

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 804 843**

51 Int. Cl.:

C12N 9/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.02.2016 PCT/EP2016/052068**

87 Fecha y número de publicación internacional: **11.08.2016 WO16124543**

96 Fecha de presentación y número de la solicitud europea: **01.02.2016 E 16703461 (0)**

97 Fecha y número de publicación de la concesión europea: **06.05.2020 EP 3253867**

54 Título: **Variantes de polimerasa**

30 Prioridad:

02.02.2015 US 201562111034 P

14.05.2015 US 201562161571 P

09.08.2015 US 201562202895 P

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

09.02.2021

73 Titular/es:

F. HOFFMANN-LA ROCHE AG (100.0%)

Grenzacherstrasse 124

4070 Basel, CH

72 Inventor/es:

AYER, ARUNA;

BIBILLO, ARKADIUSZ;

ECKERT, BARBARA;

SUKO, SHAWN;

ARNOLD, CLEOMA RENETTA;

SCHWAB, CHARLES WAYAN;

THAI, EILEEN;

LEDERMAN, ILYA;

MCGAW, COLIN ALEXANDER;

SCHULTZ, TYLER O'BRIEN;

BOENITZ-DULAT, MARA;

WOERSDOERFER, BIGNA y

WUNDERLICH, DAVID DANIEL

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 804 843 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

Variantes de polimerasa

5 SOLICITUDES RELACIONADAS

La presente solicitud es una continuación por partes de la solicitud de EE. UU. 15/012.317 presentada el 1 de febrero de 2016, titulada "Polymerase Variants", que reivindica prioridad a la solicitud de patente provisional de EE. UU. 62/161.571 que se presentó el 14 de mayo de 2015, titulada "Polymerase Variants", la solicitud de patente provisional de EE. UU. 62/202.895 que se presentó el 9 de agosto de 2015, titulada "Polymerase Variants" y la solicitud de EE. UU. 15/151.364 que se presentó el 10 de mayo de 2016, titulada "Polymerase Variants and Uses Thereof".

15 LISTADO DE SECUENCIAS

La presente solicitud contiene un listado de secuencias que se ha presentado electrónicamente en formato ASCII. Dicha copia ASCII, creada el 10 de mayo de 2016, se llama 20-04338.521WO2_SL.txt y tiene un tamaño de 17.989 octetos.

20 CAMPO TÉCNICO

En el presente documento se proporcionan, entre otras cosas, ADN polimerasas modificadas que contienen alteraciones de aminoácidos basadas en mutaciones identificadas en experimentos de evolución dirigida diseñados para seleccionar enzimas que sean más adecuadas para aplicaciones en tecnologías de ADN recombinante.

25 ANTECEDENTES

Las ADN polimerasas son una familia de enzimas que usan ADN monocatenario como molde para sintetizar la hebra de ADN complementaria. En particular, las ADN polimerasas pueden añadir nucleótidos libres al extremo 3' de una hebra recién formada, dando como resultado el alargamiento de la nueva hebra en sentido de 5' a 3'. La mayoría de las ADN polimerasas son proteínas multifuncionales que poseen actividades tanto polimerizantes como exonucleolíticas. Por ejemplo, muchas ADN polimerasas tienen actividad exonucleasa 3'→5'. Estas polimerasas pueden reconocer un nucleótido incorporado incorrectamente y la actividad exonucleasa 3'→5' de la enzima permite que se escinda el nucleótido incorrecto (esta actividad es conocida como corrección de errores). Después de la escisión de nucleótidos, la polimerasa puede reinsertar el nucleótido correcto y la replicación puede continuar. Muchas ADN polimerasas también tienen actividad exonucleasa 5'→3'.

Las polimerasas han encontrado uso en aplicaciones de ADN recombinante, incluyendo la secuenciación de nanoporos. Sin embargo, una cadena de ADN se mueve rápidamente a una tasa de 1 μs a 5 μs por base a través del nanoporo. Esto hace que el registro sea difícil y propenso al ruido de fondo, no pudiéndose obtener una resolución de un único nucleótido. Por lo tanto, el uso de marcas detectables en nucleótidos se puede usar en la secuenciación de una hebra de ADN o fragmento de la misma. Por tanto, no solo existe una necesidad de controlar la tasa a la que se secuencia el ADN, sino también proporcionar polimerasas que tengan propiedades mejoradas (en relación con la enzima natural), tales como la incorporación de nucleótidos modificados, por ejemplo, nucleótidos polifosfato con o sin marcas.

BREVE SUMARIO DE LA INVENCION

La presente invención proporciona ADN polimerasas modificadas (por ejemplo, mutantes) basadas en experimentos de evolución dirigida diseñados para seleccionar mutaciones que confieran fenotipos ventajosos en condiciones usadas en aplicaciones industriales o de investigación, por ejemplo, catalizando la incorporación de nucleótidos polifosfato modificados, por ejemplo, nucleótidos marcados, en altas concentraciones de sal.

En un aspecto, existe una polimerasa variante que comprende al menos una alteración en la posición S366 A/L y opcionalmente en una posición correspondiente a H223, N224, Y225, H227, I295, Y342, T343, I357, S360, L361, I363, S365, Y367, P368, D417, E475, Y476, F478, K518, H527, T529, M531, N535, G539, P542, N545, Q546, A547, L549, I550, N552, G553, F558, A596, G603, A610, V615, Y622, C623, D624, I628, Y629, R632, N635, M641, A643, I644, T647, I648, T651, I652, K655, W656, D657, V658, H660, F662 y L690 de la SEQ ID NO:2 (Pol6 (con marca His)).

En un modo de realización, se proporciona una ADN polimerasa modificada que tiene una actividad ADN polimerasa que comprende una secuencia de aminoácidos que tiene al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de identidad de secuencia con respecto a la secuencia de aminoácidos como se expone en la SEQ ID NO: 1 o 2.

En un segundo modo de realización, se proporciona una ADN polimerasa modificada que tiene una actividad ADN

5 polimerasa que comprende una secuencia de aminoácidos que tiene al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de identidad de secuencia con respecto a la secuencia de aminoácidos como se expone en SEQ ID NO: 1 o 2 que tiene una o más sustituciones aminoacídicas seleccionadas del grupo que consiste en H223, N224, Y225, H227, I295, Y342, T343, I357, S360, L361, I363, S365, S366, Y367, P368, D417, E475, Y476, F478, K518, H527, T529, M531, N535, G539, P542, N545, Q546, A547, L549, I550, N552, G553, F558, A596, G603, A610, V615, Y622, C623, D624, I628, Y629, R632, N635, M641, A643, I644, T647, I648, T651, I652, K655, W656, D657, V658, H660, F662 y L690 y combinaciones de las mismas. En otro modo de realización, las una o más sustituciones aminoacídicas se seleccionan de H223A, N224Y/L/Q/M/R/K, Y225L/T/I/F/A/M, H227P/E/F/Y, I295W/F/M/E, Y342L/F, T343N/F, I357G/L/Q/H/W/M/A/E/Y/P, S360G/E/Y, L361M/W/V/F, I363V/A/R/M/W, S365Q/W/M/A/G, S366A/L, Y367L/E/M/P/N/F, P368G, D417P, E475D, Y476V, F478L, K518Q, H527W/R/L/Y/T/M, T529M/F, M531H/Y/A/K/R/W/T/L/V/G, N535L/Y/M/K/I/R/W/Q, G539Y/F, P542E/S/G, N545K/D/S/L/R/Q/W, Q546W/F, A547M/Y/W/F/V/S, L549Q/Y/H/G/R/K, I550A/W/T/G/F/S, N552L/M/S/T, G553S/T/E/Q/K/R/M, F558P/T, A596S, G603T/A/L, A610T/E, V615A/T, Y622A/M, C623G/S/Y/F, D624F/K, I628Y/V/F/L/M, Y629W/H/M/K, R632L/C, N635D, M641L/Y, A643L, I644H/M/Y, T647G/A/E/K/S/Y, I648K/R/V/N/T/L, T651Y/F/M, I652Q/G/S/N/F/T/A/L/E/K/M, K655G/F/E/N/Q/M/A, W656E, D657R/P/A/E, V658L, H660A/Y, F662I/L, L690M y combinaciones de las mismas. La ADN polimerasa modificada que tiene una o más sustituciones aminoacídicas tiene una característica alterada seleccionada de actividad enzimática, fidelidad, procesividad, tasa de alargamiento, exactitud de secuenciación, capacidad de lectura continua larga, estabilidad y solubilidad en relación con la polimerasa original. En un modo de realización, la característica alterada es la actividad enzimática. En un modo de realización, la característica alterada es la fidelidad. En un modo de realización, la característica alterada es la procesividad. En un modo de realización, la característica alterada es la tasa de alargamiento. En un modo de realización, la característica alterada es la estabilidad. En un modo de realización, la característica alterada es la solubilidad. En un modo de realización, la característica alterada es la capacidad de unir y/o incorporar nucleótidos polifosfato, por ejemplo, un nucleótido tetrafosfato, pentafosfato, hexafosfato, heptafosfato u octofosfato.

25 En un tercer modo de realización, se proporciona una ADN polimerasa modificada que tiene una característica alterada seleccionada de actividad enzimática, fidelidad, procesividad, tasa de alargamiento, estabilidad o solubilidad, en comparación con la SEQ ID NO: 1 o 2. En un modo de realización, la característica alterada es la actividad enzimática. En un modo de realización, la característica alterada es la fidelidad. En un modo de realización, la característica alterada es la procesividad. En un modo de realización, la característica alterada es la tasa de alargamiento. En un modo de realización, la característica alterada es la estabilidad. En un modo de realización, la característica alterada es la solubilidad.

35 En un cuarto modo de realización, se proporciona una ADN polimerasa modificada que tiene una actividad ADN polimerasa que comprende una secuencia de aminoácidos que tiene al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de identidad de secuencia con respecto a la secuencia de aminoácidos como se expone en SEQ ID NO: 1, en la que la secuencia de aminoácidos incluye una o más sustituciones aminoacídicas, seleccionándose dichas sustituciones del grupo que consiste en H223A, N224Y/L/Q/M/R/K, Y225L/T/I/F/A, H227P/E/F/Y, I295W/F/M/E, Y342L/F, T343N/F, I357G/L/Q/H/W/M/A/E/Y/P, S360G/E/Y, L361M/W/V/F, I363V/A/R/M/W, S365Q/W/M/A/G, S366A/L, Y367L/E/M/P/N/F, P368G, D417P, E475D, Y476V, F478L, K518Q, H527W/R/L/Y/T/M, T529M/F, M531H/Y/A/K/R/W/T/L/V/G, N535L/Y/M/K/I/R/W/Q, G539Y/F, P542E/S/G, N545K/D/S/L/R/Q/W, Q546W/F, A547M/Y/W/F/V/S, L549Q/Y/H/G/R/K, I550A/W/T/G/F/S, N552L/M/S/T, G553S/T/E/Q/K/R/M, F558P/T, A596S, G603T/A/L, A610T/E, V615A/T, Y622A/M, C623G/S/Y/F, D624F/K, I628Y/V/F/L/M, Y629W/H/M/K, R632L/C, N635D, M641L/Y, A643L, I644H/M/Y, T647G/A/E/K/S/Y, I648K/R/V/N/T/L, T651Y/F/M, I652Q/G/S/N/F/T/A/L/E/K/M, K655G/F/E/N/Q/M/A, W656E, D657R/P/A/E, V658L, H660A/Y, F662I/L, L690M y combinaciones de las mismas, en la que las uno o más sustituciones aminoacídicas alteran la actividad enzimática, la fidelidad, la procesividad, la tasa de alargamiento, la exactitud de secuenciación, la capacidad de lectura continua larga, la estabilidad o la solubilidad en relación con la polimerasa original. En un modo de realización, la característica alterada es la actividad enzimática. En un modo de realización, la característica alterada es la fidelidad. En un modo de realización, la característica alterada es la procesividad. En un modo de realización, la característica alterada es la tasa de alargamiento. En un modo de realización, la característica alterada es la estabilidad. En un modo de realización, la característica alterada es la solubilidad. En un modo de realización, la característica alterada es la capacidad de unir y/o incorporar nucleótidos polifosfato, por ejemplo, un nucleótido tetrafosfato, pentafosfato, hexafosfato, heptafosfato u octofosfato.

55 En un modo de realización, la polimerasa variante que tiene actividad enzimática alterada en comparación con la SEQ ID NO: 1 o 2 se selecciona de

- 60 a. H223A;
- b. N224Y/L/Q/M/R/K;
- c. Y225L/I/T/F/A/M;
- d. H227P/E/F/Y;
- e. I295F/E/M/W;
- f. Y342L/F;
- 65 g. T343N/F;
- h. I357G/L/Q/H/W/M/A/E/Y/P;

- i. S360G/E/Y;
- j. L361M/W/V/F;
- k. I363V/A/R/M/W;
- l. S365Q/W/M/A/G;
- 5 m. S366A/L;
- n. Y367L/E/M/P/N/F;
- o. P368G;
- p. D417P;
- q. E475D;
- 10 r. Y476V;
- s. F478L;
- t. K518Q;
- u. H527W/R/L/Y/T/M;
- v. T529M/F;
- 15 w. M531H/Y/A/K/R/W/T/L/V/G;
- x. N535L/Y/M/K/I/R/W/Q;
- y. P542E/S/G;
- z. N545D/K/S/L/R/Q/W;
- aa. Q546W/F;
- 20 bb. A547F/M/W/Y/V/S;
- cc. L549H/Y/Q/G/R/K;
- dd. I550A/W;
- ee. I550T/G/F/S;
- ff. N552L/M/T;
- 25 gg. G553S/T/E/Q/K/R/M;
- hh. F558P/T;
- ii. A596S;
- jj. G603T/A/L;
- kk. A610T/E;
- 30 ll. V615A/T;
- mm. Y622A/M;
- nn. C623G/S/Y/A/F;
- oo. D624F/K;
- pp. I628Y/V/F/L/M;
- 35 qq. Y629W/H/M/K;
- rr. R632L/C;
- ss. N635D;
- tt. M641L/Y;
- uu. A643L;
- 40 vv. I644H/M/Y;
- ww. T647G/A/E/K/S/Y;
- xx. I648K/R/V/N/T/L;
- yy. T651Y/F/M;
- zz. I652Q/G/S/N/F/T/A/L/E/K/M;
- 45 aaa. K655G/F/E/N/Q/M/A;
- bbb. W656E;
- ccc. D657R/P/A/E;
- ddd. V658L;
- eee. H660A/Y;
- 50 fff. F662I/L;
- ggg. L690M;
- hhh. S366A+N535L;
- iii. T651Y+N535L;
- jjj. Y342L+E475D+F478L;
- 55 kkk. T343N+D417P+K518Q;
- lll. N535L+N545K+T651Y;
- mmm. I363V+E475D+Y476V;
- nnn. S366L+G553S+F558P;
- ooo. S366A+N535L+A547M;
- 60 ppp. S366A+P542E+N545K;
- qqq. S366A+P542E+I652Q;
- rrr. S366A+N535L+T529M;
- sss. S366A+N535L+I652Q;
- ttt. S366A+N535L+N545K;
- 65 uuu. T651Y+P542E+N545K;
- vvv. T651Y+P542E+Q546W;

www. T651Y+P542E+S366A;
 xxx. T651Y+N535L+N545K;
 yyy. S366A+N535I+I652Q;
 zzz. T651Y+S366A+A547F;
 5 aaaa. T647G+A547F+Y225T;
 bbbb. A547F+A610T+S366A;
 cccc. A547F+A610T+Y225I;
 dddd. S366A+T647G+A547F;
 eeee. T529M+S366A+A547F;
 10 ffff. T647E+S366A+A547F;
 gggg. T529M+T647G+A547F;
 hhhh. N545K+S366A+A547F;
 iii. T647G+A547F+T529M;
 jjjj. T529M+A610T+A547F;
 15 kkkk. M641Y+T529M+A547F;
 llll. T647G+C623G+A547F;
 mmmm. A610T+I295W+T651Y;
 nnnn. V615A+M531Y+T647G;
 oooo. S366L+F478L+A596S+L690M;
 20 pppp. H223A+G553S+A643L+F662I;
 qqqq. N535L+N545K+T651Y+T529M;
 rrrr. N535L+N545K+T651Y+N635D;
 ssss. N535L+N545K+T651Y+I652Q;
 tttt. S366A+N535L+I652Q+T529M;
 25 uuuu. S366A+S365A+P368G+G603T;
 vvvv. N535L+N545K+T651Y+T647G;
 wwwww. S366A+N535L+I652Q+A547Y;
 xxxx. S366A+N535L+A547M+T647G;
 yyyyy. T529M+S366A+A547F+N545K;
 30 zzzz. T529M+S366A+A547F+N545R;
 aaaaa. T529M+S366A+A547F+N552L;
 bbbbbb. T529M+S366A+A547F+Y629W;
 ccccc. N535I+N545K+T651Y+T529M;
 ddddd. N535I+N545K+T651Y+N635D;
 35 eeeee. N535I+N545K+T651Y+I652Q;
 fffff. N535L+N545K+T651Y+T647G+C623G;
 ggggg. N535L+N545K+T651Y+T647G+I628Y;
 hhhhh. S366A+N535L+A547M+T647G+S360G;
 iiiii. N535I+N545K+T651Y+I652Q+Y225I;
 40 jjjjj. N535L+N545K+T651Y+T647G+K655G;
 kkkkk. N535L+N545K+T651Y+T647G+L549Q;
 lllll. S366A+N535L+I652Q+A547Y+K655G;
 mmmmm. T529M+S366A+A547F+N545L+Y629W;
 nnnnn. T529M+S366A+A547F+N545L+Y225L;
 45 ooooo. T529M+S366A+A547F+N545L+Y225F;
 ppppp. T529M+S366A+A547F+N545L+K655F;
 qqqqq. T529M+S366A+A547F+N545L+N552L;
 rrrrr. T529M+S366A+A547F+N545R+M531A;
 sssss. T529M+S366A+A547F+N545R+G539Y;
 50 ttttt. T529M+S366A+A547F+N545R+V658L;
 uuuuu. T529M+S366A+A547F+N545L+Y225L+D657R;
 vvvvv. T529M+S366A+A547F+N545L+Y225L+N552L;
 wwwww. T529M+S366A+A547F+N545L+Y225L+I652G;
 xxxxx. T529M+S366A+A547F+N545L+Y225L+I652Q;
 55 yyyyy. T529M+S366A+A547F+N545L+Y225L+N552M
 zzzzz. T529M+S366A+A547F+N545L+Y225L+D657R+N224R;
 aaaaaa. T529M+S366A+A547F+N545L+Y225L+D657R+I628M;
 bbbbbb. T529M+S366A+A547F+N545L+Y225L+D657R+K655A; y
 ccccc. T529M+S366A+A547F+N545L+Y225L+D657R+Y629W.

60 En un modo de realización, la característica alterada es la actividad enzimática. En un modo de realización, la característica alterada es la fidelidad. En un modo de realización, la característica alterada es la procesividad. En un modo de realización, la característica alterada es la tasa de alargamiento. En un modo de realización, la característica alterada es la estabilidad. En un modo de realización, la característica alterada es la solubilidad. En un modo de realización, la característica alterada es la capacidad de unir y/o incorporar nucleótidos polifosfato, por ejemplo, un nucleótido tetrafosfato, pentafofosfato, hexafosfato, heptafosfato u octofosfato.

65

ES 2 804 843 T3

En algunos modos de realización, la polimerasa variante que tiene actividad enzimática alterada, en comparación con la SEQ ID NO: 2 que tiene las mutaciones N535I+N545K+T651Y+N635D, o la SEQ ID NO: 1 o 2 se selecciona de

- 5
- a. A547F+A610T+Y225I;
 - b. Y225T+T647G+A547F;
 - c. S366A+T647G+A547F;
 - d. S366A+A547F+A610T;
 - 10 e. T529M+S366A+A547F;
 - f. T529M+T647G+A547F;
 - g. T529M+A610T+A547F;
 - h. N545K+S366A+A547F;
 - i. N545K+T647G+A547F;
 - 15 j. A610T+I295W+T651Y;
 - k. V615A+M531Y+T647G;
 - l. M641Y+T529M+A547F;
 - m. T647E+S366A+A547F;
 - n. T647G+A547F+T529M;
 - 20 o. T647G+C623G+A547F; y
 - p. T651Y+S366A+A547F.

