

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 403 591**

51 Int. Cl.:

C12N 9/12

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.07.2008 E 08784745 (5)**

97 Fecha y número de publicación de la concesión europea: **06.03.2013 EP 2164955**

54 Título: **ADN polimerasas mutantes y métodos relacionados**

30 Prioridad:

13.07.2007 US 949732 P

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

20.05.2013

73 Titular/es:

**F. HOFFMANN-LA ROCHE AG (100.0%)
Grenzacherstrasse 124
4070 Basel, CH**

72 Inventor/es:

**BAUER, KEITH, A. y
GELFAND, DAVID HARROW**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 403 591 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

ADN polimerasas mutantes y métodos relacionados

5 **CAMPO DE LA INVENCION**

La presente invención se refiere al campo de las ADN polimerasas y su utilización en diversas aplicaciones, incluyendo la extensión y amplificación de cebadores de ácidos nucleicos.

10 **ANTECEDENTES DE LA INVENCION**

Las ADN polimerasas son responsables de la replicación y mantenimiento del genoma, un papel que es crucial para transmitir de manera fidedigna la información genética de generación en generación. Las ADN polimerasas funcionan en las células como enzimas responsables de la síntesis del ADN. Polimerizan desoxirribonucleósidos trifosfato en presencia de un activador metálico, tal como Mg^{2+} , en un orden dictado por el molde de ADN o el polinucleótido molde que se copia. In vivo, las ADN polimerasas participan en un abanico de procesos sintéticos del ADN, incluyendo la replicación del ADN, la reparación, recombinación y amplificación génica del ADN. Durante cada proceso sintético del ADN, se copia el molde de ADN una vez, o como máximo unas cuantas veces, para producir réplicas idénticas. En contraste, in vitro, la replicación del ADN puede repetirse muchas veces, tal como, por ejemplo, durante la reacción en cadena de la polimerasa (ver, por ejemplo, la patente US nº 4.683.202, de Mullis).

En los estudios iniciales con la reacción en cadena de la polimerasa (PCR), se añadió la ADN polimerasa al inicio de cada ronda de replicación del ADN (ver la patente US nº 4.683.202, supra). Posteriormente, se determinó que las ADN polimerasas termoestables podían obtenerse de bacterias que viven a temperaturas elevadas, y que estos enzimas necesitan añadirse sólo una vez (ver la patente US nº 4.889.818, de Gelfand, y la patente US nº 4.965.188, de Mullis). A las temperaturas elevadas que se utilizan durante la PCR, estos enzimas no resultan inactivados irreversiblemente. En consecuencia, pueden llevarse a cabo ciclos repetitivos de reacciones en cadena de la polimerasa sin añadir enzimas nuevos al inicio de cada procedimiento de adición sintética. Las ADN polimerasas, particularmente las polimerasas termoestables, son la clave de un gran número de técnicas en los estudios de ADN recombinante y en el diagnóstico médico de las enfermedades. Para las aplicaciones diagnósticas en particular, una secuencia de ácidos nucleicos diana puede ser sólo una parte pequeña del ADN o ARN en cuestión, de manera que puede resultar difícil detectar la presencia de una secuencia diana de ácidos nucleicos sin la amplificación.

El patrón global de plegamiento de las polimerasas es similar al de la mano derecha humana y contiene tres subdominios claros: palma, dedos y pulgar (ver Beese et al., Science 260:352-355, 1993; Patel et al., Biochemistry 34: 5351-5363, 1995). Aunque la estructura de los subdominios de los dedos y del pulgar varían mucho entre polimerasas que difieren de tamaño y funciones celulares, todos los subdominios catalíticos de palma son superponibles. Por ejemplo, el motivo A, que interactúa con el flujo de entrada de dNTP y estabiliza el estado de transición durante la catálisis química, es superponible, con una desviación promedio de aproximadamente un Å entre las familias de pol α de mamífero y pol I procariótica de ADN polimerasas (Wang et al., Cell 89:1087-1099, 1997). El motivo A se inicia estructuralmente en una cadena β antiparalela que contiene residuos predominantemente hidrofóbicos y continúa en una hélice α . La secuencia de aminoácidos primaria de los sitios activos de la ADN polimerasa se encuentra excepcionalmente conservada. En el caso del motivo A, por ejemplo, la secuencia DYSQIELR (SEC ID nº 30) se encuentra conservada en las polimerasas de organismos separados por muchos millones de años de evolución, incluyendo, por ejemplo, Thermus aquaticus, Chlamydia trachomatis y Escherichia coli. Conjuntamente, estas observaciones indican que las polimerasas funcionan mediante mecanismos catalíticos similares.

Además de encontrarse bien conservada, también se ha demostrado que el sitio activo de las ADN polimerasas es relativamente mutable, siendo capaz de incluir determinadas sustituciones de aminoácidos sin que se reduzca la actividad de ADN polimerasa significativamente (ver, por ejemplo, la patente US nº 6.602.695, de Patel et al.). Las ADN polimerasas termoestables que presentan una tasa de extensión elevada (transcripción inversa) derivada de una cepa de arqueobacteria hipertermofílica o que comprenden secuencias de motivo críticas definidas son conocidas (por ejemplo a partir de las patentes EP nº 0 745 675 y nº 1 152 062). Además, se han descrito diversas ADN polimerasas mutantes definidas que comprenden una renovación catalítica mejorada o una actividad enzimática mejorada en, por ejemplo, los documentos nº WO2005/045015 y nº WO01/051621, respectivamente. Dichas ADN polimerasas mutantes pueden proporcionar diversas ventajas selectivas en, por ejemplo, aplicaciones diagnósticas y de investigación que comprenden reacciones de síntesis de ácidos nucleicos. La presente invención, tal como se describe en la presente memoria, cumple dichas y otras necesidades.

60 **BREVE DESCRIPCIÓN RESUMIDA DE LA INVENCION**

La presente invención proporciona ADN polimerasas que presentan una actividad enzimática mejorada respecto a la polimerasa no modificada correspondiente y que resulta útil en una diversidad de aplicaciones de síntesis de ácidos

nucleicos. En algunas realizaciones, se aíslan o se purifican las polimerasas. La ADN polimerasa de la invención comprende la secuencia de aminoácidos A-G-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-Q-X₈- X₉- X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄X₁₅, en la que:

- 5 X₁ es E,
 X₂ es P,
 X₃ es N,
 X₄ es I,
 X₅ es N,
 10 X₆ es P,
 X₇ es K,
 X₈ es V,
 X₉ es S,
 X₁₀ es R,
 15 X₁₁ es I,
 X₁₂ es F,
 X₁₃ es G,
 X₁₄ es K, y
 X₁₅ es L;

20 en la que la polimerasa presenta una tasa de extensión de ácidos nucleicos mejorada respecto a una ADN polimerasa de otro modo idéntica en la que X₁₃ es E. La polimerasa mutante presenta G en la posición X₁₃ (SEC ID nº 33).

25 Las ADN polimerasas de la invención se refieren a ADN polimerasas que son versiones modificadas de una polimerasa no modificada. En la forma no modificada, la polimerasa generalmente es funcional, presentando actividad de incorporación de nucleótidos, e incluye una secuencia de aminoácidos que presenta el motivo siguiente en el dominio polimerasa:

30 A-G-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-Q-X₈-X₉-X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄-X₁₅ (SEC ID nº 29); en la que X₂, X₅, X₆, X₉ y X₁₀ son cualquier aminoácido; X₁ es H, E o Q; X₃ es N o H; X₄ es L o I; X₇ es D, K o T; X₈ es L o V; X₁₁ es V, I o L; X₁₂ es F o Y; X₁₃ es D o E; X₁₄ es K o E; X₁₅ es L o Q.

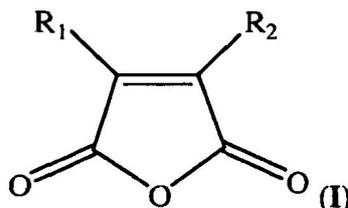
35 La polimerasa mutante (es decir, modificada a partir de SEC ID nº 29) se caracteriza adicionalmente porque incluye una sustitución de aminoácido respecto a la forma no modificada, por lo menos en la posición X₁₃, y porque presenta una tasa de extensión de ácidos nucleicos mejorada respecto a la forma no modificada. La polimerasa mutante presenta un aminoácido diferente de D o E en la posición X₁₃, en particular la polimerasa mutante presenta G en la posición X₁₃.

40 Diversas ADN polimerasas se prestan a mutación según la presente invención. Resultan particularmente adecuadas las polimerasas termoestables, incluyendo las polimerasas termoestables de tipo salvaje o naturales de diversas especies de bacterias termofílicas, así como polimerasas termoestables derivadas de dichos enzimas de tipo salvaje o naturales mediante sustitución, inserción o delección de aminoácidos u otra modificación. Entre las formas no modificadas ejemplares de polimerasa se incluyen, por ejemplo, la ADN polimerasa CS5 ó CS6, o una ADN
 45 polimerasa funcional que presenta una identidad de secuencia de por lo menos 80%, 85%, 90% ó 95% respecto a la misma. Entre otras polimerasas no modificadas se incluyen, por ejemplo, las ADN polimerasas de cualquiera de las especies siguientes de bacterias termofílicas (o una ADN polimerasa funcional que presenta una identidad de secuencia de por lo menos 90% respecto a dicha polimerasa): *Thermotoga maritima*, *Thermus aquaticus*, *Thermus thermophilus*, *Thermus flavus*, *Thermus filiformis*, *Thermus sp. sps17*, *Thermus sp. Z05*, *Thermotoga neopolitana*,
 50 *Thermosipho africanus*, *Thermus caldophilus* o *Bacillus caldotenax*. Entre las polimerasas adecuadas también se incluyen aquéllas que presentan actividad de transcriptasa inversa (RT) y/o la capacidad de incorporar nucleótidos no convencionales, tales como ribonucleótidos u otros nucleótidos modificados en la posición 2'.

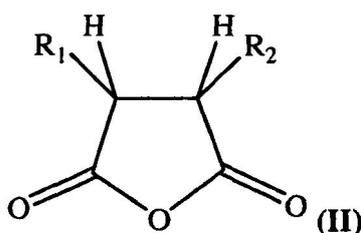
55 En algunas realizaciones, la forma no modificada de la polimerasa comprende una polimerasa quimérica. Según la invención, la forma no modificada de la polimerasa quimérica en particular es la ADN polimerasa CS5 (SEC ID nº 20), la ADN polimerasa CS6 (SEC ID nº 21) o una polimerasa que presenta una identidad de secuencia de por lo menos 90% respecto a la ADN polimerasa CS5 ó la ADN polimerasa CS6. En variaciones específicas, la forma no modificada de la polimerasa quimérica incluye una o más sustituciones de aminoácidos respecto a SEC ID nº 20 ó SEC ID nº 21 que se seleccionan de entre G46E, L329A y E678G. Por ejemplo, la forma no modificada de la
 60 polimerasa mutante puede ser G46E CS5; G46E L329A CS5; G46E E678G CS5 ó G46E L329A E678G CS5. Estas formas no modificadas se sustituyen para proporcionar una polimerasa mutante con una sustitución E558G. Por ejemplo, la ADN polimerasa mutante puede ser cualquiera de las siguientes: G46E E558G CS5; G46E L329A E558G CS5; G46E E558G E678G CS5; G46E L329A E558G E678G CS5 ó similar. En algunas realizaciones, la forma no modificada de la polimerasa quimérica incluye una o más sustituciones de aminoácidos respecto a SEC ID

n° 20 ó SEC ID n° 21 que se seleccionan de entre S671F, D640G, Q601R y I669F. Por ejemplo, la forma no modificada de la polimerasa mutante puede ser S671F CS5; D640G CS5; Q601R CS5; I669F CS5; S671F D640G CS5; S671F Q601R CS5; S671F I669F CS5; D640G Q601R CS5; D640G I669F CS5; Q601 R I669F CS5; S671F D640G Q601R CS5; S671F D640G I669F CS5; S671F Q601R I669F CS5; D640G Q601R I669F CS5 ó S671F D640G Q601R I669F CS5. Estas formas no modificadas se sustituyen para proporcionar una polimerasa mutante con una sustitución E558G. Por ejemplo, la ADN polimerasa mutante puede ser cualquiera de las siguientes: E558G S671F CS5; E558G D640G CS5; E558G Q601R CS5; E558G I669F CS5; E558G S671F D640G CS5; E558G S671F Q601R CS5; E558G S671F I669F CS5; E558G D640G Q601R CS5; E558G D640G I669F CS5; E558G Q601R I669F CS5; E558G S671F D640G Q601R CS5; E558G S671F D640G I669F CS5; E558G S671F Q601R I669F CS5; E558G D640G Q601R I669F CS5; E558G S671F D640G Q601R I669F CS5 ó similar. En algunas realizaciones, la forma no modificada de la polimerasa química incluye una o más sustituciones de aminoácidos respecto a SEC ID n° 20 ó SEC ID n° 21 que se seleccionan de entre G46E, L329A y E678G, e incluye además una o más sustituciones de aminoácidos respecto a SEC ID n° 20 ó SEC ID n° 21 que se seleccionan de entre S671F, D640G, Q601R e I669F. Por ejemplo, la forma no modificada de la polimerasa mutante puede ser G46E L329A S671F E678G CS5 ó similar. Según la invención, estas formas no modificadas se sustituyen para proporcionar una polimerasa mutante con una sustitución E558G. Por ejemplo, una polimerasa mutante según la invención puede ser E558G G46E L329A S671F E678G CS5 ó similar.

La actividad enzimática de la ADN polimerasa puede mejorarse adicionalmente con otras modificaciones no por sustitución. Una de dichas modificaciones es una modificación covalente térmicamente reversible que inactiva el enzima pero que se revierte al activar el enzima tras la incubación a una temperatura elevada, tal como una temperatura utilizada típicamente para la extensión de cebadores. En una realización, la ADN polimerasa que comprende la modificación covalente térmicamente reversible se produce mediante una reacción, llevada a cabo a pH alcalino a una temperatura que es inferior a aproximadamente 25°C, de una mezcla de una ADN polimerasa termoestable y un anhídrido de ácido dicarboxílico que presenta una de las fórmulas I ó II siguientes:



en la que R₁ y R₂ son hidrógeno o radicales orgánicos, que pueden encontrarse unidos, o:



en la que R₁ y R₂ son radicales orgánicos, que pueden encontrarse unidos, y los hidrógenos se encuentran en cis. En una variación específica de dicho enzima, la forma no modificada de la polimerasa es G64E CS5.

En otros aspectos diversos, la presente invención proporciona un ácido nucleico recombinante que codifica una ADN polimerasa según las reivindicaciones, un vector que comprende el ácido nucleico recombinante y una célula huésped transformada con el vector. En determinadas realizaciones, el vector es un vector de expresión. Las células huésped que comprenden dichos vectores de expresión resultan útiles en métodos de la invención para producir la polimerasa mediante cultivo de las células huésped bajo condiciones adecuadas para la expresión del ácido nucleico recombinante.

En todavía otro aspecto, se proporciona un método para llevar a cabo la extensión de cebadores. El método incluye generalmente poner en contacto una ADN polimerasa de la invención con un cebador, un polinucleótido molde y nucleótidos libres bajo condiciones adecuadas para la extensión del cebador, produciendo de esta manera un cebador extendido. El polinucleótido molde puede ser, por ejemplo, un molde de ARN o ADN. Los nucleótidos libres pueden incluir nucleótidos no convencionales, tales como, por ejemplo, ribonucleótidos y/o nucleótidos marcados. Además, el cebador y/o el molde pueden incluir uno o más análogos de nucleótidos. En algunas variaciones, el método de extensión de cebadores es un método para la amplificación de polinucleótidos que incluye poner en

contacto una ADN polimerasa de la invención con una pareja de cebadores, el polinucleótido molde y los nucleótidos libres bajo condiciones adecuadas para la amplificación del polinucleótido.

La presente invención proporciona además un kit que resulta útil en dicho método de extensión de cebadores. Generalmente, el kit incluye por lo menos un recipiente que proporciona una ADN polimerasa de la invención según se describe en las reivindicaciones. El kit según la invención incluye además uno o más recipientes adicionales que proporcionan uno o más reactivos adicionales. Según la invención, el recipiente o recipientes adicionales proporcionan nucleótidos libres, un tampón adecuado para la extensión de cebadores y un cebador hibridable, bajo condiciones de extensión de cebadores, con un polinucleótido molde predeterminado.

DEFINICIONES

A menos que se indique lo contrario, todos los términos técnicos y científicos utilizados en la presente memoria presentan los mismos significados entendidos comúnmente por el experto ordinario en la materia a la que se refiere la presente invención. Aunque pueden utilizarse esencialmente cualesquiera métodos y materiales similares a los descritos en la presente memoria en la práctica o ensayo de la presente invención, se describen únicamente métodos y materiales ejemplares. Para los fines de la presente invención, se define a continuación los términos siguientes.

