldes ricos en GC. Por lo tanto, las condiciones que favorecen los moldes ricos en AT incluyen condiciones de amplificación estándar. Tal como se usa en toda la memoria, la frase "condiciones de amplificación estándar" significa amplificar una secuencia de ácido nucleico en condiciones que incluyen todos los reactivos y condiciones estándar necesarios para llevar a cabo la amplificación. Las condiciones de amplificación estándar se conocen y se describen, por ejemplo, en Saiki et al., *Science*, 230:1350 (1985).

Como se usa en la presente memoria, la frase "que favorece los moldes ricos en GC" significa que la eficiencia de la amplificación de los moldes ricos en GC aumenta en relación con los moldes ricos en AT y/o la eficiencia de la amplificación de los moldes ricos en AT se reduce en relación con los moldes ricos en GC.

Los nucleótidos usados en el procedimiento de amplificación pueden ser ribonucleótidos o desoxirribonucleótidos. Los nucleótidos usados en la amplificación pueden ser nucleótidos 5' polifosfatos, por ejemplo 5' trifosfatos. Los nucleótidos usados en la reacción de amplificación pueden ser los cuatro nucleótidos trifosfatos que se encuentran típicamente en el ADN nativo: dATP, dGTP, dCTP y dTTP.

Tal como se usa en la presente memoria, las expresiones "alto", "más alto", "aumento/s", "aumentado" o "en aumento" se refieren a cualquier aumento sobre una referencia o control, a menos que se indique lo contrario. Las expresiones "bajo", "más bajo", "disminución/es", "disminuido", "en disminución", "reducir", "reducido", "en reducción" o "reducción" se refieren a cualquier disminución por debajo de una referencia o control, a menos que se indique lo contrario. A modo de ejemplo, un control incluye valores de control o niveles de control, que pueden ser valores o niveles anteriores a o en ausencia de un estímulo. Un valor de control o control incluye el nivel de eficiencia de la amplificación de secuencias de ácido nucleico en condiciones de amplificación estándar o puede comprender un valor, nivel o estándar conocido. Así, por ejemplo, un valor más alto o más bajo (por ejemplo, temperatura o concentración) en comparación con un control, se refiere a un valor que es más alto o más bajo que un valor conocido o establecido arbitrariamente.

El término "isotérmico" se refiere a procedimientos termodinámicos en los que la temperatura de un sistema permanece constante:  $DT = 0$ . Esto ocurre típicamente cuando un sistema está en contacto con un reservorio térmico exterior (por ejemplo, baños de calor y similares), y los procedimientos se producen con la suficiente lentitud como para permitir que el sistema se ajuste continuamente a la temperatura del reservorio a través del intercambio de calor.

La expresión "sustancialmente isotérmico" como se usa en la presente memoria pretende significar que el sistema se mantiene esencialmente a la misma temperatura. La expresión también pretende capturar pequeñas desviaciones en la temperatura que pueden ocurrir cuando el sistema se equilibra, por ejemplo, cuando se agregan al sistema componentes que tienen una temperatura más alta o más baja. Por lo tanto, se pretende que el término incluya desviaciones menores de la temperatura elegida inicialmente para realizar el método y aquellas en el intervalo de desviación de los termostatos comerciales. En particular, la desviación de la temperatura no será más de aproximadamente  $\pm 2$  °C, más particularmente no más de aproximadamente  $\pm 1$  °C, pero más particularmente no más de aproximadamente  $\pm 0,5$  °C, no más de aproximadamente  $\pm 0,25$  °C, no más de aproximadamente  $\pm 0,1$  °C, o no más de aproximadamente  $\pm 0,01$  °C.

El término "amplificación" como se usa en la presente memoria pretende significar el procedimiento para aumentar los números de una secuencia de polinucleótidos de molde que produce una o más copias. Por consiguiente, quedará claro que el procedimiento de amplificación puede ser exponencial o lineal. En la amplificación exponencial, el número de copias realizadas de la secuencia de polinucleótidos del molde aumenta a una velocidad exponencial. Por ejemplo, en una reacción de PCR ideal con 30 ciclos, 2 copias de ADN del molde producirán  $2^{30}$  o 1073741824 copias. En la amplificación lineal, el número de copias realizadas de las secuencias de polinucleótidos del molde aumenta a una velocidad lineal. Por ejemplo, en una reacción de amplificación lineal ideal de 4 horas cuya velocidad de copia es de 2000 copias por minuto, una molécula de ADN del molde producirá 480000 copias.

El término "copia" cuando se usa en referencia a una primera molécula de ácido nucleico, pretende significar una segunda molécula de ácido nucleico que tiene la misma secuencia que el primer ácido nucleico o la secuencia complementaria del ácido nucleico. Los ácidos nucleicos pueden ser monocatenarios o bicatenarios. Por ejemplo, una copia monocatenaria puede tener la misma secuencia de un molde monocatenario, una copia

monocatenaria puede tener la secuencia complementaria de un molde monocatenario, una copia bicatenaria puede incluir la misma secuencia y la secuencia complementaria (es decir, dos cadenas) de un molde monocatenario, o una copia bicatenaria puede incluir las mismas secuencias que un molde bicatenario. De manera similar, el término "copia" cuando se usa en referencia a una secuencia de ácido nucleico, significa la misma secuencia o la secuencia complementaria.

Como se usa en la presente memoria, las expresiones "polinucleótido", "oligonucleótido" o "ácido nucleico" pueden referirse a ácido desoxirribonucleico (ADN), ácido ribonucleico (ARN) o análogos de ADN o ARN hechos, por ejemplo, a partir de análogos de nucleótidos. Las expresiones "polinucleótido", "oligonucleótido" o "ácido nucleico" son aplicables a moléculas monocatenarias (tales como sentido o antisentido) y bicatenarias. Las expresiones "polinucleótido", "oligonucleótido" o "ácido nucleico" como se usan en la presente memoria, también abarcan el ADNc, que es un ADN complementario o de copia producido a partir de un molde de ARN, por ejemplo, por la acción de la transcriptasa inversa.

Las moléculas de polinucleótidos monocatenarios útiles en un método o composición de la invención pueden haberse originado en forma monocatenaria, como ADN o ARN, o pueden haberse originado en la forma de ADN bicatenario (dsDNA) (por ejemplo, fragmentos de ADN genómico, PCR y productos de amplificación y similares). Por lo tanto, un polinucleótido monocatenario puede ser la cadena sentido o antisentido de una doble hélice de polinucleótidos. Los métodos de preparación de moléculas de polinucleótidos monocatenarios adecuados para su uso en el método de la invención usando técnicas estándar son bien conocidos en la técnica.

El término "inmovilizado" o "unido" como se usa en la presente memoria, pretende abarcar la unión directa o indirecta, covalente o no covalente, a menos que se indique lo contrario, ya sea explícitamente o por contexto. En ciertas realizaciones de la invención, se puede preferir la unión covalente, pero en general, todo lo que se requiere es que las moléculas (por ejemplo, los ácidos nucleicos) permanezcan inmovilizadas o unidas a un soporte en condiciones en las que se pretende usar el soporte, por ejemplo en aplicaciones que requieren amplificación y/o secuenciación de ácidos nucleicos.

Los cebadores de amplificación para la amplificación en fase sólida se inmovilizan en la presente memoria mediante la unión covalente a un soporte sólido en o cerca del extremo 5' del cebador, dejando libre la porción de molde específica del cebador para hibridarse a su molde análogo y el grupo hidroxilo 3' libre para funcionar en la extensión del cebador. La química de la unión elegida dependerá de la naturaleza del soporte sólido y de cualquier funcionalización o derivación que se le aplique. El cebador en sí puede incluir un resto, que puede ser una modificación química no nucleótida para facilitar la unión. En particular, el cebador puede incluir un nucleófilo que contiene azufre tal como fosfotioato o tiofosfato en el extremo 5'. En el caso de hidrogeles de poli(acrilamida con soporte sólido, este nucleófilo puede unirse a un grupo bromoacetamida presente en el hidrogel. Preferiblemente, el medio para unir los cebadores al soporte sólido es a través de la unión del fosfotioato 5' a un hidrogel compuesto por acrilamida polimerizada y N-(5-bromoacetamidilpentil) acrilamida (BRAPA). Tal disposición se describe más detalladamente en la solicitud WO 05/065814 en tramitación con la presente.

Las moléculas de polinucleótidos monocatenarios molde pueden unirse a un soporte sólido mediante hibridación con cebadores inmovilizados, o como alternativa, las moléculas de polinucleótidos monocatenarios también pueden unirse directamente al soporte sólido en o cerca del extremo 5'. La química de la unión elegida dependerá de la naturaleza del soporte sólido y de cualquier funcionalización o derivación que se le aplique. La molécula de polinucleótido monocatenario en sí puede incluir un resto, que puede ser una modificación química no nucleotídica para facilitar la unión. En particular, una molécula de polinucleótido monocatenario puede incluir un nucleófilo que contiene azufre tal como fosforotioato o tiofosfato en el extremo 5'. En el caso de los hidrogeles de poli(acrilamida con soporte sólido, este nucleófilo también puede unirse a los grupos bromoacetamida presentes en el hidrogel. En un ejemplo, los medios para unir la molécula de polinucleótido monocatenario al soporte sólido es a través de la unión de fosforotioato 5' a un hidrogel compuesto de acrilamida polimerizada y N-(5-bromoacetamidil-pentil) acrilamida (BRAPA). Tal disposición se describe más detalladamente en la solicitud WO 05/065814 en tramitación con la presente.

La expresión "soporte sólido", como se usa en la presente memoria, se refiere a cualquier superficie, sustrato inerte o matriz a la que se pueden unir ácidos nucleicos como por ejemplo perlas, incluyendo perlas de látex o de dextrano, una superficie, como una superficie de poliestireno o polipropileno, gel de poli(acrilamida, superficies de oro, superficies de vidrio y obleas de silicio. El soporte sólido puede ser una superficie de vidrio. El soporte sólido puede ser una superficie plana, aunque la invención también funciona sobre perlas que se mueven entre recipientes de diferentes tampones, o perlas dispuestas en una superficie plana. El soporte sólido puede ser una celda de flujo, resina, gel, perla, pozo, columna, escama, membrana, matriz, placa o filtro.

El soporte sólido puede comprender un sustrato o matriz inerte que se ha "funcionalizado", por ejemplo, mediante la aplicación de una capa o recubrimiento de un material intermedio que comprende grupos reactivos que permiten la unión covalente a moléculas tales como polinucleótidos. Tales soportes pueden incluir hidrogeles de poli(acrilamida soportados sobre un sustrato inerte, como el vidrio. Las moléculas (por ejemplo, polinucleótidos) pueden unirse directamente de manera covalente al material intermedio (por ejemplo, el

hidrogel), pero el material intermedio en sí puede estar unido de manera no covalente al sustrato o matriz (por ejemplo, el sustrato de vidrio). Tal disposición se describe más detalladamente en la solicitud WO 05/065814 en tramitación con la presente.

5 Los oligonucleótidos de cebadores o los cebadores son secuencias de polinucleótidos que pueden hibridarse específicamente a uno o más moldes de polinucleótidos monocatenarios para amplificarse en las condiciones encontradas en la etapa de hibridación de cebadores de cada ciclo de una reacción de amplificación. En general, las reacciones de amplificación pueden usar al menos dos cebadores de amplificación, a menudo mencionados como cebadores "directos" e "inversos". En ciertos casos, los cebadores directos e inversos pueden ser idénticos. Los oligonucleótidos de cebadores directos pueden incluir una "porción de molde específica", siendo una secuencia de nucleótidos capaz de hibridarse a una secuencia de unión de cebador en al menos una cadena de la molécula a amplificar. Los oligonucleótidos de cebadores inversos pueden incluir una porción de molde específica capaz de hibridarse al complemento de la cadena a la que el cebador directo se hibrida durante la etapa de hibridación. Generalmente, los oligonucleótidos de cebadores son estructuras de polinucleótidos monocatenarios. También pueden contener una mezcla de bases naturales y no naturales y también, enlaces principales naturales y no naturales, siempre que las modificaciones no naturales no excluyan la función como cebador, que se define como la capacidad de hibridación a una cadena de polinucleótido de molde durante las condiciones de la reacción de amplificación y de actuar como un punto de inicio para la síntesis de una nueva cadena de polinucleótidos complementaria a la cadena de molde.

20 Los cebadores pueden comprender adicionalmente modificaciones químicas no nucleotídicas, de nuevo siempre que tales modificaciones no eviten permanentemente la función del cebador. Las modificaciones químicas pueden, por ejemplo, facilitar la unión covalente del cebador a un soporte sólido. Ciertas modificaciones químicas en sí pueden mejorar la función de la molécula como cebador, o pueden proporcionar alguna otra funcionalidad útil, como proporcionar un sitio de segmentación para permitir que el cebador (o una cadena de polinucleótidos extendida derivada de la misma) se segmente a partir de un soporte sólido.

25 Aunque la presente descripción puede abarcar métodos de amplificación en fase sólida, en los que solo un cebador de amplificación está inmovilizado sobre un soporte sólido (el otro cebador suele estar presente en solución libre), el soporte sólido puede proporcionarse con los cebadores directos e inversos inmovilizados. En la práctica, puede haber una pluralidad de cebadores directos idénticos y/o una pluralidad de cebadores inversos idénticos inmovilizados en el soporte sólido, por ejemplo, en realizaciones en las que el procedimiento de amplificación usa un exceso de cebadores para sostener la amplificación. Por lo tanto, las referencias de la presente memoria a los cebadores directos e inversos deben interpretarse en consecuencia como si abarcaran una pluralidad de dichos cebadores, a menos que el contexto indique lo contrario.

35 La expresión "amplificación en fase sólida", como se usa en la presente memoria, se refiere a cualquier reacción de amplificación de ácido nucleico llevada a cabo en o en asociación con un soporte sólido de tal manera que todos o una parte de los productos amplificados estén inmovilizados sobre el soporte sólido. En particular, la expresión abarca reacciones de amplificación en fase sólida análogas a la PCR en fase de solución estándar, excepto que uno o ambos de los cebadores de amplificación directo e inverso esté/estén inmovilizado/s en el soporte sólido.

40 Los oligonucleótidos de cebadores y las moléculas de polinucleótidos monocatenarios que se han inmovilizado en un soporte sólido a una densidad deseada se pueden usar para generar productos de extensión al llevar a cabo un número apropiado de ciclos de amplificación en las moléculas de polinucleótidos monocatenarios unidas covalentemente de manera que cada colonia, o agrupamiento, comprenda múltiples copias de la molécula de polinucleótido monocatenario original inmovilizada (y su secuencia complementaria). Un ciclo de amplificación puede incluir etapas de hibridación, extensión y desnaturalización. Tales etapas son generalmente comparables con las etapas de hibridación, extensión y desnaturalización de la PCR.

45 Cuando se usa la amplificación en fase sólida, se pueden aplicar condiciones adecuadas a una molécula de polinucleótido monocatenario y una pluralidad de oligonucleótidos de cebadores inmovilizados de modo que la secuencia Z en el extremo 3' de la molécula de polinucleótido monocatenario se hibride a una secuencia de oligonucleótidos de cebador X para formar un complejo en donde el oligonucleótido de cebador se hibrida a el molde monocatenario para crear una estructura de "puente". Las condiciones adecuadas, tales como los tampones de neutralización y/o hibridación, son bien conocidas en la técnica (véase Sambrook et al., Molecular Cloning, A Laboratory Manual, 3ª ed., Cold Spring Harbor Laboratory Press, NY; Current Protocols, eds Ausubel et al.). El tampón de neutralización y/o hibridación puede ser luego eliminado. Un tampón de hibridación adecuado se denomina "premezcla de amplificación" y contiene betaína 2 M, Tris 20 mM, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Tritón al 0,1 %, DMSO al 1,3 %, pH 8,8.

50 Aplicando condiciones adecuadas, se puede realizar una reacción de extensión para un complejo formado entre el cebador inmovilizado y el molde de polinucleótido monocatenario. El oligonucleótido de cebador del complejo puede extenderse mediante la adición secuencial de nucleótidos para generar un producto de extensión complementario a la molécula de polinucleótido monocatenario.