En algunos modos de realización, la polimerasa variante se selecciona de

- 25
- a. N535L+N545K+T651Y;
 - b. S366A+N535L+I652Q;
 - c. S366A+T529M+N535L;
 - d. S366A+N535L+N545K;
 - e. S366A+N535L+A547M;
 - 30 f. S366A+P542E+I652Q;
 - g. S366A+P542E+N545K;
 - h. S366A+P542E+T651Y;
 - i. P542E+N545K+T651Y;
 - j. P542E+Q546W+T651Y;
 - 35 k. N535L+T651Y;
 - l. S366A+N535L;
 - m. N535L+N545K+T651Y+T529M;
 - n. N535L+N545K+T651Y+N635D;
 - o. N535L+N545K+T651Y+I652Q;
 - 40 p. S366A+N535L+I652Q+T529M;
 - q. N535L+N545K+T651Y+T647G;
 - r. S366A+N535L+I652Q+A547Y;
 - s. S366A+N535L+A547M+T647G;
 - t. S366A+N535L+I652Q;
 - 45 u. N535I+N545K+T651Y+T529M;
 - v. N535I+N545K+T651Y+N635D;
 - w. N535I+N545K+T651Y+I652Q;
 - x. N535L+N545K+T651Y+T647G+C623G;
 - y. N535L+N545K+T651Y+T647G+I628Y;
 - 50 z. S366A+N535L+A547M+T647G+S360G;
 - aa. N535I+N545K+T651Y+I652Q+Y225I;
 - bb. N535L+N545K+T651Y+T647G+K655G;
 - cc. N535L+N545K+T651Y+T647G+L549Q;
 - dd. S366A+N535L+I652Q+A547Y+K655G;
 - 55 ee. T647G+A547F+Y225T;
 - ff. A547F+A610T+S366A;
 - gg. A547 F+A610T+Y225I;
 - hh. S366A+T647G+A547F;
 - ii. T651Y+S366A+A547F;
 - 60 jj. T529M+S366A+A547F;
 - kk. T647E+S366A+A547F;
 - ll. T529M+T647G+A547F;
 - mm. N545K+S366A+A547F;
 - nn. T647G+A547F+T529M;
 - 65 oo. N545K+T647G+A547F;
 - pp. T529M+A610T+A547F;

- qq. M641Y+T529M+A547F;
 rr. T647G+C623G+A547F;
 ss. A610T+I295W+T651Y;
 tt. V615A+M531Y+T647G;
 5 uu. T529M+S366A+A547F+N545K;
 vv. T529M+S366A+A547F+N545R;
 ww. T529M+S366A+A547F+N552L;
 xx. T529M+S366A+A547F+Y629W;
 yy. T529M+S366A+A547F+N545L+Y629W;
 10 zz. T529M+S366A+A547F+N545L+Y225L;
 aaa. T529M+S366A+A547F+N545L+Y225F;
 bbb. T529M+S366A+A547F+N545L+K655F;
 ccc. T529M+S366A+A547F+N545L+N552L;
 ddd. T529M+S366A+A547F+N545R+M531A;
 15 eee. T529M+S366A+A547F+N545R+G539Y;
 fff. T529M+S366A+A547F+N545R+V658L;
 ggg. T529M+S366A+A547F+N545L+Y225L+D657R;
 hhh. T529M+S366A+A547F+N545L+Y225L+N552L;
 iii. T529M+S366A+A547F+N545L+Y225L+I652G;
 20 jjj. T529M+S366A+A547F+N545L+Y225L+I652Q;
 kkk. T529M+S366A+A547F+N545L+Y225L+N552M
 ll. T529M+S366A+A547F+N545L+Y225L+D657R+N224R;
 mmm. T529M+S366A+A547F+N545L+Y225L+D657R+I628M;
 nnn. T529M+S366A+A547F+N545L+Y225L+D657R+K655A; y
 25 ooo. T529M+S366A+A547F+N545L+Y225L+D657R+Y629W..

En algunos modos de realización, la polimerasa variante tiene actividad enzimática alterada en comparación con las SEQ ID NO: 1 o 2, o la polimerasa original.

- 30 En algunos modos de realización, la polimerasa variante que tiene actividad enzimática alterada, en comparación con la SEQ ID NO: 2 que tiene las mutaciones S366A+T529M+N545L+A547F, o la SEQ ID NO: 1 o 2, se selecciona de
- a. Y225L/F/A/M;
 35 b. M531A/G;
 c. G539Y;
 d. N552L/T;
 e. Y629W/K;
 f. K655F.

- 40 En un modo de realización, la característica alterada es la actividad enzimática. En un modo de realización, la característica alterada es la fidelidad. En un modo de realización, la característica alterada es la procesividad. En un modo de realización, la característica alterada es la tasa de alargamiento. En un modo de realización, la característica alterada es la estabilidad. En un modo de realización, la característica alterada es la solubilidad. En un modo de realización, la característica alterada es la capacidad de unir y/o incorporar nucleótidos polifosfato, por ejemplo, un nucleótido tetrafosfato, pentafosfato, hexafosfato, heptafosfato u octofosfato.

- 50 En algunos modos de realización, la polimerasa original es Pol6 natural: (SEQ ID NO:1). En algunos modos de realización, la polimerasa original es Pol6 que comprende una marca His (SEQ ID NO:2). En algunos modos de realización, la polimerasa original comprende las mutaciones S366A+T529M+A547F+N545L/R. En algunos modos de realización, la polimerasa original puede ser la SEQ ID NO:1 que comprende una o más mutaciones. Por ejemplo, S366A+T529M+A547F+N545R usó S366A+T529M+A547F como la polimerasa original, a continuación se añadió N545R.

- 55 En algunos modos de realización, la polimerasa modificada tiene una $k_{quím}$ que es mayor que la polimerasa original. En algunos modos de realización, la polimerasa modificada tiene una k_{des} que es menor que la polimerasa original. En algunos modos de realización, la polimerasa modificada tiene una $k_{quím}/k_{des}$ (es decir, una proporción) que es al menos 1,5, 2,0 o 2,5 veces mayor que la polimerasa original.

60 BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **FIG. 1** ilustra un molde ejemplar usado en el ensayo de desplazamiento. Se hace referencia al ejemplo 3.

- 65 La **FIG. 2** muestra un esquema del ensayo de $k_{quím}$ usado en el presente documento para medir la tasa de incorporación de polifosfatos. Se hace referencia al ejemplo 6.

La **FIG. 3** es un resumen del ensayo de k_{des} basado en la extinción de fluorescencia usado en el presente documento para medir las propiedades cinéticas de las polimerasas variantes. Se hace referencia al ejemplo 4.

5 La **FIG. 4** es una representación del ensayo de k_{des} basado en la polarización de fluorescencia y un trazado de datos ejemplar. Se hace referencia al ejemplo 5.

La **FIG. 5** es un gráfico que muestra datos representativos del ensayo de desplazamiento para una polimerasa variante. Se hace referencia al ejemplo 3.

10 La **FIG. 6** es un gráfico de datos representativos del ensayo de k_{des} basado en polarización de fluorescencia para dos polimerasas variantes. Se hace referencia al ejemplo 5.

15 La **FIG. 7** es un trazado de una captura estática del nucleótido timina marcado a 100 mV por el complejo Pol6 (S366A+N535L+I652Q)-ADN acoplado al nanoporo de hemolisina alfa en Hepes 7,5 20 mM, NaCl 300 mM, $CaCl_2$ 3 mM y TCEP 5 mM por encima y por debajo de la bicapa. El eje vertical es el % de corriente de canal abierto (normalizada) y el eje horizontal es el tiempo en segundos. Se hace referencia al ejemplo 8.

20 La **FIG. 8** es un gráfico del tiempo de permanencia frente a la curva de corriente para un experimento de captura estática a 100 mV con el complejo Pol6 (S366A+N535L+I652Q)-ADN acoplado al nanoporo de hemolisina alfa en Hepes 20 mM pH 7,5, NaCl 300 mM, $CaCl_2$ 3 mM y TCEP 5 mM por encima y por debajo de la bicapa. El tiempo de permanencia promedio de cada captura de nucleótido marcado con dTNP es de 1,2 segundos. Se hace referencia al ejemplo 8.

25 La **FIG. 9** es un gráfico de datos representativos de un ensayo de $k_{quím}$ basado en la extinción de fluorescencia (véase la FIG. 2) para una polimerasa variante. El complejo binario preformado de polimerasa y molde de ADN-fluoresceína se mezcla rápidamente con una concentración saturante de dCnP-Alexa555 en presencia de Mg^{2+} usando un dispositivo de flujo interrumpido Kintek. Se realiza un seguimiento de la fluorescencia de la fluoresceína en el tiempo. La $k_{quím}$ estimada a partir de la etapa limitante de la velocidad es de 0,2 s⁻¹. El eje x es el tiempo (T) en segundos y el eje y son las unidades de fluorescencia relativa (UFR).

30 La **FIG. 10** es un gráfico de los datos representativos de un ensayo de k_{des} basado en la extinción de fluorescencia (véase la FIG. 3) para una polimerasa variante. El complejo ternario preformado de polimerasa, molde de ADN-fluoresceína y dCnP-Alexa555 se preincubó en presencia de Ca^{2+} y se cazó con un exceso de dCTP natural. Se realizó un seguimiento de la fluorescencia de la fluoresceína en el tiempo. La k_{des} medida a partir de esto es de 0,028 s⁻¹. El eje x es el tiempo (T) en segundos y el eje y son las unidades de fluorescencia relativa (UFR).

35 La **FIG. 11** es una imagen de un gel que muestra los productos de amplificación de un ensayo de círculo rodante. Los carriles extremos izquierdo y derecho son escaleras moleculares. El segundo carril desde la izquierda es el punto de tiempo cero. Todos los demás carriles son el punto de tiempo de 40 minutos para los diversos resultados positivos de polimerasa. Se hace referencia al ejemplo 9.

40 La **FIG. 12** es un trazado de secuenciación que muestra los cambios en la corriente que proporciona un registro de los nucleótidos marcados a medida que se incorporan en la hebra de ADN creciente. También se muestra la secuencia de ADN molde y la secuencia llamada de la hebra naciente que demuestra >70 % de exactitud (SEQ ID NO 6-8, respectivamente, en orden de aparición). Se hace referencia al ejemplo 10.

DESCRIPCIÓN DETALLADA

50 La invención se describirá ahora en detalle a modo de referencia usando solo las siguientes definiciones y ejemplos.

A menos que se defina de otro modo en el presente documento, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que se entiende comúnmente por un experto en la técnica a la que pertenece la presente invención. Singleton, *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2.^a ED., John Wiley and Sons, New York (1994), y Hale y Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) proporcionan a un experto un diccionario general de muchos de los términos usados en la presente invención. Aunque cualquier procedimiento y material similar o equivalente a los descritos en el presente documento se puede usar en la práctica o pruebas de la presente invención, se describen los procedimientos y materiales preferentes. Los profesionales sanitarios se dirigen en particular a Sambrook *et al.*, 1989, y Ausubel FM *et al.*, 1993, para consultar definiciones y términos de la técnica. Se debe entender que la presente invención no está limitada a la metodología, protocolos y reactivos particulares descritos, ya que estos pueden variar.

60 Los intervalos numéricos incluyen los números que definen el intervalo. El término aproximadamente se usa en el presente documento para querer decir más o menos un diez por ciento (10 %) de un valor. Por ejemplo, "aproximadamente 100" se refiere a cualquier número entre 90 y 110.

A menos que se indique de otro modo, los ácidos nucleicos se escriben de izquierda a derecha en una orientación de 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en una orientación de amino a carboxilo, respectivamente.

- 5 Los títulos proporcionados en el presente documento no son limitaciones de los diversos aspectos o modos de realización de la invención que se pueden tener por referencia a la memoria descriptiva como un todo. En consecuencia, los términos definidos inmediatamente a continuación se definen más completamente por referencia a la memoria descriptiva como un todo.

10 Definiciones

Aminoácido: como se usa en el presente documento, el término "aminoácido", en su sentido más amplio, se refiere a cualquier compuesto y/o sustancia que se puede incorporar en una cadena polipeptídica. En algunos modos de realización, un aminoácido tiene la estructura general $H_2N-C(H)(R)-COOH$. En algunos modos de realización, un aminoácido es un aminoácido natural. En algunos modos de realización, el aminoácido es un aminoácido sintético; en algunos modos de realización, un aminoácido es un D-aminoácido; en algunos modos de realización, un aminoácido es un L-aminoácido. "Aminoácido estándar" se refiere a cualquiera de los veinte L-aminoácidos estándar encontrados comúnmente en péptidos naturales. "Aminoácido no estándar" se refiere a cualquier aminoácido, distinto de los aminoácidos estándar, independientemente de si se prepara sintéticamente o se obtiene de una fuente natural. Como se usa en el presente documento, "aminoácido sintético" engloba aminoácidos modificados químicamente, incluyendo pero sin limitarse a sales, derivados de aminoácidos (tales como amidas) y/o sustituciones. Los aminoácidos, incluyendo aminoácidos carboxi y/o aminoterminales en péptidos, se pueden modificar por metilación, amidación, acetilación y/o sustitución con otros grupos químicos sin afectar negativamente a su actividad. Los aminoácidos pueden participar en un enlace disulfuro. El término "aminoácido" se usa de manera intercambiable con "residuo aminoacídico" y se puede referir a un aminoácido libre y/o a un residuo aminoacídico de un péptido. Será evidente a partir del contexto en el que se usa el término si se refiere a un aminoácido libre o a un residuo de un péptido. Cabe destacar que todas las secuencias de residuos aminoacídicos se representan en el presente documento por fórmulas con una orientación izquierda y derecha que está en el sentido convencional del extremo amínico al extremo carboxílico.

Par de bases (pb): como se usa en el presente documento, par de bases se refiere a una asociación de adenina (A) con timina (T), o de citosina (C) con guanina (G) en una molécula de ADN bicatenario.

Complementario: como se usa en el presente documento, el término "complementario" se refiere al concepto amplio de complementariedad de secuencia entre regiones de dos hebras polinucleotídicas o entre dos nucleótidos a través de emparejamiento de bases. Es conocido que un nucleótido adenina puede formar enlaces de hidrógeno específicos ("emparejamiento de bases") con un nucleótido que sea timina o uracilo. De forma similar, es conocido que un nucleótido citosina puede realizar emparejamiento de bases con un nucleótido guanina.

Afinidad de unión a ADN: como se usa en el presente documento, el término "afinidad de unión a ADN" típicamente se refiere a la actividad de una ADN polimerasa en la unión al ácido nucleico ADN. En algunos modos de realización, la actividad de unión a ADN se puede medir en un ensayo de desplazamiento de dos bandas. Véase, por ejemplo, Sambrook *et al.* (2001) *Molecular Cloning: A Laboratory Manual* (3.^a ed., Cold Spring Harbor Laboratory Press, NY) en 9,63-9,75 (que describe el marcaje terminal de ácidos nucleicos). Se prepara una mezcla de reacción que contiene al menos aproximadamente 0,5 µg del polipéptido en aproximadamente 10 µl de tampón de unión (tampón fosfato de sodio 50 mM (pH 8,0), glicerol al 10 %, KCl 25 mM, MgCl₂ 25 mM). Se calienta la mezcla de reacción a 37 °C durante 10 min. Se añade aproximadamente de 1×10^4 a 5×10^4 cpm (o aproximadamente 0,5-2 ng) de ácido nucleico bicatenario marcado a la mezcla de reacción y se incuba durante 10 min adicionales. Se carga la mezcla de reacción en un gel de poli(acrilamida) natural en tampón tris-borato 0,5×. Se somete la mezcla de reacción a electroforesis a temperatura ambiente. Se seca el gel y se somete a autorradiografía usando procedimientos estándar. Cualquier disminución detectable en la movilidad del ácido nucleico bicatenario marcado indica la formación de un complejo de unión entre el polipéptido y el ácido nucleico bicatenario. Dicha actividad de unión a ácido nucleico se puede cuantificar usando procedimientos densitométricos estándar para medir la cantidad de radioactividad en el complejo de unión en relación con la cantidad total de radioactividad en la mezcla de reacción inicial. Otros procedimientos de medición de afinidad de unión a ADN son conocidos en la técnica (véase, por ejemplo, Kong *et al.* (1993) *J. Biol. Chem.* 268(3):1965-1975).

Tasa de alargamiento: como se usa en el presente documento, el término "tasa de alargamiento" se refiere a la tasa promedio a la que una ADN polimerasa extiende una cadena de polímero. Como se usa en el presente documento, una alta tasa de alargamiento se refiere a una tasa de alargamiento mayor de 2 nt/s (por ejemplo, mayor de 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140 nt/s). Como se usa en la presente solicitud, los términos "tasa de alargamiento", "tasa de extensión" y "tasa de incorporación" se usan de manera intercambiable.

Actividad enzimática: como se usa en el presente documento, el término "actividad enzimática" se refiere a la especificidad y eficacia de una ADN polimerasa. La actividad enzimática de una ADN polimerasa también se

denomina "actividad polimerasa", lo que se refiere típicamente a la actividad de una ADN polimerasa al catalizar la síntesis dirigida a molde de un polinucleótido. Se puede medir la actividad enzimática de una polimerasa usando diversas técnicas y procedimientos conocidos en la técnica. Por ejemplo, se pueden preparar diluciones en serie de polimerasa en tampón de dilución (por ejemplo, Tris.Cl 20 mM, pH 8,0, KCl 50 mM, NP 40 al 0,5 % y Tween-20 al 0,5 %). Para cada dilución, se pueden retirar 5 µl y añadirse a 45 µl de una mezcla de reacción que contiene TAPS 25 mM (pH 9,25), KCl 50 mM, MgCl₂ 2 mM, dATP 0,2 mM, dGTP 0,2 mM, dTTP 0,2 mM, dCTP 0,1 mM, 12,5 µg de ADN activado, [α -³²P]dCTP 100 µM (0,05 µCi/nmol) y agua desionizada estéril. Se pueden incubar las mezclas de reacción a 37 °C (o 74 °C para ADN polimerasas termoestables) durante 10 minutos y a continuación detenerse enfriando de inmediato la reacción a 4 °C y añadiendo 10 µl de EDTA 60 mM helado. Se puede retirar una alícuota de 25 µl de cada mezcla de reacción. Se puede retirar dCTP marcado radioactivamente no incorporado de cada alícuota por filtración en gel (Centri-Sep, Princeton Separations, Adelphia, N.J.). Se puede mezclar el eluido de columna con líquido de centelleo (1 ml). Se cuantifica la radioactividad en el eluido de columna con un contador de centelleo para determinar la cantidad de producto sintetizado por la polimerasa. Una unidad de actividad polimerasa se puede definir como la cantidad de polimerasa necesaria para sintetizar 10 nmol de producto en 30 minutos (Lawyer *et al.* (1989) *J. Biol. Chem.* 264:6427-647). Otros procedimientos de medición de la actividad polimerasa son conocidos en la técnica (véase, por ejemplo, Sambrook *et al.* (2001) *Molecular Cloning: A Laboratory Manual* (3.^a ed., Cold Spring Harbor Laboratory Press, NY)).

Purificado: como se usa en el presente documento, "purificado" significa que una molécula está presente en una muestra a una concentración de al menos un 90 % en peso, o al menos un 95 % en peso, o al menos un 98 % en peso de la muestra en la que está contenida.

Aislado: una molécula "aislada" es una molécula de ácido nucleico que se separa de al menos otra molécula con la que se asocia normalmente, por ejemplo, en su entorno natural. Una molécula de ácido nucleico aislada incluye una molécula de ácido nucleico contenida en células que expresan habitualmente la molécula de ácido nucleico, pero la molécula de ácido nucleico está presente de forma extracromosómica o en una localización cromosómica que es diferente de su localización cromosómica natural.