Los términos "un", "una" y "el" o "la" incluyen los referentes plurales, a menos que el contexto indique claramente lo contrario.

Un "aminoácido" se refiere a cualquier unidad monomérica que puede incorporarse en un péptido, polipéptido o proteína. Tal como se utiliza en la presente memoria, el término "aminoácido" incluye los veinte α -aminoácidos naturales o genéticamente codificados siguientes: alanina (Ala o A), arginina (arg o R), asparagina (Asn o N), ácido aspártico (Asp o D), cisteína (Cys o C), glutamina (Gln o Q), ácido glutámico (Glu o E), glicina (Gly o G), histidina (His o H), isoleucina (Ile o I), leucina (Leu o L), lisina (Lys o K), metionina (Met o M), fenilalanina (Phe o F), prolina (Pro o P), serina (Ser o S), treonina (Thr o T), triptófano (Trp o W), tirosina (Tyr o Y) y valina (Val o V). Las estructuras de estos veinte aminoácidos naturales se muestran en, por ejemplo, Stryer et al., *Biochemistry*, 5a edición, Freeman and Company, 2002. Algunos aminoácidos adicionales, tales como la selenocisteína y la pirrolisina, también se pueden encontrarse genéticamente codificados (Stadtman, "Selenocysteine", *Annu. Rev. Biochem.* 65:83-100, 1998, e Ibbá et al., "Genetic code: introducing pyrrolysine", *Curr. Biol.* 12(13):R464-R466, 2002). El término "aminoácido" también incluye los aminoácidos no naturales, los aminoácidos modificados (por ejemplo los que presentan cadenas laterales y/o esqueletos modificados) y los análogos de aminoácidos. Ver, por ejemplo, Zhang et al., "Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells," *Proc. Natl. Acad. Sci. U.S.A.* 101(24):8882-8887, 2004; Anderson et al., "An expanded genetic code with a functional quadruplet codon" *Proc. Natl. Acad. Sci. U.S.A.* 101(20):7566-7571, 2004; Ikeda et al., "Synthesis of a novel histidine analogue and its efficient incorporation into a protein in vivo," *Protein Eng. Des. Sel.* 16 (9):699-706, 2003; Chin et al., "An Expanded Eukaryotic Genetic Code", *Science* 301(5635):964-967, 2003; James et al., "Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues," *Protein Eng. Des. Sel.* 14(12):983-991, 2001; Kohrer et al., "Import of amber and ochre suppressor tRNAs into mammalian cells: A general approach to site-specific insertion of amino acid analogues into proteins," *Proc. Natl. Acad. Sci. U.S.A.* 98(25):14310-14315, 2001; Bacher et al., "Selection and Characterization of *Escherichia coli* Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue", *J. Bacteriol.* 183(18):5414-5425, 2001; Hamano-Takaku et al., "A Mutant *Escherichia coli* Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine," *J. Biol. Chem.* 275(51):40324-40328, 2000, y Budisa et al., "Proteins with $\{\beta\}$ -(thienopyrrolyl) alanines as alternative chromophores and pharmaceutically active amino acids", *Protein Sci.* 10(7):1281-1292, 2001.

A título ilustrativo adicional, un aminoácido es típicamente un ácido orgánico que incluye un grupo amino sustituido o no sustituido, un grupo carboxi sustituido o no sustituido, y uno o más grupos o cadenas laterales, o análogos de cualquiera de dichos grupos. Entre las cadenas laterales ejemplares se incluyen, por ejemplo, tiol, seleno, sulfonilo, alquilo, arilo, acilo, ceto, azido, hidroxilo, hidrazina, ciano, halo, hidrazida, alqueno, alquino, éter, borato, boronato, fosfo, fosfono, fosfina, heterocíclico, enona, imina, aldehído, éster, tioácido, hidroxilamina o cualquier combinación de dichos grupos. Entre otros aminoácidos representativos se incluyen, aunque sin limitarse a ellos, aminoácidos que comprenden reticulantes fotoactivables, aminoácidos ligantes de metales, aminoácidos con marcador espín, aminoácidos fluorescentes, aminoácidos que contienen metal, aminoácidos con grupos funcionales nuevos, aminoácidos que interactúan covalente o no covalentemente con otras moléculas, aminoácidos fotoenjaulados y/o fotoisomerizables, aminoácidos radioactivos, aminoácidos que comprenden biotina o un análogo de biotina, aminoácidos glucosilados, aminoácidos modificados con otros carbohidratos, aminoácidos que comprenden polietilenglicol o poliéter, aminoácidos sustituidos con átomos pesados, aminoácidos escindibles químicamente y/o fotoescindibles, aminoácidos que contienen azúcar unido mediante carbono, aminoácidos redox-activos, aminoácidos que contienen amino-tioácido y aminoácidos que comprenden una o más fracciones tóxicas.

El término "mutante" en el contexto de las ADN polimerasas de la presente invención se refiere a un polipéptido,

típicamente recombinante, que comprende una o más sustituciones de aminoácidos respecto a una ADN polimerasa funcional correspondiente.

5 La expresión "forma no modificada", en el contexto de una polimerasa mutante, es una expresión utilizada en la presente memoria para definir una ADN polimerasa mutante de la presente invención: la expresión "forma no modificada" se refiere a una ADN polimerasa funcional que presenta la secuencia de aminoácidos de la polimerasa mutante, excepto en una o más posiciones de aminoácidos, especificadas como caracterizadoras de la polimerasa mutante. De esta manera, la referencia a una ADN polimerasa mutante en términos de: (a) su forma no modificada, y
 10 (b) una o más sustituciones de aminoácidos especificadas, se refiere a que, con la excepción de la sustitución o sustituciones de aminoácidos especificadas, la polimerasa mutante presenta una secuencia de aminoácidos idéntica a la forma no modificada en el motivo especificado. La polimerasa puede contener mutaciones adicionales que proporcionen la funcionalidad deseada, por ejemplo una incorporación mejorada de dideoxiribonucleótidos, ribonucleótidos, análogos de ribonucleótidos, nucleótidos marcados con pigmento; modulación de la actividad de 5'-nucleasa, modulación de la actividad de 3'-nucleasa (o correctora de errores), o similar. De acuerdo con lo anteriormente expuesto, al llevar a cabo la presente invención tal como se ha descrito en la presente memoria, la forma no modificada de una ADN polimerasa se encuentra predeterminada. La forma no modificada de una ADN polimerasa puede ser, por ejemplo, una ADN polimerasa de tipo salvaje y/o natural, o una ADN polimerasa que ya ha sido modificada intencionadamente. Una forma no modificada de la polimerasa preferentemente es una ADN polimerasa termoestable, tal como ADN polimerasas de diversas bacterias termofílicas, así como variantes
 20 funcionales de las mismas que presentan una identidad de secuencia sustancial respecto a una polimerasa termoestable de tipo salvaje o natural. Entre dichas variantes pueden incluirse, por ejemplo, ADN polimerasas quiméricas, tales como, por ejemplo, las ADN polimerasas quiméricas descritas en las patentes US nº 6.228.628 y nº 7.148.049. En determinadas realizaciones, la forma no modificada de polimerasa presenta actividad de transcriptasa inversa (RT).

25 La expresión "polimerasa termoestable" se refiere a un enzima polimerasa que es estable frente al calor, es decir, resistente al calor, y que conserva suficiente actividad para llevar a cabo posteriormente reacciones de extensión de cebador y no se desnaturaliza irreversiblemente al ser sometido a temperaturas elevadas durante el tiempo necesario para llevar a cabo la desnaturalización de los ácidos nucleicos de doble cadena. Las condiciones de calentamiento necesarias para la desnaturalización de los ácidos nucleicos son bien conocidas de la técnica y se ejemplifican en, por ejemplo, las patentes US nº 4.683.202, nº 4.683.195 y nº 4.965.188. Tal como se utiliza en la presente memoria, una polimerasa termoestable resulta adecuada para la utilización en una reacción de ciclado de la temperatura, tal como la reacción en cadena de la polimerasa ("PCR"). La desnaturalización irreversible para los fines de la presente memoria se refiere a la pérdida permanente y completa de actividad enzimática. Para una
 30 polimerasa termoestable, la actividad enzimática se refiere a la catálisis de la combinación de los nucleótidos de la manera adecuada para formar productos de extensión de cebador que sean complementarios a una cadena de ácido nucleico molde. Entre las ADN polimerasas termoestables de bacterias termofílicas se incluyen, por ejemplo, las ADN polimerasas de *Thermotoga maritima*, *Thermus aquaticus*, *Thermus thermophilus*, *Thermus flavus*, *Thermus filiformis*, especie de *Thermus sps17*, especie de *Thermus Z05*, *Thermus caldophilus*, *Bacillus caldotenax*, *Thermotoga neopolitana* y *Thermosiphon africanus*

35 Tal como se utiliza en la presente memoria, una proteína "quimérica" se refiere a una proteína cuya secuencia de aminoácidos representa un producto de fusión de subsecuencias de las secuencias de aminoácidos procedentes de por lo menos dos proteínas diferentes. Una proteína quimérica típicamente no es producida mediante manipulación directa de las secuencias de aminoácidos sino que, por el contrario, se expresa a partir de un gen "quimérico" que codifica la secuencia de aminoácidos quimérica. En determinadas realizaciones, por ejemplo, una forma no modificada de una ADN polimerasa mutante de la presente invención es una proteína quimérica que consiste de una región aminoterminal (N-terminal) derivada de una ADN polimerasa de una especie de *Thermus* y una región carboxiterminal (C-terminal) derivada de ADN polimerasa Tma. La región N-terminal se refiere a una región que se
 40 extiende entre el extremo N-terminal (posición aminoácida 1) y un aminoácido interno. De manera similar, la región C-terminal se refiere a una región que se extiende entre un aminoácido terminal y el extremo C-terminal.

45 En el contexto de las ADN polimerasas mutantes, "correspondencia" con otra secuencia (por ejemplo regiones, fragmentos, posiciones de nucleótidos o de aminoácidos, o similares) se basa en la convención de numerar según el número de la posición del nucleótido o aminoácido, alineando después las secuencias de manera que maximicen el porcentaje de identidad de las secuencias. Debido a que no todas las posiciones dentro de una "región correspondiente" dada son idénticas necesariamente, las posiciones no correspondientes dentro de una región correspondiente pueden considerarse "posiciones correspondientes". Por consiguiente, tal como se utiliza en la presente memoria, la expresión "posición aminoácida correspondiente a la posición aminoácida [X]" de una ADN polimerasa especificada se refiere a una colección de posiciones equivalentes en otras ADN polimerasas reconocidas y homólogos estructurales y familias. En realizaciones típicas de la presente invención, la "correspondencia" de posiciones aminoácidas se determina con respecto a una región de la polimerasa que comprende el motivo de la secuencia SEC ID nº 1 ó de la SEC ID nº 33, tal como se comenta en mayor detalle en la presente memoria.

El término "recombinante", tal como se utiliza en la presente memoria, se refiere a una secuencia de aminoácidos o a una secuencia de nucleótidos que ha sido modificada intencionadamente mediante métodos recombinantes. La expresión "ácido nucleico recombinante" en la presente memoria se refiere a un ácido nucleico, originalmente formando in vitro, en general, mediante la manipulación de un ácido nucleico con endonucleasas, en una forma no observada normalmente en la naturaleza. De esta manera, tanto un ácido nucleico aislado de ADN polimerasa mutante en una forma lineal, como un vector de expresión formado in vitro mediante ligación de moléculas de ADN que normalmente no se unen, se consideran recombinantes para los fines de la presente invención. Se entiende que, tras generar un ácido nucleico recombinante y reintroducirlo en una célula huésped, se replicará de manera no recombinante, es decir, utilizando la maquinaria celular in vivo de la célula huésped y no mediante manipulaciones in vitro; sin embargo, dichos ácidos nucleicos, tras producirse recombinantemente, aunque posteriormente se repliquen no recombinantemente, todavía se consideran recombinantes para los fines de la invención. Una "proteína recombinante" es una proteína generada utilizando técnicas recombinantes, es decir, mediante la expresión de un ácido nucleico recombinante tal como se ha ilustrado anteriormente. Una proteína recombinante típicamente se distingue de la proteína natural en por lo menos una o más características.

Un ácido nucleico se encuentra "operablemente ligado" en el caso de que se encuentre situado en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, un promotor o un intensificador se encuentra operablemente ligado a una secuencia codificante en el caso de que afecte a la transcripción de la secuencia, o un sitio de unión ribosómica se encuentra operablemente ligado a una secuencia codificante en el caso de que se sitúe de manera que se facilite la traducción.

La expresión "célula huésped" se refiere tanto a procariotas unicelulares como a organismos eucarióticos (por ejemplo bacterias, levaduras y actinomicetos) y a células individuales de plantas de orden superior o animales cultivadas en cultivo celular.

El término "vector" se refiere a un fragmento de ADN, típicamente de doble cadena, en el que puede haberse insertado un trozo de ADN foráneo. El vector puede ser, por ejemplo, de origen plasmídico. Los vectores contienen secuencias polinucleótidas de "replicón" que facilitan la replicación autónoma del vector en una célula huésped. El ADN foráneo se define como ADN heterólogo, que es ADN que no se encuentra naturalmente en la célula huésped, que, por ejemplo, replica la molécula del vector, codifica un marcador seleccionable o cribable, o codifica un transgén. El vector se utiliza para transportar el ADN foráneo o heterólogo a una célula huésped adecuada. Una vez se encuentra dentro de la célula huésped, el vector puede replicarse independientemente o concurrentemente con el ADN cromosómico del huésped, y pueden generarse varias copias del vector y su ADN insertado. Además, el vector también puede contener los elementos necesarios para permitir la transcripción del ADN insertado, formando una molécula de ARNm, o por el contrario causar la replicación del ADN insertado en múltiples copias de ARN. Algunos vectores de expresión contienen además elementos de secuencia contiguos al ADN insertado que incrementan la vida media del ARNm expresado y/o que permiten al traducción del ARNm en una molécula de proteína. De esta manera, muchas moléculas de ARNm y de polipéptido codificado por el ADN insertado pueden sintetizarse con rapidez.

El término "nucleótido" además de referirse a los monómeros ribonucleótidos o desoxirribonucleótidos naturales, en la presente memoria se entiende que se refiere a variantes estructurales relacionadas de los mismos, incluyendo derivados y análogos que son funcionalmente equivalentes con respecto al contexto particular en el que se utiliza el nucleótido (por ejemplo la hibridación con una base complementaria), a menos que el contexto indique claramente lo contrario.

La expresión "ácido nucleico" o "polinucleótido" se refiere a un polímero que puede corresponder a un polímero ácido ribonucleico (ARN) o ácido desoxirribonucleico (ADN), o a un análogo de los mismos. Lo anterior incluye polímeros de nucleótidos, tales como ARN y ADN, así como formas sintéticas, formas modificadas (por ejemplo modificadas química o bioquímicamente) de las mismas, y polímeros mixtos (por ejemplo que incluyen subunidades tanto de ARN como de ADN). Entre las modificaciones ejemplares se incluyen la metilación, la sustitución de uno o más nucleótidos naturales por un análogo, modificaciones internucleótidas tales como enlaces sin carga (por ejemplo metilfosfonatos, fosfotriésteres, fosfoamidatos, carbamatos y similares), fracciones colgantes (por ejemplo polipéptidos), intercalantes (por ejemplo acridina, psoralén y similares), quelantes, alquilantes y enlaces modificados (por ejemplo ácidos nucleicos α -anoméricos y similares). También se encuentran incluidas moléculas sintéticas que imitan a los polinucleótidos en su capacidad de unirse a una secuencia designada, mediante enlaces de hidrógeno y otras interacciones químicas. Típicamente, los monómeros nucleótidos se unen mediante enlaces fosfodiéster, aunque las formas sintéticas de los ácidos nucleicos pueden comprender otros enlaces (por ejemplo ácidos péptidonucleicos, tal como se describe en Nielsen et al., Science 254:1497-1500, 1991). Un ácido nucleico puede ser o puede incluir, por ejemplo, un cromosoma o segmento cromosómico, un vector (por ejemplo un vector de expresión), un casete de expresión, un polímero de ADN o ARN desnudo, el producto de una reacción en cadena de polimerasa (PCR), un oligonucleótido, una sonda y un cebador. Un ácido nucleico puede ser, por ejemplo, de cadena sencilla, de doble cadena o de triple cadena, y no se encuentra limitado a ninguna longitud particular. A

menos que se indique lo contrario, una secuencia particular de ácidos nucleicos comprende o codifica opcionalmente secuencias complementarias, además de cualquier secuencia indicada explícitamente.