60 Ejemplos de enzimas con actividad polimerasa que se pueden usar en la presente invención son ADN polimerasa (fragmento Klenow, ADN polimerasa T4, polimerasa Bst), ADN polimerasas termoestables de una

variedad de bacterias termoestables (como Taq, VENT, Pfu, Tfl, ADN polimerasa Phusion), así como sus derivados modificados genéticamente (TaqGold, VENTexo, Pfu exo). También se puede usar una combinación de ARN polimerasa y transcriptasa inversa para generar los productos de extensión. Una enzima polimerasa útil puede tener actividad de desplazamiento de cadena. La enzima polimerasa puede ser activa a un pH de aproximadamente 7 a aproximadamente 9, particularmente de pH 7,9 a pH 8,8. Las moléculas de nucleósido trifosfato usadas pueden ser trifosfatos desoxirribonucleótidos, por ejemplo dATP, dTTP, dCTP, dGTP, o pueden ser trifosfatos de ribonucleósidos, por ejemplo, ATP, UTP, CTP, GTP. Las moléculas de nucleósido trifosfato pueden ser naturales o no naturales. Una reacción de amplificación también puede contener aditivos tales como DMSO y/o betaína, por ejemplo, para normalizar las temperaturas de fusión de las diferentes secuencias en las cadenas del molde. Una solución adecuada para los ciclos iniciales de extensión se denomina "mezcla de amplificación" y contiene betaína 2 M, Tris 20 mM, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Triton al 0,1 %, DMSO al 1,3 %, pH 8,8 más 200  $\mu$ M dNTP y 80 unidades/ml de Bst polimerasa.

La desnaturalización puede llevarse a cabo usando calor o usando un tampón desnaturalizante. Los tampones desnaturalizantes adecuados son bien conocidos en la técnica (véase Sambrook et al., Molecular Cloning, A Laboratory Manual, 3<sup>a</sup> ed., Cold Spring Harbor Laboratory Press, NY; Current Protocols, eds. Ausubel et al.). A modo de ejemplo, se sabe que las alteraciones en el pH y las soluciones de baja intensidad iónica pueden desnaturalizar los ácidos nucleicos a temperaturas sustancialmente isotérmicas. La formamida y la urea pueden usarse para la desnaturalización. En particular, la concentración de formamida es del 50% o más, y se puede usar sola. Tales condiciones dan como resultado la desnaturalización de moléculas de ácido nucleico bicatenario a moléculas de ácido nucleico monocatenario. De forma alternativa o adicional, las cadenas se pueden separar por tratamiento con una solución muy baja en sal (por ejemplo, condiciones catiónicas de menos de 0,1 mM) y pH alto (> 12) o usando una sal caotrópica (por ejemplo, hidrocloreuro de guanidinio). En particular, se puede usar una base fuerte. Una base fuerte es un compuesto químico básico que puede desprotonar ácidos muy débiles en una reacción de base ácida. La intensidad de una base está indicada por su valor en  $pK_b$ . Los compuestos con un valor de  $pK_b$  de menos de aproximadamente 1 se llaman bases fuertes y son bien conocidos por un experto en la técnica. En una realización particular, la base fuerte es una solución de hidróxido de sodio (NaOH) usada a una concentración de 0,05 M a 0,25 M. Más particularmente, se usa NaOH a una concentración de 0,1 M.

Puede ser ventajoso realizar etapas de lavado opcionales entre etapas de un método de amplificación. Por ejemplo, un tampón de extensión sin enzima polimerasa con o sin dNTP podría aplicarse a un soporte sólido sobre el cual se lleva a cabo la amplificación y puede aplicarse antes de ser eliminado y reemplazado con un tampón de extensión completo (tampón de extensión que incluye todos los componentes necesarios para su extensión para proceder).

Múltiples ciclos de amplificación en una superficie sólida en las condiciones ejemplificadas anteriormente pueden dar lugar a una colonia de ácido nucleico o "agrupamiento" que comprende múltiples copias inmovilizadas de una secuencia de polinucleótidos monocatenarios particular y su secuencia complementaria. La inmovilización inicial de una molécula de polinucleótido monocatenario en condiciones ejemplificadas en la presente memoria puede dar lugar a que la molécula de polinucleótido monocatenario solo se hibride a oligonucleótidos de cebadores localizados a una distancia dentro de la longitud total de la molécula de polinucleótido monocatenario. Por lo tanto, el límite de la colonia o agrupamiento de ácido nucleico formado se puede limitar a un área relativamente local, a saber, el área en la que se inmovilizó la molécula de polinucleótido monocatenario inicial. Si se usan condiciones en las que los moldes y las copias complementarias de los mismos permanecen inmovilizadas durante todo el procedimiento de amplificación, entonces los moldes no se entremezclan más que haciéndose lo suficientemente grandes como para que se superpongan en la superficie. En ejemplos particulares, no hay ácido nucleico no inmovilizado durante ninguna parte del procedimiento de amplificación y, por lo tanto, los moldes no pueden difundir ni iniciar agrupamientos adicionales en otras partes de la superficie.

Las etapas de hibridación, extensión y desnaturalización de un método de amplificación expuesto en la presente memoria se pueden llevar a cabo a la misma temperatura sustancialmente isotérmica. Preferiblemente, la temperatura es de 37 °C a aproximadamente 75 °C, dependiendo de la elección de la enzima, más preferiblemente de 50 °C a 70 °C, y aún más preferiblemente de 60 °C a 65 °C para la polimerasa Bst. La temperatura sustancialmente isotérmica puede estar alrededor de la temperatura de fusión del o de los cebadores de oligonucleótidos. Los métodos para calcular las temperaturas de fusión apropiadas son conocidos en la técnica. Por ejemplo, la temperatura de hibridación puede ser aproximadamente 5 °C por debajo de la temperatura de fusión (Tf) de los cebadores de oligonucleótidos. En otro ejemplo particular, la temperatura sustancialmente isotérmica se puede determinar empíricamente. La temperatura puede ser aquella en la que el oligonucleótido muestra la mayor especificidad para el sitio de unión del cebador al tiempo que reduce la unión no específica.

La expresión "secuencia común", cuando se usa en referencia a una colección de moléculas de ácido nucleico, significa una secuencia que es la misma para todos los ácidos nucleicos en la colección. Los ácidos nucleicos en la colección pueden tener una región de secuencia común a pesar de la presencia de al menos otra región en cada uno de los ácidos nucleicos que difiere entre los ácidos nucleicos en la colección. Como se ejemplifica en la presente memoria, todos los moldes dentro de una genoteca modificada en 5' y 3' pueden contener

regiones de secuencia común Y y Z en (o proximales a) sus extremos 5' y 3', particularmente en donde la secuencia común en el extremo 5' de cada molde individual en la genoteca no es idéntica y no es totalmente complementaria a la secuencia común en el extremo 3' de dicho molde. El término "genoteca" se refiere a una colección o pluralidad de moléculas de molde que pueden compartir secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3'. El uso de la expresión "genoteca modificada en 5' y 3'" para referirse a una colección o una pluralidad de moléculas de molde no debe implicar que los moldes que forman la genoteca deriven de una fuente particular. A modo de ejemplo, la expresión "genoteca modificada en 5' y 3'" puede incluir moldes individuales dentro de la genoteca que tienen la misma secuencia de nucleótidos o que tienen diferentes secuencias de nucleótidos. Además, los moldes pueden, pero no necesitan estar relacionados en términos de secuencia y/o fuente.

La presente descripción incluye el uso de las denominadas genotecas "monomolde", que comprenden múltiples copias de un solo tipo de molécula de molde, cada una con secuencias comunes en sus extremos 5' y en sus extremos 3', así como genotecas "complejas" en donde muchas, si no todas las moléculas de molde individuales comprenden diferentes secuencias diana, aunque todas comparten secuencias comunes en sus extremos 5' y en sus extremos 3'. Dichas genotecas de moldes complejos pueden prepararse a partir de una mezcla compleja de polinucleótidos diana tales como (pero no limitados a) fragmentos aleatorios de ADN genómico, genotecas de ADNc, etc. Los métodos de la descripción también se pueden usar para amplificar genotecas "complejas" formadas mezclando varias genotecas individuales de "monomolde", cada una de las cuales se ha preparado por separado a partir de un único tipo de molécula diana (es decir, una monomolde). En particular, más del 50%, o más del 60%, o más del 70%, o más del 80%, o más del 90%, o más del 95% de los moldes de polinucleótidos individuales en una genoteca compleja pueden comprender diferentes secuencias diana, aunque todos los moldes de una genoteca determinada pueden compartir una secuencia común en sus extremos 5' y una secuencia común en sus extremos 3'.

Los aditivos descritos en la presente memoria están relacionados, al menos, por su capacidad para normalizar la amplificación de los moldes de ácido nucleico de diferentes secuencias. Los métodos incluyen opcionalmente el uso de diferentes concentraciones de nucleótidos y/o análogos de nucleótidos como se describe en la presente memoria. Los aditivos pueden ser, por ejemplo, etilenglicol, polietilenglicol, 1,2-propanodiol, dimetilsulfóxido (DMSO), glicerol, formamida, 7-deaza-GTP, acetamida, cloruro de tetrametil amonio (TMACl), sal o betaína. Por ejemplo, puede añadirse betaína (carboximetil trimetil amonio  $((\text{CH}_3)_3\text{N}^+\text{CH}_2\text{COO}^-)$ ) a la mezcla de amplificación para normalizar la amplificación de diferentes secuencias molde. Opcionalmente, se usa una combinación de betaína y DMSO o una combinación de betaína, DMSO y 7-deaza-dGTP. En particular, las concentraciones de betaína pueden estar por encima de 2 Molar (M), por ejemplo, entre 2 M y 5 M, entre 2,5 M y 4 M o entre 2,75 M y 3,75 M.

### Descripción

La presente invención proporciona métodos para reducir el sesgo de GC dependiente de la densidad en la amplificación isotérmica en puente usada para la generación de agrupamientos. En una realización, el método de la invención usa la adición de un aditivo a una primera solución en un protocolo de amplificación isotérmica en puente. La primera solución en el protocolo de amplificación en puente es una solución de formamida usada para desnaturar ADN bicatenario. En una realización, la primera solución comprende, además, EDTA. En un ejemplo, el aditivo es una mezcla de dNTP. En otro ejemplo, el aditivo es un solo dNTP (es decir, dATP, dCTP, dGTP o dTTP). También se describe el aditivo que es citrato trisódico 10 mM, citrato trisódico 1 mM o citrato trisódico de menos de 1 mM.

En otra realización, el método de la invención usa la adición de un aditivo a una segunda solución en un protocolo de amplificación isotérmica en puente. En un ejemplo, la segunda solución en el protocolo de amplificación en puente es una solución de premezcla. En un ejemplo, el aditivo es una mezcla de dNTP. En otro ejemplo, la segunda solución en el protocolo de amplificación en puente es agua. En este ejemplo, el agua se puede usar sola o con la adición de dNTP, citrato trisódico y/u otro aditivo. Reemplazar con agua la solución de premezcla, que contiene sales y tampones, mejora la amplificación de los moldes ricos en GC. Un protocolo de amplificación isotérmica en puente para la generación de agrupamientos en una celda de flujo generalmente incluye ciclos repetidos de desnaturación, hibridación y extensión. La Figura 1 ilustra un diagrama de flujo de un ejemplo de un protocolo de amplificación isotérmica en puente 100 para la amplificación de agrupamientos en una superficie de celda de flujo. En un ejemplo, el protocolo de amplificación en puente 100 se realiza a aproximadamente 60 °C y se repite cualquier número de veces (por ejemplo, típicamente 35 veces) para generar una población clonal de cadenas de ADN idénticas para cada molde de ADN sembrado. Por ejemplo, el protocolo de amplificación en puente puede repetirse de uno a cincuenta, de uno a veinticinco, de uno a quince o de uno a diez ciclos de amplificación. El protocolo de amplificación en puente 100 incluye pero no se limita a las siguientes etapas.

En una etapa 110, una primera solución, por ejemplo, de formamida (por ejemplo, de aproximadamente 28  $\mu\text{l}$ ), se bombea a través de cada carril de la celda de flujo sembrada con molde de ADN. La formamida desnatura el ADN bicatenario del agrupamiento en la celda de flujo dando ADN monocatenario.

En una etapa 115, se elimina la formamida bombeando una segunda solución, por ejemplo, una solución de

premezcla de agrupamiento (por ejemplo, de aproximadamente 28  $\mu$ l), a través de cada carril de la celda de flujo. En un ejemplo, la solución de premezcla es una solución de premezcla estándar que comprende Tris-HCl 20 mM, pH 8,8 a 25 °C, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Triton X-100 al 0,1 %, DMSO al 1,3 % y betaína 2 M. Tras la eliminación de formamida mediante el lavado con la solución de premezcla, las cadenas de ADN desnaturalizadas se unen y se hibridan a los cebadores de oligonucleótidos unidos a la superficie de la celda de flujo.

En la etapa 120, se bombea una tercera solución, por ejemplo, una solución de mezcla Bst (por ejemplo, de aproximadamente 36  $\mu$ l), a cada carril de la celda de flujo. La mezcla Bst contiene dNTP y Bst ADN polimerasa. En un ejemplo, una solución de mezcla Bst estándar es Tris-HCl 20 mM, pH 8,8 a 25 °C, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Triton X-100 al 0,1 %, DMSO al 1,3 %, betaína 2 M, 200  $\mu$ M de cada nucleótido, y 79 u 80 U/ml de polimerasa Bst. En otro ejemplo, la mezcla Bst es una mezcla Bst sin sulfato de amonio (sin  $\text{NH}_4$ ) de Tris-HCl 20 mM, pH 8,8 a 25 °C, sulfato de Mg 2 mM, Triton X-100 al 0,1 %, DMSO al 1,3 %, betaína 2 M, 200  $\mu$ M de cada nucleótido y 79 u 80 U/ml de polimerasa Bst. En presencia de la mezcla Bst, los cebadores de oligonucleótidos unidos a la superficie que se hibridan a moléculas de molde están completamente extendidos.

El protocolo de amplificación en puente 100 se puede repetir cualquier número de veces. Desde un sitio de siembra, una molécula de ADN genera un agrupamiento que aumenta de tamaño en cada ciclo de amplificación. Sin embargo, para cualquier longitud de inserto en particular (longitud del molde de ADN), el tamaño del agrupamiento y su intensidad pueden verse afectados por la composición de GC del inserto. Por ejemplo, los moldes de ADN que contienen una alta proporción de nucleótidos A y T (o T y U para el ARN) tienden a producir agrupamientos más grandes y más brillantes en comparación con los moldes que tienen una mayor proporción de nucleótidos G y C.

Las Figuras 2A y 2B muestran un gráfico 200 del tamaño de los agrupamientos y un gráfico 250 de la intensidad de los agrupamientos, respectivamente, de agrupamientos generados a partir de monomoldes con diferentes composiciones de GC. Un monomolde puede definirse como una secuencia molde lista para una secuenciación única (por ejemplo, cebador P5, cebador de secuenciación, inserto de molde y cebador P7) con un cierto tamaño de inserto. En este ejemplo, se sembraron monomoldes CT151 (31 % de GC), CT161 (50 % de GC), CT159 (55 % de GC), CT152 (64 % de GC) y CT180 (80 % de GC) con un inserto de 400 pb de tamaño en diferentes carriles de la misma celda de flujo. Después de una primera extensión, se generaron grupos usando 35 ciclos de amplificación isotérmica. La celda de flujo luego se tiñó con SYBR® Green y se tomaron imágenes de dos placas representativas para cada carril usando un microscopio. Se determinó el tamaño del agrupamiento y las intensidades usando el programa informático de análisis de imagen Firecrest.

La Figura 3 muestra una captura de pantalla 300 de una imagen de agrupamientos que muestra agrupamientos ricos en GC y agrupamientos ricos en AT. En este ejemplo, se sembraron cuatro monomoldes a una densidad relativamente baja. Los agrupamientos se generaron amplificando en exceso cada molécula para generar agrupamientos más grandes. Después de la linealización, el bloqueo del extremo 3' con ddNTP y la hibridación del cebador de secuenciación, se realizó un primer ciclo de amplificación en puente. Luego, se produjo una imagen de superposición de 3 colores. De los cuatro monomoldes sembrados, tres fueron ricos en AT (no se muestran los agrupamientos del tercer monomolde rico en AT) y uno fue rico en GC (80 % de GC). Los agrupamientos están etiquetados en la imagen en escala de grises como "G" y "B" ricos en AT, y "R" rico en GC. Debido a la inhibición del contacto, los agrupamientos no se superponen entre sí; dejan de crecer cuando entran en contacto. En el protocolo de amplificación actual, los moldes ricos en AT tienden a producir agrupamientos grandes, mientras que los moldes ricos en GC dan origen a agrupamientos más pequeños. Como se muestra en la Figura 3, los agrupamientos ricos en GC (R) se han comprimido entre los agrupamientos ricos en AT (G y B) y su crecimiento ha sido restringido por los agrupamientos ricos en AT que crecen más rápido.

Una disminución en la eficiencia de la amplificación de agrupamientos de los agrupamientos ricos en GC puede deberse al daño en el ADN causado por los reactivos de amplificación de agrupamientos. Cuando una cadena de ADN está dañada (por ejemplo, ya sea la estructura principal de fosfato-azúcar o las bases dañadas que detienen la extensión del cebador), puede que no actúe como molde en los ciclos de amplificación posteriores y, por lo tanto, se puede observar una amplificación de agrupación menos eficiente (es decir, agrupaciones más débiles).