% de homología: el término "% de homología" se usa de manera intercambiable en el presente documento con el término "% de identidad" en el presente documento y se refiere al nivel de identidad de secuencia de ácido nucleico o aminoácidos entre la secuencia de ácido nucleico que codifica uno cualquiera de los polipéptidos según la invención o la secuencia de aminoácidos del polipéptido según la invención, cuando se alinea usando un programa de alineación de secuencias.

Por ejemplo, como se usa en el presente documento, un 80 % de homología quiere decir lo mismo que un 80 % de identidad de secuencia determinada por un algoritmo definido y, en consecuencia, un homólogo de una secuencia dada tiene más de un 80 % de identidad de secuencia en una longitud de la secuencia dada. Los niveles ejemplares de identidad de secuencia incluyen, pero no se limitan a, 80, 85, 90, 95, 98 % o más de identidad de secuencia para una secuencia dada, por ejemplo, la secuencia codificante para uno cualquiera de los polipéptidos según la invención, como se describe en el presente documento.

Los programas informáticos ejemplares que se pueden usar para determinar la identidad entre dos secuencias incluyen, pero no se limitan a, el paquete de programas BLAST, por ejemplo, BLASTN, BLASTX y TBLASTX, BLASTP y TBLASTN, disponibles públicamente en Internet. Véanse también, Altschul, *et al.*, 1990 y Altschul, *et al.*, 1997.

Las búsquedas de secuencias se llevan a cabo típicamente usando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico dada relativa a las secuencias de ácido nucleico en las secuencias de ADN de GenBank y otras bases de datos públicas. El programa BLASTX es preferente para buscar secuencias de ácido nucleico que se han traducido en todos los marcos de lectura frente a secuencias de aminoácidos en las secuencias de proteínas de GenBank y otras bases de datos públicas. Tanto BLASTN como BLASTX se ejecutan usando parámetros predeterminados de una penalización por hueco abierto de 11,0 y una penalización por hueco extendido de 1,0, y utilizan la matriz BLOSUM-62. (Véase, por ejemplo, Altschul, S. F., *et al.*, *Nucleic Acids Res.* 25:3389-3402, 1997).

Se realiza una alineación preferente de secuencias seleccionadas para determinar el "% de identidad" entre dos o más secuencias, usando por ejemplo, el programa CLUSTAL-W en MacVector versión 13.0.7, operado con parámetros predeterminados, incluyendo una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitud BLOSUM 30.

ADN polimerasa modificada: como se usa en el presente documento, el término "ADN polimerasa modificada" se refiere a una ADN polimerasa originada a partir de otra ADN polimerasa (es decir, original) y contiene una o más alteraciones de aminoácidos (por ejemplo, sustitución, delección o inserción de aminoácidos) en comparación con la ADN polimerasa original. En algunos modos de realización, una ADN polimerasa modificada de la invención se origina o modifica a partir de una ADN polimerasa natural. En algunos modos de realización, una ADN polimerasa modificada de la invención se origina o modifica a partir de una ADN polimerasa recombinante o genomanipulada

que incluye, pero sin limitarse a, ADN polimerasa quimérica, ADN polimerasa de fusión u otra ADN polimerasa modificada. Típicamente, una ADN polimerasa modificada tiene al menos un fenotipo cambiado en comparación con la polimerasa original.

5 **Mutación:** como se usa en el presente documento, el término "mutación" se refiere a un cambio introducido en una secuencia original incluyendo, pero sin limitarse a, sustituciones, inserciones, deleciones (incluyendo truncamientos). Las consecuencias de una mutación incluyen, pero no se limitan a, la creación de un nuevo carácter, propiedad, función, fenotipo o rasgo no encontrado en la proteína codificada por la secuencia original.

10 **Mutante:** como se usa en el presente documento, el término "mutante" se refiere a una proteína modificada que muestra características alteradas cuando se compara con la proteína original. Los términos "variante" y "mutante" se usan de manera intercambiable en el presente documento.

15 **Natural:** como se usa en el presente documento, el término "natural" se refiere a un gen o producto génico que tiene las características de ese gen o producto génico cuando se aísla de una fuente natural.

20 **Fidelidad:** como se usa en el presente documento, el término "fidelidad" se refiere a la exactitud de la polimerización de ADN por ADN polimerasa dependiente de molde o bien la diferencia medida en la k_{des} del nucleótido correcto frente a la unión de nucleótido incorrecto al ADN molde. La fidelidad de una ADN polimerasa se mide típicamente por la tasa de error (la frecuencia de incorporar un nucleótido inexacto, es decir, un nucleótido que no se incorpora de una manera dependiente de molde). La exactitud o fidelidad de la polimerización de ADN se mantiene tanto para la actividad polimerasa como para la actividad 3'-5' exonucleasa de una ADN polimerasa. El término "alta fidelidad" se refiere a una tasa de error menor de $4,45 \times 10^{-6}$ (por ejemplo, menor que $4,0 \times 10^{-6}$, $3,5 \times 10^{-6}$, $3,0 \times 10^{-6}$, $2,5 \times 10^{-6}$, $2,0 \times 10^{-6}$, $1,5 \times 10^{-6}$, $1,0 \times 10^{-6}$, $0,5 \times 10^{-6}$) mutaciones/nt/duplicación. La fidelidad o tasa de error de una ADN polimerasa se puede medir usando ensayos conocidos para la técnica. Por ejemplo, se pueden someter a prueba las tasas de error de las ADN polimerasas como se describe en el presente documento o como se describe en Johnson, *et al.*, Biochim Biophys Acta, mayo 2010; 1804(5): 1041-1048.

30 **Nanoporo:** el término "nanoporo", como se usa en el presente documento, se refiere en general a un poro, canal o paso formado o proporcionado de otro modo en una membrana. Una membrana puede ser una membrana orgánica, tal como una bicapa lipídica, o una membrana sintética, tal como una membrana formada de un material polimérico. La membrana puede ser un material polimérico. El nanoporo se puede disponer adyacente o próximo a un circuito sensor o un electrodo acoplado a un circuito sensor, tal como, por ejemplo, un semiconductor de óxido metálico complementario (CMOS) o circuito de transistor de efecto de campo (FET). En algunos ejemplos, un nanoporo tiene un ancho o diámetro característico del orden de 0,1 nanómetros (nm) a aproximadamente 1000 nm. Algunos nanoporos son proteínas. La hemolisina alfa, la MspA son ejemplos de un nanoporo proteico.

40 **Nucleótido:** como se usa en el presente documento, una unidad monomérica de ADN o ARN que consiste en un resto glucídico (pentosa), un fosfato y una base heterocíclica nitrogenada. La base se une al resto glucídico por el carbono glucosídico (carbono 1' de la pentosa) y esa combinación de base y glúcido es un nucleósido. Cuando el nucleósido contiene un grupo fosfato enlazado a la posición 3' o 5' de la pentosa se denomina nucleótido. Una secuencia de nucleótidos unidos de forma funcional se denomina típicamente en el presente documento "secuencia de bases" o "secuencia de nucleótidos," y se representa en el presente documento por una fórmula con una orientación de izquierda a derecha que está en el sentido convencional del extremo 5' al extremo 3'. Como se usa en el presente documento, un "nucleótido modificado" se refiere a un nucleótido polifosfato, por ejemplo, 3, 4, 5, 6, 7 u 8 fosfatos.

50 **Oligonucleótido o polinucleótido:** como se usa en el presente documento, el término "oligonucleótido" se define como una molécula que incluye dos o más desoxirribonucleótidos y/o ribonucleótidos, preferentemente más de tres. Su tamaño exacto dependerá de muchos factores, que a su vez dependen de la función definitiva o el uso del oligonucleótido. El oligonucleótido se puede derivar sintéticamente o por clonación. Como se usa en el presente documento, el término "polinucleótido" se refiere a una molécula de polímero compuesta de monómeros nucleotídicos enlazados de forma covalente en una cadena. ADN (ácido desoxirribonucleico) y ARN (ácido ribonucleico) son ejemplos de polinucleótidos.

55 **Polimerasa:** como se usa en el presente documento, una "polimerasa" se refiere a una enzima que cataliza la polimerización de nucleótidos (es decir, la actividad polimerasa). En general, la enzima iniciará la síntesis en el extremo 3' del cebador hibridado a una secuencia molde de polinucleótido, y avanzará hacia el extremo 5' de la hebra molde. Una "ADN polimerasa" cataliza la polimerización de desoxinucleótidos.

60 **Cebador:** como se usa en el presente documento, el término "cebador" se refiere a un oligonucleótido, ya sea natural o producido sintéticamente, que puede actuar como un punto de inicio de la síntesis de ácido nucleico cuando se dispone en condiciones en las que la síntesis de un producto de extensión del cebador que es complementario a una hebra de ácido nucleico se induce, por ejemplo, en presencia de cuatro nucleótidos trifosfatos diferentes y enzima termoestable en un tampón apropiado ("tampón" incluye pH, fuerza iónica, cofactores, etc.) y a una temperatura adecuada. El cebador es preferentemente monocatenario para eficacia máxima en la amplificación,

65

pero puede ser de forma alternativa bicatenario. Si es bicatenario, en primer lugar se trata el cebador para separar sus hebras antes de usarse para preparar productos de extensión. Preferentemente, el cebador es un oligodesoxirribonucleótido. El cebador debe ser suficientemente largo para cebar la síntesis de los productos de extensión en presencia de la enzima termoestable. Las longitudes exactas de los cebadores dependerán de muchos factores, incluyendo temperatura, fuente de cebador y uso del procedimiento. Por ejemplo, dependiendo de la complejidad de la secuencia diana, el cebador oligonucleotídico contiene típicamente 15-25 nucleótidos, aunque puede contener más o menos nucleótidos. En general, las moléculas cebadoras cortas requieren temperaturas más bajas para formar complejos híbridos suficientemente estables con el molde.

Procesividad: como se usa en el presente documento, "procesividad" se refiere a la capacidad de una polimerasa para mantenerse unida al molde y realizar múltiples reacciones de modificación. Las "reacciones de modificación" incluyen pero no se limitan a polimerización y escisión exonucleolítica. En algunos modos de realización, "procesividad" se refiere a la capacidad de una ADN polimerasa para realizar una secuencia de etapas de polimerización sin disociación intermedia de la enzima de las cadenas de ADN crecientes. Típicamente, la "procesividad" de una ADN polimerasa se mide por la longitud de nucleótidos (por ejemplo 20 nts, 300 nts, 0,5-1 kb, o más) que se polimerizan o se modifican sin disociación intermedia de la ADN polimerasa de la cadena de ADN creciente. La "procesividad" puede depender de la naturaleza de la polimerasa, la secuencia de un molde de ADN, y condiciones de reacción, por ejemplo, concentración de sal, temperatura o presencia de proteínas específicas. Como se usa en el presente documento, el término "alta procesividad" se refiere a una procesividad mayor a 20 nts (por ejemplo, mayor a 40 nts, 60 nts, 80 nts, 100 nts, 120 nts, 140 nts, 160 nts, 180 nts, 200 nts, 220 nts, 240 nts, 260 nts, 280 nts, 300 nts, 320 nts, 340 nts, 360 nts, 380 nts, 400 nts o más) por asociación/disociación con el molde. La procesividad se puede medir de acuerdo con los procedimientos definidos en el presente documento y en el documento WO 01/92501 A1 (MJ Bioworks, Inc., Improved Nucleic Acid Modifying Enzymes, publicado el 6 de diciembre de 2001).

Síntesis: como se usa en el presente documento, el término "síntesis" se refiere a cualquier procedimiento *in vitro* para preparar una nueva hebra de polinucleótido o alargar el polinucleótido existente (es decir, ADN o ARN) de una manera dependiente de molde. La síntesis, de acuerdo con la invención, incluye amplificación, lo que incrementa el número de copias de una secuencia de molde de polinucleótido con el uso de una polimerasa. La síntesis de polinucleótidos (por ejemplo, amplificación) da como resultado la incorporación de nucleótidos en un polinucleótido (es decir, un cebador) formando de este modo una nueva molécula de polinucleótido complementaria al molde de polinucleótido. La molécula de polinucleótido formada y su molde se pueden usar como moldes para sintetizar moléculas de polinucleótidos adicionales. "Síntesis de ADN," como se usa en el presente documento, incluye, pero no se limita a, PCR, el marcaje de polinucleótido (es decir, para sondas y cebadores oligonucleotídicos), secuenciación de polinucleótidos.

Molécula de ADN molde: como se usa en el presente documento, el término "molécula de ADN molde" se refiere a una hebra de un ácido nucleico a partir de la que se sintetiza una hebra de ácido nucleico complementaria por una ADN polimerasa, por ejemplo, en una reacción de extensión de cebador.

Manera dependiente de molde: como se usa en el presente documento, el término "manera dependiente de molde" se refiere a un procedimiento que implica la extensión dependiente de molde de una molécula cebadora (por ejemplo, síntesis de ADN por ADN polimerasa). El término "manera dependiente de molde" se refiere típicamente a la síntesis de polinucleótidos de ARN o ADN en la que la secuencia de la hebra de polinucleótido recién sintetizada se dicta por las normas bien conocidas del emparejamiento de bases complementarias (véase, por ejemplo, Watson, J. D. *et al.*, en: Molecular Biology of the Gene, 4.^a ed., W. A. Benjamin, Inc., Menlo Park, Calif. (1987)).

Marca: como se usa en el presente documento, el término "marca" se refiere a un resto detectable que puede ser átomos o moléculas, o una colección de átomos o moléculas. Una marca puede proporcionar una firma óptica, electroquímica, magnética o electrostática (por ejemplo, inductiva, capacitiva), firma que se puede detectar con la ayuda de un nanoporo.

Nucleótido marcado: como se usa en el presente documento, el término "nucleótido marcado" se refiere a un nucleótido o nucleótido modificado que tiene una marca unida. La marca se puede unir de forma covalente al azúcar, el fosfato (o polifosfato) o base. La marca puede estar en el fosfato terminal.

Vector: como se usa en el presente documento, el término "vector" se refiere a una construcción de ácido nucleico diseñada para la transferencia entre diferentes células huésped. Un "vector de expresión" se refiere a un vector que tiene la capacidad de incorporar y expresar fragmentos de ADN heterólogos en una célula exógena. Muchos vectores de expresión procariotas y eucariotas están disponibles comercialmente. La selección de vectores de expresión apropiados está dentro del conocimiento de los expertos en la técnica.

Las variantes de polimerasa proporcionadas en el presente documento son útiles en la secuenciación de polinucleótidos basada en chip como se describe en el documento WO2013/188841 (Genia Technologies, Inc., Chip Set-Up and High-Accuracy Nucleic Acid Sequencing, publicado el 19 de diciembre de 2013).

Las características deseadas de una polimerasa que encuentra uso en la secuenciación del ADN son:

- 5 a. k_{dis} lenta (para nucleótido modificado)
- b. k_{aso} rápida (para nucleótido modificado)
- c. Alta fidelidad
- d. Baja actividad de exonucleasa
- e. Desplazamiento de la hebra de ADN
- f. k_{quim} más rápida (para sustratos de nucleótidos modificados)
- 10 g. Estabilidad incrementada
- h. Procesividad
- i. Tolerancia a la sal
- j. Compatible con unión a nanoporo
- k. Capacidad para incorporar un polifosfato que tenga 4, 5, 6, 7 u 8 fosfatos, por ejemplo, nucleótido cuadrafosfato, pentafosfato, hexafosfato, heptafosfato u octofosfato
- 15 l. Exactitud de secuenciación
- m. Longitudes de lectura largas, es decir, lecturas continuas largas.

Nomenclatura

20 En la presente descripción y reivindicaciones, se usan los códigos de una letra y tres letras convencionales para los residuos aminoacídicos.

Para facilidad de referencia, las variantes de polimerasa de la solicitud se describen por el uso de la siguiente nomenclatura:

25 Aminoácido(s) original(es): posición/posiciones: aminoácido(s) sustituido(s). De acuerdo con esta nomenclatura, por ejemplo, la sustitución de serina por una alanina en la posición 242 se muestra como:

30 Ser242Ala o S242A

Las mutaciones múltiples se separan por signos de más, es decir:

Ala30Asp+Glu34Ser o A30N+E34S

35 representando mutaciones en las posiciones 30 y 34 sustituyendo alanina y ácido glutámico por asparagina y serina, respectivamente.

40 Cuando uno o más residuos aminoacídicos alternativos se pueden insertar en una posición dada, se indica como: A30N/E o A30N o A30E.

A menos que se exprese de otro modo, el número de residuos corresponde a la numeración de residuos de SEQ ID NO:2.

Mutagénesis de la polimerasa dirigida a sitio

45 En el presente documento se proporcionan las secuencias naturales de fago de *Clostridium* phiCPV4 (SEQ ID NO:3, región codificante de ácido nucleico más una marca His; SEQ ID NO:1, región codificante de proteína) y están disponibles en otros lugares (National Center for Bioinformatics o GenBank con números de acceso AFH27113).

50 Se pueden introducir mutaciones puntuales usando el kit QuikChange Lightning 2 (Stratagene/Agilent) siguiendo las instrucciones del fabricante.

Se pueden encargar los cebadores a empresas comerciales, por ejemplo, IDT DNA.

55 Ensamblaje e inserción de nanoporos

Los procedimientos descritos en el presente documento pueden usar un nanoporo que tiene una polimerasa unida al nanoporo. En algunos casos, es deseable tener una y solo una polimerasa por nanoporo (por ejemplo, de modo que solo se secuencie una molécula de ácido nucleico en cada nanoporo). Sin embargo, muchos nanoporos, incluyendo, por ejemplo, hemolisina alfa (aHL), pueden ser proteínas multiméricas que tienen una pluralidad de subunidades (por ejemplo, 7 subunidades para aHL). Las subunidades pueden ser copias idénticas del mismo polipéptido. En el presente documento se proporcionan proteínas multiméricas (por ejemplo, nanoporos) que tienen una proporción definida de subunidades modificadas (por ejemplo, variantes de a-HL) con respecto a subunidades no modificadas (por ejemplo, a-HL). En el presente documento también se proporcionan procedimientos para producir proteínas multiméricas (por ejemplo, nanoporos) que tienen una proporción definida de subunidades modificadas con respecto a subunidades no modificadas.

Con referencia a la figura 27 del documento WO2014/074727 (Genia Technologies, Inc.), un procedimiento para ensamblar una proteína que tiene una pluralidad de subunidades comprende proporcionar una pluralidad de primeras subunidades 2705 y proporcionar una pluralidad de segundas subunidades 2710, donde las segundas subunidades se modifican en comparación con las primeras subunidades. En algunos casos, las primeras subunidades son naturales (por ejemplo, purificadas de fuentes naturales o producidas de forma recombinante). Las segundas subunidades se pueden modificar de cualquier forma adecuada. En algunos casos, las segundas subunidades tienen una proteína (por ejemplo, una polimerasa) unida (por ejemplo, como una proteína de fusión).

Las subunidades modificadas pueden comprender un resto químicamente reactivo (por ejemplo, una acida o un grupo alquino adecuado para formar un enlace). En algunos casos, el procedimiento comprende además realizar una reacción (por ejemplo, una cicloadición química *Click*) para unir una entidad (por ejemplo, una polimerasa) al resto químicamente reactivo.

El procedimiento puede comprender además poner en contacto las primeras subunidades con las segundas subunidades 2715 en una primera proporción para formar una pluralidad de proteínas 2720 que tienen las primeras subunidades y las segundas subunidades. Por ejemplo, una parte de las subunidades de aHL modificadas que tienen un grupo reactivo adecuado para unir una polimerasa se puede mezclar con seis partes de subunidades de aHL naturales (es decir, siendo la primera proporción 1:6). La pluralidad de proteínas puede tener una pluralidad de proporciones de las primeras subunidades con respecto a las segundas subunidades. Por ejemplo, las subunidades mixtas pueden formar varios nanoporos que tienen una distribución de estequiometrías de subunidades modificadas con respecto a no modificadas (por ejemplo, 1:6, 2:5, 3:4).