5 El término "oligonucleótido" se refiere a un ácido nucleico que incluye por lo menos dos unidades monoméricas de ácidos nucleicos (por ejemplo nucleótidos). Un oligonucleótido típicamente incluye entre aproximadamente seis y aproximadamente 175 unidades monoméricas de ácidos nucleicos, más típicamente entre aproximadamente ocho y aproximadamente 100 unidades monoméricas de ácidos nucleicos, y todavía más típicamente entre aproximadamente 10 y aproximadamente 50 unidades monoméricas de ácidos nucleicos (por ejemplo aproximadamente 15, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 35 ó
10 más unidades monoméricas de ácidos nucleicos). El tamaño exacto de un oligonucleótido dependerá de muchos factores, incluyendo la función o uso último del oligonucleótido. Los oligonucleótidos se preparan opcionalmente mediante cualquier método adecuado, incluyendo, aunque sin limitarse a ellos, el aislamiento de una secuencia existente o natural, la replicación o amplificación del ADN, la transcripción inversa, la clonación y digestión de restricción de secuencias apropiadas o la síntesis química directa mediante un método tal como el método del fosfotriéster de Narang et al. (Meth. Enzymol. 68:90-99, 1979); el método del fosfodiéster de Brown et al. (Meth. Enzymol. 68:109-151, 1979); el método de la dietilfosforamida de Beaucage et al. (Tetrahedron Lett. 22:1859-1862, 1981); el método del triéster de Matteucci et al. (J. Am. Chem. Soc. 103:3185-3191, 1981); los métodos de síntesis automatizada, o el método en soporte sólido de la patente US nº 4.458.066 titulado "Process for preparing polynucleotides", publicada el 3 de julio de 1984, de Caruthers et al., u otros métodos conocidos por el experto en la materia.

25 El término "cebador" tal como se utiliza en la presente memoria se refiere a un polinucleótido capaz de actuar como punto de inicio de la síntesis de ácidos nucleicos dirigida por un molde en el caso de que se someta a condiciones en las que se inicia la extensión de cebadores (por ejemplo bajo condiciones que comprenden la presencia de nucleósidos trifosfato necesarios (tal como dicta el molde que se copia) y una polimerasa en un tampón apropiado y a una temperatura adecuada o uno o más ciclos de temperaturas (por ejemplo tal como en una reacción en cadena de polimerasa)). A título ilustrativo adicional, los cebadores también pueden utilizarse en una diversidad de otros procesos de síntesis mediada por oligonucleótidos, incluyendo como iniciadores de la síntesis de novo de ARN y procesos relacionados con la transcripción in vitro (por ejemplo la amplificación basada en secuencias de ácidos nucleicos (NASBA), la amplificación mediada por la transcripción (TMA), etc.). Un cebador típicamente es un oligonucleótido de cadena sencilla (por ejemplo un oligodesoxirribonucleótido). La longitud apropiada de un cebador depende del uso pretendido para el cebador, aunque típicamente se encuentra comprendida entre 6 y 40 nucleótidos, más típicamente entre 15 y 35 nucleótidos. Las moléculas de cebador cortas generalmente requieren temperaturas más bajas para formar complejos híbridos suficientemente estables con el molde. Un cebador no refleja necesariamente la secuencia exacta del molde sino que debe ser suficientemente complementaria para hibridarse con un molde para que se produzca el alargamiento del cebador. En determinadas realizaciones, la expresión "pareja de cebadores" se refiere a un conjunto de cebadores que incluye un cebador sentido 5' (en ocasiones denominado "directo") que se hibrida con el complemento del extremo 5' de la secuencia de ácidos nucleicos que debe amplificarse, y un cebador antisentido 3' (en ocasiones denominado "inverso") que se hibrida
40 con el extremo 3' de la secuencia que debe amplificarse (por ejemplo en el caso de que la secuencia diana se exprese como ARN o sea un ARN). Un cebador puede marcarse, si se desea, mediante la incorporación de un marcaje detectable por medios espectroscópicos, fotoquímicos, bioquímicos, inmuoquímicos o químicos. Por ejemplo, entre los marcajes útiles se incluyen ³²P, pigmentos fluorescentes, reactivos electrodenso, enzimas (utilizados comúnmente en los ensayos de ELISA), biotina o haptenos y proteínas para los que se dispone de anticuerpos o anticuerpos monoclonales.

50 El término "convencional" o "natural" en referencia a las bases de ácidos nucleicos, a los nucleósidos trifosfato o a los nucleótidos se refiere a los que se encuentran naturalmente en el polinucleótido que se describe (es decir, para el ADN estos son dATP, dGTP, dCTP y dTTP). Además, dITP y 7-deaza-dGTP se utilizan frecuentemente en lugar de dGTP y puede utilizarse 7-deaza-dATP en lugar de dATP en reacciones in vitro de síntesis de ADN, tal como la secuenciación. Colectivamente éstas pueden denominarse dNTP.

55 La expresión "no convencional" o "modificado" en referencia a una base de ácidos nucleicos, a un nucleósido o a un nucleótido incluye modificaciones, derivaciones o análogos de bases convencionales, nucleósidos o nucleótidos que se encuentran naturalmente en un polinucleótido particular. Determinados nucleótidos no convencionales se encuentran modificados en la posición 2' del azúcar ribosa en comparación con los dNTP convencionales. De esta manera, aunque para ARN los nucleótidos naturales son ribonucleótidos (es decir, ATP, GTP, CTP, UTP, colectivamente, rNTP), debido a que estos nucleótidos presentan un grupo hidroxilo en la posición 2' del azúcar, que, por el contrario, se encuentra ausente en los dNTP, tal como se utiliza en la presente memoria los ribonucleótidos son nucleótidos no convencionales que son sustratos de las ADN polimerasas. Tal como se utiliza en la presente memoria, entre los nucleótidos no convencionales se incluyen, aunque sin limitarse a ellos, los compuestos utilizados como terminadores de la secuenciación de los ácidos nucleicos. Entre los compuestos terminadores ejemplares se incluyen, aunque sin limitarse a ellos, aquellos compuestos que presentan una estructura 2',3'-dideoxi y se denominan dideoxinucleósidos trifosfato. Los dideoxinucleósidos trifosfato ddATP,
60

ddTTP, ddCTP y ddGTP se denominan colectivamente ddNTP. Entre los ejemplos adicionales de compuestos terminadores se incluyen los análogos 2'-PO₄ de ribonucleótidos (ver, por ejemplo, las solicitudes publicadas de patente US n° 2005/0037991 y n° 2005/0037398). Entre otros nucleótidos no convencionales se incluyen los fosforotioato dNTP ([[α]-S]dNTP), los 5'-[α]-borano-dNTP, los [α]-metilfosfonato dNTP y los ribonucleósidos trifosfato (rNTP). Las bases no convencionales pueden marcarse con isótopos radioactivos tales como ³²P, ³³P ó ³⁵S; marcajes fluorescentes; marcajes quimioluminiscentes; marcajes bioluminiscentes; marcajes hapteno tales como biotina; o marcajes enzimáticos tales como estreptavidina o avidina. Entre los marcajes fluorescentes pueden incluirse pigmentos que se encuentran cargados negativamente, tales como pigmentos de la familia de la fluoresceína, o pigmentos que presentan una carga neutra, tales como pigmentos de la familia de la rodamina, o pigmentos que presentan una carga positiva, tales como los pigmentos de la familia de la cianina. Entre los pigmentos de la familia de la fluoresceína se incluyen, por ejemplo, FAM, HEX, TET, JOE, NAN y ZOE. Entre los pigmentos de la familia de la rodamina se incluyen Rojo Texas, ROX, R110, R6G y TAMRA. Diversos pigmentos o nucleótidos marcados con FAM, HEX, TET, JOE, NAN, ZOE, ROX, R110, R6G, Rojo Texas TAMRA son comercializados por Perkin-Elmer (Boston, MA), Applied Biosystems (Foster City, CA) o Invitrogen/Molecular Probes (Eugene, OR). Entre los pigmentos de la familia de la cianina se incluyen Cy2, Cy3, Cy5 y Cy7, y son comercializados por GE Healthcare UK Limited (Amersham Place, Little Chalfont, Buckinghamshire, Inglaterra).

Tal como se utiliza en la presente memoria, el "porcentaje de identidad de secuencias" se determina mediante comparación de dos secuencias alineadas óptimamente en una ventana de comparación, en el que la parte de la secuencia en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula mediante la determinación del número de posiciones en las que se encuentra una base de ácido nucleico o residuo aminoácido idéntico en ambas secuencias, rindiendo el número de posiciones correspondientes, dividiendo el número de posiciones correspondientes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencias.

Los términos "idéntico" o porcentaje de "identidad", en el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son iguales o que presentan un porcentaje especificado de nucleótidos o de residuos aminoácidos que son iguales (por ejemplo una identidad de 60%, opcionalmente de 65%, 70%, 75%, 80%, 85%, 90% ó 95% en una región especificada), al comparar y alinear para la máxima correspondencia en una ventana de comparación, o a una región designada que se ha medido utilizando uno de los algoritmos de comparación de secuencias siguientes o mediante alineación manual e inspección visual. Las secuencias son "sustancialmente idénticas" si son idénticas por lo menos al 20%, por lo menos al 25%, por lo menos al 30%, por lo menos al 35%, por lo menos al 40%, por lo menos al 45%, por lo menos al 50% o por lo menos al 55%. Estas definiciones se refieren también al complemento de una secuencia de ensayo. Opcionalmente, existe identidad en una región que presenta una longitud de por lo menos aproximadamente 50 nucleótidos, o más típicamente en una región que presenta una longitud de 100 a 500 ó 1.000 ó más nucleótidos.

Los términos "similitud" o "porcentaje de similitud", en el contexto de dos o más secuencias polipeptídicas, se refieren a dos o más secuencias o subsecuencias que presentan un porcentaje especificado de residuos aminoácidos que son iguales o similares según definen unas sustituciones conservadoras de aminoácidas (por ejemplo una identidad de 60%, opcionalmente de 65%, 70%, 75%, 80%, 85%, 90% ó 95% en una región especificada), al comparar y alinear para la máxima correspondencia en una ventana de comparación, o a una región designada que se ha medido utilizando uno de los algoritmos de comparación de secuencias siguientes o mediante alineación manual e inspección visual. Las secuencias son "sustancialmente similares" si son similares por lo menos al 20%, por lo menos al 25%, por lo menos al 30%, por lo menos al 35%, por lo menos al 40%, por lo menos al 45%, por lo menos al 50% o por lo menos al 55%. Opcionalmente, esta similitud se produce en una región que presenta una longitud de por lo menos aproximadamente 50 nucleótidos, o más típicamente en una región que presenta una longitud de 100 a 500 ó 1.000 ó más aminoácidos.

Para la comparación entre secuencias, típicamente una secuencia actúa como secuencia de referencia, con la que se comparan las secuencias de ensayo. Al utilizar un algoritmo de comparación de secuencias, se introducen en un ordenador las secuencias de ensayo y de referencia, se designan las coordenadas de las subsecuencias, en caso necesario, y se fijan los parámetros del programa algorítmico de secuencias. Habitualmente se utilizan parámetros por defecto del programa, o pueden fijarse parámetros alternativos. A continuación, el algoritmo de comparación de secuencias calcula el porcentaje de identidad o similitud de secuencias para las secuencias de ensayo respecto a la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", tal como se utiliza en la presente memoria, incluye la referencia a un segmento de entre cualquiera de las posiciones contiguas seleccionadas de entre el grupo que consiste de 20 a 600, habitualmente de entre aproximadamente 50 y aproximadamente 200, más habitualmente de entre aproximadamente 100 y aproximadamente 150, en el que una secuencia puede compararse con una secuencia de referencia del mismo número de posiciones contiguas tras la alineación óptima de las dos secuencias. Los métodos

de alineación de secuencias para la comparación son bien conocidos de la técnica. La alineación óptima de secuencias para la comparación puede llevarse a cabo, por ejemplo, con el algoritmo de homología local de Smith y Waterman (Adv. Appl. Math. 2:482, 1970), mediante el algoritmo de alineación de homologías de Needleman y Wunsch (J. Mol. Biol. 48:443, 1970), mediante el método de búsqueda de similitud de Pearson y Lipman (Proc. Natl. Acad. Sci. USA 85: 2444, 1988), mediante implementaciones computerizadas de dichos algoritmos (por ejemplo GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.) o mediante alineación manual e inspección visual (ver, por ejemplo, Ausubel et al., Current Protocols in Molecular Biology (suplemento de 1995)).

Un ejemplo de un algoritmo que resulta adecuado para determinar el porcentaje de identidad de secuencias y de similitud de secuencias son los algoritmos de BLAST y de BLAST 2.0, los cuales se describen en Altschul et al. (Nuc. Acids Res. 25:3389-402, 1977), y en Altschul et al. (J Mol. Biol. 215:403-10, 1990), respectivamente. El software para llevar a cabo los análisis de BLAST se encuentra disponible públicamente en el National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica identificar en primer lugar parejas de secuencias de alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia de pregunta, que se corresponden o satisfacen alguna puntuación umbral T de valor positivo al alinearse con una palabra de la misma longitud en una secuencia de base de datos. T se denomina umbral de puntuación de palabra vecina (Altschul et al., supra). Estos aciertos de palabra vecina actúan como núcleos para iniciar búsquedas de HSP más largas que dichos núcleos. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia hasta donde se incrementa la puntuación acumulada de alineación. Las puntuaciones acumuladas se calculan utilizando, para las secuencias de nucleótidos, los parámetros M (puntuación de premio para una pareja de residuos correspondientes; en todos los casos >0) y N (puntuación de penalización para residuos no correspondientes; en todos los casos <0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuaciones para calcular la puntuación acumulada. La extensión de los aciertos de palabra en cada dirección se detiene cuando: la puntuación acumulada de alineación cae en una cantidad X respecto al valor máximo alcanzado; la puntuación acumulada cae hasta cero o menos debido a la acumulación de una o más alineaciones de residuos de puntuación negativa, o se alcanza el final de cualquiera de las dos secuencias. Los parámetros de algoritmo de BLAST, W, T y X, determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para las secuencias de nucleótidos) utiliza como valores por defecto la longitud de palabra (W) de 11, un valor esperado (E) de 10, M=5, N=-4 y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza como valores por defecto una longitud de palabra de 3, un valor esperado (E) de 10, y las alineaciones de la matriz de puntuaciones BLOSUM62 (ver Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) (B) de 50, valor esperado (E) de 10, M=5, N=-4 y una comparación de ambas cadenas.

El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre dos secuencias (ver, por ejemplo, Karlin y Altschul, Proc. Natl. Acad. Sci. USA 90:5873-87, 1993). Una medida de similitud proporcionada por el algoritmo de BLAST es la probabilidad de suma mínima (P(N)), que proporciona una indicación de la probabilidad con la que se produciría al azar una correspondencia entre dos secuencias de nucleótidos o de aminoácidos. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia en el caso de que la probabilidad de suma mínima en una comparación del ácido nucleico de ensayo y el ácido nucleico de referencia se inferior a aproximadamente 0,2, típicamente inferior a aproximadamente 0,01, y más típicamente inferior a aproximadamente 0,001.

La expresión "tasa de extensión de ácidos nucleicos" se refiere a la tasa a la que un biocatalizador (por ejemplo un enzima, tal como una polimerasa, ligasa o similar) extiende un ácido nucleico (por ejemplo un cebador u otro oligonucleótido) de una manera dependiente de molde o independiente de molde mediante la unión (por ejemplo covalente) de uno o más nucleótidos al ácido nucleico. A título ilustrativo, determinadas ADN polimerasas mutantes descritas en la presente memoria presentan tasas de extensión de ácidos nucleicos mejoradas respecto a las formas no modificadas de dichas ADN polimerasas, de manera que pueden extender cebadores a tasas más altas que dichas formas no modificadas bajo un conjunto dado de condiciones de reacción.

La expresión "eficiencia de transcripción inversa" se refiere a la fracción de moléculas de ARN que son transcritas inversamente como ADNc en una reacción de transcripción inversa dada.