La modificación de los reactivos de amplificación de agrupamientos mediante la adición de aditivos y/o la eliminación de algunos componentes de las soluciones de reactivo de amplificación estándar se puede usar para mejorar los efectos del daño en el ADN. Esto se puede lograr mediante la reducción del daño en el ADN a través de formamida usando un aditivo en la solución de formamida, la solución de premezcla o ambas. Además, reducir la concentración de sal puede potenciar el crecimiento del agrupamiento rico en GC a través de un mecanismo que no implica daño en el ADN. Estos resultados fueron sorprendentes. Anteriormente no se conocía que la formamida causaba daños en el ADN.

La Figura 4 ilustra un diagrama de flujo de un ejemplo de un ensayo 400 de daño en el ADN para evaluar el efecto de los reactivos de amplificación de agrupamientos y/o la adición de aditivos en la generación de

agrupamientos durante la amplificación isotérmica en puente. La Figura 5 muestra gráficamente las etapas del ensayo 400 de daño en el ADN de la Figura 4. El método 400 incluye pero no está limitado a las siguientes etapas.

5 En una etapa 410, se siembran moléculas de ADN molde en la celda de flujo y se las hibrida a cebadores de oligonucleótidos unidos a la superficie. Esta etapa también se muestra gráficamente en la Figura 5.

10 En la etapa 415, se copian las moléculas de ADN molde (primera extensión) usando ADN polimerasa (por ejemplo, Taq ADN polimerasa o ADN polimerasa Phusion) y luego, se retira la cadena de molde sembrada haciendo fluir NaOH 0,1 M a través de la celda de flujo (el NaOH también desnatura la ADN polimerasa). Las moléculas de ADN monocatenario permanecen unidas covalentemente a la superficie de la celda de flujo. Esta etapa también se muestra gráficamente en la Figura 5.

15 En una etapa 420, las moléculas de ADN monocatenario unidas covalentemente a la superficie de la celda de flujo se someten a un tratamiento de amplificación isotérmica "simulada". La amplificación isotérmica simulada comprende el ciclo de los reactivos que se usan en la amplificación de agrupamientos (por ejemplo, formamida, premezcla y mezcla de amplificación con la omisión de Bst ADN polimerasa para evitar la amplificación de ADN durante el tratamiento). En un ejemplo, la amplificación isotérmica simulada comprende hacer fluir la formamida a través de un carril en la celda de flujo, eliminar la formamida haciendo fluir la premezcla a través del carril de la celda de flujo y hacer fluir una premezcla que contiene dNTP en el carril de la celda de flujo. Incluido en la celda de flujo, hay un carril de control que sirve como control negativo (sin daño en el ADN). Para el carril de control, un tampón de lavado (0.3X SSC que contiene 0,1 % de Tween-20) se hace fluir continuamente a través del carril; el tampón de lavado no daña el ADN. Esta etapa también se muestra gráficamente en la Figura 5.

20 En una etapa 425, todos los carriles se enjuagan con tampón de lavado y luego, se los somete a amplificación isotérmica. Solo las moléculas no dañadas pueden actuar como molde durante esta etapa de amplificación isotérmica y, por lo tanto, conducir a la formación de agrupamientos. Las moléculas de ADN dañadas o las moléculas rotas no se amplificarán. Esta etapa también se muestra gráficamente en la Figura 5.

25 En la etapa 430, los agrupamientos se visualizan mediante tinción con un tinte de unión al ADN (por ejemplo, SYBR® Green) y se toman imágenes (por ejemplo, tres placas para cada carril) usando un microscopio y una cámara. El programa informático de análisis de imágenes (por ejemplo, el programa informático Firecrest) se usa para determinar la cantidad de agrupamientos. Al contar el número de agrupamientos de cada carril, es posible inferir cuántas moléculas han sido dañadas por un tratamiento en particular.

30 Se evaluó el daño potencial al ADN causado por los reactivos de amplificación isotérmica estándar usando el ensayo 400 de daño en el ADN de la Figura 4 y los monomoldes sembrados. En un ejemplo, se puede usar un monomolde CT180 en el ensayo 400 de daño en el ADN para evaluar el daño potencial en el ADN causado por los reactivos de amplificación isotérmica. El monomolde CT180 comprende un 80 % de GC (es decir que es rica en GC) y tiene un tamaño de inserción de 400 pb.

35 La Figura 6 muestra un gráfico de barras 600 del porcentaje de moléculas de ADN no dañadas en muestras tratadas con formamida, premezcla y/o premezcla más dNTP y evaluadas usando el ensayo 400 de daño en el ADN de la Figura 4. En este experimento, el monomolde CT180 se sembró a 1 pM. Después de la primera extensión de la cadena con Taq ADN polimerasa y la desnaturación del molde con NaOH, se trataron las moléculas de ADN individuales con 26 ciclos de amplificación isotérmica simulada (etapa 420 del ensayo 400 de daño en el ADN de la Figura 4) usando una primera solución de 28 µl de formamida, una segunda solución de 28 µl de premezcla con o sin la adición de dNTP, y una tercera solución de 36 µl de premezcla con la adición de dNTP. Cada barra en el gráfico representa un carril en la celda de flujo. La línea de control (primera barra del gráfico) se trató con tampón de lavado durante todo el tratamiento (26 ciclos). Se dividieron el número de agrupamientos para cada carril tratado (barras 2 y 3 del gráfico) en la celda de flujo por el número de agrupamientos del carril de control y se expresaron como un porcentaje del carril de control (referencia). El carril 2, que se trató con 26 ciclos de amplificación isotérmica simulada usando las condiciones estándar (formamida/premezcla/premezcla con dNTP) muestra una disminución significativa en el número de agrupamientos lo que demuestra que los reactivos de amplificación isotérmica causan daño en el ADN (un experimento de ensayo de densidad de cebador mostró que no hubo una pérdida significativa de superficie u oligonucleótidos de la superficie; no se muestran los datos). Cuando se incluyeron los dNTP en la solución de premezcla (premezcla/dNTP) que se bombeó inmediatamente después de la formamida, el daño en el ADN fue menos grave (comparar los carriles 2 y 3). Estos resultados sugieren que los dNTP protegen el ADN de los daños causados por los reactivos de amplificación de agrupamientos (por ejemplo, formamida).

55 La Figura 7 muestra un gráfico de barras 700 del porcentaje de moléculas de ADN no dañadas en muestras tratadas con formamida ± dNTP, premezcla ± dNTP o agua ± dNTP y evaluadas usando el ensayo 400 de daño en el ADN de la Figura 4. En este experimento, el monomolde CT180 se sembró a 1 pM. Después de la primera extensión de la cadena con Taq ADN polimerasa y la desnaturación del molde con NaOH, se trataron las moléculas de ADN individuales con 26 ciclos de amplificación isotérmica simulada (etapa 420 del ensayo 400 de daño de ADN de la Figura 4) usando tres soluciones. La primera solución fue de 28 µl de tampón de lavado (TL), formamida (F) o de formamida más dNTP (F/nt). La segunda solución fue de 28 µl de tampón de lavado (TL), formamida (F), premezcla, premezcla más dNTP (premezcla/nt), agua (H<sub>2</sub>O) o agua más dNTP (H<sub>2</sub>O/nt).

60

La tercera solución fue de 36  $\mu$ l de tampón de lavado (TL) o premezcla/nt. Los dNTP se usaron a una concentración final de 200  $\mu$ M cada uno. Cada barra en el gráfico representa un carril de la celda de flujo. Los datos muestran el efecto protector de los dNTP cuando están presentes en la solución de premezcla (comparando los carriles 3 y 4), en formamida (comparando los carriles 4 y 5) y en agua (comparando los carriles 6 y 7). Asimismo, el carril 8 muestra que el tampón de lavado también protege las moléculas de ADN del daño causado por la formamida (comparar los carriles 2 y 8).

Se evaluó el efecto de la concentración de dNTP en la reducción del daño en el ADN inducido por formamida en un experimento de titulación en el que se añadieron diferentes cantidades de dNTP a la formamida. La Figura 8 muestra una curva 800 de respuesta a la dosis del efecto de la concentración de dNTP en formamida en el número de agrupamientos durante la amplificación isotérmica simulada. En este experimento, se sembró CT180 a 0,6 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Luego, se trataron las moléculas individuales con 26 ciclos de amplificación isotérmica simulada de acuerdo con el ensayo 400 de daño en el ADN de la Figura 4 usando 28  $\mu$ l de una primera solución de formamida con diferentes concentraciones de dNTP, 28  $\mu$ l de una segunda solución de agua con las mismas concentraciones de dNTP, y 36  $\mu$ l de una tercera solución de premezcla con las mismas concentraciones de dNTP. Después de la amplificación isotérmica simulada, se generaron poblaciones clonales de moléculas (agrupamientos) con 26 ciclos de amplificación isotérmica usando 28  $\mu$ l de formamida seguido de 64  $\mu$ l de mezcla Bst en cada ciclo (sin usar premezcla). Los resultados muestran una respuesta a la dosis de la anulación del daño en el ADN con concentraciones más altas de dNTP que son más efectivas para evitar el daño en el ADN.

Se evaluó el efecto de los dNTP individuales en la reducción del daño en el ADN inducido por formamida usando el ensayo 400 de daño en el ADN de la Figura 4. La Figura 9 muestra un gráfico de barras 900 de la efectividad de cada dNTP para evitar daños en el ADN cuando se usa como un aditivo de formamida. En este experimento, se sembró CT180 a 0,6 pM. La primera extensión de la cadena se realizó usando Taq ADN polimerasa seguida de desnaturalización con NaOH. Luego, se trataron las moléculas individuales con tampón de lavado (TL), formamida (F) o formamida más dATP, dCTP, dGTP, dTTP o una mezcla de dNTP durante 80 minutos a 60 °C (un período de tiempo similar a la duración de 26 ciclos de amplificación isotérmica). Después del tratamiento, se realizó una amplificación del agrupamiento usando 26 ciclos isotérmicos de 28  $\mu$ l de formamida seguidos de 64  $\mu$ l de la mezcla Bst para cada ciclo. La concentración final de los nucleótidos individuales fue de 800  $\mu$ M. La concentración de cada nucleótido en la mezcla de dNTP fue de 200  $\mu$ M. Los datos muestran que los cuatro nucleótidos son igualmente efectivos para proteger el ADN del daño. Los resultados también muestran que los dNTP evitan el daño en el ADN causado por la formamida cuando este producto químico se bombea constantemente (en lugar de ser ciclado junto con agua y la premezcla).

Para determinar si la adición de dNTP puede rescatar la amplificación clonal cuando se usa un volumen relativamente grande de formamida en un procedimiento de amplificación isotérmica, se llevó a cabo una amplificación usando 120  $\mu$ l de formamida con o sin la adición de dNTP. La Figura 10 muestra paneles de agrupamientos generados por amplificación isotérmica usando formamida con y sin dNTP añadidos. En este ejemplo, se usaron dos moldes, CT180 (un molde rico en GC) y CT3164 (un molde rico en AT). Los paneles de la izquierda, de arriba a abajo, representan los carriles 1 a 4 de la celda de flujo. Los paneles de la derecha, de arriba a abajo, representan los carriles 5 a 8 de la celda de flujo. Se sembraron los carriles 1 a 4 en la celda de flujo con 0,6 pM de CT180; se sembraron los carriles 5 a 8 en la celda de flujo con 0,6 pM de CT3164. Después de la primera extensión de la cadena usando Taq ADN polimerasa, seguida de la desnaturalización con NaOH, se realizó una amplificación de agrupamientos usando 26 ciclos de amplificación isotérmica de 120  $\mu$ l de formamida (un volumen relativamente grande en comparación con los 28  $\mu$ l estándar de formamida), seguido de 28  $\mu$ l de agua y 36  $\mu$ l de la mezcla Bst estándar. Los agrupamientos se visualizaron mediante tinción con SYBR® Green y se tomaron imágenes. En la Figura 9, se muestra una placa representativa de cada carril. Los datos muestran que, si bien tanto el molde CT180 como la CT3164 no se amplifican en un procedimiento de amplificación isotérmica que usa un volumen relativamente grande de formamida, la amplificación de agrupamientos se rescata mediante la adición de dNTP en la solución de formamida.

La Figura 11 muestra una tabla de datos 1100 de métrica de secuenciación para una ejecución de secuenciación que evalúa el "método de los dNTP" y el "método del agua". El método de los dNTP incluye la adición de los dNTP tanto a la primera solución de formamida como a la segunda solución de agua. El método del agua reemplaza una segunda solución de premezcla con agua. En este experimento, se sembró CT4008 (genoteca BCG estándar, genoma rico en GC; tamaño de inserción promedio de 300 pb) en baja densidad (1,5 pM) o alta densidad (7,5 pM). Después de la primera extensión estándar de la cadena con ADN polimerasa Phusion, los agrupamientos se amplificaron con 26 ciclos de amplificación isotérmica usando diferentes condiciones: sin premezcla = 28  $\mu$ l de formamida y 64  $\mu$ l de mezcla Bst; premezcla = 28  $\mu$ l de formamida, 28  $\mu$ l de premezcla y 36  $\mu$ l de mezcla Bst; método del agua = 28  $\mu$ l de formamida, 28  $\mu$ l de agua y 36  $\mu$ l de mezcla Bst; método de los dNTP = 28  $\mu$ l de formamida que contiene 200  $\mu$ M de cada nucleótido, 28  $\mu$ l de agua que contiene 200  $\mu$ M de cada nucleótido y 36  $\mu$ l de mezcla Bst. Esta ejecución fue analizada con el programa informático Pipeline 1.8. La celda de flujo se secuenció con 36 ciclos de SBS usando la química 95G en un analizador de genoma llx.

Para el genoma de BCG, hubo un aumento significativo en términos de los agrupamientos que pasaron el filtro

(véase la tabla de datos 1100 de la columna "% de agrupamientos PF") tanto para el "método del agua" como para el "método de los dNTP" en comparación con los métodos sin premezcla y de amplificación estándar (etiquetado "premezcla" en la Figura). Los agrupamientos más brillantes se obtuvieron con el "método de los dNTP" (comparar los carriles 6 y 8 con los carriles 4 y 7).

5 Las Figuras 12A y 12B muestran curvas 1200 del sesgo de GC y curvas 1250 del sesgo de GC normalizadas a la línea de baja densidad 1, respectivamente, de la ejecución de secuencia descrita con referencia a la Figura 11.

10 La Figura 13A muestra un diagrama esquemático 1300 resumido de soluciones de amplificación en puente de acuerdo con los métodos de la invención, mientras que la Figura 13B muestra un diagrama esquemático 1350 resumido de soluciones de amplificación en puente de acuerdo con un método estándar. En una realización, el citrato trisódico se usa en una concentración final de 1 mM en formamida.

15 La amplificación en fase sólida se puede realizar de manera eficiente en una celda de flujo, ya que es una característica de la presente descripción que los cebadores, el molde y los productos amplificados (de extensión) pueden permanecer inmovilizados en el soporte sólido durante la amplificación. Por consiguiente, se describe un aparato que puede permitir que los ácidos nucleicos inmovilizados se amplifiquen isotérmicamente. Un aparato también puede incluir una fuente de reactivos y medios de detección para detectar una señal que puede generarse una vez que se hayan aplicado uno o más reactivos a las moléculas de ácido nucleico inmovilizadas. También puede disponerse un aparato con una superficie que comprende moléculas de ácido nucleico inmovilizadas en forma de colonias.

20 Se proporciona un aparato que comprende uno o más de los siguientes:

a) al menos una entrada;

b) medios para inmovilizar cebadores en una superficie (aunque esto no es necesario si ya se proporcionan cebadores inmovilizados);

25 c) medios para la amplificación sustancialmente isotérmica de ácidos nucleicos (por ejemplo, solución desnaturalizante, solución de hibridación, solución de extensión, una o más soluciones de lavado);

d) al menos una salida; y

e) medios de control para coordinar las diferentes etapas requeridas para el método de la presente invención.

30 Cuando se elimina un volumen de una solución adecuada particular en contacto con un soporte sólido, se la reemplaza con un volumen similar de una solución igual o diferente. Por lo tanto, las soluciones aplicadas a una celda de flujo a través de una entrada pueden eliminarse a través de una salida mediante un procedimiento de intercambio de soluciones.

35 Resulta deseable que un medio para detectar una señal tenga una resolución suficiente para permitirle distinguir entre señales generadas por colonias diferentes. Los instrumentos que son útiles para detectar una señal fluorescente se describen, por ejemplo, en los documentos WO 2007/123744, US 2010/0111768 y en la patente de EE. UU. N.º 7.329.860.