En algunos casos, las proteínas se forman simplemente mezclando las subunidades. En el caso de nanoporos de aHL, por ejemplo, un detergente (por ejemplo, ácido desoxicólico) puede provocar que el monómero de aHL adopte la conformación de poro. Los nanoporos también se pueden formar usando un lípido (por ejemplo, 1,2-difitanoil-sn-glicero-3-fosfolina (DPhPC) o 1,2-di-O-fitanil-sn-glicero-3-fosfolina (DoPhPC)) y temperatura moderada (por ejemplo, menos de aproximadamente 100 °C). En algunos casos, mezclar DPhPC con una solución tampón crea grandes vesículas multilaminares (LMV), y añadir subunidades de aHL a esta solución e incubar la mezcla a 40 °C durante 30 minutos da como resultado la formación de poro.

Si se usan dos tipos diferentes de subunidades (por ejemplo, la proteína natural y un segundo monómero de aHL que puede contener una única mutación puntual), las proteínas resultantes pueden tener una estequiometría mixta (por ejemplo, proteínas naturales y mutantes). La estequiometría de estas proteínas puede seguir una fórmula que es dependiente de la proporción de las concentraciones de las dos proteínas usadas en la reacción de formación de poros. Esta fórmula es como sigue:

$$100 P_m = 100 \frac{n!}{m!(n-m)!} \cdot f_{mut}^m \cdot f_{wt}^{n-m}, \text{ donde}$$

P_m = probabilidad de que un poro tenga un número m de subunidades mutantes
 n = número total de subunidades (por ejemplo, 7 para aHL)
 m = número de subunidades "mutantes"
 f_{mut} = fracción o proporción de subunidades mutantes mezcladas entre sí
 f_{wt} = fracción o proporción de subunidades naturales mezcladas entre sí

El procedimiento puede comprender además fraccionar la pluralidad de proteínas para enriquecer proteínas que tienen una segunda proporción de las primeras subunidades con respecto a las segundas subunidades 2725. Por ejemplo, se pueden aislar proteínas de nanoporo que tienen una y solo una subunidad modificada (por ejemplo, una segunda proporción de 1:6). Sin embargo, cualquier segunda proporción es adecuada. También se puede fraccionar una distribución de segundas proporciones, tal como enriquecer proteínas que tienen una o bien dos subunidades modificadas. El número total de subunidades que forman la proteína no siempre es 7 (por ejemplo, se puede usar un nanoporo diferente o se puede formar un nanoporo de hemolisina alfa que tiene seis subunidades) como se representa en la figura 27 del documento WO2014/074727. En algunos casos, se enriquecen las proteínas que tienen solo una subunidad modificada. En dichos casos, la segunda proporción es 1 segunda subunidad por $(n-1)$ primeras subunidades donde n es el número de subunidades que comprenden la proteína.

La primera proporción puede ser la misma que la segunda proporción, sin embargo, esto no se requiere. En algunos casos, las proteínas que tienen monómeros mutados se pueden formar menos eficazmente que las que no tienen subunidades mutadas. Si este es el caso, la primera proporción puede ser mayor que la segunda proporción (por ejemplo, si se desea una segunda proporción de 1 subunidad mutada con respecto a 6 no mutadas en un nanoporo, formar un número adecuado de 1:6 proteínas puede requerir mezclar las subunidades en una proporción mayor que 1:6).

Las proteínas que tienen segundas proporciones de subunidades diferentes se pueden comportar de manera diferente (por ejemplo, tener diferentes tiempos de retención) en una separación. En algunos casos, las proteínas se fraccionan usando cromatografía, tal como cromatografía de intercambio iónico o cromatografía de afinidad. Puesto que las primeras y segundas subunidades pueden ser idénticas aparte de la modificación, el número de

modificaciones en la proteína puede servir como base para la separación. En algunos casos, las primera o bien segunda subunidades tienen una marca de purificación (por ejemplo, además de la modificación) para permitir o mejorar la eficacia del fraccionamiento. En algunos casos, se usa una marca polihistidina (marca His), una marca estreptavidina (marca Strep) u otra marca peptídica. En algunos casos, las primera y segunda subunidades comprenden cada una diferentes marcas y la etapa de fraccionamiento fracciona en base a cada marca. En el caso de una marca His, se crea una carga en la marca a pH bajo (los residuos de histidina se cargan positivamente por debajo del pKa de la cadena lateral). Con una diferencia significativa de carga en una de las moléculas de aHL en comparación con las otras, se puede usar cromatografía de intercambio iónico para separar los oligómeros que tienen 0, 1, 2, 3, 4, 5, 6 o 7 de las subunidades de aHL "marcadas con carga". En principio, esta marca de carga puede ser una cadena de cualquier aminoácido que porte una carga uniforme. La figura 28 y figura 29 muestran ejemplos de fraccionamiento de nanoporos en base a una marca His. La figura 28 muestra una curva de absorbancia ultravioleta a 280 nanómetros, absorbancia ultravioleta a 260 nanómetros y conductividad. Los picos corresponden a nanoporos con diversas proporciones de subunidades modificadas y no modificadas. La figura 29 del documento WO2014/074727 muestra el fraccionamiento de nanoporos de aHL y mutantes de los mismos usando tanto marca His como marcas Strep.

En algunos casos, una entidad (por ejemplo, una polimerasa) se une a la proteína después del fraccionamiento. La proteína puede ser un nanoporo y la entidad puede ser una polimerasa. En algunos casos, el procedimiento comprende además insertar las proteínas que tienen las segundas subunidades de proporción en una bicapa.

En algunas situaciones, un nanoporo puede comprender una pluralidad de subunidades. Se puede unir una polimerasa a una de las subunidades y al menos una y menos que todas las subunidades comprenden una primera marca de purificación. En algunos ejemplos, el nanoporo es hemolisina alfa o una variante de la misma. En algunos casos, todas las subunidades comprenden una primera marca de purificación o una segunda marca de purificación. La primera marca de purificación puede ser una marca polihistidina (por ejemplo, en la subunidad que tiene la polimerasa unida).

Polimerasa unida a nanoporo

En algunos casos, una polimerasa (por ejemplo, ADN polimerasa) se une y/o se localiza próxima al nanoporo. La polimerasa se puede unir al nanoporo antes o después de que el nanoporo se incorpore en la membrana. En algunos casos, el nanoporo y la polimerasa son una proteína de fusión (es decir, una única cadena polipeptídica).

La polimerasa se puede unir al nanoporo de cualquier forma adecuada. En algunos casos, la polimerasa se une al monómero proteico de nanoporo (por ejemplo, hemolisina) y a continuación se ensambla el heptámero de nanoporo completo (por ejemplo, en una proporción de un monómero con una polimerasa unida con respecto a 6 monómeros de nanoporo (por ejemplo, hemolisina) sin una polimerasa unida). El heptámero de nanoporo se puede insertar a continuación en la membrana.

Otro procedimiento para unir una polimerasa a un nanoporo implica unir una molécula conectora a un monómero de hemolisina o mutar un monómero de hemolisina para tener un sitio de unión y a continuación ensamblar el heptámero de nanoporo completo (por ejemplo, en una proporción de un monómero con conector y/o sitio de unión con respecto a 6 monómeros de hemolisina sin conector y/o sitio de unión). A continuación se puede unir una polimerasa al sitio de unión o conector de unión (por ejemplo, en bloque, antes de insertarlo en la membrana). La polimerasa también se puede unir al sitio de unión o conector de unión después de que el nanoporo (por ejemplo, heptámero) se forme en la membrana. En algunos casos, se inserta una pluralidad de pares nanoporo-polimerasa en una pluralidad de membranas (por ejemplo, dispuestas sobre los pocillos y/o electrodos) del biochip. En algunos casos, la unión de la polimerasa al complejo de nanoporo se produce en el biochip encima de cada electrodo.

La polimerasa se puede unir al nanoporo con cualquier química adecuada (por ejemplo, enlace covalente y/o conector). En algunos casos, la polimerasa se une al nanoporo con grapas moleculares. En algunos casos, las grapas moleculares comprenden tres secuencias de aminoácidos (indicados como conectores A, B y C). El conector A se puede extender desde un monómero de hemolisina, el conector B se puede extender desde la polimerasa, y el conector C puede unir a continuación los conectores A y B (por ejemplo, envolviendo ambos conectores A y B) y por tanto, la polimerasa al nanoporo. El conector C también se puede construir para formar parte del conector A o conector B, reduciendo por tanto el número de moléculas conectoras.

En algunos casos, la polimerasa se une al nanoporo usando la química Solulink™. Solulink™ puede ser una reacción entre HyNic (ácido 6-hidracino-nicotínico, una hidracina aromática) y 4FB (4-formilbenzoato, un aldehído aromático). En algunos casos, la polimerasa se une al nanoporo usando la química Click (disponible en Life Technologies, por ejemplo). En algunos casos, se introducen mutaciones de dedos de cinc en la molécula de hemolisina y a continuación se usa una molécula (por ejemplo, una molécula intermedia de ADN) para unir la polimerasa a los sitios de dedos de cinc en la hemolisina.

Otros conectores que pueden ser útiles en la unión de la polimerasa a un nanoporo son el enlace genético directo (por ejemplo, el conector de aminoácidos (GGGGS)₁₋₃ (SEQ ID NO: 4)), el enlace mediado por transglutaminasa

(por ejemplo, RSKLG (SEQ ID NO: 5)), el enlace mediado por sortasa y enlace químico a través de modificaciones de cisteína. Los conectores específicos contemplados como útiles en el presente documento son (GGGGS)₁₋₃ (SEQ ID NO: 4), marca K (RSKLG (SEQ ID NO: 5)) en el extremo N, sitio ΔTEV (12-25), sitio ΔTEV + extremo N de SpyCatcher (12-49).

5

Configuración del aparato

El nanoporo se puede formar o incluirse de otro modo en una membrana dispuesta adyacente a un electrodo sensor de un circuito sensor, tal como un circuito integrado. El circuito integrado puede ser un circuito integrado específico de la aplicación (ASIC). En algunos ejemplos, el circuito integrado es un transistor de efecto de campo o un semiconductor de óxido metálico complementario (CMOS). El circuito sensor puede estar situado en un chip u otro dispositivo que tenga el nanoporo, o fuera del chip o dispositivo, tal como en una configuración fuera de chip. El semiconductor puede ser cualquier semiconductor, incluyendo, sin limitación, los semiconductores del grupo IV (por ejemplo, silicio) y grupo III-V (por ejemplo, arseniuro de galio). Véase, por ejemplo, el documento WO 2013/123450, para la configuración del aparato y dispositivo para detectar un nucleótido o marca.

10

15

Se pueden usar sensores basados en poros (por ejemplo, biochips) para consulta electrónica de moléculas individuales. Un sensor basado en poros puede incluir un nanoporo de la presente divulgación formado en una membrana que se dispone adyacente o próxima a un electrodo sensor. El sensor puede incluir un contraelectrodo. La membrana incluye un lado *trans* (es decir, un lado orientado hacia el electrodo sensor) y un lado *cis* (es decir, un lado orientado hacia el contraelectrodo).

20

En la divulgación experimental que sigue, se aplican las siguientes abreviaturas: eq (equivalentes); M (Molar); μM (micromolar); N (normal); mol (moles); mmol (milimoles); μmol (micromoles); nmol (nanomoles); g (gramos); mg (miligramos); kg (kilogramos); μg (microgramos); l (litros); ml (mililitros); μl (microlitros); cm (centímetros); mm (milímetros); μm (micrómetros); nm (nanómetros); °C (grados centígrados); h (horas); min (minutos); s (segundos); ms (milisegundos).

25

30

EJEMPLOS

Las figuras adjuntas se deben considerar como partes integrales de la memoria descriptiva y la descripción de la invención. Todas las referencias citadas se incorporan específicamente en el presente documento por referencia para todo lo que se describe en el mismo. Se ofrecen los siguientes ejemplos para ilustrar, pero no para limitar la invención reivindicada.

35

Ejemplo 1

Mutagénesis dirigida

Este ejemplo ilustra la introducción de una mutación en una polimerasa pol6 en una posición deseada.

40

El ADN que codifica la pol6 natural marcada con His se adquirió de una fuente comercial (DNA 2.0, Menlo Park, California). Se verificó la secuencia por secuenciación.

Para el cribado de mutantes, se expresó la polimerasa tal cual (N-ter His-Pol6). Para someter a prueba los resultados positivos de pol en el chip, se genomanipuló en un dominio de SpyCatcher en el extremo N o extremo C de Pol6.

45

Las posiciones racionales para afectar a la unión de Pol6-nucleótido se identificaron en base a modelos de homología de estructuras cristalinas conocidas.

50

Para el cribado principal, cada una de las posiciones racionales se mutaron en Gly, Ala, Leu, Glu, Gln, Lys, His, Tyr, Pro, Trp, Thr o Met usando el protocolo de mutagénesis Q5.

Los cebadores para cada reacción de mutagénesis se diseñaron usando el protocolo de NEBaseChanger y se encargaron en formato de placa de 96 pocillos de IDT.

55

Los cebadores directo e inverso se fosforilaron en 5' en formato de alto rendimiento (HTP) usando la polinucleotidocinasa (PNK) T4 adquirida en NEB. Una reacción típica de 25 μl contenía 15 μl de cebador a 10 μM, 5 μl de tampón de reacción 5X (de NEB), 1,25 μl de enzima PNK, 3,75 μl de agua. Se realizó la reacción a 37 °C durante 30 min y se termoinactivó la enzima a 65 °C durante 20 min.

60

Se realizó la mutagénesis por PCR usando ADN polimerasa Q5 de NEB. Una reacción típica de 25 μl contenía 5 μl de tampón Q5, 5 μl de potenciador de GC, 0,5 μl de dNTP 10 mM, 1,25 μl de cebadores directo e inverso de mutagénesis fosforilada 10 μM, 0,25 μl de polimerasa Q5 y 1 μl de 5 ng/ml de molde de Pol6 natural, es decir, His-Pol6, y 10,75 μl de H₂O.

65

Una vez que se completa la PCR, se añadieron 0,5 µl de Dpn1 a 25 µl de la mezcla de PCR y se incubó a 37 °C durante 1 h.

5 Añadir 2,5 µl de producto de PCR tratado con Dpn1 con 2,5 µl de Blunt/TA Ligase Master Mix. Incubar a temperatura ambiente durante 1 h.

Añadir 1 µl de mezcla de fijación a 20 µl de células BL21DE3 en 96 pocillos (EMD Millipore) e incubar en hielo durante 5 min.

10 Choque térmico a 42 °C durante exactamente 30 s usando el dispositivo de PCR y poner en hielo durante 2 min.

Añadir 80 µl de SOC e incubar en una estufa de incubación a 37 °C durante 1 h sin agitar.

15 Añadir 100 µl de SOC o agua ultra pura y sembrarlas en placas con LB-agar de 48 pocillos con 50-100 µg/ml de kanamicina.

Ejemplo 2 Expresión y purificación

20 El siguiente ejemplo detalla cómo se expresaron y purificaron las variantes de pol6 usando un procedimiento de alto rendimiento.

25 El ADN que codifica las variantes en el vector pD441 (plásmido de expresión) se transformó en *E. coli* competente y se prepararon reservas de glicerol. Comenzando a partir de una pequeña selección de las reservas de glicerol, cultivar 1 ml de cultivo iniciador en LB con glucosa al 0,2 % y 100 µg/ml de kanamicina durante aproximadamente 8 h. Transferir 25 µl de cultivo iniciador en fase logarítmica a 1 ml de medio de expresión (medio de autoinducción *Terrific Broth* (TB) complementado con glucosa al 0,2 %, fosfato de potasio 50 mM, MgCl₂ 5 mM y 100 µg/ml de kanamicina) en placas de 96 pocillos profundos. Se incubaron las placas con agitación a 250-300 rpm durante 36-40 h a 28 °C.

35 A continuación, se recogieron las células por medio de centrifugación a 3200 × g durante 30 minutos a 4 °C. Se decantó el medio y se resuspendió el sedimento celular en 200 µl de tampón de lisis preenfriado (fosfato de potasio 20 mM pH 7,5, NaCl 100 mM, Tween20 al 0,5 %, TCEP 5 mM, imidazol 10 mM, PMSF 1 mM, BugBuster 1X, 100 µg/ml de inhibidores de lisozima y proteasa) e incubar a temperatura ambiente durante 20 min con agitación suave. A continuación, añadir 20 µl de una reserva 10x a una concentración final de 100 µg/ml de DNasa, MgCl₂ 5 mM, 100 µg/ml de RNasa I e incubar en hielo durante 5-10 min para producir un lisado. Complementar el lisado con 200 µl de fosfato de potasio 1 M, pH 7,5 (la concentración final será de aproximadamente 0,5 M de fosfato de potasio en 400 µl de lisado) y filtrar a través de placas de filtro Pall (pieza n.º 5053, filtros de 3 micrómetros) por medio de centrifugación a aproximadamente 1500 rpm a 4 °C durante 10 minutos. A continuación, se aplicaron los lisados clarificados a placas de cobalto His-Pur de 96 pocillos equilibradas (Pierce pieza n.º 90095) y se unieron durante 15-30 min.

45 Se recogió el flujo pasante (FT) por centrifugación a 500xg durante 3 min. A continuación, se lavó 3 veces el FT con 400 µl de tampón de lavado 1 (fosfato de potasio 0,5 M pH 7,5, NaCl 1 M TCEP 5 mM, imidazol 20 mM + Tween20 al 0,5 %). A continuación, se lavó el FT dos veces en 400 µl de tampón de lavado 2 (Tris 50 mM pH 7,4, KCl 200 mM, TCEP 5 mM, Tween20 al 0,5 %, imidazol 20 mM).

50 Se eluyó la Pol6 usando 200 µl de tampón de elución (Tris 50 mM pH 7,4, KCl 200 mM, TCEP 5 mM, Tween20 al 0,5 %, imidazol 300 mM, glicerol al 25 %) y se recogió después de una incubación de 1-2 min. Volver a aplicar eluido a la misma placa con His-Pur 2-3 veces para obtener Pol6 concentrada en el eluido. La polimerasa purificada es >95 % pura como se evalúa por SDS-PAGE. La concentración de proteína es de ~3 µM (0,35 mg/ml) con una proporción 260/280 de 0,6 como se evalúa por Nanodrop.

55 La actividad polimerasa se verifica por ensayo de desplazamiento de fluorescencia (véase el ejemplo 3).

Ejemplo 3 Determinación de actividad

60 Este ejemplo proporciona procedimientos para determinar la actividad de las polimerasas variantes.

Protocolo de ensayo de desplazamiento

65 Este ensayo caracteriza la capacidad de la polimerasa mutante de incorporar nucleótidos polifosfato en una hebra de ADN, así como su capacidad para desenrollar y desplazar el ADN bicatenario.