55 BREVE DESCRIPCIÓN DE LOS DIBUJOS

La figura 1 ilustra una alineación de secuencias de aminoácidos de una región del dominio de polimerasa de las ADN polimerasas termoestables ejemplares de diversas especies de bacterias termófilas y del bacteriófago T7: Thermus thermophilus (Tth) (SEC ID nº 3), Thermus caldophilus (Tca) (SEC ID nº 4), Thermus sp. Z05 (Z05) (SEC ID nº 5), Thermus aquaticus (Taq) (SEC ID nº 6), Thermus flavus (Tfl) (SEC ID nº 7), Thermus filiformis (Tfi) (SEC ID nº 8), Thermus sp. sps 17 (Sps17) (SEC ID nº 9), Deinococcus radiodurans (Dra) (SEC ID nº 10), Hot Spring family B1clone 7 (HspB) (SEC ID nº11), Bacillus stearothermophilus (Bst) (SEC ID nº 12), Bacillus caldotenax (Bca) (SEC ID nº 13), Escherichia coli (Eco) (SEC ID nº 14), Thermotoga maritime (Tma) (SEC ID nº 15), Thermotoga neapolitana

(Tne) (SEC ID n° 16), Thermosipho africanus (Taf) (SEC ID n° 17), Hot Spring family A (HspA) (SEC ID n°18) y el bacteriófago T7 (T7) (SEC ID n° 19). La alineación de secuencias de aminoácidos también incluye una reacción (SEC ID n° 31 y n° 32) del dominio de polimerasa de ADN polimerasas termoestables químicas representativas, es decir, CS5 y CS6. Además, también se incluye una secuencia (Cons) (SEC ID n° 24) que muestra residuos aminoácidos de consenso entre estas secuencias ejemplares. Además, las regiones de polipéptido mostradas comprenden el motivo de aminoácidos AGXXFXXXSXXQXXXXLXXXX (SEC ID n° 1), las posiciones variables del cual se definen adicionalmente en la presente memoria. Estos motivos se encuentran subrayados en negrita para las secuencias de polimerasa CS5 y CS6. La posición aminoácida susceptible de mutación según la presente invención o mutaciones relacionadas se indica con un asterisco (*). Los huecos en las alineaciones se indican con un punto (.).

La figura 2A presenta la secuencia de aminoácidos de la ADN polimerasa termoestable química CS5 (SEC ID n° 20).

La figura 2B presenta una secuencia de ácidos nucleicos codificante de la ADN polimerasa termoestable química CS5 (SEC ID n° 22).

La figura 3A presenta la secuencia de aminoácidos de la ADN polimerasa termoestable química CS6 (SEC ID n° 21).

La figura 3B presenta una secuencia de ácidos nucleicos codificante de la ADN polimerasa termoestable química CS6 (SEC ID n° 23).

La figura 4 es un gráfico de columnas que muestra las tasas de extensión normalizadas de diversos mutantes de una ADN polimerasa CS5 G46E L329A E678G (GLE). El eje y representa las tasas de extensión relativas, mientras que el eje x representa las ADN polimerasa que presentan mutaciones puntuales especificadas (GLE = ADN polimerasa CS5 G46E L329A E678G, GLDE = ADN polimerasa CS5 G46E L329A D640G E678G, GLEE = ADN polimerasa CS5 G46E L329A E558G E678G, GLEQDSE = ADN polimerasa CS5 G46E L329A E558G Q601R D640G S671F E678G, GLQDSE = ADN polimerasa CS5 G46E L329A Q601R E678G, GLSE = ADN polimerasa CS5 G46E L329A S671F E678G). Los valores de la tasa de extensión obtenidos para las polimerasas mutantes se normalizan respecto al valor obtenido para la ADN polimerasa CS5 GLE, que se fija en 1,00.

DESCRIPCIÓN DETALLADA

La presente invención proporciona nuevas ADN polimerasas que presentan tasas mejoradas de extensión de cebador. Las ADN polimerasas de la invención pueden utilizarse a concentraciones más bajas para un rendimiento superior o equivalente al de los enzimas parentales. En vista de las actividades similares de otros mutantes identificados previamente, se prevé que las ADN polimerasas de la presente invención en determinadas realizaciones presenten incrementos concomitantes de la actividad de transcriptasa inversa y/o la actividad de amplificación. Por lo tanto, las ADN polimerasas de la invención resultan útiles en una diversidad de aplicaciones que implican la extensión de los cebadores, así como la transcripción inversa o la amplificación de moldes polinucleótidos, incluyendo, por ejemplo, aplicaciones en estudios de ADN recombinante y en el diagnóstico médico de enfermedades.

Según la invención, las ADN polimerasas de la invención comprenden el motivo de aminoácidos siguiente:

Ala-Gly-X_{aa}-X_{aa}-Phe-X_{aa}-X_{aa}-X_{aa}-Ser-X_{aa}-X_{aa}-Gln-X_{aa}-X_{aa}-X_{aa}-X_{aa}-Leu-X_{aa}-X_{aa}-X_{aa}-X_{aa} (también denominado en la presente memoria con el código de una letra como A-G-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-Q-X₈-X₉-X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄-X₁₅ (SEC ID n° 1)), en el que:

X₁ es Glu (E);
 X₂ es Pro (P);
 X₃ es Asn (N);
 X₄ es Ile (I);
 X₅ es Asn (N);
 X₆ es Pro (P);
 X₇ es Lys (K);
 X₈ es Val (V);
 X₉ es Ser (S);
 X₁₀ es Arg (R);
 X₁₁ es Ile (I);
 X₁₂ es Phe (F);
 X₁₃ es Gly (G);
 X₁₄ es Lys (K); y

X₁₅ es Leu (L),

en el que la polimerasa presenta una tasa mejorada de extensión de ácidos nucleicos en comparación con una ADN polimerasa de otro modo idéntica en la que X₁₃ es Glu (E). Según la invención, X₁₃ es G (SEC ID n° 33).

En otras realizaciones de SEC ID n° 1, X₂, X₅, X₆, X₉ y X₁₀ son cualesquiera aminoácidos presentes en posiciones correspondientes en cualquier ADN polimerasa. Entre las ADN polimerasas ejemplares se incluyen las de *Thermus thermophilus*, *Thermus caldophilus*, *Thermus sp. Z05*, *Thermus aquaticus*, *Thermus flavus*, *Thermus filiformis*, *Thermus sp. sps17*, *Deinococcus radiodurans*, Hot Spring family B/clone 7, *Bacillus stearothermophilus*, *Bacillus caldotenax*, *Escherichia coli*, *Thermotoga maritima*, *Thermotoga neapolitana*, *Thermosiphon africanus*, Hot Spring family A y el bacteriófago T7. Según la invención, X₂ es Pro (P), X₅ es Asn (N), X₆ es Pro (P), X₉ es Ser (S) y X₁₀ es Arg (R).

La invención se refiere además a ADN polimerasas que comprenden el motivo de aminoácidos siguiente:

X_{aa}-X_{aa}-X_{aa}-X_{aa}-Phe-X_{aa}-X_{aa}-X_{aa}-Ser-X_{aa}-X_{aa}-X_{aa}-X_{aa}-X_{aa}-X_{aa}-X_{aa}-Leu-X_{aa}-X_{aa}-X_{aa}-X_{aa} (también denominado en la presente memoria con el código de una letra como T₁-T₂-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-T₃-X₈-X₉-X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄-X₁₅ (SEC ID n° 2));

en el que:

T₁ es Ala (A) o Val (V),

T₂ es Gly (G) o Val (V),

T₃ es Gln (L) o Hes (H),

X₂, X₅, X₆, X₉, y X₁₀ son cualquier aminoácido o ningún aminoácido,

X₁ es Hes (H), Glu (E), Gln (Q) o ningún aminoácido,

X₃ es Asn (N) o Hes (H),

X₄ es Leu (L), Ile (I) o Pro (P),

X₇ es Asp (D), Lys (K) o Thr (T),

X₈ es Leu (L), Val (V) o Ile (I),

X₁₁ es Val (V), Ile (I), Leu (L) o Lys (K),

X₁₂ es Phe (F), Tyr (Y) o Gln (Q),

X₁₃ es un aminoácido diferente de Asp (D) o Glu (E),

X₁₄ es Lys (K), Glu (E) o Ala (A), y

X₁₅ es Leu (L), Gln (Q) o Gly (G),

en el que la polimerasa presenta una tasa mejorada de extensión de ácidos nucleicos en comparación con una ADN polimerasa de otros idéntica en la que X₁₃ es D o E. En realizaciones especialmente preferentes, X₁₃ es G. El motivo anteriormente indicado (SEC ID n° 2) se generó mediante alineación de todas las secuencias mostradas en la figura 1, mientras que el motivo mostrado en SEC ID n° 1 se generó mediante alineación de todas las secuencias mostradas en la figura 1 excepto la de los residuos aminoácidos de la ADN polimerasa de T7.

En algunas otras realizaciones de SEC ID n° 2, X₂, X₅, X₆, X₉ y X₁₀ son cualesquiera aminoácidos presentes en posiciones correspondientes en cualquier ADN polimerasa, por ejemplo las ADN polimerasas de *Thermus thermophilus*, *Thermus caldophilus*, *Thermus sp. Z05*, *Thermus aquaticus*, *Thermus flavus*, *Thermus filiformis*, *Thermus sp. sps17*, *Deinococcus radiodurans*, clon 7/ familia B de Hot Spring, *Bacillus stearothermophilus*, *Bacillus caldotenax*, *Escherichia coli*, *Thermotoga maritima*, *Thermotoga neapolitana*, *Thermosiphon africanus*, familia A de Hot Spring y bacteriófago T7. Por ejemplo, X₂ se selecciona de entre el grupo que consiste de P, A, E, T y V, o se encuentra ausente; X₅ se selecciona de entre el grupo que consiste de N, R, G y S; X₆ se selecciona de entre el grupo que consiste de R, P, S y T; X₉ se selecciona de entre el grupo que consiste de E, G, Q, S y A; y X₁₀ se selecciona de entre el grupo que consiste de R, T, A, V, Y, S, N y K.

Tal como se ha comentado anteriormente, X₂, X₅, X₆, X₉ y X₁₀ pueden ser cualquier aminoácido. Por ejemplo, X₂ se selecciona de entre el grupo que consiste de Pro (P), Ala (A), Glu (E), Thr (T) y Val (V). X₅ se selecciona de entre el grupo que consiste de Asn (N), Arg (R), Gly (G) y Ser (S). En otras realizaciones, X₆ se selecciona de entre el grupo que consiste de Arg (R), Pro (P), Ser (S) y Thr (T). X₉ se selecciona de entre el grupo que consiste de Glu (E), Gly (G), Gln (Q), Ser (S) y Ala (A). X₁₀ se selecciona de entre el grupo que consiste de Arg (R), Thr (T), Ala (A), Val (V), Tyr (Y), Ser (S) y Asn (N).

La presente invención se refiere además a formas no modificadas de ADN polimerasas susceptibles de mutación que presentan un dominio de polimerasa funcional que comprende el motivo de aminoácidos siguiente:

Ala-Gly-X_{aa}-X_{aa}-Phe-X_{aa}-X_{aa}-X_{aa}-Ser-X_{aa}-X_{aa}-Gln-X_{aa}-X_{aa}-X_{aa}-X_{aa}-Leu-X_{aa}-X_{aa}-X_{aa}-X_{aa} (también denominado en la presente memoria con el código de una letra como A-G-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-Q-X₈-X₉-X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄-

X₁₅ (SEC ID nº 29));

en el que:

- 5 X₂, X₅, X₆, X₉ y X₁₀ son cualquier aminoácido;
 X₁ es Hes (H), Glu (E) o Gln (Q),
 X₃ es Asn (N) o Hes (H),
 X₄ es Leu (L) o Ile (I),
 X₇ es Asp (D), Lys (K) o Thr (T),
 10 X₈ es Leu (L) o Val (V),
 X₁₁ es Val (V), Ile (I) o Leu (L),
 X₁₂ es Phe (F) o Tyr (Y),
 X₁₃ es es Asp (D) o Glu (E),
 X₁₄ es Lys (K) o Glu (E) y
 15 X₁₅ es Leu (L) o Gln (Q).

El motivo presentado anteriormente (SEC ID nº 29) se encuentra presente dentro del subdominio "thumb" en el sitio activo de muchas ADN polimerasas de familia A dependientes de ADN, particularmente ADN polimerasas termoestables de bacterias termófilas y del bacteriófago T7. Por ejemplo, la figura 1 muestra una alineación de secuencias de aminoácidos de una región del dominio de polimerasa de las ADN polimerasas de varias especies de bacterias: *Thermus thermophilus*, *Thermus caldophilus*, *Thermus sp. Z05*, *Thermus aquaticus*, *Thermus flavus*, *Thermus filiformis*, *Thermus sp. sps17*, *Deinococcus radiodurans*, clon 7 familia B Hot Spring, *Bacillus stearothermophilus*, *Bacillus caldotenax*, *Escherichia coli*, *Thermotoga maritima*, *Thermotoga neapolitana*, *Thermosiphon africanus*, Familia A Hot Spring y el bacteriófago T7. La alineación de secuencias de aminoácidos mostrada en la figura 1 también incluye una región del dominio de polimerasa de ADN polimerasas termoestables químicas representativas. Tal como se muestra, el motivo de SEC ID nº 29 se encuentra presente en cada una de dichas polimerasas, lo que indica una función conservada para esta región del sitio activo.

De acuerdo con lo anteriormente expuesto, en algunas realizaciones la forma no modificada de la ADN polimerasa es una ADN polimerasa de tipo salvaje o una ADN polimerasa natural, tal como, por ejemplo, una polimerasa de cualquiera de las especies de bacterias indicadas anteriormente. En una variación, la polimerasa no modificada es de una especie del género *Thermus*. En otras realizaciones de la invención, la polimerasa no modificada es de una especie termófila diferente de *Thermus*. Se encuentran disponibles las secuencias completas de ácidos nucleicos y de aminoácidos de numerosas ADN polimerasas termoestables. Las secuencias de cada una de las polimerasas de *Thermus aquaticus* (Taq), *Thermus thermophilus* (Tth), *Thermus species Z05*, *Thermus species sps17*, *Thermotoga maritima* (Tma) y *Thermosiphon africanus* (Taf) han sido publicadas en la solicitud publicada de patente PCT nº WO 92/06200. La secuencia de la ADN polimerasa de *Thermus flavus* ha sido publicada en Akhmetzjanov y Vakhitov (Nucleic Acids Research 20:5839, 1992). La secuencia de la ADN polimerasa termoestable de *Thermus caldophilus* se encuentra en EMBL/GenBank nº de acceso U62584. La secuencia de la ADN polimerasa termoestable de *Thermus filiformis* puede recuperarse del depósito de la ATCC nº 42380 utilizando, por ejemplo, los métodos proporcionados en la patente US nº 4.889.818, así como la información de secuencia proporcionada en la misma. La secuencia de la ADN polimerasa de *Thermotoga neapolitana* procede del nº de acceso de la base de datos de patentes GeneSeq y del documento nº WO 97/09451. La secuencia de la ADN polimerasa termoestable de *Bacillus caldotenax* se describe en, por ejemplo, Uemori et al. (J Biochem (Tokyo) 113(3):401-410, 1993; ver también el nº de acceso de la base de datos Swiss-Prot Q04957 y los nº de acceso de GenBank D 12982 y BAA02361). La secuencia de la ADN polimerasa de *Bacillus stearothermophilus* ha sido publicada en la patente US nº 6.066.483. También se describen ejemplos de formas no modificadas de ADN polimerasa que pueden modificarse tal como se indica en la presente memoria en, por ejemplo, las patentes US nº 6.228.628, titulada "Mutant chimeric DNA polimerasa", publicada el 8 de mayo de 2001, de Gelfand et al.; nº 6.346.379, titulada "Thermostable DNA polymerases incorporating nucleoside triphosphates labeled with fluorescein family dyes", publicada el 12 de febrero de 2002, de Gelfand et al.; nº 7.030.220, titulada "Thermostable enzyme promoting the fidelity of thermostable DNA polymerases-for improvement of nucleic acid synthesis and amplification in vitro", publicada el 18 de abril de 2006, de Ankenbauer et al.; nº 6.881.559, titulada "Mutant B-type DNA polymerases exhibiting improved performance in PCR", publicada el 19 de abril de 2005, de Sobek et al.; nº 6.794.177, titulada "modified DNA-polymerase from *Carboxydotherrmus hydrogenoformans* and its use for coupled reverse transcription and polymerase chain reaction", publicada el 21 de septiembre de 2004, de Markau et al.; nº 6.468.775, titulada "Thermostable DNA polymerase from *Carboxydotherrmus hydrogenoformans*", publicada el 22 de octubre de 2002, de Ankenbauer et al.; nº 7.148.049, titulada "Thermostable or thermoactive DNA polymerase molecules with attenuated 3'-5' exonuclease activity", publicada el 12 de diciembre de 2007, de Schoenbrunner et al.; nº 7.378.262, titulada "Reversibly modified thermostable enzymes for DNA synthesis and amplification in vitro", publicada el 27 de mayo de 2008, de Sobek et al., y las solicitudes de patente US nº 2002/0012970, titulada "High temperature reverse transcription using mutant DNA polymerases", presentada el 30 de marzo de 2001, de Smith et al.; nº 2006/0078928, titulada "Thermostable enzyme promoting the fidelity of thermostable DNA polymerases-for improvement of nucleic acid synthesis and amplification in vitro", presentada el 29 de septiembre de 2005, de Ankenbauer et al.