Los aparatos de la presente invención se proporcionan preferiblemente en forma automatizada de modo que una vez que se activan, pueden repetirse automáticamente las etapas individuales del procedimiento.

**REIVINDICACIONES**

1. Un método para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en la amplificación en puente de un molde de ácido nucleico bicatenario en una superficie que comprende:
- 5 a) desnaturalizar el molde de ácido nucleico con una primera solución que comprende formamida y un solo tipo de dNTP y/o un solo tipo de rNTP, o una mezcla de diferentes dNTP y/o rNTP para producir cadenas de molde de ácido nucleico monocatenario;
- b) opcionalmente, reemplazar la primera solución por una segunda solución;
- c) hibridar las cadenas del molde de ácido nucleico monocatenario a cebadores de oligonucleótidos unidos a la superficie; y
- 10 d) reemplazar la primera o la segunda solución por una solución que comprende una polimerasa, por lo que los cebadores de oligonucleótidos unidos a la superficie se extienden completamente;
- en donde se reduce el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en la amplificación del molde de ácido nucleico bicatenario.
2. El método de acuerdo con la reivindicación 1, en donde al menos una de la primera y segunda soluciones comprende al menos un aditivo, opcionalmente en donde:
- 15 i) el aditivo comprende un agente quelante;
- ii) el aditivo comprende al menos un citrato, en donde opcionalmente el aditivo comprende al menos un citrato seleccionado de entre: citrato monosódico, citrato disódico, citrato trisódico, citrato de potasio;
- iii) el aditivo comprende EDTA;
- 20 iv) el aditivo comprende betaína;
- v) el aditivo comprende DMSO.
3. El método de acuerdo con la reivindicación 1 o 2, en donde se genera un agrupamiento de cadenas de ácido nucleico sustancialmente idénticas.
4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde la segunda solución comprende agua y/o una solución de premezcla.
- 25 5. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en donde la segunda solución comprende menos de aproximadamente 100 mM de sal, opcionalmente en donde la segunda solución está sustancialmente libre de sal.
6. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en donde la solución que comprende la polimerasa comprende, además, una mezcla de diferentes dNTP y/o rNTP.
- 30 7. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en donde la solución que comprende la polimerasa comprende menos de aproximadamente 100 mM de sal.
8. El método de acuerdo con la reivindicación 7, en donde la solución que comprende la polimerasa está sustancialmente libre de cualquier sal.
- 35 9. Un sistema para la amplificación en puente que comprende un aparato que tiene al menos una entrada y al menos una salida; y medios para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en una amplificación en puente de un molde de ácido nucleico bicatenario en una superficie en donde el aparato está programado para coordinar las etapas del método de acuerdo con una cualquiera de las reivindicaciones 1 a 8.
- 40 10. El sistema de acuerdo con la reivindicación 9, en donde el aparato comprende medios para inmovilizar cebadores sobre una superficie.
11. El sistema de acuerdo con una cualquiera de las reivindicaciones 9 o 10, en donde el aparato comprende una celda de flujo y las soluciones se aplican a través de la entrada y se eliminan a través de la salida mediante un procedimiento de intercambio de solución.
- 45 12. El sistema de acuerdo con la reivindicación 11, en donde el sistema comprende medios de detección para detectar una señal fluorescente.
13. Un kit para reducir el sesgo de GC dependiente de la densidad y/o del daño del ácido nucleico en una amplificación en puente de un molde de ácido nucleico bicatenario en una superficie que comprende:

una primera solución desnaturizante para producir cadenas molde de ácido nucleico monocatenario que comprenden formamida y al menos uno de un grupo que comprende i) un solo tipo de dNTP, ii) un solo tipo de rNTP, iii) una mezcla de dNTP, y iv) una mezcla de rNTP; opcionalmente al menos un aditivo; y una polimerasa.

- 5 14. El kit de acuerdo con la reivindicación 13, en donde el al menos un aditivo comprende: un agente quelante, EDTA, un citrato, una sal de citrato, citrato monosódico, citrato disódico, citrato trisódico, citrato de potasio, betaína y/o DMSO.
15. El kit de acuerdo con la reivindicación 14, que comprende, además, instrucciones para realizar el método de acuerdo con cualquiera de las reivindicaciones 1 a 8.