ES 2 804 843 T3

Los reactivos de partida son como sigue:

Bajo contenido de sal

	Reactivo A	Reactivo B
Reactivo	Concentración	Concentración
KCl	21,4 mM	20 mM
Bicina 7,9	26,75 mM	25 mM
EDTA	0,284 mM	N/A
Triton X-100	0,0535 %	0,05 %
DTT	5,35 mM	5 mM
BSA	26,75 µg/ml	25 µg/ml
Molde de FRET de ADN	71 nM	N/A
MgSO ₄	N/A	20 mM

5

N/A = no aplicable

Alto contenido de sal

	Reactivo A	Reactivo B
Reactivo	Concentración	Concentración
NaCl	75 mM	300 o 150 mM
HEPES 7,5	32,6 mM	25 mM
EDTA	0,3 mM	N/A
Triton X-100	0,065 %	0,05 %
TCEP	6,5 mM	5 mM
BSA	32,6 µg/ml	25 µg/ml
Molde de FRET de ADN	87 nM	N/A
MgCl	N/A	20 mM

10

N/A = no aplicable

Para el cribado de mutantes únicos y dobles:

15 Usando el reactivo A como diluyente, preparar 4 condiciones de nucleótidos diferentes a 1,42X:

1		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	28,4 µM	20 µM
dATP	21,3 µM	15 µM
dCTP	21,3 µM	15 µM
dGTP	21,3 µM	15 µM

2		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	2,84 µM	2 µM
dATP	21,3 µM	15 µM
dCTP	21,3 µM	15 µM
dGTP	21,3 µM	15 µM

3		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	0 µM	0 µM
dATP	21,3 µM	15 µM
dCTP	21,3 µM	15 µM
dGTP	21,3 µM	15 µM

4		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	0	0
dATP	0	0
dCTP	0	0
dGTP	0	0

Para el cribado de mutantes triples:

20 Usando el reactivo A como diluyente, preparar 4 condiciones de nucleótidos diferentes a 1,42X:

1		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	28,4 µM	20 µM
dAnP-NH ₂	28,4 µM	20 µM
dCnP-NH ₂	28,4 µM	20 µM
dGnP-NH ₂	28,4 µM	20 µM

2		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	1,42 µM	1 µM
dAnP-NH ₂	1,42 µM	1 µM
dCnP-NH ₂	1,42 µM	1 µM
dGnP-NH ₂	1,42 µM	1 µM

3		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	0 μM	0 μM
dATP	21,3 μM	15 μM
dCTP	21,3 μM	15 μM
dGTP	21,3 μM	15 μM

4		
Nucleótido	[1,42X]	[Final]
dTnP-NH ₂	0	0
dAnP-NH ₂	0	0
dCnP-NH ₂	0	0
dGnP-NH ₂	0	0

dNnP es un nucleótido polifosfato donde N es el nucleótido (es decir, A, T, C o G) y nP es 4-8 fosfatos

5 La condición de nucleótido 1 somete a prueba la actividad a alta concentración de hexafosfato.

La condición de nucleótido 2 somete a prueba la actividad a baja concentración de hexafosfato.

10 La condición de nucleótido 3 somete prueba la tasa de incorporación incorrecta (es decir, la fidelidad). Si una polimerasa mutante muestra actividad significativa con solo 3 de los 4 nucleótidos necesarios, entonces se concluye que no discrimina entre nucleótidos correctos o incorrectos mientras extiende una hebra de ADN.

15 La condición de nucleótido 4 somete a prueba la actividad exonucleasa. Si una polimerasa muestra actividad significativa sin nucleótidos presentes, entonces se concluye que la polimerasa presenta actividad exonucleasa.

Añadir a cada pocillo de reacción en una placa transparente de mitad de área de 96 pocillos:
23 μl de reactivo A/mezcla de nucleótidos
2 μl de polimerasa (1-10 μM)

20 Agitar a 800 rpm en un agitador de placas durante ~10 min.

Añadir 5 μl de NaCl 1,4 M a cada pocillo para llevar la concentración de NaCl hasta 300 mM o 5 μl de NaCl 525 mM a cada pocillo para llevar la concentración de NaCl hasta 150 mM.

25 Incubar durante 30 minutos.

En el lector de placas BMG LABTECH, inyectar 10 μl de reactivo B y leer la señal de fluorescencia durante de 2 a 10 min.

30 Los datos representativos del ensayo de desplazamiento para una polimerasa variante se muestran en la **FIG. 5**. La actividad de la polimerasa se midió usando el ensayo de desplazamiento en presencia de A. dTnP 20 μM + dA,C,G3P 15 μM (cuadrados rojos; ■), B. dTnP 5 μM + dA,C,G3P 15 μM (rombos azules; ◆), C. dA,C,G3P 15 μM (triángulos verdes; ▲) o D. en ausencia de nucleótidos (X moradas). A y B muestran que una variante mutante puede incorporar y extender a lo largo de un molde de ADN con un nucleótido polifosfato. C muestra que la variante no ha perdido su fidelidad y no incorpora incorrectamente nucleótidos aleatorios en ausencia de un nucleótido T. D. muestra que la señal generada no es el resultado de la actividad exonucleasa de la polimerasa en ausencia de todos los nucleótidos. Las cuatro curvas son representativas de una única variante sometida a prueba en 4 placas de ensayo diferentes como parte del cribado de polimerasa.

40 **Ejemplo 4** **Determinación de k_{dis}**

Se usó el siguiente ensayo de flujo interrumpido para determinar la velocidad k_{dis} de las polimerasas variantes.

45 Para el reactivo A, la polimerasa se enlaza a un cebador-molde de ADN marcado con fluoresceína con un nucleótido polifosfato unido a Cy3 (o Alexa555) en presencia de un metal divalente no catalítico como Ca²⁺. Esto forma un par FRET, siendo fluoresceína el fluoróforo donante y siendo Cy3 el fluoróforo aceptor. El reactivo B contiene el nucleótido de caza. Para los propósitos de este ensayo, el primer nucleótido que se va a incorporar en el molde/cebador es citosina.

50 Se preparó recientemente el reactivo A (NaCl 75 mM, HEPES 25 mM (pH 7,5), CaCl₂ 2 mM, fluoresceína-molde/cebador 250 nM, dCnP-Cy3 20 uM y polimerasa >250 nM) mezclando los componentes para garantizar que la polimerasa se añade al final. Permitir que la polimerasa se incube en el reactivo A durante 10 minutos.

55 Se preparó el reactivo B (NaCl 75 mM, HEPES 25 mM (pH 7,5), CaCl₂ 2 mM y dCTP 200 mM).

Cuando se mezclan el reactivo A y B, dCTP compite con dCnP-Cy3 por la asociación, se observa un incremento en la fluorescencia dado que la concentración de dCTP está en exceso. El ensayo se puede realizar con un dispositivo de interrupción de flujo (Kintek Corp) o bien un lector de placas fluorescente. El incremento de la fluorescencia frente

al tiempo se ajustó a una exponencial de primer orden o de segundo orden para proporcionar la constante cinética k_{dis} para esa polimerasa particular.

Los rendimientos de purificación y k_{dis} para variantes seleccionadas se presentan en la tabla 1.

5

Tabla 1 - Rendimientos de purificación y k_{dis} para variantes de Pol6 seleccionadas

	k_{dis} (s ⁻¹)	Rendimiento de prep. de 2,5 ml (µg)
Buen resultado positivo de k_{dis} y moderadamente buena actividad a 20 µM		
S366A+N535L+A547M	0,0039	77,58
Resultados positivos de 20 mod (buena actividad a 20 µM de nucleótido hexafosfato y muy baja o nula actividad a 0 µM)		
T651Y+P542E+N545K	0,0562	5,16
T651Y+P542E+Q546W	0,0201	64,58
S366A+P542E+N545K	0,0295	125,24
Resultados positivos de 20 MSR (buena actividad a 20 µM, actividad a 0 µM)		
S366A+P542E+I652Q	0,0440	196,63
T651Y+P542E+S366A	0,0583	10,11
Resultados positivos de 1 MSR/1 mod y también muestran una moderadamente buena actividad a 20 µM		
S366A+N535L+N545K	0,0103	60,25
T651Y+N535L+N545K	0,0273	210,66
S366A+N535L+T529M	0,0327	153,09
S366A+N535L+I652Q	0,0097	161,55
Resultado positivo de mutante doble		
T651Y+N535L	0,0312	281,84
S366A+N535L	0,0134	666,95

FP = polarización fluorescente

10

Véase la **FIG. 3** para una representación esquemática del ensayo y un gráfico de una reacción ejemplar. Véase la **FIG. 10** para datos representativos.

Ejemplo 5

15 Determinación de K_{DIS}

Este ejemplo proporciona un procedimiento alternativo usando polarización de fluorescencia para determinar la k_{dis} .

20 Se usó un tampón de ensayo que comprendía Tris 25 mM pH 7,0, KCl 75 mM, Triton-X100 al 0,01 %, BSA 1X (100 µg/ml), EDTA 0,5 mM, CaCl₂ 2 mM, DTT 2 mM, para preparar una mezcla maestra de ensayo que contenía molde de ADN marcado con fluoresceína de horquilla 250 nM y nucleótido marcado con dC6P-C6-Cy3 250 nM. Se añadieron cincuenta y cinco microlitros de la mezcla maestra a cada uno de los pocillos de una placa Costar negra de 96 pocillos; y se añadieron 20 µl de mutantes de polimerasa, que se habían purificado de cultivos de 1 ml, en un formato de alto rendimiento (HTP). Se agitó la placa en un agitador de placas durante 1 minuto para permitir la

25 formación de complejos ternarios homogéneos de polimerasa-molde de ADN-nucleótido. Se puso la placa en un lector de placa Polarstar BMG (BMG LABTECH Inc., Carolina del Norte) y se ajustó la milipolarización objetivo a 200 mP y al 10 % para obtener una ganancia de alrededor de 2000. Se estableció el filtro de excitación en 485 nM y se estableció el filtro de emisión en 590-20 nM. Se cebó el inyector con 1 ml de solución de nucleótido de caza dCTP 1 mM. Se recopilaron los datos con un mínimo de 30 destellos por pocillo por intervalo y un tiempo de lectura total de 60 s para el comienzo. Los destellos se incrementaron a 50 o más y se tomaron tiempos de lectura más

30 largos para los mutantes con resultado positivo que mostraron una disociación lenta. La recopilación de datos empezó con la inyección de 25 µl de dCTP 1 mM.

35

Véase la **FIG. 4** para una representación esquemática del ensayo y un gráfico de una reacción ejemplar.

Véase la **FIG. 6** para datos representativos del ensayo de k_{dis} basado en polarización de fluorescencia para dos polimerasas variantes (S366A+N535L+I652Q (B6) y S366A+P542E+I652Q (C6)). mP es milipolarización. El

complejo ternario preformado de polimerasa-molde de ADN-dCnP-Alexa555 se caza con dCTP natural y se realizó un seguimiento de dCnP-Alexa555 de polarización en el tiempo.

Ejemplo 6

5 Determinación de K_{quim}

Este ejemplo proporciona un ensayo basado en FRET para determinar la k_{quim} para las polimerasas variantes.

10 Para el reactivo A, la polimerasa se enlaza al cebador-molde de ADN marcado con fluoresceína. El reactivo B contiene nucleótido polifosfato unido a Cy3 (o Alexa555) en presencia de un metal divalente catalítico como Mg^{2+} . Para los propósitos de este protocolo, el primer nucleótido que se va a incorporar en el molde/cebador es citosina.

15 Se preparó el reactivo A (NaCl 75 mM, HEPES 25 mM (pH 7,5), fluoresceína-molde/cebador 250 nM, polimerasa >250 nM). Se dejó incubar la polimerasa en el reactivo A durante 10 min.

Se preparó el reactivo B (NaCl 75 mM, HEPES 25 mM (pH 7,5), $MgCl_2$ 10 mM y dCnP-Cy3 20 μ M).

20 Cuando se mezclan el reactivo A y B, el complejo polimerasa-fluoresceína-molde-cebador se une a dCnP-Cy3 y se extingue la fluorescencia. Mg^{2+} permite que la polimerasa incorpore el nucleótido que libera el producto de escisión, pirofosfato con Cy3 unido, nP-Cy3. Puesto que se libera la molécula de extinción, se incrementa la fluorescencia. El ensayo se puede realizar con un dispositivo de interrupción de flujo (Kintek Corp) o bien un lector de placas fluorescente.

25 Véase la FIG. 2 para una representación esquemática del ensayo y un gráfico de una reacción ejemplar. Véase la FIG. 9 para datos representativos.

Ejemplo 7

Unión a nanoporo

30 Este ejemplo proporciona procedimientos de unión de una polimerasa variante a un nanoporo, por ejemplo, hemolisina α .

35 Se puede acoplar la polimerasa al nanoporo por cualquier medio adecuado. Véanse, por ejemplo, los documentos PCT/US2013/068967 (publicado como WO2014/074727; Genia Technologies, Inc.), PCT/US2005/009702 (publicado como W02006/028508; Presidente y miembros de la Universidad de Harvard) y PCT/US2011/065640 (publicado como WO2012/083249; Universidad de Columbia).

40 La polimerasa, por ejemplo, una ADN polimerasa pol6 variante, se acopló a un nanoporo proteico (por ejemplo, hemolisina alfa), a través de una molécula conectora. Específicamente, se usó el sistema SpyTag y SpyCatcher, que forma espontáneamente enlaces isopeptídicos covalentes en condiciones fisiológicas. Véase, por ejemplo, Li *et al.*, J Mol Biol., 23 de enero de 2014;426(2):309-17.

45 Se expresó la marca His de variante de pol6 SpyCatcher de acuerdo con el ejemplo 2 y se purificó usando una columna de afinidad de cobalto. Se incubaron la polimerasa SpyCatcher y la proteína de nanoporo oligomerizada SpyTag durante la noche a 4 °C en $SrCl_2$ 3 mM. A continuación se purificó el complejo 1:6-polimerasa-molde usando cromatografía de exclusión por tamaño.

50 Se unió el conector al extremo N o bien al extremo C de la variante de pol6. Se descubrió que las variantes unidas de forma N terminal eran más robustas, por ejemplo, más estables. Por lo tanto, se usaron conectores unidos de forma N terminal.

Ejemplo 8

Actividad en un biochip

55 Este ejemplo demuestra la capacidad de una polimerasa variante enlazada a nanoporo para unir nucleótidos marcados y, por lo tanto, permite la detección de corrientes de canal bloqueado en el nanoporo al que está unida la polimerasa.

60 La polimerasa se unió a un nanoporo y se incluyó en una bicapa lipídica sobre un pocillo en un chip sensor semiconductor, también llamado biochip. Se formó la bicapa lipídica y se insertó el nanoporo con polimerasa unida como se describe en el documento PCT/US2014/061853 (titulado "Methods for Forming Lipid Bilayers on Biochips" y presentado el 22 de octubre de 2014).

Las polimerasas variantes se complejaron con ADN molde en condiciones de bajo contenido de sal.

65 Se determinó la capacidad de la polimerasa variante enlazada a nanoporo de unir nucleótidos marcados en

experimentos de captura estática con lo que los nucleótidos marcados se enlazan por la polimerasa, y se mide la corriente de canal bloqueado a medida que el nucleótido marcado se presenta al nanoporo. Los experimentos de captura estática se realizan en presencia de Ca^{2+} , lo que evita la catálisis y el alargamiento del ADN, y permite la detección de capturas repetidas del mismo tipo de nucleótido marcado. En este experimento, el nucleótido marcado usado fue dTnP-marca.

Una Pol6 variante de polimerasa ejemplar (S366A+N535L+I652Q) se acopló a un nanoporo de hemolisina alfa en el biochip, y se complejó con ADN molde.

Se registró la captura estática del nucleótido timidina marcado (la marca es T30 (SEQ ID NO: 9)) por el complejo Pol6 (S366A+N535L+I652Q)-ADN a 100 mV en presencia de Hepes 7,5 20 mM, NaCl 300 mM, CaCl_2 3 mM y TCEP 5 mM por encima y por debajo de la bicapa.

[0019] Los resultados se muestran en las FIG. 7 y 8. Los trazados en la FIG. 7 muestran la corriente electrolítica medida a 100 mV a través del poro como función del tiempo. La corriente de poro abierto a esta tensión fue de aproximadamente 1 nA (trazado superior); y la corriente de poro bloqueado a la misma tensión fue de aproximadamente 0,33 nA (trazado del medio). La corriente de canal abierto se normalizó a 1 de acuerdo con el programa informático del sistema, y la corriente de canal bloqueado disminuyó por el dTnP-T30 (SEQ ID NO: 9) a un 33 % de la corriente de canal abierto. Los bloqueos de corriente mostrados en este trazado están asociados con la unión de timidina polifosfato por el complejo Pol6 variante-ADN, que se produce próximo al nanoporo. El histograma superior correspondiente en la FIG. 7 (derecha) muestra la frecuencia de los bloqueos de corriente observados a 100 mV con un cambio en la corriente normalizada a la corriente de poro abierto en el mismo poro; y el histograma (derecha) correspondiente al trazado del medio muestra la frecuencia de los bloqueos de corriente observados a 100 mV con un cambio en la corriente normalizada a la corriente de poro bloqueado en el mismo poro.

La FIG. 8 muestra el tiempo de permanencia para la captura estática de la timidina marcada mostrada en la FIG. 7. La FIG. 8 (izquierda) muestra un histograma del número de veces que se enlazó dTNP marcado por Pol6 variante como función de la corriente como se normaliza con respecto a la corriente de canal abierto. Se determinó que el tiempo de permanencia promedio de cada captura de nucleótido marcado con dTNP era de 1,2 segundos. La captura de fondo (es decir, no mediada por polimerasa) de la marca en el poro tiene un tiempo de permanencia en el intervalo de unos pocos milisegundos (datos no mostrados). Un objetivo en la evolución de la enzima era mejorar la velocidad de disociación del nucleótido polifosfato marcado, para que se puedan ver tiempos de permanencia lo suficientemente largos para registrar una captura mediada por polimerasa que se distinga bien del fondo. Como se muestra en la FIG. 8. el tiempo de permanencia promedio de 1,2 s está muy por encima del fondo. El índice celular es un esquema basado en colores para las aproximadamente 8000 células presentes en el chip usado en este experimento.

Los datos muestran que la polimerasa variante ejemplar, Pol6 (S366A+N535L+I652Q), puede unir nucleótidos marcados y permitir la detección del cambio de corriente a través del nanoporo al que está unida la polimerasa.

Los resultados proporcionan pruebas de que las polimerasas variantes unidas a nanoporos en un biochip pueden unir nucleótidos marcados con alta fidelidad, y presentar los nucleótidos marcados a los nanoporos para tiempos de permanencia que proporcionan tiempo suficiente para la detección de incorporación de nucleótidos, y posiblemente para disminuir la probabilidad de errores de secuenciación, por ejemplo, inserciones, deleciones, etc., durante la secuenciación de nanoporos.

Ejemplo 9 Ensayo de amplificación de círculo rodante

Este ejemplo describe la amplificación de un molde de polinucleótido en un ensayo basado en círculo rodante.

El molde usado fue un molde interno, HFcirc10. Es un molde circular simple de ~150 pb de largo.

El ensayo se procesó en un volumen de reacción total de 40 μl (28 μl de reactivo A + 2 μl de polimerasa 2 μM + 10 μl de reactivo B).

Reactivo A:

Kglu	75mM
HEPES 7,5	25mM
EDTA	0,2mM
Triton X-100	0,05%
TCEP	5mM
BSA	25 $\mu\text{g/ml}$
Molde circular cebado	100nM

dNTP/dN6P/marcas	25µM
------------------	------

Reactivo B:

HEPES 7,5	25mM
Kglu	Variado mM
Triton X-100	0,05%
TCEP	5mM
BSA	25µg/ml
MgCl ₂	40mM

5 Se añadieron dos µl de polimerasa 2 µM a 28 µl de reactivo A para dar una proporción molar 1:1 de ADN con respecto a polimerasa (100 nM cada uno) en la mezcla de ensayo de 40 µl final. Se incubó la mezcla de reactivo A/polimerasa durante 10 min en esta condición de sal 75 mM para permitir que la polimerasa se uniera al ADN.

Seguidamente, se añadieron 10 µl de reactivo B a la mezcla de reactivo A/polimerasa para comenzar la reacción.

10 En puntos de tiempo predeterminados, se retiraron 10 µl de muestras de la reacción y se añadieron a 10 µl de formamida con EDTA 50 mM para desactivar la reacción. Se tomaron muestras en puntos de tiempo de 0 min, 10 min, 30 min y 40 minutos.

15 Se calentaron las muestras de formamida hasta 94 °C durante aproximadamente 3 min para desnaturalizar proteínas y estructuras secundarias de ADN. No se dejaron enfriar las muestras hasta 4 °C. Añadir 2 µl de tinte 100X SYBR GREEN o GOLD.

Se procesaron 15 µl de cada muestra en un gel de agarosa al 1,2 % durante 1 hora y 15 minutos a 100 V.

20 Se obtuvo una imagen del gel usando la bandeja azul para el generador de imágenes GEL DOC EZ de Biorad.

La **FIG. 11** muestra los resultados de puntos de tiempo de 40 minutos del ensayo. Se muestra una escalera molecular en los carriles 1 y 19, numerando de izquierda a derecha. El carril 2 tiene una muestra de t=0; ningún producto es visible. Cada uno de los carriles 3-18 es una variante diferente de polimerasa pol6.