También susceptibles a las mutaciones indicadas en la presente memoria son las ADN polimerasas funcionales que han sido modificadas previamente (por ejemplo mediante sustitución, adición o delección de aminoácidos), con la condición de que la polimerasa previamente modificada conserve el motivo de aminoácidos de SEC ID nº 1 ó SEC ID nº 33. De esta manera, las ADN polimerasas no modificadas adecuadas también incluyen las variantes funcionales de polimerasas de tipo salvaje o naturales. Dichas variantes típicamente presentan una identidad o similitud de secuencias sustancial respecto a la polimerasa de tipo salvaje o natural, típicamente una identidad de secuencia de por lo menos 80%, y más típicamente una identidad de secuencia de por lo menos 90%, 95% ó 98%. En determinadas realizaciones, la ADN polimerasa no modificada presenta actividad de transcriptasa inversa (RT) y/o la capacidad de incorporar ribonucleótidos u otros nucleótidos modificados en la posición 2'.

Entre las polimerasas adecuadas también se incluyen, por ejemplo, determinadas ADN polimerasas quiméricas que comprende regiones polipeptídicas procedentes de dos o más enzimas. Se describen ejemplos de dichas ADN polimerasas quiméricas en, por ejemplo, la patente US nº 6.228.628. Resultan particularmente adecuadas las ADN polimerasas de la familia CS quiméricas, entre las que se incluyen las polimerasas CS5 (SEC ID nº 20) y CS6 (SEC ID nº 21) y las variantes de las mismas que presentan una identidad o similitud de secuencias sustancial respecto a la SEC ID nº 20 ó a SEC ID nº 21 (típicamente una identidad de secuencia de por lo menos 80%, y más típicamente una identidad de secuencia de por lo menos 90%). Las ADN polimerasas CS5 y CS6 son enzimas quiméricas derivados de las ADN polimerasas de *Thermus sp. Z05* y de *Thermotoga maritima* (Tma). Comprenden el dominio N-terminal de 5'-nucleasa del enzima de *Thermus* y los dominios C-terminal de exonucleasa 3'-5' y de polimerasa del enzima Tma. Estos enzimas presentan una actividad eficiente de transcriptasa inversa, pueden extender los cebadores que contienen análogos de nucleótidos y pueden incorporar dNTP α -fosforotioato, dUTP, dITP y también dNTP marcados de las familias de la fluoresceína y de la cianina. Las polimerasas CS5 y CS6 también son enzimas de PCR activados por Mg^{2+} eficientes. Las secuencias de ácidos nucleicos codificantes de las polimerasas CS5 y CS6 se proporcionan en las figuras 2B y 3B, respectivamente. Las polimerasas quiméricas CS5 y CS6 se describen adicionalmente en, por ejemplo, la patente US nº 7.148.049.

En algunas realizaciones, la forma no modificada de la ADN polimerasa es una polimerasa que ha sido modificada previamente, típicamente por medios recombinantes, para proporcionar cierta ventaja selectiva. Entre dichas modificaciones se incluyen, por ejemplo, las sustituciones de aminoácidos G46E, L329A y/o E678G en la ADN polimerasa CS5, en la ADN polimerasa CS6 o una o más mutaciones correspondientes en otras polimerasas. De acuerdo con lo anteriormente expuesto, en variaciones específicas, la forma no modificada de la ADN polimerasa es una de las siguientes (presentando cada una la secuencia de aminoácidos SEC ID nº 20 ó SEC ID nº 21, excepto por la sustitución o sustituciones designadas: G46E; G46E L329A; G46E E678G ó G46E L329A E678G. La sustitución E678G, por ejemplo, permite la incorporación de ribonucleótidos y otros nucleótidos modificados en la posición 2', aunque esta mutación aparentemente también resulta en una capacidad deteriorada de extender moldes con cebadores. En determinadas realizaciones, las mutaciones según la presente invención que resultan en una tasa de extensión más rápida de la polimerasa mutante, mejoran la capacidad deteriorada de la mutación E678G de extender los moldes con cebadores.

Las ADN polimerasas mutantes de la presente invención comprenden una o más sustituciones de aminoácidos respecto a la polimerasa no modificada, es decir, en la posición X_{13} de SEC ID nº 1 ó SEC ID nº 33. La sustitución de aminoácido en esta posición, es decir G en la posición X_{13} , proporciona una actividad mejorada de incorporación de nucleótidos, rindiendo una ADN polimerasa con una tasa de extensión de ácidos nucleicos mejorada (más rápida) en comparación con la ADN polimerasa correspondiente, que de otra manera es idéntica pero incluye una E en la posición X_{13} . Aunque sin pretensión de limitarse a ninguna teoría en particular, los presentes inventores creen que la tasa mejorada de extensión de ácidos nucleicos de las polimerasas mutantes de la invención es una consecuencia de la unión más fuerte a un molde, es decir, una disociación menos frecuente del molde, resultando en un enzima de "procesividad" más elevada. Estas características permiten utilizar concentraciones más bajas de la polimerasa mutante en, por ejemplo, reacciones de extensión de cebadores en comparación con las reacciones en las que participa la ADN polimerasa no modificada. De esta manera, a una concentración de enzima suficientemente alta, la tasa de extensión de la polimerasa no modificada (es decir, que no presenta las mutaciones específicas que son el objeto de la invención) podría concebirse que se aproximase a la del enzima mutante. También se espera que las polimerasas mutantes presenten un rendimiento mucho mejor que las formas no modificadas a fuerza iónica elevada. Sin embargo, a una concentración de enzima suficientemente alta, el rendimiento de la polimerasa no modificada a fuerza iónica baja se aproximaría a la de la polimerasa mutante.

Debido a que las formas no modificadas de la ADN polimerasa son únicas, la posición aminoácida correspondiente a X_{13} típicamente es diferente para cada polimerasa mutante. Los programas de alineación de secuencias de aminoácidos y de ácidos nucleicos se encuentran fácilmente disponibles (ver, por ejemplo, los indicados supra) y, dado el motivo particular identificado en la presente memoria, sirven para ayudar en la identificación de los aminoácidos exactos (y codones correspondientes) para la modificación según la presente invención. Las posiciones correspondientes a X_{13} se muestran en la Tabla 1 para las ADN polimerasas termoestables quiméricas representativas y para las ADN polimerasas termoestables de especies termofílicas ejemplares.

Tabla 1. Posiciones aminoácidas correspondientes a las posiciones de motivo X₁₃ en las polimerasas termoestables ejemplares.

Organismo o secuencia quimérica	Posición X ₁₃
Consenso	
T. thermophilus	498
T. caldophilus	498
T. sp. Z05	498
T. aquaticus	496
T. flavus	495
T. filiformis	494
T. sp. sps17	494
D. radiodurans	586
Clon 7/familia B Hot Spring	546
B. stearothermophilus	540
B. caldotenax	540
E. coli	592
T. maritima	558
T. neapolitana	558
T. africanus	558
Familia A Hot Spring	595
CS5	558
CS6	558

5 Tal como se ha comentado anteriormente, en algunas realizaciones la ADN polimerasa mutante de la presente invención se deriva de la ADN polimerasa CS5 (SEC ID nº 20), de la ADN polimerasa CS6 (SEC ID Nº 21) o de una variante de dichas polimerasas (por ejemplo G46E, G46E L329A, G46E E678G, G46E L329A E678G o similar). Tal como se ha indicado anteriormente, en la ADN polimerasa CS5 ó en la ADN polimerasa CS6, la posición X₁₃ corresponde a ácido glutámico (E) en la posición 558. De esta manera, según la invención, la polimerasa mutante comprende una sustitución de aminoácido en la posición X₁₃ respecto a una ADN polimerasa CS5 ó una ADN polimerasa CS6 que de otra manera es idéntica. Entre los mutantes ejemplares de la ADN polimerasa CS5 y de la ADN polimerasa CS6 se incluyen aquellos que comprenden la sustitución o sustituciones de aminoácidos E558G. Entre otros mutantes ejemplares de las ADN polimerasas CS5 y CS6 se incluyen los siguientes (presentando cada uno la secuencia de aminoácidos SEC ID nº 20 ó SEC ID nº 21, excepto por las sustituciones designadas):

15 G46E E558G,
 G46E L329A E558G,
 G46E E558G E678G,
 L329A E558G E678G y
 20 G46E L329A E558G E678G.

En algunas realizaciones, las sustituciones de aminoácidos son sustituciones de aminoácidos individuales en la posición X₁₃. Alternativamente, la polimerasa mutante comprende la sustitución de aminoácido de la posición X₁₃ en combinación con sustituciones de aminoácidos en otras posiciones, particularmente aquellas sustituciones de aminoácidos que es conocido que mejoran la tasa de extensión de ácidos nucleicos de la ADN polimerasa, por ejemplo sustituciones de aminoácidos (diferentes de aquellos residuos mostrados bajo X_{a8}) en la posición X_{a8} de las ADN polimerasas que presentan un dominio de polimerasa funcional que comprende el motivo de aminoácidos siguiente:

30 Xaa-Xaa-Xaa-Xaa-Arg-Xaa-Xaa-Xaa-Lys-Leu-Xaa-Xaa-Thr-Tyr-Xaa-Asp (también denominado en la presente memoria en el código de una letra como X_{a1}-X_{a2}-X_{a3}-X_{a4}-R-X_{a6}-X_{a7}-X_{a8}-K-L-X_{a11}-X_{a12}-T-Y-X_{a15}-X_{a16} (SEC ID nº 25)), en el que:

35 X_{a1} es Ile (I) o Leu (L),
 X_{a2} es Gln (Q) o Leu (L),
 X_{a3} es Gln (Q), Hes (H) o Glu (E),
 X_{a4} es Tyr (Y), Hes (H), o Phe (F),
 X_{a6} es Glu (E), Gln (Q) o Lys (K),
 X_{a7} es Ile (I), Leu (L) o Tyr (Y),
 40 X_{a8} es Gln (Q), Thr (T), Met (M), Gly (G) o Leu (L),
 X_{a11} es Lys (K) o Gln (Q),
 X_{a12} es Ser (S) o Asn (N),
 X_{a15} es Ile (I) o Val (V) y

X_{a16} es Glu (E) o Asp (D).

En algunas realizaciones, la polimerasa mutante comprende la sustitución del aminoácido de la posición X₁₃ en combinación con las sustituciones de aminoácidos (diferentes de aquellos residuos mostrados posteriormente para X_{b8}) en la posición X_{b8} de las ADN polimerasas que presentan un dominio de polimerasa funcional que comprende el motivo de aminoácidos siguiente:

Thr-Gly-Arg-Leu-Ser-Ser-Xaa-Xaa-Pro-Asn-Leu-Gln-Asn
(al que también se hace referencia en la presente memoria con el código de una letra como:
T-G-R-L-S-S-X_{b7}-X_{b8}-P-N-L-Q-N (SEC ID n° 26));
en el que:

X_{b7} es Ser (S) o Thr (T),
X_{b8} es Asp (D), Glu (E) o Asn (N).

En algunas realizaciones, la polimerasa mutante comprende la sustitución del aminoácido de la posición X₁₃ en combinación con las sustituciones de aminoácidos (diferentes de aquellos residuos mostrados posteriormente para X_{c4} y/o X_{c6}) en la posición X_{c4} y/o X_{c6} de las ADN polimerasas que presentan un dominio de polimerasa funcional que comprende el motivo de aminoácidos siguiente:

Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Asp-Tyr-Ser-Gln-Ile-Glu-Leu-Arg (también denominado en la presente memoria en el código de una letra como X_{c1}-X_{c2}-X_{c3}-X_{c4}-X_{c5}-X_{c6}-X_{c7}-D-Y-S-Q-I-E-L-R (SEC ID n° 27), en el que:

X_{c1} es Gly (G), Asn (N) o Asp (D),
X_{c2} es Trp (W) o Hes (H); es Trp (W), Ala (A), Leu (L) o Val (V),
X_{c3} es Trp (W), Ala (A), Leu (L) o Val (V),
X_{c4} es Ile (I) o Leu (L),
X_{c5} es Val (V), Phe (F) o Leu (L),
X_{c6} es Ser (S), Ala (A), Val (V) o Gly (G) y
X_{c7} es Ala (A) o Leu (L).

En determinadas variaciones de la invención, la polimerasa mutante comprende una sustitución de aminoácido en la posición X₁₃ respecto a una ADN polimerasa CS5 ó una ADN polimerasa CS6 que, de otro modo, es idéntica. Entre los mutantes ejemplares de la ADN polimerasa CS5 y de la ADN polimerasa CS6 se incluyen aquellos que comprenden la sustitución de aminoácido E558G. Entre otros mutantes ejemplares de las ADN polimerasas CS5 y CS6 se incluyen los siguientes (presentando cada uno la secuencia de aminoácidos SEC ID n° 20 ó SEC ID n° 21, excepto por las sustituciones designadas):

E558G Q601 R,
E558G D640G,
E558G I669F,
E558G S671F,
E558G D640G S671 F,
E558G Q601R S671F,
E558G I669F S671F,
E558G Q601 R D640G,
E558G D640G I669F,
E558G Q601 R I669F,
E558G S671F D640G Q601R,
E558G S671 F D640G I669F,
E558G S671 F Q601R I669F,
E558G D640G Q601R I669F y
E558G Q601R D640G I669F S671F,

en los que la sustitución de aminoácido Q601R corresponde a una sustitución de aminoácido en la posición X_{a8}; la sustitución de aminoácido D640G corresponde a una sustitución de aminoácido en la posición X_{b8}; la sustitución de aminoácido I669F corresponde a una sustitución de aminoácido en la posición X_{c4}; la sustitución de aminoácido S671F corresponde a una sustitución de aminoácido en la posición X_{c6}.

En algunas realizaciones, la forma no modificada de la polimerasa quimérica incluye una o más sustituciones de aminoácidos respecto a SEC ID n° 20 ó SEC ID n° 21 que se seleccionan de entre G46E, L329A y E678G, e incluye además una o más sustituciones de aminoácidos respecto a SEC ID n° 20 ó SEC ID n° 21 que se seleccionan de entre S671F, D640G, Q601R e I669F. Por ejemplo, la forma no modificada de la polimerasa mutante puede ser G46E L329A S671F E678G CS5 ó similar. En realizaciones ejemplares, dichas formas no modificadas se sustituyen

para proporcionar una polimerasa mutante con una sustitución E558G. Por ejemplo, una ADN polimerasa mutante puede ser E558G G46E L329A S671F E678G CS5 ó similar.

5 De acuerdo con lo anteriormente expuesto, la mutación del motivo de SEC ID nº 1 o, más preferentemente, SEC ID nº 33, solo o en combinación con otras sustituciones de aminoácidos en otras posiciones, proporciona una tasa mejorada de extensión de ácidos nucleicos en comparación con una ADN polimerasa de otro modo idéntica. Pueden utilizarse diversos ensayos bien conocidos por el experto en la materia para medir la tasa de extensión de ácidos nucleicos. En algunas realizaciones, dichos ensayos se llevan a cabo para comparar una ADN polimerasa que comprende el motivo de SEC ID nº 1 ó de SEC ID nº 33 con otra ADN polimerasa que presenta la misma secuencia de aminoácidos en cada posición excepto por una única sustitución en la posición X₁₃. En algunas realizaciones, dichos ensayos se llevan a cabo para comparar una ADN polimerasa que comprende el motivo de SEC ID nº 1 ó de SEC ID nº 33 con otra ADN polimerasa que presenta la misma secuencia de aminoácidos en cada posición excepto por una única sustitución en la posición X₁₃ y sustituciones en otras posiciones, tal como se indica en la presente memoria.