10

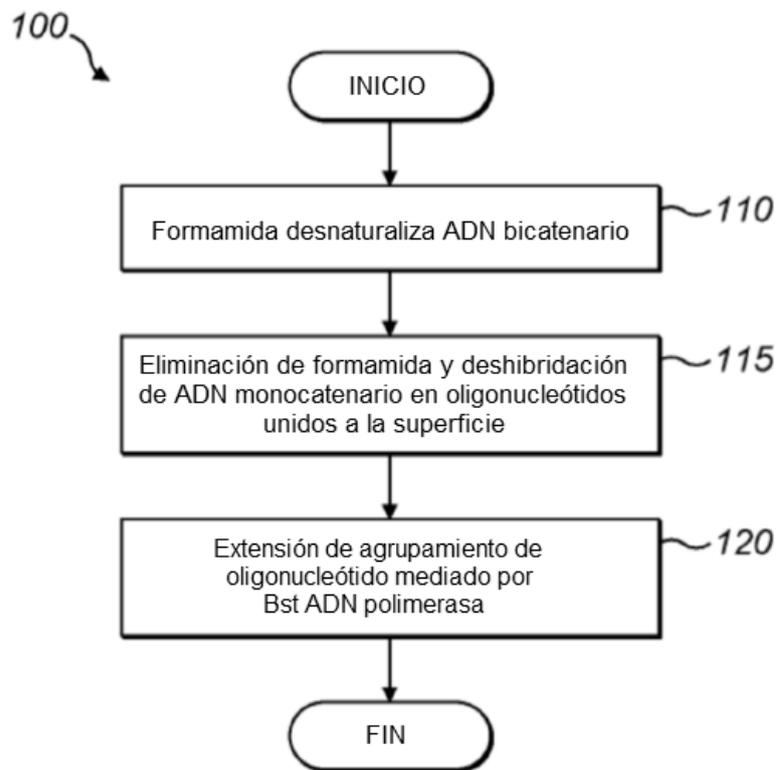


FIG.1

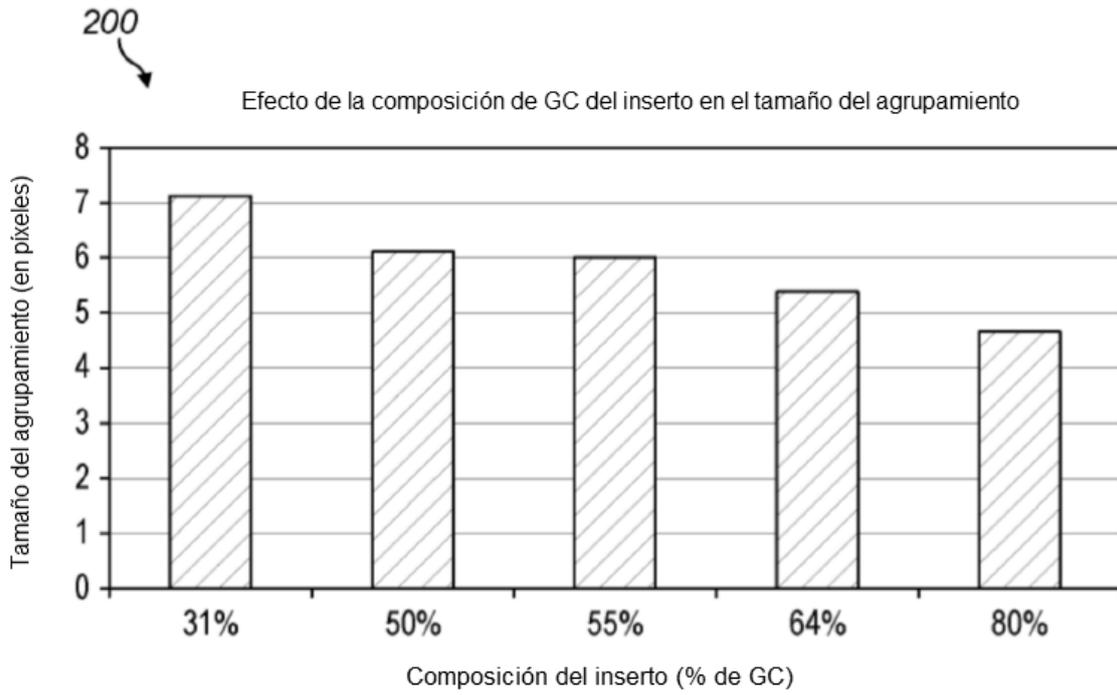


FIG.2A

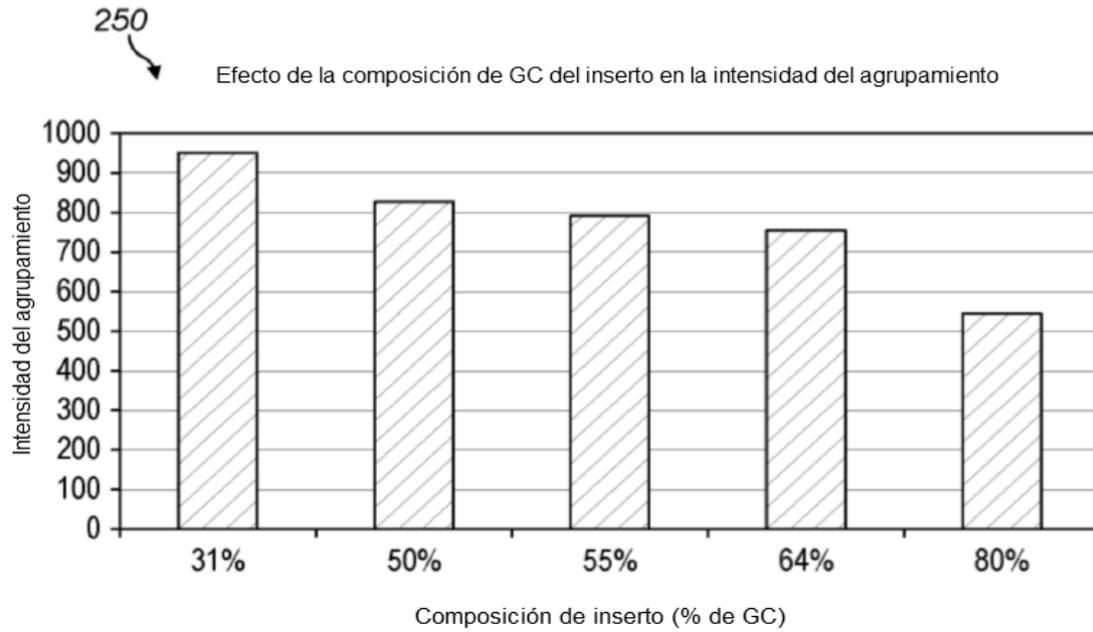


FIG.2B

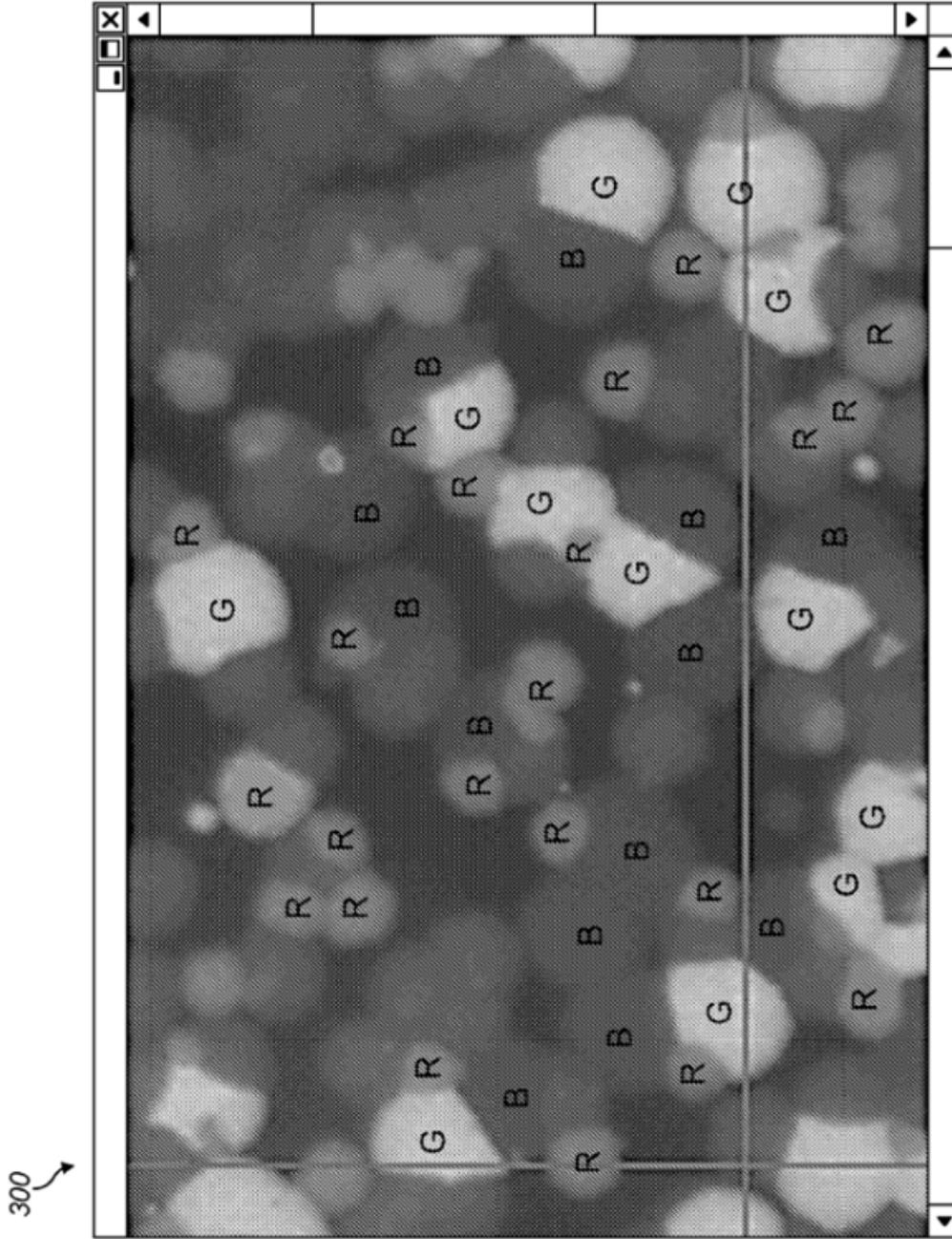


FIG.3

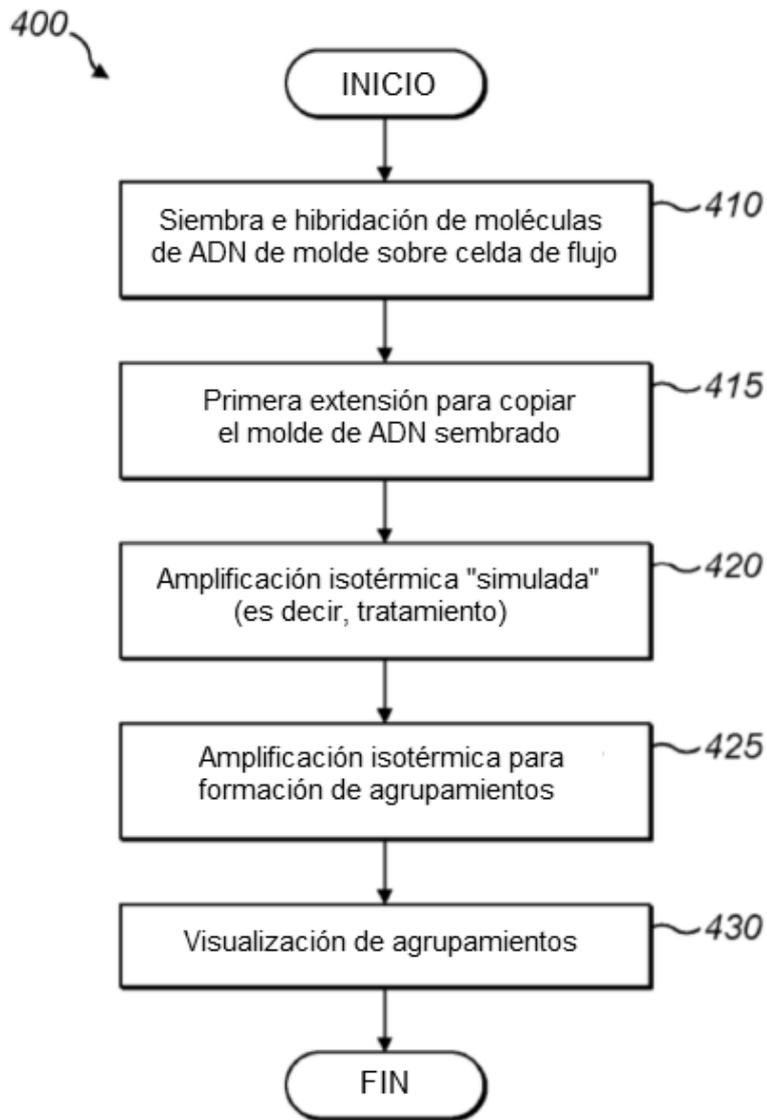


FIG.4

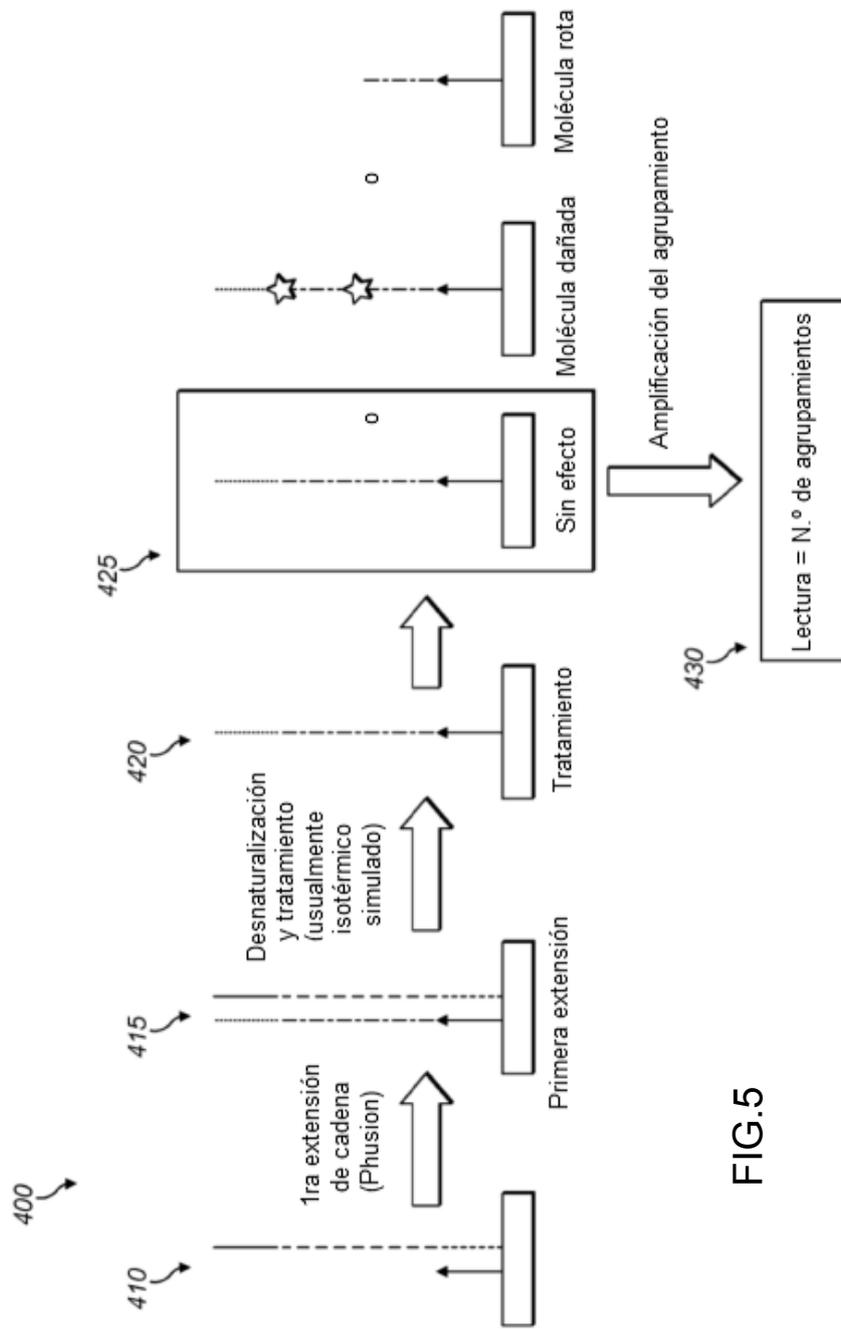


FIG.5

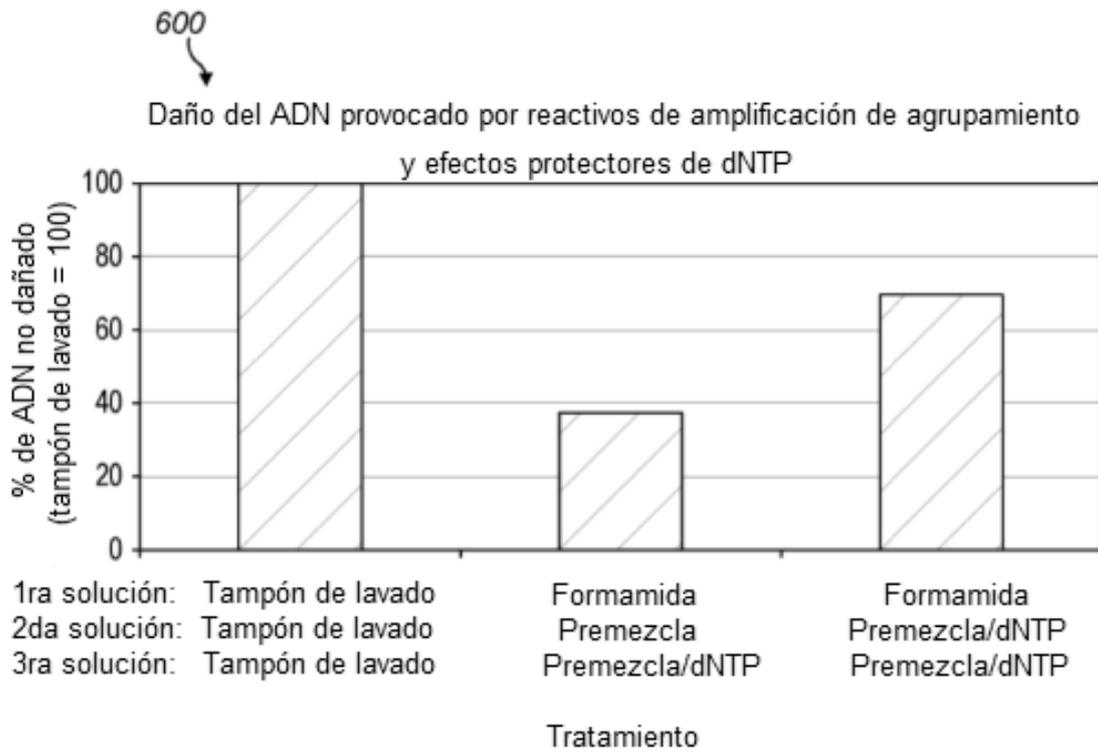
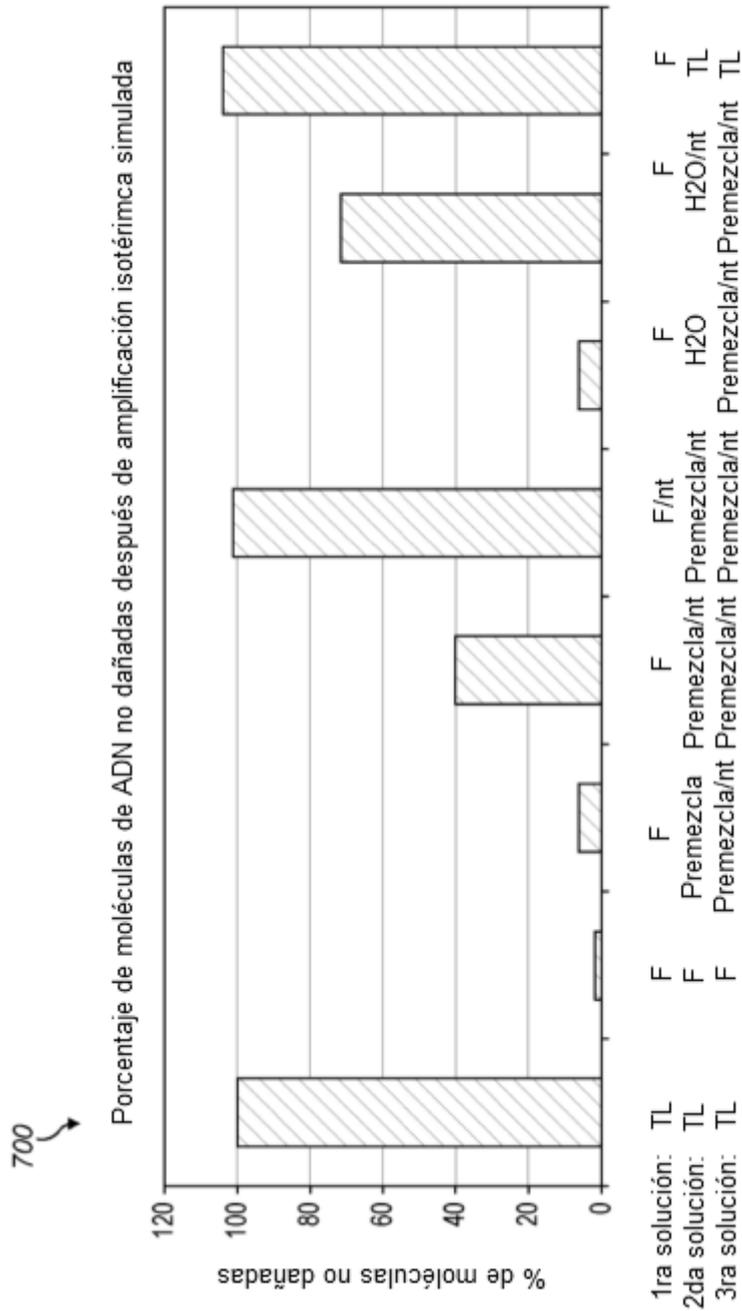