Todas las variantes mostradas pueden realizar desplazamiento de hebra y generar productos de ADN de kilobases de largo con todos los nucleótidos hexafosfato.

30 **Ejemplo 10**
Secuenciación de ADN molde usando nucleótidos marcados

Este ejemplo demuestra que la polimerasa variante es funcional en una secuenciación por procedimiento de síntesis en un biochip.

35 Secuenciación con AC de un molde de heteropolímero usando polimerasa Pol6-26i-D44A en Hepes 20 mM, pH 8, glutamato de potasio 500 mM y MgCl₂ 3 mM a temperatura ambiente. Se incluyó un nanoporo con polimerasa unida en la bicapa lipídica como se describe en el presente documento. Se añadió ADN cebado y se dejó que se complejara con la polimerasa. Se añadieron cuatro nucleótidos marcados diferentes a una concentración de 25 µM.
40 La secuenciación por síntesis puede proceder como se describe en el documento WO 2014/074727 titulado "Nucleic Acid Sequencing Using Tags". El trazado en la **FIG. 12** muestra una exactitud de secuenciación de un 76 % para un molde de heteropolímero y una longitud de lectura de 96 pb.

TEXTO LIBRE DEL LISTADO DE SECUENCIAS

SEQ ID NO:1 - Pol6 natural (ADN polimerasa [fago de *Clostridium* phiCPV4]; GenBank: AFH27113.1)

1 mdkhtqyvke hsfnydeykk anfdkiecli fdtesctnye ndntgarvyg
wglgvtrnhn 061 miyggqnlqf wevcqnifnd wyhdnkhtik
itkttkkgfpk rkyikfpiav hnlgdvfl 121 kyslvengfn ydkgllktvf
skgapqvtv dveepktfhi vqnnnivygc nvymdkffev 181 enkdgsttei
glclldffdsy kiitcaesqf hnyvhvdvpm fykmgeeydy dtwrspthkq 241
ttlelryqyn diymlrevie qfyidglcgg elpltgmrtta ssiavnvlkk mtfgeektee
301 gyinyfeldk ktkfeflrkr iemesytggy thanhkavgk tinkigcsld
inssypsma 361 ykvfpygkp rktwgrkpkt eknevyliev gfdvpepkhe
eyaldifkig avnskalspi 421 tgavsgqeyf ctnikdgkai pvykelkdtk
ltnynvvlv sveyefwikh fnfgvfkde 481 ydcfevdnle ftglkigsil
yykaekgkfk pyvdhftkmm venkklgnkp ltnqakliln 541 gaygkfgtkq
nkeekdlimd knlltftgs vteyegkefy rpyasfvtay grlqlwnaii 601
yavgvenfly ctdtsiycnr evnslidmn aigetidkti lgkwdvehvf dkfkvlqgkk
661 ymyhdckedk tdlkccglps darkiiiggg fdefylgknv egkkqrkkvi
5 ggcllldtlf 721 tikkimf

SEQ ID NO:2 - Pol6 (con marca His)

MHHHHHHHHS GGSDKHTQYV KEHSFNYDEY KKANFDKIEC LIFDTECTN 50
YENDNTGARV YGWGLGVTRN HNMIYGQNLN QFWEVCQNIF NDWYHDNKHT
100

10

ES 2 804 843 T3

IKITKTKKGF PKRKYIKFPI AVHNLGWDVE FLKYSLVENG FNYDKGLLKT
 150
 VFSKGAPYQT VTDVEEPKTF HIVQNNNIVY GCNVYMDKFF EVENKDGSTT
 200
 EIGLCLDFFD SYKIITCAES QFHNYVHDVD PMFYKMGEY DYDTWRSPH
 250
 KQTTLELRYQ YNDIYMLREV IEQFYIDGLC GGELPLTGMR TASSIAFNVL
 300
 KKMTFGEEKT EEGYINYFEL DKKTKFEFLR KRIEMESYTG GYTHANHKAV
 350
 GKTINKIGCS LDINSSYPSQ MAYKVFPY GK PVRKTWGRKP KTEKNEVYLI
 400
 EVGFDFVEPK HEEYALDIFK IGAVNSKALS PITGAVSGQE YFCTNIKD GK
 450
 AIPVYKELKD TKLTTNYNV LTSVEYEFWI KHFNFVGFKK DEYDCFEVDN
 500
 LEFTGLKIGS ILYYKAEK GK FKP YVDHFTK MKVENKKLGN KPLTNQAKLI
 550
 LNGAYGKFGT KQNKEEKDLI MDKNGLLTFT GSVTEYEGKE FYR PYASFVT
 600
 AYGRLQLWNA I IYAVGVENF LYCDTDSIYC NREVNSLIED MNAIGETIDK
 650
 TILGKWDVEH VFDKFKVLGQ KKYMYHDCKE DKTDLKCCGL PSDARKIIIG
 700
 QGFDEFYLGK NVEGKKQRKK VIGGCLLLDT LFTIKKIMF*

73

9

SEQ ID NO:3 - Pol6 con marca His (secuencia de ADN)

ATG CATCACC ATCATCATCA CCACCAC AGC	GGCGGTTCCG	ACAAACACAC	50
GCAGTACGTC AAAGAGCATA GCTTCAATTA	TGACGAGTAT	AAGAAAGCGA	100
ATTTTCGACAA GATCGAGTGC CTGATCTTTG	ACACCGAGAG	CTGCACGAAT	150
TATGAGAACG ATAATACCGG TGCACGTGTT	TACGGTTGGG	GTCTTGGCGT	200
CACCCGCAAC CACAATATGA TCTACGGCCA	AAATCTGAAT	CAGTTTTGGG	250
AAGTATGCCA GAACATTTTC AATGATTGGT	ATCACGACAA	CAAACATACC	300
ATTAAGATTA CCAAGACCAA GAAAGGCTTC	CCGAAACGTA	AGTACATTAA	350

5

ES 2 804 843 T3

GTTTCCGATT	GCAGTTCACA	ATTTGGGCTG	GGATGTTGAA	TTCCTGAAGT	400
ATAGCCTGGT	GGAGAATGGT	TTCAATTACG	ACAAGGGTCT	GCTGAAAAC	450
GTTTTTAGCA	AGGGTGCGCC	GTACCAAACC	GTGACCGATG	TTGAGGAACC	500
GAAAACG TTC	CATATCGTCC	AGAATAACAA	CATCGTTTAT	GGTTGTAACG	550
TGTATATGGA	CAAATTC TTT	GAGGTCGAGA	ACAAAGACGG	CTCTACCACC	600
GAGATTGGCC	TGTGCTTGGA	TTTCTTCGAT	AGCTATAAGA	TCATCACGTG	650
TGCTGAGAGC	CAGTTC CACA	ATTACGTTCA	TGATGTGGAT	CCAATGTTCT	700
ACAAAATGGG	TGAAGAGTAT	GATTACGATA	CTTGGCGTAG	CCCGACGCAC	750
AAGCAGACCA	CCCTGGAGCT	GCGCTACCAA	TACAATGATA	TCTATATGCT	800
GCGTGAAGTC	ATCGAACAGT	TTTACATTGA	CGGTTTATGT	GGCGGCGAGC	850
TGCCGCTGAC	CGGCATGCGC	ACCGCTTCCA	GCATTGCGTT	CAACGTGCTG	900
AAAAAGATGA	CCTTTGGTGA	GGAAAAGACG	GAAGAGGGCT	ACATCAACTA	950
TTTTGAATTG	GACAAGAAAA	CCAAATTCGA	GTTTCTGCGT	AAGCGCATTG	1000
AAATGGAATC	GTACACCGGT	GGCTATACGC	ACGCAAATCA	CAAAGCCGTT	1050
GGTAAGACTA	TTAACAAGAT	CGGTTGCTCT	TTGGACATTA	ACAGCTCATA	1100
CCCTTCGCAG	ATGGCGTACA	AGGTCTTTCC	GTATGGCAA	CCGGTTCGTA	1150
AGACCTGGGG	TCGTAAACCA	AAGACCGAGA	AGAACGAAGT	TTATCTGATT	1200
GAAGTTGGCT	TTGACTTCGT	GGAGCCGAAA	CACGAAGAAT	ACGCGCTGGA	1250
TATCTTTAAG	ATTGGTGCGG	TGAACTCTAA	AGCGCTGAGC	CCGATCACCG	1300
GCGCTGTCAG	CGGTCAAGAG	TATTTCTGTA	CGAACATTAA	AGACGGCAA	1350
GCAATCCCGG	TTTACAAAGA	ACTGAAGGAC	ACCAAATTGA	CCACTAACTA	1400
CAATGTCGTG	CTGACCAGCG	TGGAGTACGA	GTTCTGGATC	AAACACTTCA	1450
ATTTTGGTGT	GTTTAAGAAA	GACGAGTACG	ACTGTTTCGA	AGTTGACAAT	1500
CTGGAGTTTA	CGGGTCTGAA	GATTGGTTCC	ATTCTGTACT	ACAAGGCAGA	1550
GAAAGGCAAG	TTTAAACCTT	ACGTGGATCA	CTTCACGAAA	ATGAAAGTGG	1600
AGAACAAGAA	ACTGGGTAAT	AAGCCGCTGA	CGAATCAGGC	AAAGCTGATT	1650
CTGAACGGTG	CGTACGGCAA	ATTCGGCACC	AAACAAAACA	AAGAAGAGAA	1700
AGATTTGATC	ATGGATAAGA	ACGGTTTGCT	GACCTTCACG	GGTAGCGTCA	1750
CGGAATACGA	GGGTAAAGAA	TTCTATCGTC	CGTATGCGAG	CTTCGTTACT	1800
GCCTATGGTC	GCCTGCAACT	GTGGAACGCG	ATTATCTACG	CGGTTGGTGT	1850
GGAGAATTTT	CTGTACTGCG	ACACCGACAG	CATCTATTGT	AACCGTGAAG	1900
TTAACAGCCT	CATTGAGGAT	ATGAACGCCA	TTGGTGAAAC	CATCGATAAA	1950
ACGATTCTGG	GTAATGGGA	CGTGGAGCAT	GTCTTTGATA	AGTTTAAGGT	2000
CCTGGGCCAG	AAGAAGTACA	TGTATCATGA	TTGCAAAGAA	GATAAAACGG	2050
ACCTGAAGTG	TTGCGGTCTG	CCGAGCGATG	CCCGTAAGAT	TATCATTGGT	2100
CAAGGTTTCG	ACGAGTTTTA	TCTGGGCAAA	AATGTCGAAG	GTAAGAAGCA	2150
ACGCAAAAAA	GTGATCGGCG	GTTGCCTGCT	GCTGGACACC	CTGTTTACGA	2200
TCAAGAAAAT	CATGTTCTAA				2220

Lista de citasLiteratura de patentes

- 5 [1] Documento PCT/US2005/009702 (publicado como WO2006/028508 el 16 de marzo de 2006; Presidente y miembros de la Universidad de Harvard; titulado METHODS AND APPARATUS FOR CHARACTERIZING POLYNUCLEOTIDES).
- 10 [2] Documento PCT/US2011/065640 (publicado como WO2012/083249 el 21 de junio de 2012; Universidad de Columbia; titulado DNA SEQUENCING BY SYNTHESIS USING MODIFIED NUCLEOTIDES AND NANOPORE DETECTION).
- [3] Documento PCT/US2013/068967 (publicado como WO2014/074727 el 15 de mayo de 2014; Genia Technologies; titulado NUCLEIC ACID SEQUENCING USING TAGS).
- [4] Documento PCT/US2013/046012 (Genia Technologies, Inc., titulado CHIP SET-UP AND HIGH-ACCURACY NUCLEIC ACID SEQUENCING, publicado el 19 de diciembre de 2013 como WO2013/188841).
- 15 [5] Documento US 2013/0053544 (Isis Innovation Limited) titulado Peptide Tag Systems That Spontaneously Form an Irreversible Link to Protein Partners via Isopeptide Bonds, publicado el 28 de febrero de 2013.

Literatura distinta de patentes

- 20 [1] Altschul, S. F. *et al.*, *J. Mol. Biol.* (1990) 215:403-410.
- [2] Altschul, S. F., *et al.*, *Nucleic Acids Res.* 25:3389-3402, 1997.
- [3] Ausubel, Frederick *et al.*, (1992) *Short Protocols in Molecular Biology, Current Protocols in Molecular Biology*, 2.^a ed., Greene Publishing Associates & John Wiley & Sons. New York, N.Y.
- 25 [4] Gardner *et al.*, *Nucleic Acids Res.* (2012) páginas 1-12 (*doi: 10.1093/nar/gks330*; Publicado en primer lugar en línea: 8 de mayo de 2012).
- [5] Hale y Marham, *THE HARPER COLLINS DICTIONARY OF BIOLOGY*, Harper Perennial, NY (1991).
- [6] Johnson, *et al.*, *Biochim Biophys Acta.* Mayo 2010; 1804(5):1041-1048.
- [7] Kong *et al.*, (1993) *J. Biol. Chem.* 268(3):1965-1975.
- [8] Lawyer *et al.*, (1989) *J. Biol. Chem.* 264:6427-647.
- 30 [9] Li *et al.*, *Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag; J Mol Biol.* (23 de enero de 2014) 426(2):309-17.
- [10] Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual* (2.^a ed., Cold Spring Harbor Laboratory Press, NY).
- [11] Sambrook *et al.*, (2001) *Molecular Cloning: A Laboratory Manual* (3.^a ed., Cold Spring Harbor Laboratory Press, NY) en 9,63-9,75 (que describe el marcaje terminal de ácidos nucleicos).
- 35 [12] Singleton, *et al.*, *DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY*, 2.^a ed., John Wiley and Sons, New York (1994).
- [13] Watson, J. D. *et al.*, en: *Molecular Biology of the Gene*, 4.^a ed., W. A. Benjamin, Inc., Menlo Park, Calif. (1987)).
- 40 [14] Zakari y Howarth, (2010) *Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting*, *J. Am. Chem. Soc.*, 132(13):4526-4527.
- [15] Zakari, B. *et al.*, (2012) *Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesion*, *PNAS* 109 (12):E690-E697.

LISTADO DE SECUENCIAS

- 5 <110> GENIA TECHNOLOGIES, INC.
ROCHE MOLECULAR SYSTEMS, INC.
ROCHE DIAGNOSTICS GMBH
- <120> VARIANTES DE POLIMERASA
- 10 <130> P32651-WO
- <140>
<141>
- 15 <150> 62/202.895
<151> 09/08/2015
- <150> 62/161.571
<151> 14/05/2015
- 20 <150> 62/111.034
<151> 02/02/2015
- <160> 9
- 25 <170> PatentIn versión 3.5
- <210> 1
<211> 727
<212> PRT
- 30 <213> Fago de *Clostridium*
- <220>
<221> fuente
<223> /nota=" fago de *Clostridium* phiCPV4"
- 35 <400> 1

```

Met Asp Lys His Thr Gln Tyr Val Lys Glu His Ser Phe Asn Tyr Asp
 1           5           10           15

Glu Tyr Lys Lys Ala Asn Phe Asp Lys Ile Glu Cys Leu Ile Phe Asp
 20           25           30

Thr Glu Ser Cys Thr Asn Tyr Glu Asn Asp Asn Thr Gly Ala Arg Val
 35           40           45

Tyr Gly Trp Gly Leu Gly Val Thr Arg Asn His Asn Met Ile Tyr Gly
 50           55           60

Gln Asn Leu Asn Gln Phe Trp Glu Val Cys Gln Asn Ile Phe Asn Asp
 65           70           75           80

Trp Tyr His Asp Asn Lys His Thr Ile Lys Ile Thr Lys Thr Lys Lys
 85           90           95

Gly Phe Pro Lys Arg Lys Tyr Ile Lys Phe Pro Ile Ala Val His Asn
 100          105          110
    
```

ES 2 804 843 T3

Leu Gly Trp Asp Val Glu Phe Leu Lys Tyr Ser Leu Val Glu Asn Gly
 115 120 125
 Phe Asn Tyr Asp Lys Gly Leu Leu Lys Thr Val Phe Ser Lys Gly Ala
 130 135 140
 Pro Tyr Gln Thr Val Thr Asp Val Glu Glu Pro Lys Thr Phe His Ile
 145 150 155 160
 Val Gln Asn Asn Asn Ile Val Tyr Gly Cys Asn Val Tyr Met Asp Lys
 165 170 175
 Phe Phe Glu Val Glu Asn Lys Asp Gly Ser Thr Thr Glu Ile Gly Leu
 180 185 190
 Cys Leu Asp Phe Phe Asp Ser Tyr Lys Ile Ile Thr Cys Ala Glu Ser
 195 200 205
 Gln Phe His Asn Tyr Val His Asp Val Asp Pro Met Phe Tyr Lys Met
 210 215 220
 Gly Glu Glu Tyr Asp Tyr Asp Thr Trp Arg Ser Pro Thr His Lys Gln
 225 230 235 240
 Thr Thr Leu Glu Leu Arg Tyr Gln Tyr Asn Asp Ile Tyr Met Leu Arg
 245 250 255
 Glu Val Ile Glu Gln Phe Tyr Ile Asp Gly Leu Cys Gly Gly Glu Leu
 260 265 270
 Pro Leu Thr Gly Met Arg Thr Ala Ser Ser Ile Ala Phe Asn Val Leu
 275 280 285
 Lys Lys Met Thr Phe Gly Glu Glu Lys Thr Glu Glu Gly Tyr Ile Asn
 290 295 300
 Tyr Phe Glu Leu Asp Lys Lys Thr Lys Phe Glu Phe Leu Arg Lys Arg
 305 310 315 320
 Ile Glu Met Glu Ser Tyr Thr Gly Gly Tyr Thr His Ala Asn His Lys
 325 330 335
 Ala Val Gly Lys Thr Ile Asn Lys Ile Gly Cys Ser Leu Asp Ile Asn
 340 345 350
 Ser Ser Tyr Pro Ser Gln Met Ala Tyr Lys Val Phe Pro Tyr Gly Lys

ES 2 804 843 T3

Leu Tyr Cys Asp Thr Asp Ser Ile Tyr Cys Asn Arg Glu Val Asn Ser
610 615 620

Leu Ile Glu Asp Met Asn Ala Ile Gly Glu Thr Ile Asp Lys Thr Ile
625 630 635 640

Leu Gly Lys Trp Asp Val Glu His Val Phe Asp Lys Phe Lys Val Leu
645 650 655

Gly Gln Lys Lys Tyr Met Tyr His Asp Cys Lys Glu Asp Lys Thr Asp
660 665 670

Leu Lys Cys Cys Gly Leu Pro Ser Asp Ala Arg Lys Ile Ile Ile Gly
675 680 685

Gln Gly Phe Asp Glu Phe Tyr Leu Gly Lys Asn Val Glu Gly Lys Lys
690 695 700

Gln Arg Lys Lys Val Ile Gly Gly Cys Leu Leu Leu Asp Thr Leu Phe
705 710 715 720

Thr Ile Lys Lys Ile Met Phe
725

<210> 2
<211> 739
5 <212> PRT
<213> Secuencia artificial

<220>
<221> fuente
10 <223> /nota="Descripción de secuencia artificial: polipéptido sintético"

<400> 2
Met His His His His His His His His Ser Gly Gly Ser Asp Lys His
1 5 10 15