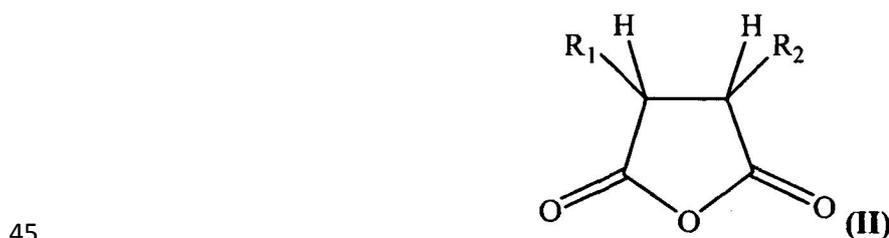
15 Además de la mutación del motivo de SEC ID nº 1 ó SEC ID nº 33 tal como se indica en la presente memoria, las ADN polimerasas mutantes también pueden incluir otra u otras modificaciones no realizadas por sustitución. Entre dichas modificaciones pueden incluirse, por ejemplo, modificaciones covalentes conocidas de la técnica para proporcionar una ventaja adicional en aplicaciones que comprenden la extensión de cebadores. Por ejemplo, en determinadas realizaciones, la ADN polimerasa mutante incluye además una modificación covalente térmicamente reversible. En estas realizaciones, se une covalentemente un grupo modificador a la proteína, resultando en una pérdida de la totalidad, o de prácticamente la totalidad, de la actividad enzimática. El grupo modificador se selecciona de manera que la modificación se invierta mediante incubación a una temperatura elevada. Las ADN polimerasas que comprenden dichas modificaciones térmicamente reversibles resultan particularmente adecuadas para las aplicaciones de inicio en caliente, tales como, por ejemplo, diversas técnicas de PCR de inicio en caliente. Los reactivos modificadores térmicamente reversibles que pueden utilizarse de acuerdo con las ADN polimerasas mutantes de la presente invención se describen en, por ejemplo, la patente US nº 5.773.258, de Birch et al. Entre las modificaciones ejemplares se incluyen, por ejemplo, el bloqueo reversible de los residuos lisina mediante modificación química del grupo ε-amino de los residuos lisina (ver Birch et al., supra). En determinadas variaciones, la modificación covalente térmicamente reversible incluye la unión covalente con el grupo ε-amino de los residuos lisina, de un anhídrido dicarboxílico tal como se describe en Birch et al., supra.

35 Por ejemplo, se producen polimerasas mutantes particularmente adecuadas que comprenden una modificación covalente térmicamente reversible mediante una reacción, llevada a cabo a pH alcalino a una temperatura que es inferior a aproximadamente 25°C, de una mezcla de un enzima termoestable y un anhídrido de ácido dicarboxílico que presenta una fórmula general seleccionada de entre el grupo que consiste de:

(a) fórmula I:



en la que R₁ y R₂ son hidrógeno o radicales orgánicos, que pueden encontrarse unidos, y (b) fórmula II:



50 en la que R₁ y R₂ son radicales orgánicos, que pueden encontrarse unidos, y los hidrógenos se encuentran en cis, esencialmente tal como se describe en Birch et al., supra. En realizaciones específicas que comprenden una modificación covalente térmicamente reversible, la forma no modificada de la polimerasa es la ADN polimerasa CS5 G64E.

5 Las ADN polimerasas mutantes de la presente invención pueden construirse mediante la mutación de las secuencias de ADN que codifican la polimerasa no modificada correspondiente (por ejemplo una polimerasa de tipo salvaje o una variante correspondiente a partir de la que se deriva la polimerasa mutante de la invención), tal como mediante la utilización de técnicas comúnmente denominadas mutagénesis sitio-dirigida. Las moléculas de ácidos nucleicos codificantes de la forma no modificada de la polimerasa pueden mutarse mediante una diversidad de técnicas de reacción en cadena de la polimerasa (PCR) bien conocidas por el experto ordinario en la materia (ver, por ejemplo, PCR Strategies (M.A. Innis, D.H. Gelfand y J.J. Sninsky, editores, 1995, Academic Press, San Diego, CA) en el capítulo 14; PCR Protocols: A Guide to Methods and Applications (M.A. Innis, D. H. Gelfand, J.J. Sninsky y T.J. White, editores, Academic Press, NY, 1990).

15 A título de ejemplo no limitativo, el sistema de dos cebadores utilizado en el kit transformante de mutagénesis sitio-dirigida de Clontech, puede utilizarse para introducir mutaciones sitio-dirigidas en un polinucleótido codificante de una forma no modificada de la polimerasa. Tras la desnaturalización del plásmido diana en este sistema, se hibridan simultáneamente dos cebadores con el plásmido; uno de estos cebadores contiene la mutación sitio-dirigida deseada, y el otro contiene una mutación en otro punto del plásmido, resultando en la eliminación de un sitio de restricción. A continuación, se lleva a cabo la síntesis de la segunda cadena, ligando íntimamente estas dos mutaciones, y los plásmidos resultantes se transforman en una cepa mutS de E. coli. Se aísla el plásmido de ADN a partir de las bacterias transformadas, se corta con el enzima de restricción relevante (linearizando de esta manera los plásmidos no mutados) y después se transforma nuevamente en E. coli. Este sistema permite la generación de mutaciones directamente en un plásmido de expresión, sin la necesidad de subclonar o de generar fagémidos de cadena sencilla. El ligamiento íntimo de las dos mutaciones y la posterior linearización de los plásmidos no mutados resulta en una eficiencia de mutación elevada y permite un cribado mínimo. Tras la síntesis del cebador de sitio de restricción inicial, este método requiere la utilización de sólo un nuevo tipo de cebador por cada sitio de mutación. En lugar de preparar cada mutante posicional por separado, puede sintetizarse un conjunto de cebadores oligonucleótidos "degenerados de diseño" con el fin de introducir todas las mutaciones deseadas en un sitio dado simultáneamente. Los transformantes pueden cribarse mediante secuenciación del ADN plasmídico a través de la región mutagenizada con el fin de identificar y clasificar los clones mutantes. A continuación, puede cortarse cada ADN mutante y analizarse mediante electroforesis, tal como, por ejemplo, en un gel de incremento de la detección de mutaciones (Mallinckrodt Baker, Inc., Phillipsburg, NJ) para confirmar que no se ha producido ninguna otra alteración de la secuencia (mediante comparación del desplazamiento de bandas con el control no mutagenizado). Alternativamente, puede secuenciarse la región de ADN completa para confirmar que no se han producido sucesos mutacionales adicionales fuera de la región diana.

35 Los dúplex mutantes verificados en vectores de sobreexpresión pET (u otros) pueden utilizarse para transformar E. coli, tales como, por ejemplo, la cepa E. coli BL21 (DE3) pLysS, para una producción de alto nivel de la proteína mutante y la purificación mediante protocolos estándares. El método de mapeado por BAR-EM, por ejemplo, puede utilizarse para comprobar rápidamente la fidelidad de la expresión de mutantes. Esta técnica permite la secuenciación de segmentos en toda la proteína completa y proporciona la confianza necesaria en la asignación de secuencias. En un experimento de mapeado de este tipo, la proteína se digiere con una proteasa (la elección dependerá de la región específica que debe modificarse, ya que este segmento resulta de interés primordial y el mapa restante debería ser idéntico al mapa de la proteína no mutagenizada). El conjunto de fragmentos de corte se fracciona mediante, por ejemplo, HPLC Microbore (fase inversa o intercambio iónico, nuevamente dependiendo de la región específica que debe modificarse), proporcionando varios péptidos en cada fracción, y se determinan los pesos moleculares de los péptidos mediante métodos estándares, tales como BAR-EM. La masa determinada de cada fragmento seguidamente se compara con los pesos moleculares de los péptidos esperados de la digestión de la secuencia predicha y se determina rápidamente el grado de corrección de la secuencia. Debido a que este enfoque de mutagénesis a la modificación de las proteínas es dirigido, no debería resultar necesaria la secuenciación del péptido alterado en el caso de que los datos de EM concuerden con las predicciones. En caso necesario para verificar un residuo modificado, puede utilizarse la DIC (disociación inducida por colisión)-EM en tándem/EM para secuenciar los péptidos de la mezcla en cuestión, o el péptido diana puede purificarse para la degradación Edman sustractiva o la digestión con carboxipeptidasa Y, dependiendo de la localización de la modificación.

55 Las ADN polimerasas de la invención con más de una sustitución de aminoácido pueden generarse de diversas maneras. En el caso de los aminoácidos situados estrechamente próximos en la cadena polipeptídica, pueden mutarse simultáneamente utilizando un oligonucleótido que codifica la totalidad de las sustituciones de aminoácidos deseadas. Sin embargo, si los aminoácidos se encuentran situados a cierta distancia entre sí (separados por más de diez aminoácidos, por ejemplo), resulta más difícil generar un solo oligonucleótido que codifique la totalidad de los cambios deseados. De esta manera, puede utilizarse cualquiera de dos métodos alternativos. En el primer método, se genera un oligonucleótido separado para cada aminoácido que debe sustituirse. A continuación, los oligonucleótidos se hibridan con el ADN molde de cadena sencilla simultáneamente, y la segunda cadena de ADN que se sintetiza a partir del molde codificará todas las sustituciones de aminoácidos deseadas. Un método alternativo implica dos o más rondas de mutagénesis para producir el mutante deseado. La primera ronda es tal

como se describe para los mutantes únicos: se utiliza ADN codificante de la polimerasa no modificada para el molde, un oligonucleótido codificante de la primera o primeras sustituciones de aminoácidos deseadas se hibrida con este molde y seguidamente se genera la molécula de ADN heterodúplex. La segunda ronda de mutagénesis utiliza el ADN mutado producido en la primera ronda de mutagénesis como el molde. De esta manera, este molde ya contiene una o más mutaciones. El oligonucleótido codificante de la sustitución o sustituciones de aminoácidos deseadas adicionales seguidamente se hibrida con este molde, y la cadena resultante de ADN codifica ahora mutaciones tanto de la primera como de la segunda rondas de mutagénesis. Este ADN resultante puede utilizarse como molde en una tercera ronda de mutagénesis, y de esta manera sucesivamente. Alternativamente, puede utilizarse el método de mutagénesis multisitio de Seyfang y Jin (Anal. Biochem. 324:285-291, 2004).

De acuerdo con lo anteriormente expuesto, también se proporcionan ácidos nucleicos recombinantes que codifica cualquiera de las ADN polimerasas de la presente invención. Utilizando un ácido nucleico de la presente invención, codificante de una ADN polimerasa de la invención, puede prepararse una diversidad de vectores. En la práctica de la invención puede utilizarse cualquier vector que contenga replicón y secuencias de control derivadas de una especie compatible con la célula huésped. Generalmente, entre los vectores de expresión se incluyen regiones de ácidos nucleicos reguladoras de la transcripción y la traducción operablemente ligadas al ácido nucleico codificante de la ADN polimerasa mutante. La expresión "secuencias de control" se refiere a secuencias de ADN necesarias para la expresión de una secuencia codificante operablemente ligada en un organismo huésped particular. Entre las secuencias de control que resultan adecuadas para los procariotas se incluyen, por ejemplo, un promotor, opcionalmente una secuencia de operador, y un sitio de unión ribosómica. Además, el vector puede contener un elemento retrorregulador positivo (ERP) para incrementar la vida media del ARNm transcrito (ver Gelfand et al., patente US nº 4.666.848). Las regiones de ácidos nucleicos reguladoras de la transcripción y la traducción generalmente serán apropiadas para la célula huésped utilizada para expresar la polimerasa. Son conocidas de la técnica numerosos tipos de vectores de expresión apropiados, y secuencias reguladoras adecuadas, para una diversidad de células huésped. En general, entre las secuencias reguladoras de la transcripción y de la traducción pueden incluirse, por ejemplo, secuencias de promotor, sitios de unión ribosómica, secuencias de inicio y de parada de la transcripción, secuencias de inicio y de parada de la traducción y secuencias intensificadoras o activadoras. En realizaciones típicas, entre las secuencias reguladoras se incluyen un promotor y secuencias de inicio y de parada de la transcripción. Entre los vectores también se incluye típicamente una región policonectora que contiene varios sitios de restricción para la inserción de ADN foráneo. En determinadas realizaciones, se utilizan "etiquetas de fusión" para facilitar la purificación y, si se desea, la eliminación posterior de la secuencia de etiqueta/Flag, por ejemplo la "etiqueta His". Sin embargo, éstas generalmente resultan innecesarias al purificar una proteína termoactiva y/o termoestable a partir de un huésped mesofílico (por ejemplo *E. coli*), en donde puede utilizarse una "etapa de calentamiento". La construcción de vectores adecuados que contienen ADN codificante de secuencias de replicación, secuencias reguladoras, genes de selección fenotípica y la polimerasa mutante de interés se preparan utilizando procedimientos estándares de ADN recombinante. Los plásmidos, vectores víricos y fragmentos de ADN aislados se corta, se ajustan y se ligan entre sí en un orden específico para generar los vectores deseados, tal como es bien conocido de la técnica (ver, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, NY, 2a edición, 1989)).

En determinadas realizaciones, el vector de expresión contiene un gen de marcador seleccionable que permite la selección de las células huésped transformadas. Los genes de selección son bien conocidos de la técnica y varían según la célula huésped utilizada. Entre los genes de selección adecuados pueden incluirse, por ejemplo, genes codificantes de resistencia a la ampicilina y/o a la tetraciclina, lo que permite que las células transformadas con estos vectores crezcan en presencia de estos antibióticos.

En un aspecto de la presente invención, se introduce un ácido nucleico codificante de una ADN polimerasa mutante en una célula, solo o en combinación con un vector. La expresión "se introduce" o equivalentes gramaticales en la presente memoria se refieren a que los ácidos nucleicos se introducen en las células de una manera adecuada para la integración, amplificación y/o expresión posteriores del ácido nucleico. El método de introducción está dictado en gran medida por el tipo celular diana. Entre los métodos ejemplares se incluyen la precipitación con CaPO_4 , la fusión de liposomas, LIPOFECTIN[®], la electroporación, la infección vírica y similares.

Los procariotas son utilizados típicamente como células huésped para las etapas de clonación iniciales de la presente invención. Resultan particularmente útiles para la producción rápida de grandes cantidades de ADN, para la producción de moldes de ADN de cadena sencilla utilizados para la mutagénesis sitio-dirigida, para el cribado de muchos mutantes simultáneamente y para la secuenciación de ADN de los mutantes generados. Entre las células huésped procarióticas adecuadas se incluyen *E. coli* K12 cepa 94 (ATCC nº 31.446), *E. coli* cepa W3110 (ATCC nº 27.325), *E. coli* K12 cepa DG116 (ATCC nº 53.606), *E. coli* X1776 (ATCC nº 31.537) y *E. coli* B; sin embargo, muchas otras cepas de *E. coli*, tales como HB101, JM101, NM522, NM538, NM539 y muchas otras especies y géneros de procariotas, incluyendo bacilos tales como *Bacillus subtilis*, otras enterobacteriáceas tales como *Salmonella typhimurium* o *Serratia marcescens*, y diversas especies de *Pseudomonas*, pueden utilizarse como huéspedes. Las células huésped procarióticas u otras células huésped con paredes celulares rígidas se transforman típicamente utilizando el método del cloruro de calcio, tal como se describe en la sección 1.82 de Sambrook et al.,

supra. Alternativamente, puede utilizarse la electroporación para la transformación de estas células. Las técnicas de transformación de procarióticas se describen en, por ejemplo, Dower, en: *Genetic Engineering, Principles and Methods* 12:275-296 (Plenum Publishing Corp., 1990); Hanahan et al., *Meth. Enzymol.* 204:63, 1991. Entre los plásmidos utilizados típicamente para la transformación de *E. coli* se incluyen pBR322, pUC18, pUC19, pUC118, pUC119 y Bluescript M13, la totalidad de los cuales se describe en las secciones 1.12-1.20 de Sambrook et al., supra. Sin embargo, también se encuentran disponibles muchos otros vectores adecuados.

Las ADN polimerasas de la presente invención se producen típicamente mediante cultivo de una célula huésped transformada con un vector de expresión que contiene un ácido nucleico codificante de la ADN polimerasa mutante, bajo las condiciones apropiadas para inducir o provocar la expresión de la ADN polimerasa mutante. Los métodos de cultivo de células huésped transformadas bajo condiciones adecuadas para la expresión de proteínas son bien conocidos de la técnica (ver, por ejemplo, Sambrook et al., supra). Entre las células huésped adecuadas para la producción de las polimerasas mutantes a partir de vectores plásmidos que contienen el promotor pL de lambda se incluyen *E. coli* cepa DG116 (ATCC n° 53606) (ver la patente US n° 5.079.352, y Lawyer F.C. et al., *PCR Methods and Applications* 2:275-87, 1993). Tras la expresión, la polimerasa mutante puede recolectarse y aislarse. Los métodos para purificar la ADN polimerasa termoestable se describen en, por ejemplo, Lawyer et al., supra.