FIG.6



Tratamiento

FIG.7

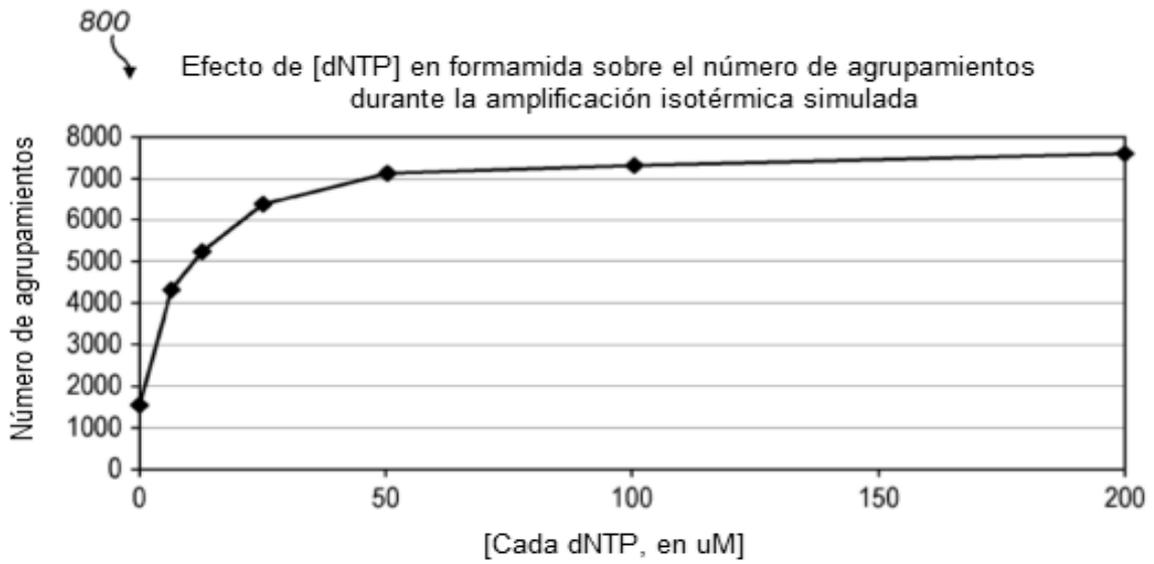


FIG.8

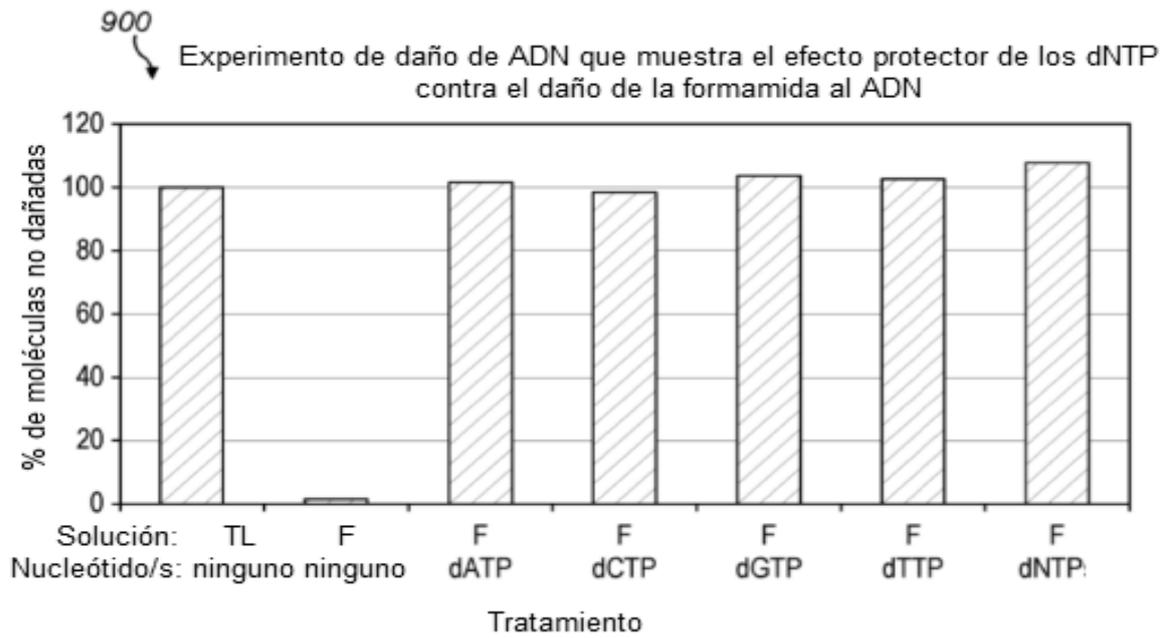


FIG.9

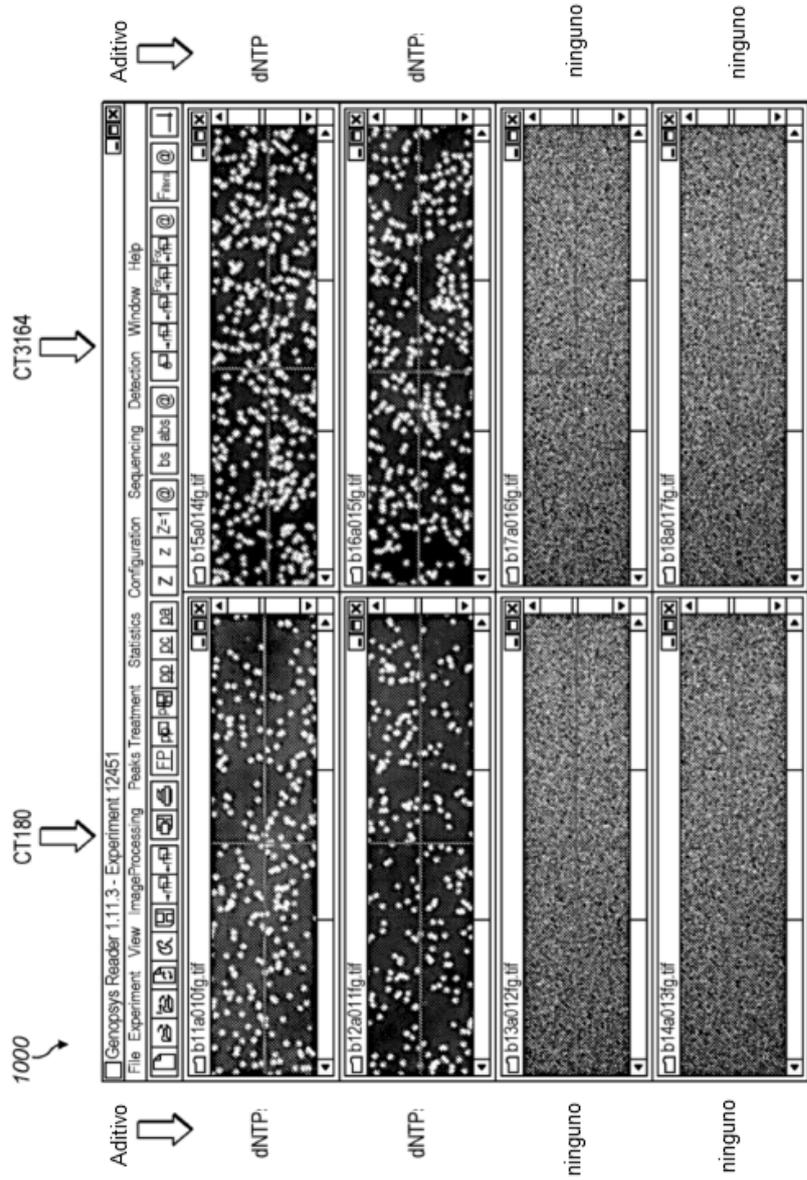


FIGURA 10

1100 - Resumen de resultados de carril

Información de carril		Media de placa +/- desviación estándar para el carril									
Carril	Rendimiento del carril (kb)	Agrupamientos (originales)	Agrupamientos (PF)	Int. del 1.º ciclo (PF)	% de int. tras 20 ciclos (PF)	% de agrupamientos PF	% de alineación (PF)	Puntuación de alineación (PF)	% de tasa de errores (PF)		
1	94438	167783 +/- 9102	138067 +/- 32957	202 +/- 51	106,98 +/- 175,77	82,34 +/- 19,53	88,74 +/- 21,66	116,15 +/- 33,44	0,71 +/- 1,25		1,5 pM sin premezcla
2	190665	405672 +/- 3215	278750 +/- 10091	143 +/- 12	78,18 +/- 4,51	68,71 +/- 2,46	95,59 +/- 0,18	132,15 +/- 0,93	0,15 +/- 0,09		7,5 pM sin premezcla
3	224249	447815 +/- 17841	311457 +/- 14663	192 +/- 16	80,63 +/- 2,96	69,66 +/- 4,31	95,58 +/- 0,00	132,47 +/- 0,00	0,10 +/- 0,01		7,5 pM sin premezcla
4	254968	472258 +/- 9014	354123 +/- 14414	211 +/- 20	79,35 +/- 4,02	75,00 +/- 2,97	95,60 +/- 0,00	132,52 +/- 0,29	0,09 +/- 0,02		7,5 pM protocolo de agua
5	116525	171872 +/- 12264	161840 +/- 11748	274 +/- 16	83,96 +/- 5,08	94,16 +/- 1,23	95,69 +/- 0,21	133,39 +/- 0,28	0,05 +/- 0,00		1,5 pM protocolo de dNTP
6	252790	475383 +/- 17830	351097 +/- 16518	236 +/- 15	79,62 +/- 4,25	73,92 +/- 3,77	95,56 +/- 0,00	132,43 +/- 0,25	0,09 +/- 0,02		7,5 pM protocolo de dNTP
7	257008	470337 +/- 11913	356956 +/- 15717	206 +/- 22	81,19 +/- 3,19	75,92 +/- 3,39	95,63 +/- 0,13	132,71 +/- 0,00	0,09 +/- 0,01		7,5 pM protocolo de agua
8	256960	479137 +/- 16277	356889 +/- 13421	264 +/- 28	81,45 +/- 5,03	74,56 +/- 3,54	95,57 +/- 0,15	132,37 +/- 0,38	0,09 +/- 0,02		7,5 pM protocolo de dNTP

FIG.11

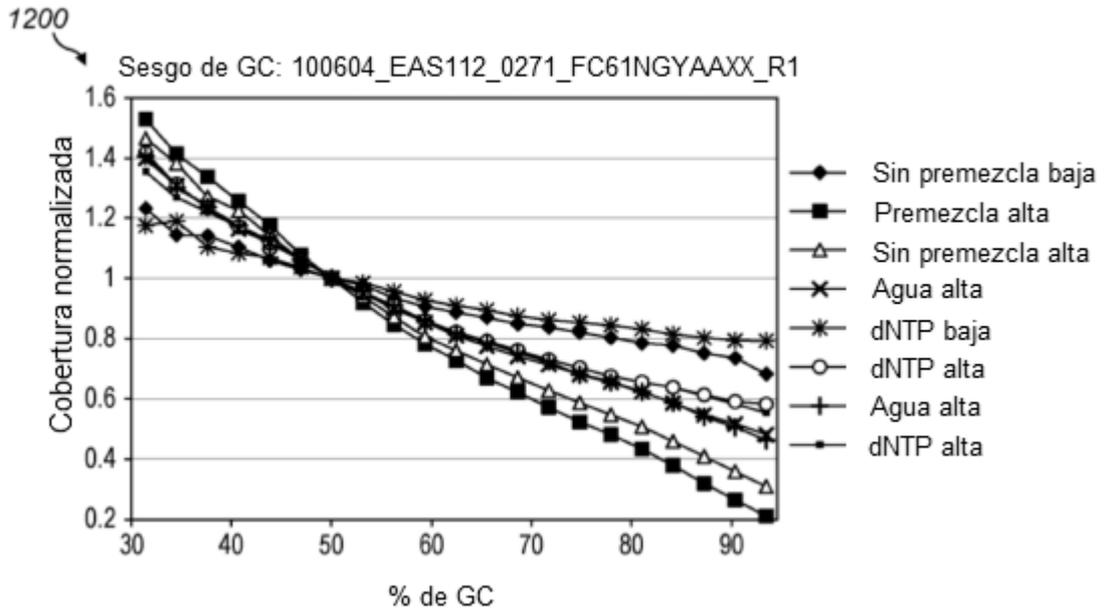


FIG.12A

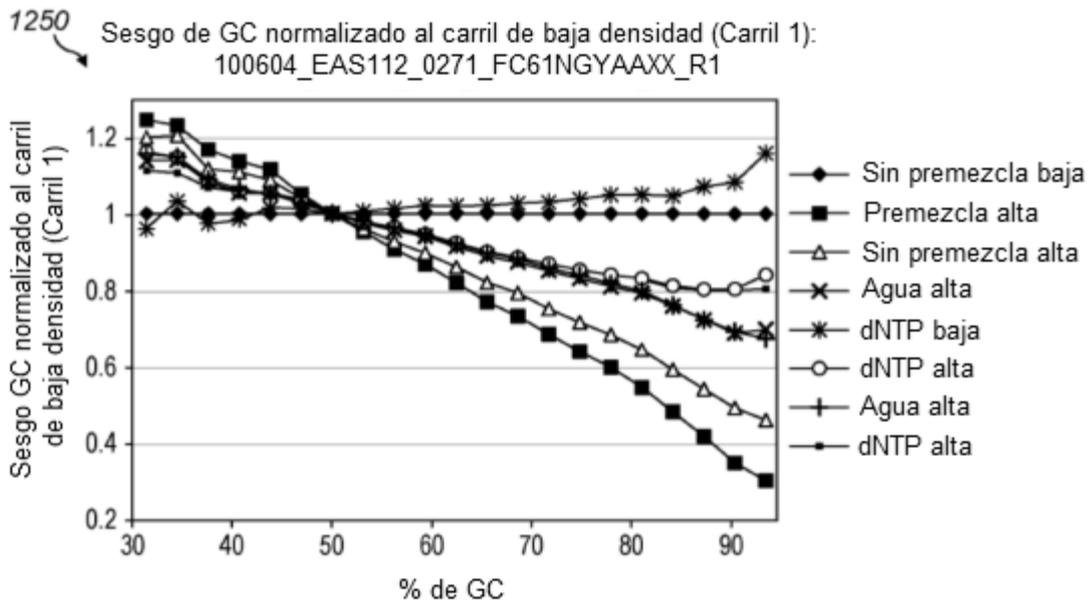


FIG.12B

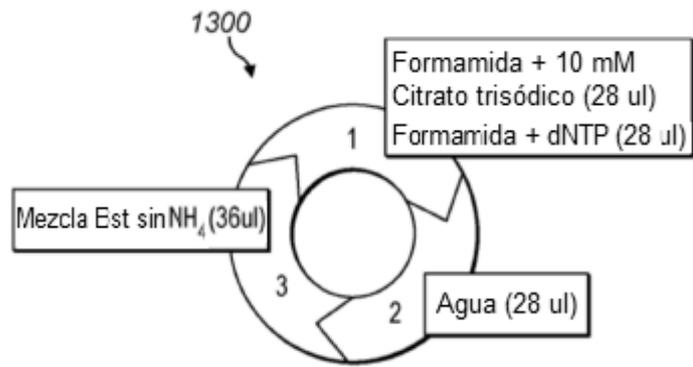


FIG.13A

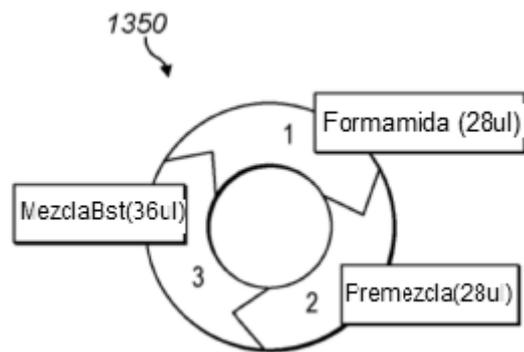


FIG.13B

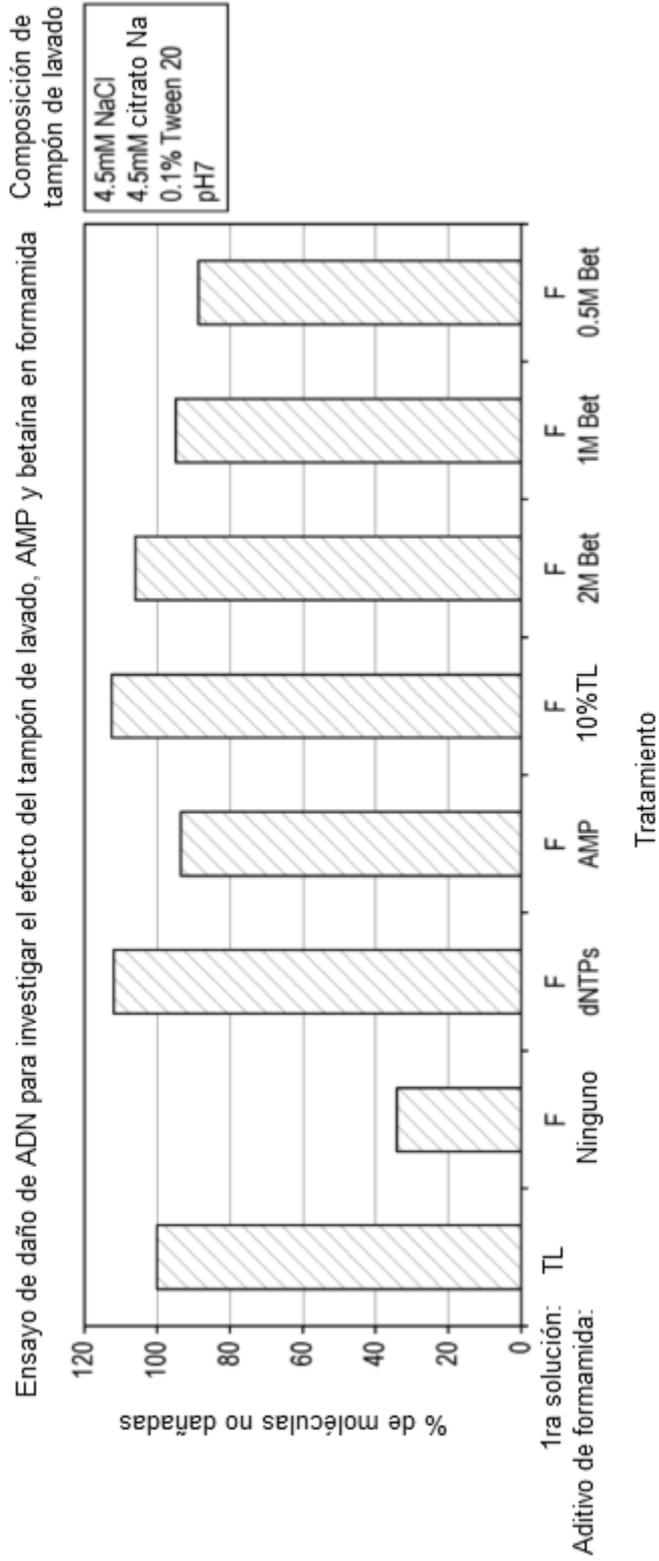


FIG.14

Ensayo de daño de ADN para investigar el efecto protector de los componentes individuales del tampón de lavado

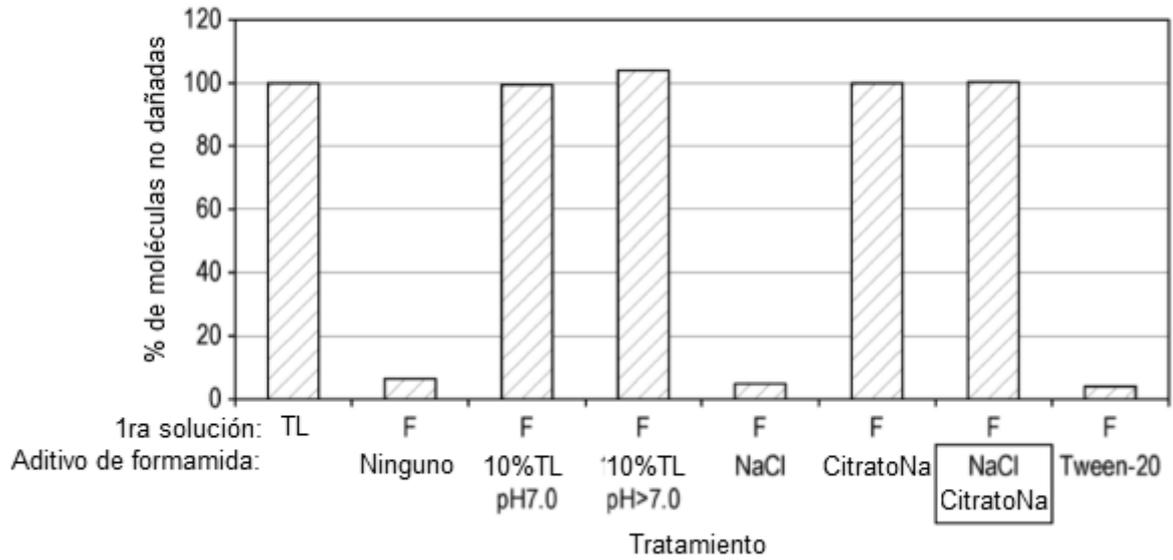


FIG.15

Número de moldes CT180 no dañados como una función de [citrato de Na] en formamida

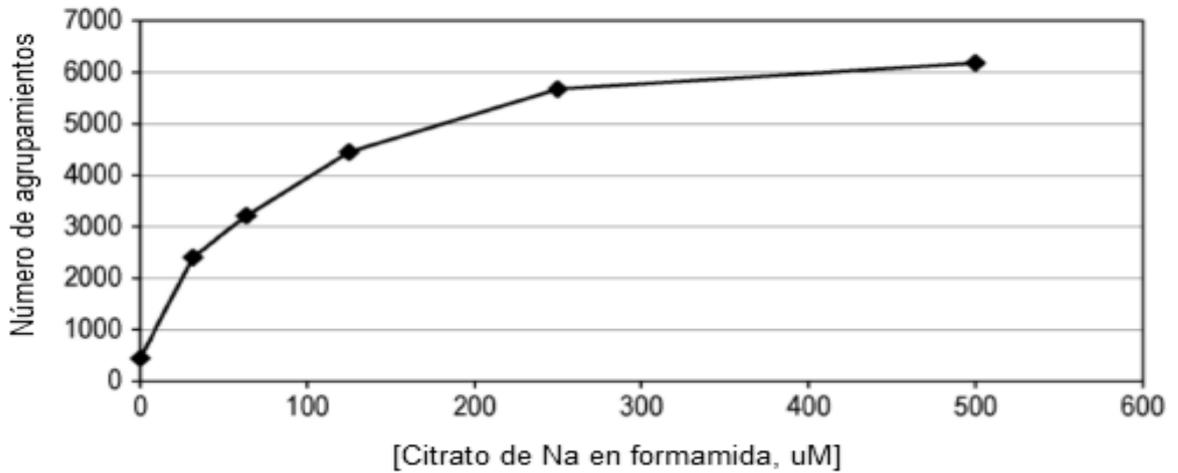


FIG.16

Carril	1.º solución (120 µl)	2.º solución (28 µl)	3.º solución (36 µl)
1	tampón de lavado	tampón de lavado	tampón de lavado
2	formamida	agua	premezcla/dNTP
3	formamida	agua	premezcla sin NH <sub>4</sub> /dNTP
4	formamida + citrato de Na 1 mM	agua	premezcla sin NH <sub>4</sub> /dNTP
5	formamida + 200 µM de cada nucleótido	agua	premezcla sin NH <sub>4</sub> /dNTP
6	formamida + 200 µM de cada nucleótido	agua	premezcla/dNTP
7	formamida + 200 µM de cada nucleótido + MgSO <sub>4</sub> 3 mM	agua	premezcla sin NH <sub>4</sub> /dNTP
8	formamida + EDTA 1 mM	agua	premezcla/dNTP