Thr Gln Tyr Val Lys Glu His Ser Phe Asn Tyr Asp Glu Tyr Lys Lys
20 25 30

Ala Asn Phe Asp Lys Ile Glu Cys Leu Ile Phe Asp Thr Glu Ser Cys
35 40 45

Thr Asn Tyr Glu Asn Asp Asn Thr Gly Ala Arg Val Tyr Gly Trp Gly
50 55 60

Leu Gly Val Thr Arg Asn His Asn Met Ile Tyr Gly Gln Asn Leu Asn
65 70 75 80

ES 2 804 843 T3

Gln Phe Trp Glu Val Cys Gln Asn Ile Phe Asn Asp Trp Tyr His Asp
85 90 95

Asn Lys His Thr Ile Lys Ile Thr Lys Thr Lys Lys Gly Phe Pro Lys
100 105 110

Arg Lys Tyr Ile Lys Phe Pro Ile Ala Val His Asn Leu Gly Trp Asp
115 120 125

Val Glu Phe Leu Lys Tyr Ser Leu Val Glu Asn Gly Phe Asn Tyr Asp
130 135 140

Lys Gly Leu Leu Lys Thr Val Phe Ser Lys Gly Ala Pro Tyr Gln Thr
145 150 155 160

Val Thr Asp Val Glu Glu Pro Lys Thr Phe His Ile Val Gln Asn Asn
165 170 175

Asn Ile Val Tyr Gly Cys Asn Val Tyr Met Asp Lys Phe Phe Glu Val
180 185 190

Glu Asn Lys Asp Gly Ser Thr Thr Glu Ile Gly Leu Cys Leu Asp Phe
195 200 205

Phe Asp Ser Tyr Lys Ile Ile Thr Cys Ala Glu Ser Gln Phe His Asn
210 215 220

Tyr Val His Asp Val Asp Pro Met Phe Tyr Lys Met Gly Glu Glu Tyr
225 230 235 240

Asp Tyr Asp Thr Trp Arg Ser Pro Thr His Lys Gln Thr Thr Leu Glu
245 250 255

Leu Arg Tyr Gln Tyr Asn Asp Ile Tyr Met Leu Arg Glu Val Ile Glu
260 265 270

Gln Phe Tyr Ile Asp Gly Leu Cys Gly Gly Glu Leu Pro Leu Thr Gly
275 280 285

Met Arg Thr Ala Ser Ser Ile Ala Phe Asn Val Leu Lys Lys Met Thr
290 295 300

Phe Gly Glu Glu Lys Thr Glu Glu Gly Tyr Ile Asn Tyr Phe Glu Leu
305 310 315 320

Asp Lys Lys Thr Lys Phe Glu Phe Leu Arg Lys Arg Ile Glu Met Glu
325 330 335

ES 2 804 843 T3

Ser Tyr Thr Gly Gly Tyr Thr His Ala Asn His Lys Ala Val Gly Lys
340 345 350

Thr Ile Asn Lys Ile Gly Cys Ser Leu Asp Ile Asn Ser Ser Tyr Pro
355 360 365

Ser Gln Met Ala Tyr Lys Val Phe Pro Tyr Gly Lys Pro Val Arg Lys
370 375 380

Thr Trp Gly Arg Lys Pro Lys Thr Glu Lys Asn Glu Val Tyr Leu Ile
385 390 395 400

Glu Val Gly Phe Asp Phe Val Glu Pro Lys His Glu Glu Tyr Ala Leu
405 410 415

Asp Ile Phe Lys Ile Gly Ala Val Asn Ser Lys Ala Leu Ser Pro Ile
420 425 430

Thr Gly Ala Val Ser Gly Gln Glu Tyr Phe Cys Thr Asn Ile Lys Asp
435 440 445

Gly Lys Ala Ile Pro Val Tyr Lys Glu Leu Lys Asp Thr Lys Leu Thr
450 455 460

Thr Asn Tyr Asn Val Val Leu Thr Ser Val Glu Tyr Glu Phe Trp Ile
465 470 475 480

Lys His Phe Asn Phe Gly Val Phe Lys Lys Asp Glu Tyr Asp Cys Phe
485 490 495

Glu Val Asp Asn Leu Glu Phe Thr Gly Leu Lys Ile Gly Ser Ile Leu
500 505 510

Tyr Tyr Lys Ala Glu Lys Gly Lys Phe Lys Pro Tyr Val Asp His Phe
515 520 525

Thr Lys Met Lys Val Glu Asn Lys Lys Leu Gly Asn Lys Pro Leu Thr
530 535 540

Asn Gln Ala Lys Leu Ile Leu Asn Gly Ala Tyr Gly Lys Phe Gly Thr
545 550 555 560

Lys Gln Asn Lys Glu Glu Lys Asp Leu Ile Met Asp Lys Asn Gly Leu
565 570 575

Leu Thr Phe Thr Gly Ser Val Thr Glu Tyr Glu Gly Lys Glu Phe Tyr
580 585 590

ES 2 804 843 T3

Arg Pro Tyr Ala Ser Phe Val Thr Ala Tyr Gly Arg Leu Gln Leu Trp
595 600 605

Asn Ala Ile Ile Tyr Ala Val Gly Val Glu Asn Phe Leu Tyr Cys Asp
610 615 620

Thr Asp Ser Ile Tyr Cys Asn Arg Glu Val Asn Ser Leu Ile Glu Asp
625 630 635 640

Met Asn Ala Ile Gly Glu Thr Ile Asp Lys Thr Ile Leu Gly Lys Trp
645 650 655

Asp Val Glu His Val Phe Asp Lys Phe Lys Val Leu Gly Gln Lys Lys
660 665 670

Tyr Met Tyr His Asp Cys Lys Glu Asp Lys Thr Asp Leu Lys Cys Cys
675 680 685

Gly Leu Pro Ser Asp Ala Arg Lys Ile Ile Ile Gly Gln Gly Phe Asp
690 695 700

Glu Phe Tyr Leu Gly Lys Asn Val Glu Gly Lys Lys Gln Arg Lys Lys
705 710 715 720

Val Ile Gly Gly Cys Leu Leu Leu Asp Thr Leu Phe Thr Ile Lys Lys
725 730 735

Ile Met Phe

<210> 3

<211> 2220

5 <212> ADN

<213> Secuencia artificial

<220>

<221> fuente

10 <223> /nota="Descripción de secuencia artificial: polinucleótido sintético"

<400> 3

atgcatcacc atcatcatca ccaccacagc ggcggttccg acaaacacac gcagtagctc 60

aaagagcata gcttcaatta tgacgagtat aagaaagcga atttcgacaa gatcgagtgc 120

ctgatctttg acaccgagag ctgcacgaat tatgagaacg ataataccgg tgcacgtggt 180

tacggttggg gtcttggcgt caccgcgaac cacaatatga tctacggcca aaatctgaat 240

cagttttggg aagtatgccca gaacattttc aatgattggt atcacgacaa caaacatacc 300

ES 2 804 843 T3

attaagatta ccaagaccaa gaaaggcttc ccgaaacgta agtacattaa gtttccgatt 360
 gcagttcaca atttgggctg ggatggtgaa ttctctgaagt atagcctggt ggagaatggt 420
 ttcaattacg acaagggctc gctgaaaact gtttttagca aggggtgcgcc gtaccaaacc 480
 gtgaccgatg ttgaggaacc gaaaacgctc catatcgtcc agaataacaa catcgtttat 540
 ggttgtaacg tgtatatgga caaattcttt gaggtcgaga acaagacgg ctctaccacc 600
 gagattggcc tgtgcttggg tttcttcgat agctataaga tcatcacgtg tgctgagagc 660
 cagttccaca attacgttca tgatgtggat ccaatgttct acaaaatggg tgaagagtat 720
 gattacgata cttggcgtag cccgacgcac aagcagacca ccctggagct gcgctaccaa 780
 tacaatgata tctatatgct gcgtgaagtc atogaacagt tttacattga cggtttatgt 840
 ggcggcgagc tgccgctgac cggcatgcgc accgcttcca gcattgcgtt caacgtgctg 900
 aaaaagatga cctttggtga ggaaaagacg gaagagggct acatcaacta ttttgaattg 960
 gacaagaaaa ccaaattcga gtttctgcgt aagcgcattg aatggaatc gtacaccggt 1020
 ggctatacgc acgcaaatca caaagccgtt ggtaagacta ttaacaagat cggttgctct 1080
 ttggacatta acagctcata cccttcgcag atggcgtaca aggtctttcc gtatggcaaa 1140
 ccggttcgta agacctgggg tcgtaaacca aagaccgaga agaacgaagt ttatctgatt 1200
 gaagttggct ttgacttcgt ggagccgaaa cacgaagaat acgcgctgga tatctttaag 1260
 attggtgagg tgaactctaa agcgtgagc ccgatcaccg gcgctgtcag cggtaagag 1320
 tatttctgta cgaacattaa agacggcaaa gcaatcccgg tttacaaaga actgaaggac 1380
 accaaattga ccactaacta caatgtcgtg ctgaccagcg tggagtacga gttctggatc 1440
 aaacacttca attttggtgt gttaagaaa gacgagtacg actgtttcga agttgacaat 1500
 ctggagttta cgggtctgaa gattggttcc attctgtact acaaggcaga gaaaggcaag 1560
 tttaaacctt acgtggatca cttcacgaaa atgaaagtgg agaacaagaa actgggtaat 1620
 aagccgctga cgaatcaggc aaagctgatt ctgaacggtg cgtacggcaa attcggcacc 1680
 aaacaaaaca aagaagagaa agatttgatc atggataaga acggtttgct gaccttcacg 1740
 ggtagcgtca cggaaatcga gggtaaagaa ttctatcgtc cgtatgcgag ctctggtact 1800
 gcctatggtc gcctgcaact gtggaacgcg attatctacg cggttggtgt ggagaatfff 1860
 ctgtactgcg acaccgacag catctattgt aaccgtgaag ttaacagcct cattgaggat 1920
 atgaacgcca ttggtgaaac catcgataaa acgattctgg gtaaatggga cgtggagcat 1980
 gtctttgata agtttaaggt cctgggccag aagaagtaca tgtatcatga ttgcaaagaa 2040
 gataaaacgg acctgaagtg ttgcggtctg ccgagcgtg cccgtaagat tatcattggt 2100
 caaggtttcg acgagtttta tctgggcaaa aatgtcgaag gtaagaagca acgcaaaaaa 2160
 gtgatcggcg gttgcctgct gctggacacc ctgtttacga tcaagaaaat catgttctaa 2220

<210> 4
 <211> 15
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: péptido sintético"
 10
 <220>
 <221> MISC_FEATURE
 <222> (1)..(15)
 <223> /nota="Esta secuencia puede englobar de 1 a 3 unidades repetitivas 'Gly Gly Gly Gly Ser', en las que algunas posiciones pueden estar ausentes"
 15
 <400> 4

 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15
 20
 <210> 5
 <211> 5
 <212> PRT
 <213> Secuencia artificial
 25
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: péptido sintético"
 <400> 5
 30
Arg Ser Lys Leu Gly
1 5

 <210> 6
 <211> 131
 <212> ADN
 <213> Secuencia artificial
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: polinucleótido sintético"
 <400> 6

 tatgatgata gtagtagtgc attcaggaac agatctccag tggaacagat ctccagttga 60
 acagatctcc ttactggaga tctgttcaac tggagatctg ttccactgga gatctgttcc 120
 tgaatcaaaa a 131
 45
 <210> 7
 <211> 84
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"
 55
 <400> 7

ES 2 804 843 T3

tatgatgata gtagtagtgcg attcaggaac agatctccag tggaacagat ctccagttga 60
acagatctcc ttactggaga tctg 84

5 <210> 8
<211> 69
<212> ADN
<213> Secuencia artificial

10 <220>
<221> fuente
<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

<400> 8

tatgatatgt agtaagtcga tttaggcaga tctgtgaaac agactccagt tgaacgatct 60
tgagattcg 69

15 <210> 9
<211> 30
<212> ADN
<213> Secuencia artificial

20 <220>
<221> fuente
<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

25 <400> 9

tttttttttt tttttttttt tttttttttt 30

REIVINDICACIONES

1. Una ADN polimerasa modificada que tiene una actividad ADN polimerasa que comprende una secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia con respecto a la secuencia de aminoácidos como se expone en SEQ ID NO: 2, secuencia de aminoácidos que tiene una sustitución de acuerdo con S366 A/L y que opcionalmente incluye una o más sustituciones aminoacídicas seleccionadas del grupo que consiste en H223, N224, Y225, H227, I295, Y342, T343, I357, S360, L361, I363, S365Q, Y367, P368, D417, E475, Y476, F478, K518, H527, T529, M531, N535, G539, P542, N545, Q546, A547, L549, I550, N552, G553, F558, A596, G603, A610, V615, Y622, C623, D624, I628, Y629, R632, N635, M641, A643, I644, T647, I648, T651, I652, K655, W656, D657, V658, H660, F662, L690 y combinaciones de las mismas.
2. La ADN polimerasa modificada de la reivindicación 1, en la que dicha polimerasa tiene la capacidad de incorporar un nucleótido polifosfato que tiene 4, 5, 6, 7 u 8 fosfatos en una hebra de ADN creciente.
3. La ADN polimerasa modificada de la reivindicación 1 seleccionada de
 S366L+G553S+F558P;
 S366A+N535L+A547M;
 S366A+P542E+N545K;
 S366A+P542E+I652Q;
 S366A+N535L+T529M;
 S366A+N535L+I652Q;
 S366A+N535L+N545K;
 S366A+N535I+I652Q;
 T651Y+S366A+A547F;
 A547F+A610T+S366A;
 S366A+T647G+A547F;
 T529M+S366A+A547F;
 T647E+S366A+A547F;
 N545K+S366A+A547F;
 S366A+N535L+I652Q+T529M;
 S366A+S365A+P368G+G603T;
 S366A+N535L+I652Q+A547Y;
 S366A+N535L+A547M+T647G;
 T529M+S366A+A547F+N545K;
 T529M+S366A+A547F+N545R;
 T529M+S366A+A547F+N552L;
 T529M+S366A+A547F+Y629W;
 S366A+N535L+I652Q+A547Y+K655G;
 T529M+S366A+A547F+N545L+Y629W;
 T529M+S366A+A547F+N545L+Y225L;
 T529M+S366A+A547F+N545L+Y225F;
 T529M+S366A+A547F+N545L+K655F;
 T529M+S366A+A547F+N545L+N552L;
 T529M+S366A+A547F+N545R+M531A;
 T529M+S366A+A547F+N545R+G539Y;
 T529M+S366A+A547F+N545R+V658L;
 T529M+S366A+A547F+N545L+Y225L+D657R;
 T529M+S366A+A547F+N545L+Y225L+N552L;
 T529M+S366A+A547F+N545L+Y225L+I652G;
 T529M+S366A+A547F+N545L+Y225L+I652Q; y
 T529M+S366A+A547F+N545L+Y225L+N552M.
4. La ADN polimerasa modificada de la reivindicación 1 que comprende S366A+T529M+N545L+A547F.
5. La ADN polimerasa modificada de la reivindicación 4 que comprende además al menos una mutación seleccionada de
 a. Y225L/F/A;
 b. M531A;
 c. G539Y;
 d. N552L;
 e. Y629W; y
 f. K655F.
6. La ADN polimerasa modificada de la reivindicación 1 que comprende S365A+P368G+G603T.

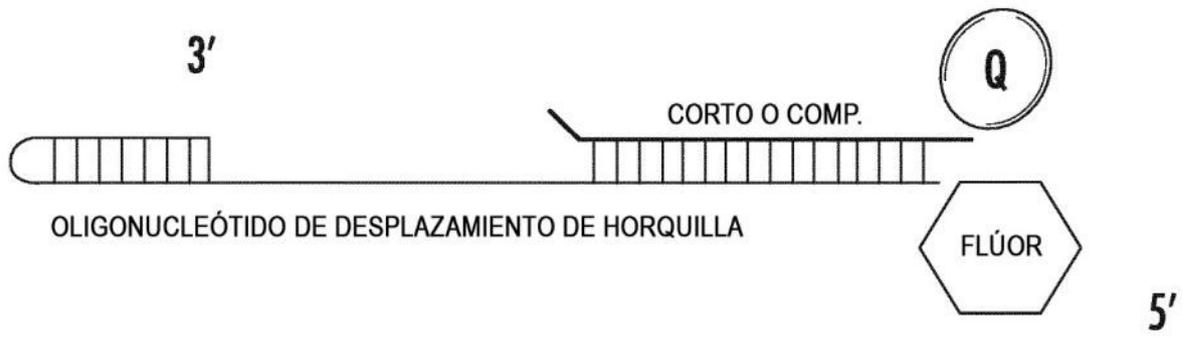


FIG. 1

Ensayo de kquím

LA REACCIÓN SE PUEDE EJECUTAR USANDO FLUJO INTERRUPTIDO O LECTOR DE PLACAS

JERINGUILLA A:

- POLIMERASA
- CEBADOR/MOLDE MARCADO CON FLUORESCÉINA
- KCl 75 mM, HEPES pH 7,5

JERINGUILLA B:

- MgCl
- dAnP-Cy3 (MOLÉCULA DE EXTINCIÓN)
- KCl 75 mM, pH 7,5

CUANDO SE MEZCLA O INYECTA:

- POLIMERASA FORMA COMPLEJO TERNARIO, Cy3 EXTINGUE FLUORESCÉINA
- POLIMERASA INCORPORA NUCLEÓTIDO QUE LIBERA PRODUCTO DE ESCISIÓN, POLIFOSFATO CON MOLÉCULA DE EXTINCIÓN UNIDA (nP-Cy3) QUE LIBERA SIMULTÁNEAMENTE EXTINGCIÓN

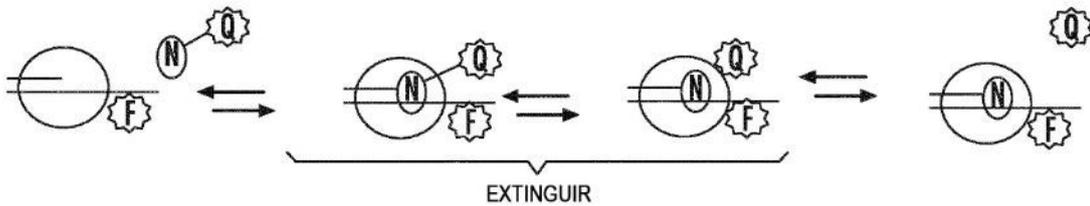
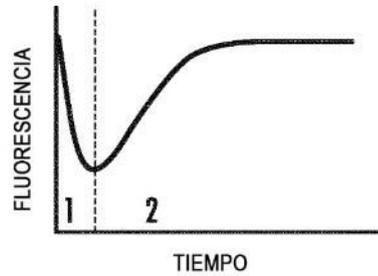


FIG. 2

Ensayo de kdis

LA REACCIÓN SE PUEDE EJECUTAR USANDO FLUJO INTERRUPTIDO O LECTOR DE PLACAS

PREINCUBAR ~10 min

JERINGUILLA A:

- POLIMERASA
- CEBADOR/MOLDE MARCADO CON FLUORESCÉINA
- KCl 75 mM, HEPES 25 pH 7,5
- CaCl₂
- dNnP-Cy3 (MOLÉCULA DE EXTINCIÓN)

JERINGUILLA B:

- CaCl₂
- dNTP (EN EXCESO)
- KCl 75 mM, HEPES 25 mM pH 7,5

PREINCUBACIÓN:

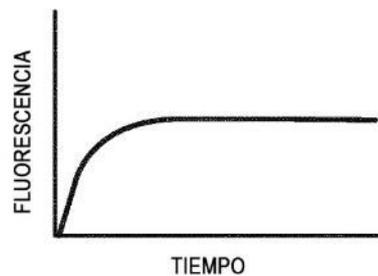
- AÑADIR dNnP-Cy3, POLIMERASA FORMA COMPLEJO TERNARIO (EXTINGUIDA). NO PUEDE INCORPORAR

CUANDO SE MEZCLA:

- CAZA DE dATP: dNTP COMPITE CON dNnP-Cy3 PARA SU ASOCIACIÓN (INCR. DE FLUORESCENCIA)