Tras la purificación, la capacidad de las ADN polimerasas mutantes para extender los moldes con cebadores puede someterse a ensayo en cualquiera de entre diversos ensayos conocidos para medir la extensión. Por ejemplo, en presencia de molécula de molde con cebadores (por ejemplo ADN de M13, etc.), un tampón apropiado, un juego completo de dNTP (por ejemplo dATP, dCTP, dGTP y dTTP) e iones metálicos, las ADN polimerasas extenderán los cebadores, convirtiendo el ADN de cadena sencilla (ADNmc) en ADN de doble cadena (ADNdc). Esta conversión puede detectarse y cuantificarse mediante, por ejemplo, la adición de un pigmento de unión a ADNdc, tal como SYBR verde I. Utilizando un termociclador cinético (ver Watson et al., *Anal. Biochem.* 329:58-67, 2004, y también disponible de, por ejemplo, Applied Biosystems, Stratagene y BioRAD), pueden obtenerse imágenes digitales de las placas de reacción (por ejemplo a intervalos de 10 a 30 segundos), permitiendo de esta manera realizar un seguimiento del progreso de las reacciones. El nivel de fluorescencia detectado puede convertirse fácilmente en tasas de extensión. Utilizando dichos ensayos rutinarios, pueden determinarse las tasas de extensión de los mutantes respecto a las formas no modificadas de la polimerasa.

Las ADN polimerasas de la presente invención pueden utilizarse para cualquier fin en el que dicha actividad enzimática resulte necesaria o deseada. De acuerdo con lo anteriormente expuesto, en otro aspecto de la invención se proporcionan métodos de extensión de cebadores utilizando las ADN polimerasas de la invención. Las condiciones adecuadas para la extensión de cebadores son conocidas de la técnica (ver, por ejemplo, Sambrook et al., supra). Ver también Ausubel et al., *Short Protocols in Molecular Biology* (4a edición, John Wiley & Sons, 1999). Generalmente se aparean un cebador, es decir, se hibrida, con un ácido nucleico diana, formando un complejo de cebador-molde. El complejo de cebador-molde se pone en contacto con la ADN polimerasa mutante y nucleótidos libres en un medio adecuado para permitir la adición de uno o más nucleótidos al extremo 3' del cebador, produciendo de esta manera un cebador extendido que es complementario al ácido nucleico diana. El cebador puede incluir, por ejemplo, uno o más análogos de nucleótido. Además, los nucleótidos libres pueden ser nucleótidos convencionales, nucleótidos no convencionales (por ejemplo ribonucleótidos o nucleótidos marcados) o una mezcla de los mismos. En algunas variaciones, la reacción de extensión de cebador comprende la amplificación de un ácido nucleico diana. Las condiciones adecuadas para la amplificación de los ácidos nucleicos utilizando una ADN polimerasa y una pareja de cebadores también son conocidas de la técnica (por ejemplo los métodos de amplificación por PCR). Ver, por ejemplo, Sambrook et al., supra; Ausubel et al., supra; *PCR Applications: Protocols for Functional Genomics* (Innis et al., editores, Academic Press, 1999). En otras realizaciones no mutuamente exclusivas, la reacción de extensión de cebadores comprende la transcripción inversa de un molde de ARN (por ejemplo la RT-PCR). La utilización de las presentes polimerasas mutantes, que proporcionan una tasa de extensión mejorada, proporcionan, por ejemplo, la capacidad de llevar a cabo dichas reacciones de extensión de cebadores con tiempos de incubación relativamente cortos, concentraciones de enzima reducidas y/o un rendimiento de producto incrementado.

En todavía otras realizaciones, las ADN polimerasas de la invención se utilizan para la extensión de cebadores en el contexto de la secuenciación del ADN, el marcaje del ADN o el marcaje de los productos de extensión de cebadores. Por ejemplo, la secuenciación del ADN mediante el método del dideoxinucleótido de Sanger (Sanger et al., *Proc. Natl. Acad. Sci. USA* 74: 5463, 1977) es mejorada por la presente invención para las polimerasas capaces de incorporar nucleótidos de terminación de cadena no convencionales. Los avances en el método básico de Sanger et al. han proporcionado nuevos vectores (Yanisch-Perron et al., *Gene* 33:103-119, 1985) y análogos de bases (Mills et al., *Proc. Natl. Acad. Sci. USA* 76:2232-2235, 1979, y Barr et al., *Biotechniques* 4:428-432, 1986). En general, la secuenciación del ADN requiere la extensión de cebadores dependiente de molde en presencia de análogos de bases terminadores de cadena, resultando en una distribución de fragmentos parciales que posteriormente se separan según tamaño. El procedimiento básico de secuenciación dideoxi implica: (i) la hibridación de un cebador oligonucleótido, opcionalmente marcado, con un molde, (ii) la extensión del cebador con ADN polimerasa en cuatro reacciones separadas, conteniendo cada uno una mezcla de dNTP no marcados y una cantidad limitante de un

agente terminador de cadena, tal como un ddNTP, opcionalmente marcado, e (iii) la resolución de los cuatro conjuntos de productos de reacción en un gel desnaturante de alta resolución de poliacrilamida/urea. Los productos de reacción pueden detectarse en el gel mediante autorradiografía o mediante detección de fluorescencia, dependiendo del marcaje utilizado, y la imagen puede examinarse para inferir la secuencia de nucleótidos. Estos métodos utilizan la ADN polimerasa, tal como el fragmento Klenow de Pol I de *E. coli* o una ADN polimerasa de T7 modificada.

La disponibilidad de las polimerasas termoestables, tales como la ADN polimerasa Taq, resultó en métodos mejorados de secuenciación con ADN polimerasa termoestable (ver Innis et al., Proc. Natl. Acad. Sci. USA 85:9436, 1988) y modificaciones de los mismos denominadas "secuenciación cíclica" (Murray, Nuc. Acids Res. 17:8889, 1989). De acuerdo con lo anteriormente expuesto, las polimerasas termoestables mutantes de la presente invención pueden utilizarse conjuntamente con dichos métodos. A modo de alternativa a la secuenciación dideoxi básica, la secuenciación cíclica es una amplificación asimétrica lineal de secuencias diana complementarias a la secuencia del molde en presencia de terminadores de cadena. Un único ciclo produce una familia de productos de extensión de todas las longitudes posibles. Tras la desnaturación del producto de reacción de extensión a partir del ADN molde, se producen múltiples ciclos de hibridación del cebador y de extensión del cebador en presencia de terminadores tales como los ddNTP. La secuenciación cíclica requiere menos ADN molde que la secuenciación convencional de terminación de cadena. Las ADN polimerasas termoestables presentan varias ventajas en la secuenciación cíclica: toleran las temperaturas restrictivas de hibridación que resultan necesarias para la hibridación específica de los cebadores con las dianas de ácidos nucleicos, así como toleran los múltiples ciclos de desnaturación a alta temperatura que se producen en cada ciclo, por ejemplo a 90-95°C. Por ello, la ADN polimerasa AMPLITAQ® y sus derivados y descendientes, por ejemplo la ADN polimerasa CS AmpliTaq y la ADN polimerasa FS AmpliTaq se han incluido en los kits de secuenciación cíclica Taq comercializados por compañías tales como Perkin-Elmer (Norwalk, CT) y Applied Biosystems (Foster City, CA).

Entre las variaciones de los métodos de secuenciación por terminación de cadena se incluyen la secuenciación de pigmento-cebador y la secuenciación de pigmento-terminador. En la secuenciación de pigmento-cebador, los terminadores ddNTP no se encuentran marcados y se utiliza un cebador marcado para detectar los productos de extensión (Smith et al., Nature 32:674-679, 1986). En la secuenciación de ADN de pigmento-terminador, se utiliza una ADN polimerasa para incorporar los dNTP y los ddNTP marcados fluorescentemente en el extremo de un cebador de ADN (Lee et al., Nuc. Acids Res. 20:2471, 1992). Este procedimiento ofrece la ventaja de que no resulta necesario sintetizar cebadores marcados con pigmento. Además, las reacciones de pigmento terminador resultan más convenientes en el aspecto de que la totalidad de las cuatro reacciones puede llevarse a cabo en el mismo tubo.

Los métodos tanto de pigmento-cebador como de pigmento-terminador pueden automatizarse utilizando un instrumento de secuenciación automatizado producido por Applied Biosystems (Foster City, CA) (patente US nº 5.171.534). Al utilizar el instrumento, la mezcla de reacción de secuenciación completada se fracciona en un gel de poliacrilamida desnaturante o en capilares montados en el instrumento. Un láser en la parte inferior del instrumento detecta los productos fluorescentes al someterlos a electroforesis según tamaño a través del gel.

Comúnmente se utilizan dos tipos de pigmentos fluorescentes para marcar los terminadores utilizados para la secuenciación de pigmento-terminador: pigmentos fluorescentes con carga negativa y zwitteriónicos. Entre los pigmentos fluorescentes con carga negativa se incluyen los de la familia de la fluoresceína y de BODIPY. Los pigmentos BODIPY (4,4-difloro-4-bora-3a,4a-diaza-s-indaceno) se describen en la publicación de patente internacional WO nº 97/00967. Entre los pigmentos fluorescentes zwitteriónicos se incluyen los de la familia de la rodamina. Los kits de secuenciación cíclica comercialmente disponibles utilizan terminadores marcados con derivados de rodamina. Sin embargo, los terminadores marcados con rodamina son bastante caros y el producto debe separarse de pigmento no incorporado-ddNTP antes de cargarlos en el gel, ya que comigran con los productos de secuenciación. Los terminadores de la familia del pigmento rodamina aparentemente estabilizan las estructuras de horquilla en las regiones ricas en GC, lo que provoca que los productos migren anómalamente. Lo anterior requiere la utilización de dTTP, que relaja la estructura secundaria, aunque también afecta a la eficiencia de la incorporación del terminador.

En contraste, los terminadores marcados con fluoresceína eliminan la etapa de separación previamente a la carga en el gel, ya que presentan una carga negativa neta mayor y migran más rápidamente que los productos de secuenciación. Además, los productos de secuenciación marcados con fluoresceína presentan una migración electroforética mejor que los productos de secuenciación marcados con rodamina. Aunque la ADN polimerasa Taq de tipo salvaje no incorpora eficientemente los terminadores marcados con pigmentos de la familia de la fluoresceína, lo anterior ahora puede llevarse a cabo eficientemente mediante la utilización de los enzimas modificados que se indican en la solicitud publicada de patente US nº 2002/0142333. De acuerdo con lo anteriormente expuesto, las modificaciones descritas en el documento US nº 2002/142333 pueden utilizarse en el contexto de la presente invención para producir polimerasas termoestables que incorporan pigmento de la familia de la fluoresceína que presentan tasas mejoradas de extensión de cebador. Por ejemplo, en determinadas

realizaciones, la ADN polimerasa no modificada según la presente invención es una polimerasa termoestable modificada tal como se describe en el documento US nº 2002/0142333 y que presenta el motivo indicado en SEC ID nº 1 ó de SEC ID nº 33, respectivamente.

5 Entre otros formatos ejemplares de secuenciación de ácidos nucleicos en los que pueden utilizarse las ADN polimerasas de la invención se incluyen aquellos que implican compuestos terminadores que incluyen análogos 2'-PO₄ de ribonucleótidos (ver, por ejemplo, las solicitudes publicadas de patente US nº 2005/0037991 y nº 2005/0037398, y la solicitud de patente US nº 11/583,605, titulada "SYNTHESIS AND COMPOSITIONS OF NUCLEIC ACIDS COMPRISING 2'-TERMINATOR NUCLEOSIDES", presentada el 19 de octubre de 2006, de Bodepudi et al. y las solicitud de patente US nº 11/583,606, titulada "2'-TERMINATOR RELATED PYROPHOSPHOROLYSIS ACTIVATED POLYMERIZATION", presentada el 19 de octubre de 2006, de Gelfand et al.). Las ADN polimerasas indicadas en la presente memoria generalmente mejoran dichos métodos de secuenciación, por ejemplo mediante la reducción del tiempo necesario para las reacciones cicladas de extensión y/o mediante la reducción de la cantidad o concentración de enzima que se utiliza para un rendimiento satisfactorio.

15 En otro aspecto de la presente invención, se proporcionan kits para la utilización en métodos de extensión de cebadores descritos en la presente memoria. Típicamente, el kit se compartimentaliza para facilitar el uso y contiene por lo menos un recipiente que proporciona una ADN polimerasa mutada según la presente invención. También pueden incluirse uno o más recipientes adicionales que proporcionan uno o más reactivos adicionales. Dichos recipientes adicionales pueden incluir cualesquiera reactivos u otros elementos reconocidos por el experto en la materia para la utilización en procedimiento de extensión de cebadores según los métodos descritos anteriormente, incluyendo reactivos para la utilización en, por ejemplo, procedimientos de amplificación de ácidos nucleicos (por ejemplo PCR, RT-PCR), procedimientos de secuenciación de ADN o procedimientos de marcaje de ADN. En particular, el kit incluye además un recipiente que proporciona un cebador sentido 5' hibridable, bajo condiciones de extensión de cebadores, con un polinucleótido molde predeterminado, o una pareja de cebadores que comprende el cebador sentido 5' y un cebador antisentido 3' correspondiente. El kit incluye además uno o más recipientes que proporcionan nucleótidos libres (convencionales y/o no convencionales). En particular, el kit incluye dNTP α -fosforotioato, dUTP, dITP y/o dNTP marcados, tales como, por ejemplo, los dNTP de la familia del pigmento fluoresceína o del pigmento cianina. El kit según la invención incluye además uno o más recipientes que proporcionan un tampón adecuado para una reacción de extensión de cebadores.

Ejemplos

35 EJEMPLO 1: IDENTIFICACIÓN Y CARACTERIZACIÓN DE UNA ADN POLIMERASA MUTANTE CON ACTIVIDAD DE EXTENSIÓN MEJORADA

Se identificó una mutación en las polimerasas de la familia CS que proporcionaba, por ejemplo, una capacidad mejorada de extender los moldes de ADN cebados en presencia de nucleótidos libres. En breve, entre las etapas de este procedimiento de cribado se incluían la generación de una biblioteca, la expresión y purificación parcial de enzimas mutantes, el cribado de los enzimas para la propiedad deseada, la secuenciación del ADN, la purificación clonal y la caracterización adicional de los mutantes seleccionados. Se describe en mayor detalle a continuación cada una de estas etapas.

45 La mutación identificada mediante este procedimiento era E558G. La mutación resultó en una capacidad mejorada de extender los moldes cebados. En el contexto particular de la mutación E678G, que permite la incorporación de ribonucleótidos y otros nucleótidos modificados en la posición 2', pero que también resultan en una capacidad deteriorada de extender moldes cebados, la mutación E558G mejora dicha propiedad de capacidad deteriorada de extensión de cebadores.

50 Generación de biblioteca clonal: un ácido nucleico codificante del dominio de polimerasa de la ADN polimerasa CS5 E678G se sometió a PCR mutagénica propensa a errores.

Se llevó a cabo una PCR utilizando un abanico de concentraciones de Mg²⁺ de entre 1,8 y 3,5 mM con el fin de generar bibliotecas con un abanico correspondiente de tasas de mutación. Las condiciones del tampón eran: bicina 50 mM, pH 8,2, KOAc 115 mM, glicerol al 8% p/v, 0,2 mM de cada dNTP y 0,2X de SYBR Verde I. Se utilizó un enzima de PCR de inicio en caliente GeneAmp[®] AccuRT a razón de 0,15 U/ml. Partiendo de 5x10⁵ copias de ADN de plásmido CS5 E678G/volumen de reacción de 50 ml, se llevaron a cabo 30 ciclos de amplificación utilizando una temperatura de hibridación de 60°C durante 15 segundos, una temperatura de extensión de 72°C durante 45 segundos y una temperatura de desnaturalización de 95°C durante 15 segundos.

60 Se purificó el amplicón resultante en una columna de centrifugación Qiaquick (Qiagen, Inc., Valencia, CA, USA) y se cortó con Bg1 II e Hind III, y después se purificó nuevamente. Se preparó un vector plásmido, una modificación de CS5 G46E L329A que portaba una delección grande en el dominio de polimerasa entre los sitios BgIII e HindIII, mediante el corte con los mismos dos enzimas de restricción y el tratamiento con fosfatasa intestinal bovina (FIB). El

vector cortado y la inserción mutada se mezclaron en diferentes proporciones y se trataron con ligasa de T4 durante la noche a 15°C. Se purificaron las ligaciones y se transformaron en *E. coli* cepa LK3 mediante electroporación.

5 Se sembraron alícuotas en una placa con medio selectivo con ampicilina con el fin de determinar el número de transformantes únicos en cada transformación. Las transformaciones con los transformantes más únicos a cada tasa de mutagénesis se almacenaron a una temperatura de entre -70°C y -80°C en presencia de glicerol como crioprotector.

10 A continuación, se extendió cada biblioteca sobre placas de agar de gran formato selectivas con ampicilina. Se transfirieron colonias individuales a placas de 384 pocillos que contenía 2X caldo Luria con ampicilina y glicerol al 10% p/v utilizando un recolector automático de colonias (QPix2, Genetic Ltd.). Estas placas se incubaron durante la noche a 30°C para permitir el crecimiento de los cultivos y después se almacenaron entre -70°C y -80°C. El glicerol añadido al 2X caldo Luria fue suficientemente poco para permitir el crecimiento del cultivo, aunque suficientemente alto para proporcionar crioprotección. Se prepararon varios miles de colonias a varios niveles de mutagénesis (Mg^{2+}) de esta manera para el uso posterior.