FIG.17A

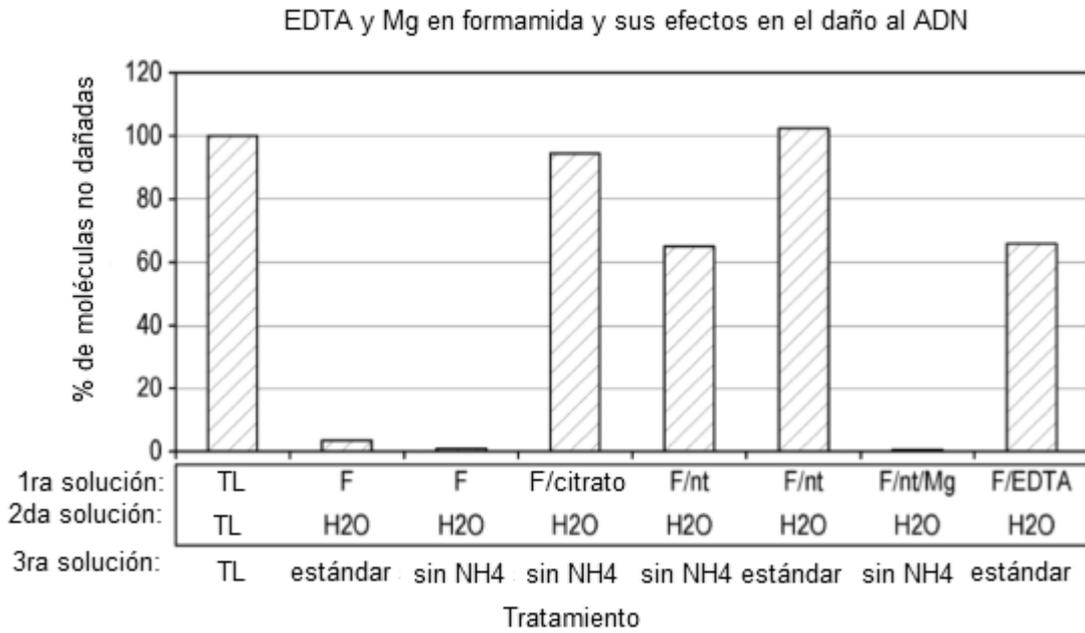


FIG.17B

Resumen de resultados de carril: Lectura 1

Carril	Medida de placa +/- desviación estándar para el carril											% de tasa de errores (PF)
	Información de carril	Rendimiento del carril (kb)	Agrupamientos (originales)	Agrupamientos (PF)	Int. del 1.º ciclo (PF)	% de int. tras 20 ciclos (PF)	% de agrupamientos PF	% de alineación (PF)	Puntuación de alineación (PF)	de	% de tasa de errores (PF)	
1	61710	90908 +/- 2650	85709 +/- 2979	126 +/- 22	80,63 2,97	+/-	94,27 +/- 1,34	85,45 +/- 0,33	81,76 +/- 1,70		0,30 +/- 0,42	1,0 pM amonio mM
2	179644	301159 +/- 2379	249505 +/- 6820	82 +/- 29	75,40 4,90	+/-	82,84 +/- 1,70	85,69 +/- 0,03	81,54 +/- 0,66		0,29 +/- 0,08	5,5 pM amonio mM
3	186381	308498 12251	+/- 258863 11406	94 +/- 36	76,78 3,49	+/-	83,92 +/- 1,83	85,70 +/- 0,22	81,45 +/- 1,82		0,26 +/- 0,05	5,5 pM amonio mM
4	179460	300064 +/- 4288	249250 +/- 6709	98 +/- 35	78,03 2,51	+/-	83,06 +/- 1,48	85,72 +/- 0,00	82,04 +/- 0,42		0,25 +/- 0,04	5,5 pM amonio mM
5	159158	275462 +/- 3248	221052 +/- 3295	89 +/- 26	77,50 2,04	+/-	80,24 +/- 1,44	85,68 +/- 0,21	81,95 +/- 0,18		0,25 +/- 0,02	5,5 pM amonio mM
6	178529	295354 18927	+/- 247957 14332	82 +/- 27	78,75 2,78	+/-	84,01 +/- 1,75	85,70 +/- 0,00	81,92 +/- 0,37		0,25 +/- 0,04	5,5 pM amonio mM
7	176046	294977 +/- 9470	244509 10690	85 +/- 30	79,06 2,19	+/-	82,88 +/- 1,96	85,71 +/- 0,00	82,03 +/- 0,24		0,25 +/- 0,03	5,5 pM amonio mM

FIG.18A

Curvas de sesgo de GC: 100526\_EAS20\_0170\_FC61RUAAXX\_R1 Pipeline 1.6

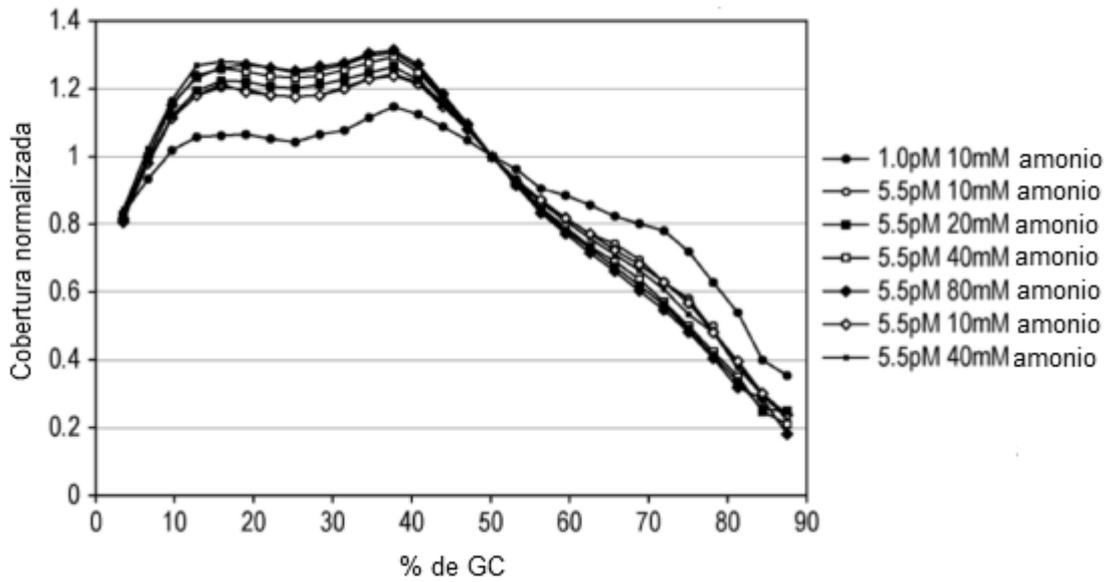


FIG.18B

Curvas de sesgo de GC: 100526\_EAS20\_0170\_FC61LRUAAXX\_R1 Pipeline 1.6

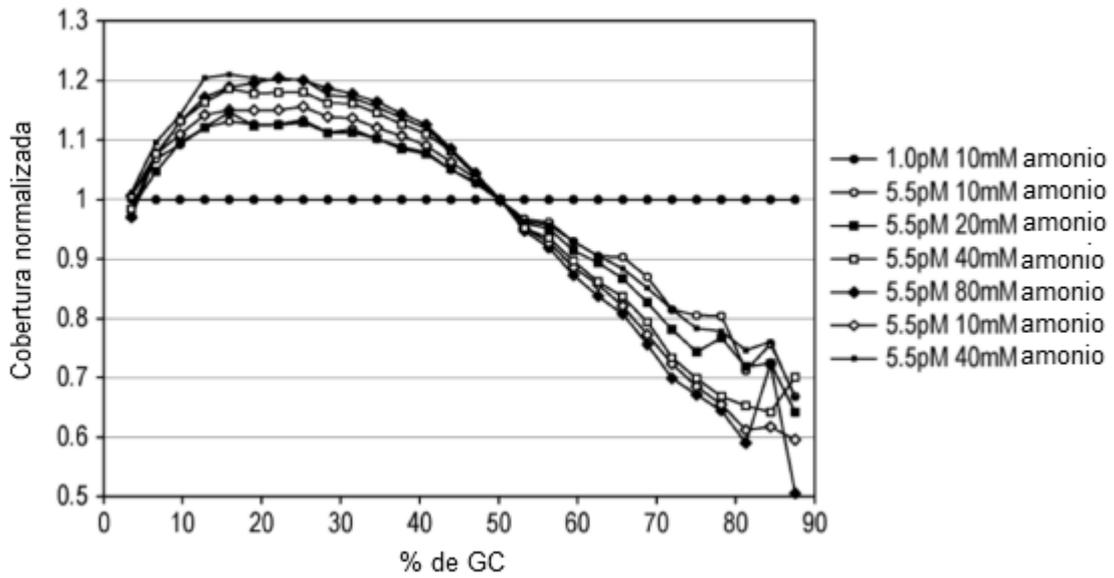


FIG.18C

ES 2 705 423 T3

Carril	1.º solución (28 µl)	2.º solución (28 µl)	3.º solución (36 µl)
1	formamida	Mezcla Bst (amonio 10 mM estándar)	Mezcla Bst (amonio 10 mM estándar)
2	formamida	Mezcla Bst (amonio 10 mM estándar)	Mezcla Bst (amonio 10 mM estándar)
3	formamida + dNTP	agua + dNTP	Mezcla Bst (amonio 10 mM estándar)
4	formamida + dNTP	agua + dNTP	Mezcla Bst (amonio 5,0 mM estándar)
5	formamida + dNTP	agua + dNTP	Mezcla Bst (amonio 2,5 mM estándar)
6	formamida + dNTP	agua + dNTP	Mezcla Bst (amonio 0,0 mM estándar)
7	formamida + betaína 2 M	agua + dNTP	Mezcla Bst (amonio 10 mM estándar)
8	formamida + betaína 1 M	agua + dNTP	Mezcla Bst (amonio 10 mM estándar)

formamida + betaína 2 M = 60 % de formamida + betaína 2 M + H<sub>2</sub>O

formamida + betaína 1 M = 80% de formamida + betaína 1 M +

FIG.19A

FIG.19B

Resumen de resultados de carril

Información de carril		Media de placa +/- desviación estándar para el carril										Método de los dNTP
Carril	Rendimiento del carril (kb)	Agrupamientos (originales)	Agrupamientos (PF)	Int. del 1.º ciclo (PF)	% de int. tras 20 ciclos (PF)	% agrupamientos PF	% de alineación (PF)	Puntuación de alineación (PF)	% de tasa de errores (PF)			
1	69820	101210 +/- 6320	96973 +/- 5937	290 +/- 10	90,74 +/- 1,59	95,83 +/- 1,00	81,26 +/- 0,19	86,23 +/- 0,15	0,16 +/- 0,01	10 mP sin premezcla		
2	267985	471469 +/- 16182	372201 +/- 10617	268 +/- 8	86,62 +/- 1,30	78,97 +/- 1,39	81,42 +/- 0,00	85,59 +/- 0,20	0,20 +/- 0,01	5,5 pM sin premezcla		
3	289332	493599 +/- 12105	401850 +/- 8629	302 +/- 14	86,98 +/- 2,37	81,43 +/- 1,68	81,22 +/- 0,22	85,35 +/- 0,00	0,18 +/- 0,00	5,0 pM amonio 10 mM		
4	283095	483942 +/- 15419	393188 +/- 10969	269 +/- 6	87,13 +/- 1,92	81,27 +/- 1,51	81,20 +/- 0,17	85,31 +/- 0,13	0,18 +/- 0,00	5,0 pM amonio 0 mM		
5	274016	468625 +/- 14523	380578 +/- 8580	248 +/- 10	87,01 +/- 1,82	81,24 +/- 1,45	81,22 +/- 0,12	85,30 +/- 0,12	0,18 +/- 0,00	5,0 pM amonio 2,5 mM		
6	254869	437317 +/- 20015	353985 +/- 15382	210 +/- 7	86,96 +/- 1,80	80,97 +/- 1,51	81,21 +/- 0,12	85,27 +/- 0,00	0,19 +/- 0,00	5,0 pM amonio 0,0 mM		
7	285092	492840 +/- 16785	395961 +/- 9373	278 +/- 12	87,40 +/- 1,64	80,38 +/- 1,55	81,24 +/- 0,00	85,35 +/- 0,24	0,18 +/- 0,00	betaina 2 M en form.		
8	289208	494008 +/- 13950	401678 +/- 9425	316 +/- 27	89,58 +/- 4,30	81,33 +/- 1,54	81,16 +/- 0,19	85,05 +/- 0,70	0,22 +/- 0,12	betaina 1 M en form.		

Curvas de GC normalizadas al carril de baja densidad:  
100604\_EAS20\_0179\_FC61WCBAAAXX\_R1 Pipeline 1.8.0a5

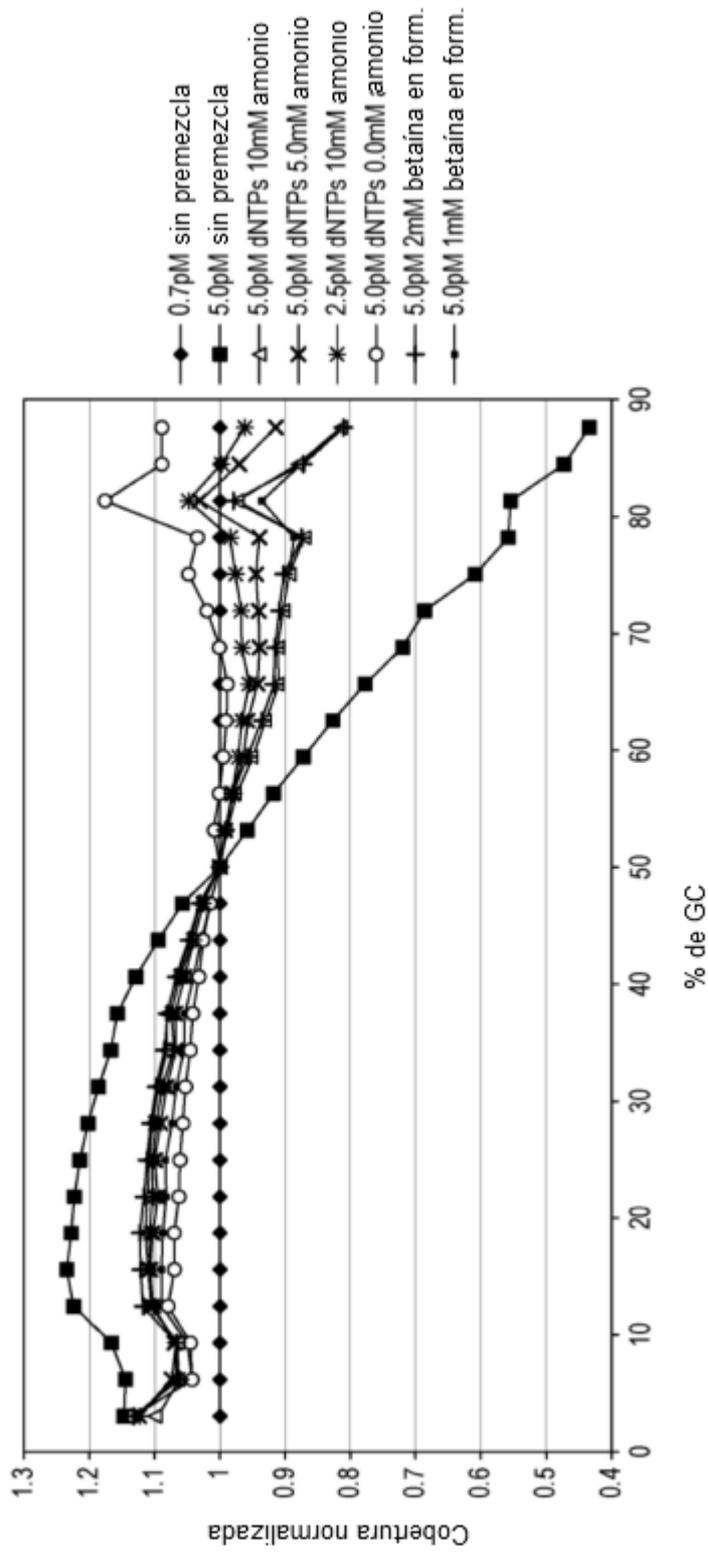


FIGURA 19C

ES 2 705 423 T3

FC61W56AAXX

Carril	1.º solución (28 µl)	2.º solución (28 µl)	3.º solución (36 µl)	N.º de ciclos
1	formamida	mezcla Bst (amonio 10 mM estándar)	mezcla Bst (amonio 10 mM estándar)	26
2	formamida	mezcla Bst (amonio 10 mM estándar)	mezcla Bst (amonio 10 mM estándar)	26
3	formamida + dNTP	agua + dNTP	mezcla Bst (amonio 10 mM estándar)	26
4	formamida + dNTP	agua + dNTP	mezcla Bst (amonio 0,0 mM estándar)	26
5	formamida + citrato de Na 1 mM	agua	mezcla Bst (amonio 0,0 mM estándar)	26
6	formamida + citrato de Na 1 mM	agua + citrato de Na 1 mM	mezcla Bst (amonio 0,0 mM estándar)	26
7	formamida + dNTP	agua + dNTP	mezcla Bst (amonio 0,0 mM estándar)	32
8	formamida + citrato de Na 1 mM	agua	mezcla Bst (amonio 0,0 mM estándar)	32