FIG. 3



ENSAYO DE KDIS BASADO EN POLARIZACIÓN DE FLUORESCENCIA

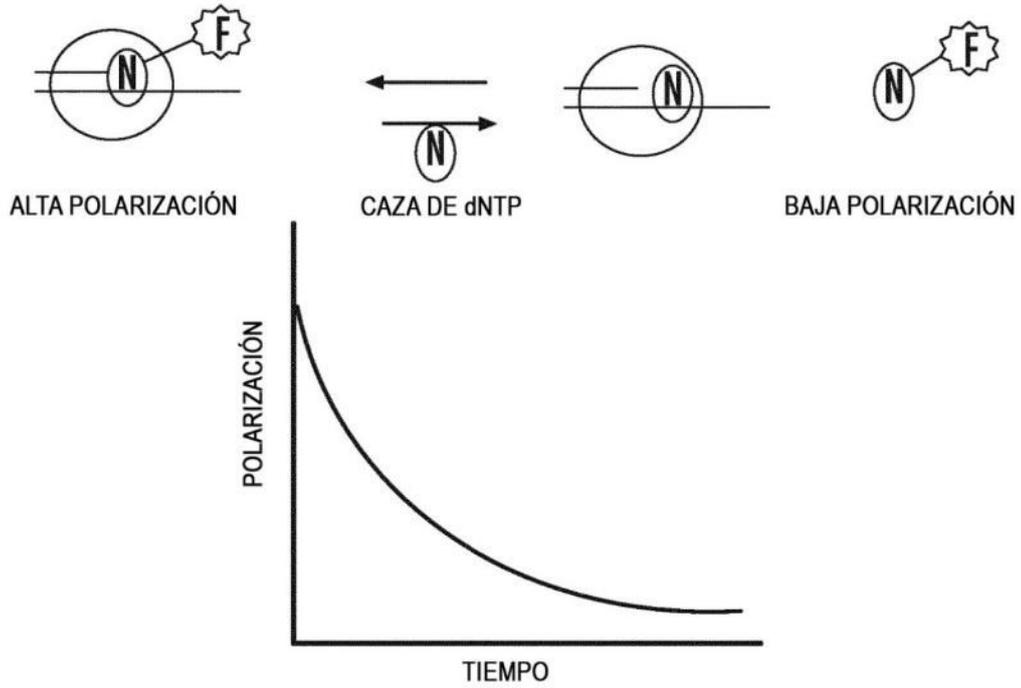


FIG. 4

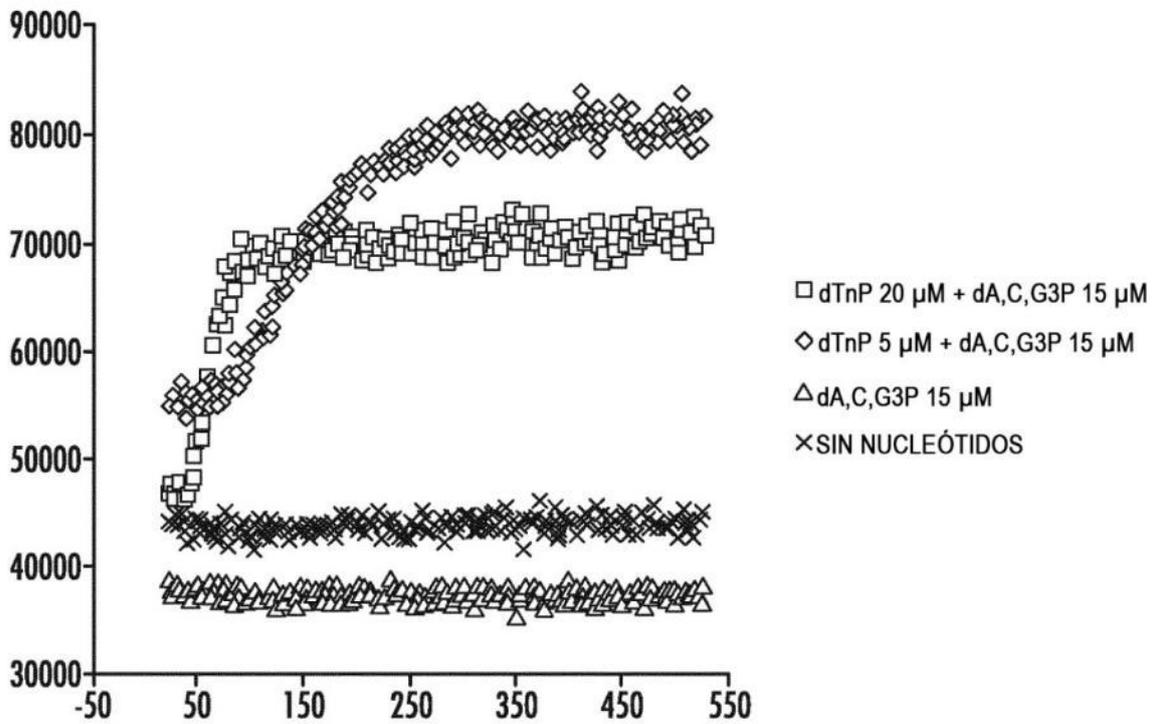


FIG. 5

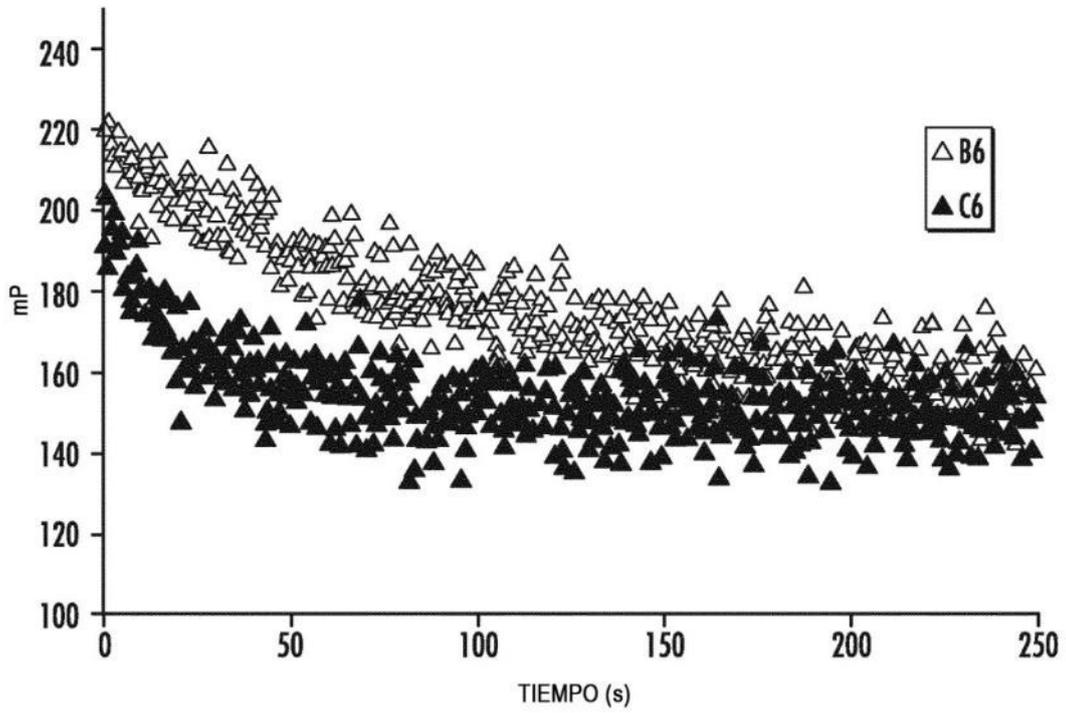


FIG. 6

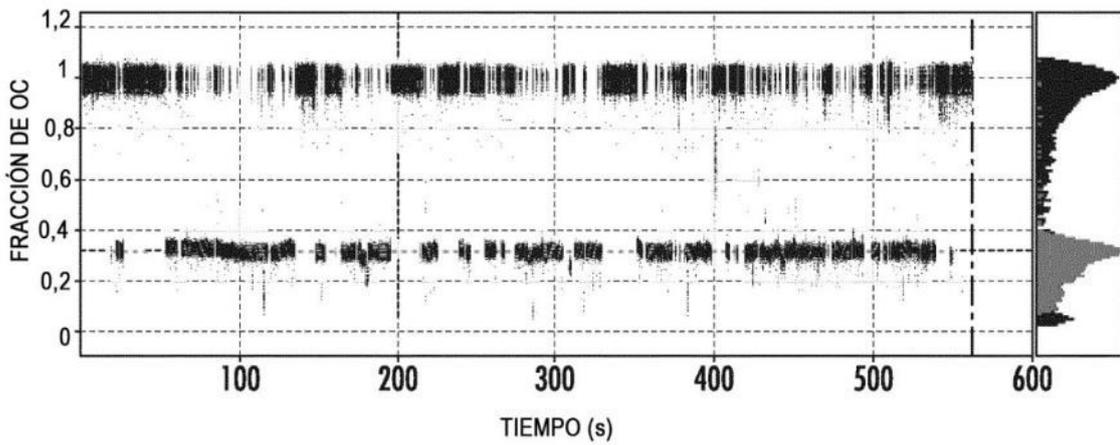


FIG. 7

MEDIANA DE LA CORRIENTE FRENTE A TIEMPO DE PERMANENCIA (P_2015-02-13-15-25-58_KB-SERVER2_dc-ANALYSIS_v1.5.0)

N_CELLS=8192.N_CLEAN_SC=734.N_QUALITY_SEQUENCING=252. VALOR DE CORTE DE PERMANENCIA = 10 ms

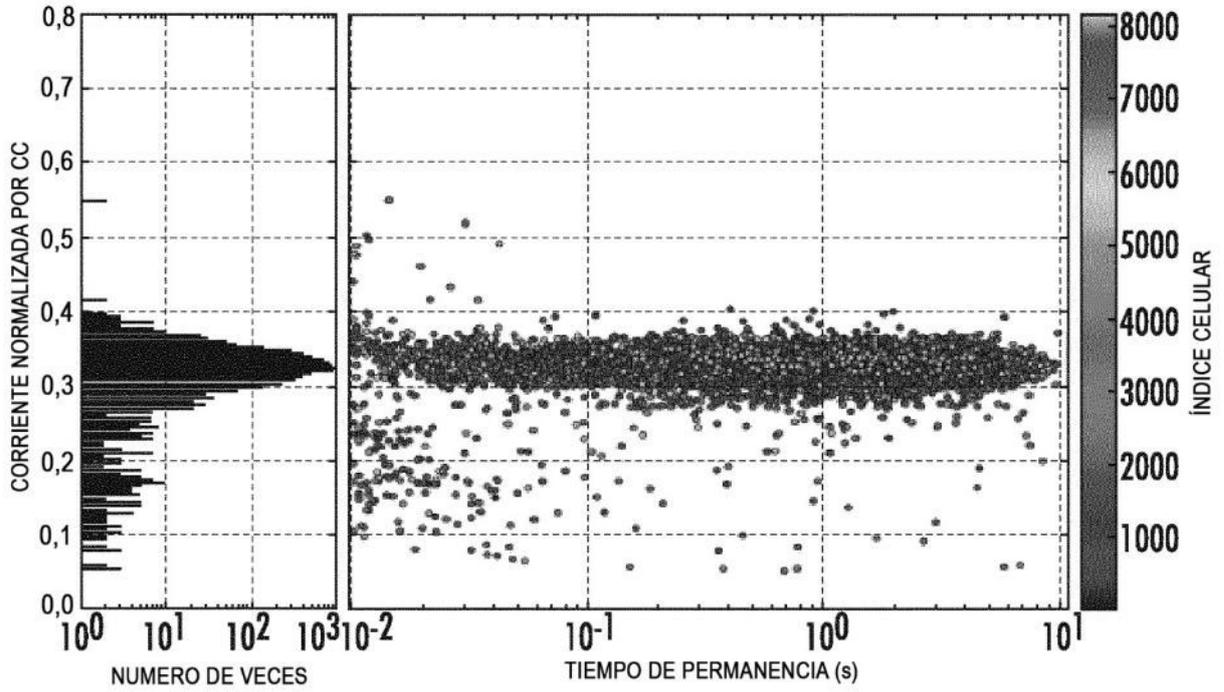


FIG. 8

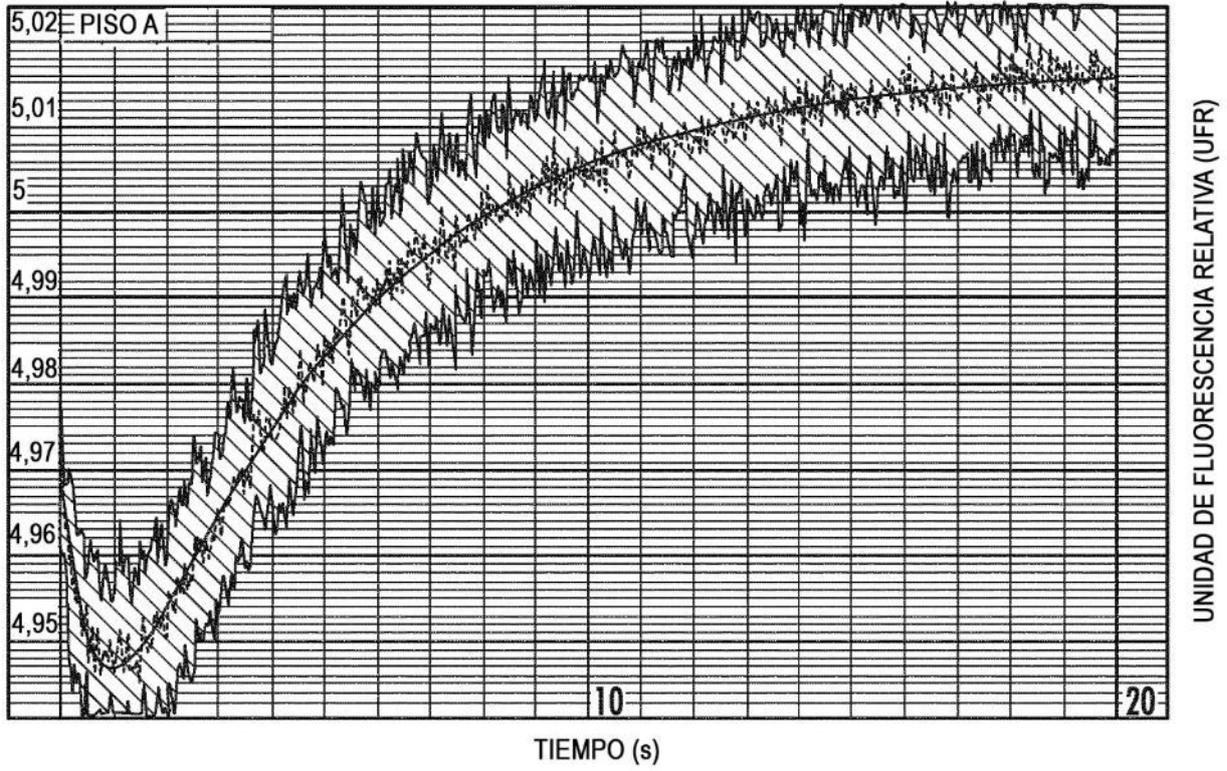


FIG. 9

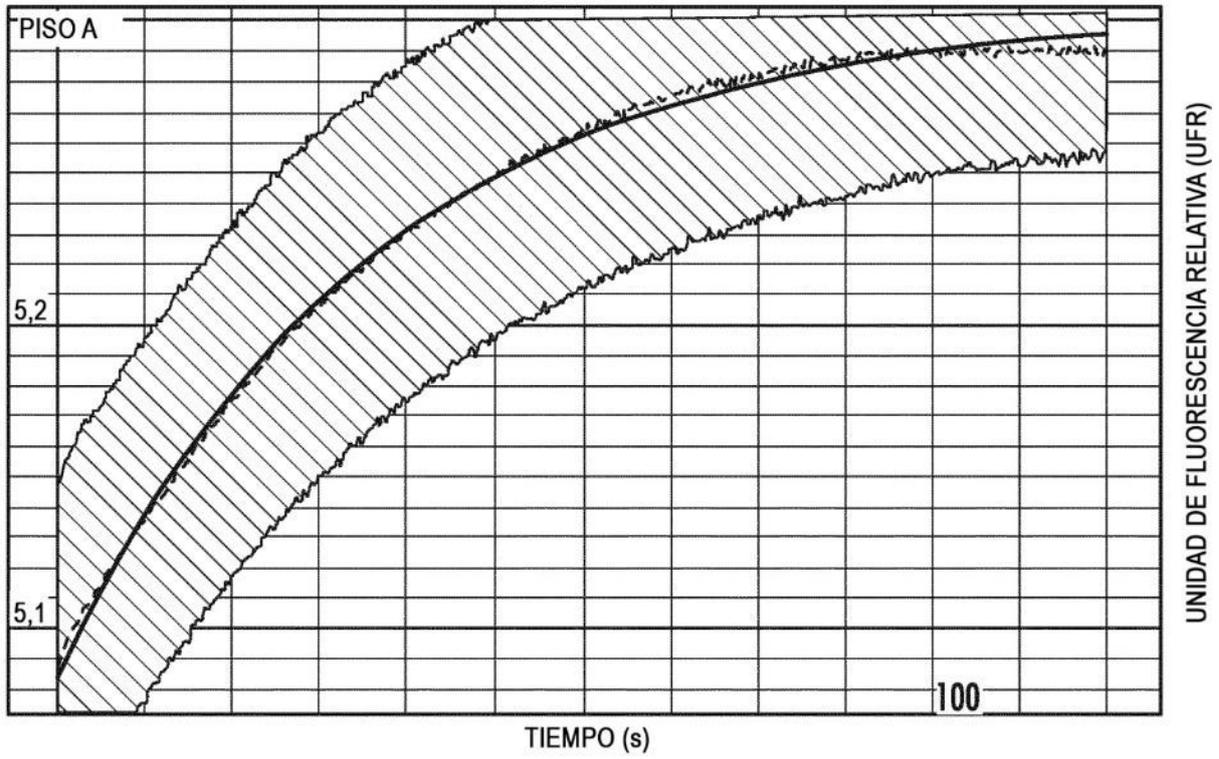


FIG. 10

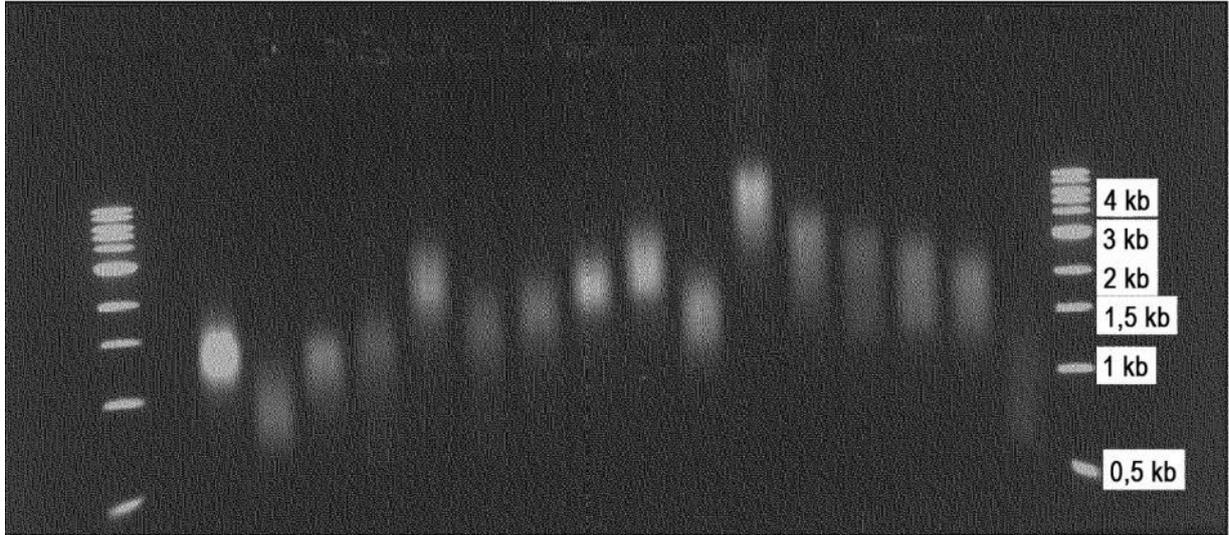


FIG. 11