FIG.20A

Resultados de carril: Lectura 1

Información de carril		Media de placa +/- desviación estándar para el carril									
Carril	Rendimiento del carril (bases M)	Agrupamientos (originales)	Agrupamientos (PF)	Int. del 1.º ciclo (PF)	% de int. tras 20 ciclos (PF)	% agrupamientos PF	% de alineación (PF)	Puntuación de alineación (PF)	% de tasa de errores (PF)		
1	60	82426 +/- 9349	75623 +/- 2408	292 +/- 34	92,30 +/- 17,83	92,55 +/- 7,63	85,42 +/- 0,87	80,79 +/- 4,30	0,41 +/- 0,58	0,7 pM sin premezcla	
2	272	399931 +/- 12424	343252 +/- 10154	269 +/- 9	86,68 +/- 2,41	85,84 +/- 1,53	85,67 +/- 0,57	81,17 +/- 2,17	0,24 +/- 0,06	5,0 pM sin premezcla	
3	296	422726 +/- 7121	373315 +/- 5079	307 +/- 5	85,60 +/- 2,13	88,32 +/- 0,63	85,74 +/- 0,11	81,82 +/- 0,14	0,21 +/- 0,00	5,0 pM dNTP 10 mM	
4	267	380804 +/- 8469	336566 +/- 7047	224 +/- 5	86,79 +/- 2,19	88,39 +/- 0,38	85,71 +/- 0,00	81,78 +/- 0,00	0,22 +/- 0,01	5,0 M dNTP 0,0 mM	
5	283	403808 +/- 8493	347415 +/- 5896	232 +/- 5	87,47 +/- 2,00	88,52 +/- 0,62	85,71 +/- 0,13	81,80 +/- 0,00	0,21 +/- 0,00	5,0 pM citrato en form.	
6	258	370978 +/- 22164	326232 +/- 16963	214 +/- 13	87,25 +/- 2,49	87,98 +/- 0,73	85,69 +/- 0,20	81,74 +/- 0,21	0,22 +/- 0,00	5,0 pM citrato en form. y H <sub>2</sub> O	sin NH <sub>4</sub>
7	294	415625 +/- 19497	371202 +/- 15753	261 +/- 11	87,36 +/- 1,82	89,33 +/- 0,61	85,66 +/- 0,18	81,67 +/- 0,25	0,22 +/- 0,00	5,0 M dNTP 10 mM hasta completar	
8	293	420844 +/- 8715	369664 +/- 6938	308 +/- 52	86,88 +/- 2,11	87,84 +/- 0,69	85,65 +/- 0,10	81,50 +/- 0,24	0,23 +/- 0,05	5,0 pM citrato en form. hasta completar	

FIG.20B

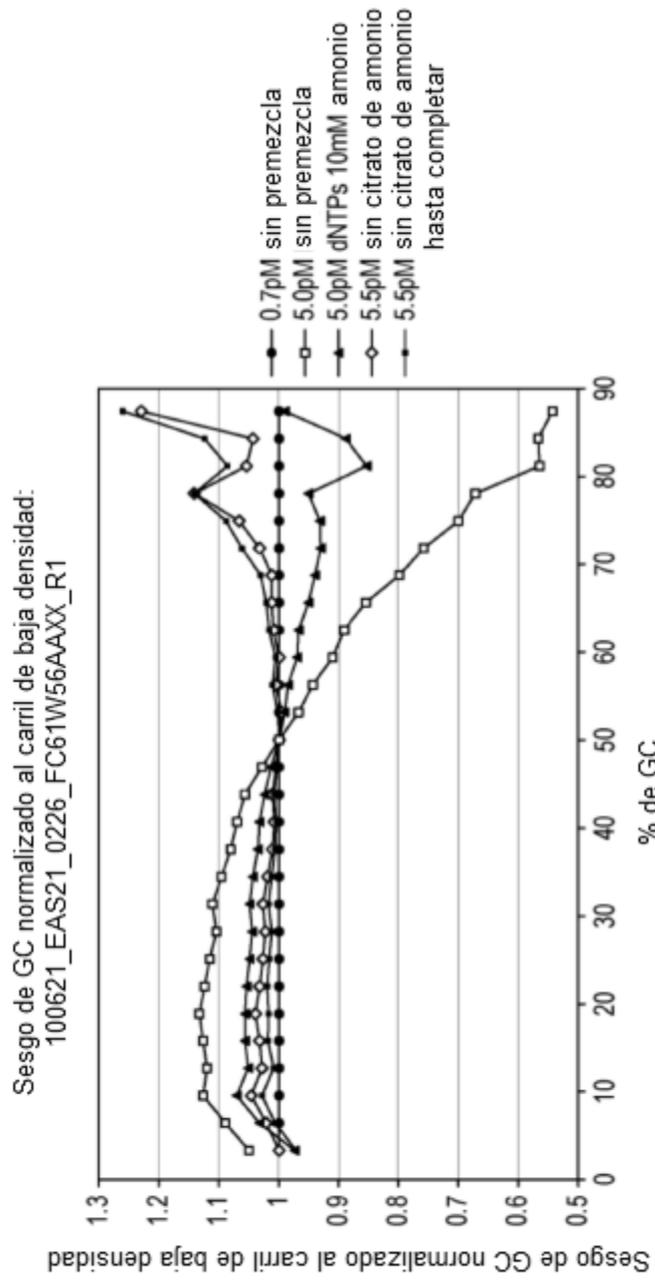


FIG.20C

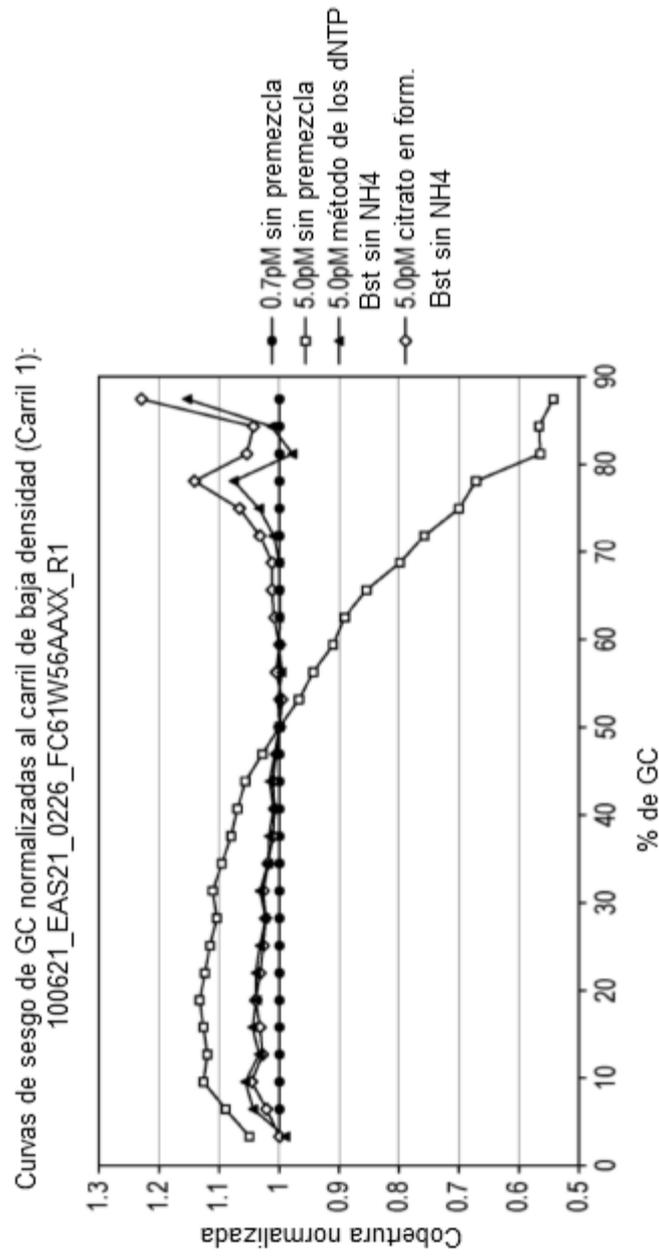


FIG.20D

ES 2 705 423 T3

Carril	1.º solución (28 µl)	2.º solución (28 µl)	3.º solución (36 µl)
1	formamida	premezcla	mezcla Bst (amonio 10 mM estándar)
2	formamida	premezcla	mezcla Bst (amonio 10 mM estándar)
3	formamida + citrato de Na 1 mM	agua	mezcla Bst (sin NH <sub>4</sub> )
4	formamida + citrato de Na 5 mM	agua	mezcla Bst (sin NH <sub>4</sub> )
5	formamida + citrato de Na 1 mM	agua	mezcla Bst (sin NH <sub>4</sub> )
6	formamida + citrato de Na 10 mM	agua	mezcla Bst (sin NH <sub>4</sub> )
7	formamida + betaína 3,2 M	agua	mezcla Bst (sin NH <sub>4</sub> )
8	formamida + betaína 3,2 M	agua	mezcla Bst (sin NH <sub>4</sub> )

FIG.21A

Resumen de resultados de carril

Información de carril		Media de placa +/- desviación estándar para el carril									
Carril	Rendimiento del carril (bases M)	Agrupamientos (originales)	Agrupamientos (PF)	Int. del 1.º ciclo (PF)	% de int. tras 20 ciclos (PF)	% agrupamientos PF	% de alineación (PF)	Puntuación de alineación (PF)	% de tasa de errores (PF)		
1	81887	124672 9803	+/- 113732 +/- 9485	68 +/- 3	70,84 +/- 2,03	91,20 +/- 0,55	80,50 +/- 0,30	78,95 +/- 0,92	0,38 +/- 0,09	0,7 pM amplific. estándar (35 ciclos)	
2	252389	462767 17904	+/- 350541 +/- 4040	85 +/- 6	68,94 +/- 2,29	75,84 +/- 2,51	81,04 +/- 0,19	79,88 +/- 0,70	0,35 +/- 0,05	5,0 pM amplific. estándar (35 ciclos)	
3	271010	485091 20033	+/- 376403 +/- 7387	89 +/- 3	70,40 +/- 1,64	77,69 +/- 2,79	80,81 +/- 0,20	79,72 +/- 0,47	0,31 +/- 0,03	5,0 pM citrato de Na 1 mM (32 ciclos)	
4	275411	483346 19966	+/- 382516 +/- 6915	88 +/- 3	71,56 +/- 1,70	79,23 +/- 2,67	80,90 +/- 0,21	78,85 +/- 0,45	0,29 +/- 0,02	5,0 pM citrato de Na 5 mM (32 ciclos)	
5	82983	123128 11659	+/- 115255 +/- 1256	69 +/- 4	70,08 +/- 2,58	93,58 +/- 0,45	80,67 +/- 0,24	79,56 +/- 0,89	0,31 +/- 0,08	0,7 pM citrato de Na 1 mM (32 ciclos)	
6	281421	484869 16717	+/- 390862 +/- 5813	80 +/- 6	71,59 +/- 2,15	80,69 +/- 2,55	80,85 +/- 0,30	79,95 +/- 0,64	0,28 +/- 0,04	5,0 pM citrato de Na 10 mM (32 ciclos)	
7	272759	487812 15356	+/- 378832 +/- 6871	83 +/- 5	71,53 +/- 2,30	77,75 +/- 3,32	80,94 +/- 0,07	80,06 +/- 0,55	0,28 +/- 0,04	5,0 pM betaina 3,2 M (32 ciclos)	
8	272522	484751 15993	+/- 378503 +/- 9792	89 +/- 8	72,99 +/- 2,35	78,16 +/- 3,07	80,95 +/- 0,12	80,25 +/- 0,53	0,26 +/- 0,04	5,0 pM betaina 3,2 M (32 ciclos)	

FIG.21B

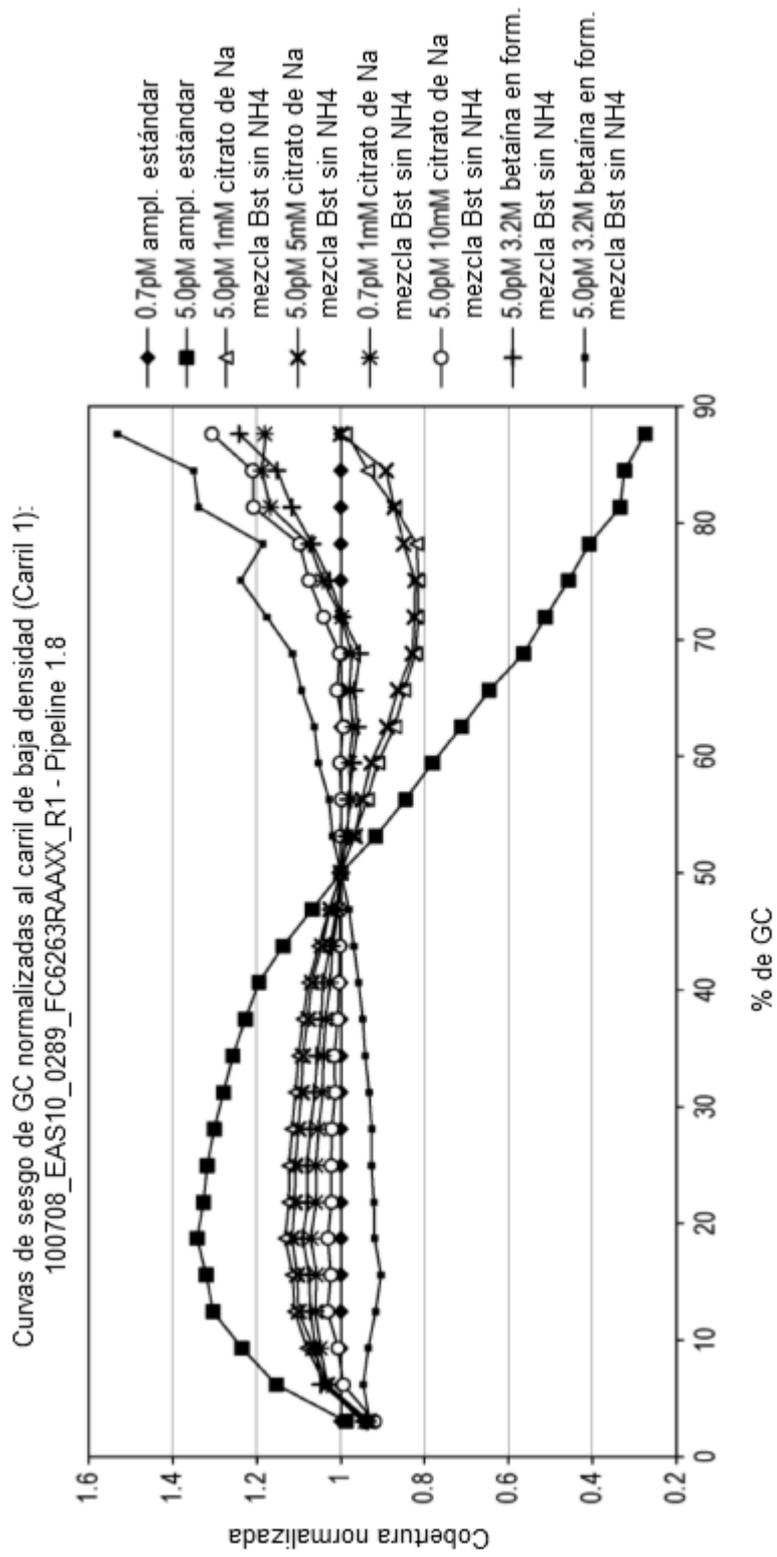


FIG.21C