15 Preparación de biblioteca de extractos Parte 1 - Fermentación: A partir de las bibliotecas clonales indicadas anteriormente, se preparó una biblioteca correspondiente de extractos parcialmente purificados adecuados para los fines de cribado. La primera etapa de este procedimiento fue preparar cultivos de expresión a pequeña escala de cada clon. Estos cultivos se cultivaron en formato de 96 pocillos; por lo tanto, había 4 placas de cultivo de expresión por cada placa de biblioteca de 384 pocillos. Se transfirió un ml de cada pocillo de la placa de biblioteca clonal a un pocillo de una placa de inóculos de 96 pocillos, que contenía 150 ml de medio A (ver la Tabla 3, posteriormente). Esta placa de inóculos se agitó durante la noche a 1.150 rpm a 30°C en un incubador/agitador de placas iEMS (ThermoElectron). A continuación, estos cultivos de inóculos se utilizaron para inocular el mismo medio, esta vez inoculando 10 ml en 300 ml de medio A en placas de 96 pocillos de gran formato (Nunc nº 267334). Estas placas se incubaron durante la noche a 37°C. El plásmido de expresión contenía elementos de control transcripcional que permitían la expresión a 37°C pero no a 30°C. Tras la incubación durante la noche, los cultivos expresaban la proteína de clon típicamente a razón de 1-10% de las proteínas celulares totales. Las células de estos cultivos se recolectaron mediante centrifugación. Estas células se congelaron (-20°C) o se procesaron inmediatamente, tal como se indica a continuación.

Tabla 3. Medio A (esterilizado mediante filtración previamente a la utilización)

Componente	Concentración
MgSO ₄ .7H ₂ O	0,2 g/l
Ácido cítrico.H ₂ O	2 g/l
K ₂ HPO ₄	10 g/l
NaNH ₄ PO ₄ .4H ₂ O	3,5 g/l
MgSO ₄	2 mM
Casaminoácidos	2,5 g/l
Glucosa	2 g/l
HCl de tiamina	10 mg/l
Ampicilina	100 mg/l

35 Preparación de biblioteca de extractos Parte 2 - Extracción: Se resuspendieron pellets celulares de la etapa de fermentación en 30 ml de tampón de lisis (Tabla 4, posteriormente) y se transfirieron a placas del termociclador de 384 pocillos. Observar que el tampón contenía lisozima para ayudar a la lisis celular, y dos nucleasas para eliminar tanto el ARN como el ADN del extracto. Las placas se sometieron a tres rondas de congelación-descongelación (-70°C congelación, 37°C descongelación, no menos de 15 minutos en cada etapa) para lisar las células. Se añadió sulfato amónico (5 ml de una solución 0,75 M) y las placas se incubaron a 75°C durante 15 minutos con el fin de precipitar e inactivar las proteínas contaminantes, incluyendo las nucleasas añadidas exógenamente. Las placas se centrifugaron a 3.000xg durante 15 minutos y los sobrenadantes se transfirieron a una placa de termociclador nueva de 384 pocillos. Estas placas de extractos se congelaron a -20°C para la utilización posterior en cribados. Cada pocillo contenía aproximadamente 0,5 a 3 mM del enzima polimerasa mutante.

45 Tabla 4. Tampón de lisis

Componente	Concentración o porcentaje
Tris pH 8,0	20 mM
EDTA	1 mM
MgCl ₂	5 mM
TLCK	1 mM
Leupeptina	1 mg/ml

Pefabloc	0,5 mg/ml
Tween-20	al 0,5% v/v
Lisozima (procedente de polvos)	2 mg/ml
ARNasa	0,025 mg/ml
ADNasa I	0,075 unidades/ml

Cribado de bibliotecas de extractos para tasa de extensión mejorada: se utilizó ADN de cadena sencilla M13mp18 (ADN de M13), cebado con un oligonucleótido que presentaba la secuencia siguiente:

5 5'-GGGAAGGGCGATCGGTGCGGCCTCTTCGC-3' (SEC ID nº 28)

10 como la molécula de molde en el cribado del ensayo de extensión. En este cribado, las placas de extractos indicadas anteriormente se diluyeron 10 veces en Tris 10 mM, pH 8,0 / EDTA 1 mM/KCl 100 mM/Tween-20 al 0,2% y se trató por calor a 90°C durante 10 minutos, con el fin de incrementar la pureza. Se añadió 1,0 ml de extracto a 13 ml de mezcla maestra de reacción que contenía 1 nM de molde de M13 cebado en placas de PCR de 384 pocillos. Se realizó un seguimiento de la extensión del molde cebado a 64°C cada 20 segundos en un ciclador térmico cinético modificado utilizando una cámara de CCD. A continuación se lista una mezcla maestra de reacción típica. La mezcla de reacción contenía además Tricina 100 mM, pH 8,0, KOAc 20 mM, MgCl₂ 3 mM, tampón de almacenamiento de enzima al 2,5% v/v que contenía Tween-20 al 0,5%, 0,1 mM de cada uno de dATP, dCTP, dGTP y dTTP y SYBR Verde I a 0,6X (Molecular Probes), que permitió la detección fluorescente de la extensión de la cadena de cebador. Con el fin de distinguir la fluorescencia derivada de la extensión de la fluorescencia de fondo, se incluyeron pocillos paralelos en el experimento en el que se había bloqueado la extensión de la cadena de cebador, por ejemplo mediante adición de un quelante metálico tal como EDTA, o excluyendo los nucleótidos de la mezcla maestra de reacción.

20 En este cribado se identificaron extractos mutantes que mostraban una tasa incrementada de extensión. El cribado primario se llevó a cabo a la escala de miles de extractos. Los pocillos de cultivo correspondientes a los mejores extractos se seleccionaron para ensayos posteriores. En primer lugar se estrieron en placas de agar selectivas para garantizar la pureza clonal. Se purificó el enzima mutante a partir de 100 ml de cultivos en matraz de agitación y se determinó la concentración mediante densitometría en gel. Estas preparaciones de enzima cuantificadas se compararon con el enzima parental bajo las condiciones utilizadas en el cribado, aunque a igual concentración de proteína. Este cribado final garantizó que las diferencias observadas no eran simplemente efectos de la concentración de proteína.

30 Tras esta ronda final de cribado, aparentemente cinco clones todavía presentaban tasas de extensión mejoradas. Se determinó que las secuencias de estos clones codificaban para los cambios de aminoácidos siguientes respecto a la cepa parental:

35 clon 1: S671 F
clon 2: S671F
clon 3: Q610R E779K I812L M844I
clon 4: E558G I829V
clon 5: E558G K861M

40 En el caso de los clones 1 y 2, resulta evidente que la mutación S671F debe haber sido responsable del fenotipo observado, ya que era la única mutación de aminoácido en el clon. Para el clon 3, el fenotipo con toda probabilidad era el resultado de la mutación Q601R, basándose en otros resultados. Debido a que tanto el clon 4 como el 5 portan la misma mutación E558G, aparentemente esta mutación era con toda probabilidad responsable del fenotipo observado. Para confirmarlo, se mutó un clon parental (ADN polimerasa CS5 G46E L329A E678G, "GLE") para que incluyese una mutación E558G adicional (ADN polimerasa CS5 G46E L329A E558G E678G, "GLEE") utilizando la bien conocida técnica de la mutagénesis in vitro mediante PCR solapante. El plásmido resultante se secuenció para confirmar que portaba la mutación deseada y no otra mutaciones no deseadas que ocasionalmente se generan durante las etapas de PCR de este procedimiento.

50 El nuevo plásmido se transformó en el huésped E. coli cepa LK3 y la proteína polimerasa se expresó, se purificó a homogeneidad y se cuantificó. Estos nuevos enzimas mutantes resultantes se compararon con los tipos parentales y con otros mutantes, bajo condiciones similares a las del cribado original. Se muestran los resultados en la figura 4. La cepa portadora de la mutación E558G, "GLEE", era más de 12 veces más rápida en la extensión de M13 cebado que el clon parental "GLE" bajo las condiciones del presente ensayo. Resulta evidente a partir de estos datos que la mutación E558G era única responsable del fenotipo mejorado de los clones mutantes 4 y 5. La figura también demuestra la tasa relativa de otras mutaciones determinadas en el esqueleto GLE, tales como D640G ("GLDE"), D573G ("997-OI"), Q601R ("GLQE"), S671F ("GLSE"), así como clones portadores de múltiples mutaciones, tales

como la combinación de Q601R, S671F y D640G ("GLQDSE") y finalmente E558G en combinación con Q601R, S671F y D640G ("GLEQDSE").

LISTADO DE SECUENCIAS

- 5 <110> F. Hoffmann-La Roche AG and Roche Diagnostics GmbH
- <120> ADN polimerasas mejoradas y métodos relacionados
- 10 <130> 24304 WO
- <140> Todavía no asignado
<141> Todavía no asignado
- 15 <150> US 60/949,732
<151> 2007-07-13
- <160> 33
- 20 <170> FastSEQ para Windows versión 4.0
- <210> 1
<211> 21
<212> PRT
- 25 <213> Secuencia artificial
- <220>
<223> región del dominio polimerasa de ADN polimerasa termoestable mejorada modificada, motivo de dominio polimerasa
- 30 <221> MOD_RES 45
<222> (3)...(3)
<223> Xaa = His, Glu o Gln
- 35 <221> MOD_RES
<222> (4)...(4)
<223> Xaa = cualquier aminoácido
- 40 <221> MOD_RES
<222> (6)...(6)
<223> Xaa = Asn o His
- <221> MOD_RES
<222> (7)...(7)
- 45 <223> Xaa = Leu o Ile
- <221> MOD_RES
<222> (8)...(8)
<223> Xaa = cualquier aminoácido
- 50 <221> MOD_RES
<222> (10)...(10)
<223> Xaa = cualquier aminoácido
- 55 <221> MOD_RES
<222> (11)...(11)
<223> Xaa = Asp, Lys o Thr
- <221> MOD_RES
<222> (13)...(13)
<223> Xaa = Leu o Val
- 60 <221> MOD_RES
<222> (14)...(15)

<223> Xaa = cualquier aminoácido

<221> MOD RES
 <222> (16)...(16)
 5 <223> Xaa = Val, Ile o Leu

<221> MOD_RES
 <222> (18)...(18)
 10 <223> Xaa = Phe o Tyr

<221> MOD RES
 <222> (19)...(19)
 <223> Xaa = Ala, Cys, Phe, Gly, His, Ile, Lys, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp o Tyr, un aminoácido diferente de Asp o Glu

15 <221> MOD RES
 <222> (20)...(20)
 <223> Xaa = Lys o Glu

20 <221> MOD RES
 <222> (21)...(21)
 <223> Xaa = Leu o Gln

<400> 1

25

Ala	Gly	Xaa	Xaa	Phe	Xaa	Xaa	Xaa	Ser	Xaa	Xaa	Gln	Xaa	Xaa	Xaa	Xaa
1				5				10						15	
Leu	Xaa	Xaa	Xaa	Xaa											
			20												

<210> 2
 <211> 21
 30 <212> PRT
 <213> Secuencia artificial

<220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable mejorada modificada, motivo de dominio polimerasa

35

<221> MOD RES
 <222> (1)...(1)
 <223> Xaa = Ala o Val

40

<221> MOD_RES
 <222> (2)...(2)
 <223> Xaa = Gly o Val

45

<221> MOD_RES
 <222> (3)...(3)
 <223> Xaa = His, Glu o Gln

50

<221> MOD_RES
 <222> (4)...(4)
 <223> Xaa = cualquier aminoácido

55

<221> MOD_RES
 <222> (6)...(6)
 <223> Xaa = Asn o His

60

<221> MOD_RES
 <222> (7)...(7)
 <223> Xaa = Leu, Ile o Pro

<221> MOD_RES

- <222> (8)...(8)
 <223> Xaa = cualquier aminoácido
- 5 <221> MOD RES
 <222> (10)...(10)
 <223> Xaa = cualquier aminoácido
- 10 <221> MOD RES
 <222> (11)...(11)
 <223> Xaa = Asp, Lys o Thr
- 15 <221> MOD RES
 <222> (12)...(12)
 <223> Xaa = Gln o His
- 20 <221> MOD RES
 <222> (13)...(13)
 <223> Xaa = Leu, Val o Ile
- 25 <221> MOD RES
 <222> (14)...(15)
 <223> Xaa = cualquier aminoácido
- 30 <221> MOD RES
 <222> (16)...(16)
 <223> Xaa = Val, Ile, Leu o Lys
- 35 <221> MOD RES
 <222> (18)...(18)
 <223> Xaa = Phe, Tyr o Gln
- 40 <221> MOD RES
 <222> (19)...(19)
 <223> Xaa = Ala, Cys, Phe, Gly, His, Ile, Lys, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp or Tyr, un aminoácido diferente de Asp o Glu
- 45 <221> MOD RES
 <222> (20)...(20)
 <223> Xaa = Lys, Glu o Ala
- 50 <221> MOD RES
 <222> (21)...(21)
 <223> Leu, Gln o Gly
- 45 <400> 2
- | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Xaa | Xaa | Xaa | Phe | Xaa | Xaa | Xaa | Ser | Xaa |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Xaa | Xaa | Xaa | Xaa | | | | | | | | | | | |
| | | | | 20 | | | | | | | | | | | |
- 50 <210> 3
 <211> 26
 <212> PRT
 <213> Secuencia artificial
- 55 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermus thermophilus (Tth)
- <400> 3

ES 2 403 591 T3

Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Arg Leu
 20 25

- <210> 4
- <211> 26
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica *Thermus caldophilus* (Tca)
- <400> 4
- Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Arg Leu
 20 25
- 15 <210> 5
- <211> 26
- <212> PRT
- <213> Secuencia artificial
- 20 <220>
- <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica *Thermus* sp. Z05 (Z05)
- 25 <400> 5
- Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Arg Leu
 20 25
- 30 <210> 6
- <211> 26
- <212> PRT
- <213> Secuencia artificial
- <220>
- 35 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica *Thermus aquaticus* (Taq)
- <400> 6
- Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Gly Leu
 20 25
- 40 <210> 7
- <211> 26
- <212> PRT
- <213> Secuencia artificial
- 45 <220>
- <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica *Thermus flavus* (Tfl)
- 50 <400> 7

ES 2 403 591 T3

Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Gly Leu
 20 25

<210> 8
 <211> 26
 <212> PRT
 <213> Secuencia artificial

5
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermus filiformis (Tfi)
 10

<400> 8

His Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Gly Leu
 20 25

15
 <210> 9
 <211> 26
 <212> PRT
 <213> Secuencia artificial

20
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermus sp. sps17 (Sps17)

25 <400> 9

His Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 1 5 10 15
 Glu Arg Val Leu Phe Asp Glu Leu Gly Leu
 20 25

30
 <210> 10
 <211> 26
 <212> PRT
 <213> Secuencia artificial

35
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Deinococcus radiodurans (Dra)

<400> 10

His Glu Tyr Ala Gly Glu Glu Phe His Ile Arg Ser Pro Lys Gln Leu
 1 5 10 15
 Glu Thr Val Leu Tyr Asp Lys Leu Glu Leu
 20 25

40
 <210> 11
 <211> 26
 <212> PRT
 <213> Secuencia artificial

45
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica familia Hot Spring/clon 7 (HspB)

50 <400> 11

ES 2 403 591 T3

Tyr Thr Leu Ala Gly Glu Ala Phe Asn Ile Gly Ser Pro Lys Gln Leu
 1 5 10 15
 Gly Ala Ile Leu Phe Glu Lys Leu Gly Leu
 20 25

5 <210> 12
 <211> 26
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Bacillus stearothermophilus (Bst)

<400> 12

Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile Asn Ser Pro Lys Gln Leu
 1 5 10 15
 Gly Thr Val Leu Phe Asp Lys Leu Gln Leu
 20 25

15 <210> 13
 <211> 26
 <212> PRT
 20 <213> Secuencia artificial

<220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Bacillus caldotenax (Bca)

25 <400> 13

Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile Asn Ser Pro Lys Gln Leu
 1 5 10 15
 Gly Val Ile Leu Phe Glu Lys Leu Gln Leu
 20 25

30 <210> 14
 <211> 26
 <212> PRT
 <213> Secuencia artificial

35 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Escherichia coli (Eco)

<400> 14

40 His Glu Ile Ala Gly Glu Glu Phe Asn Leu Ser Ser Thr Lys Gln Leu
 1 5 10 15
 Gln Thr Ile Leu Phe Glu Lys Gln Gly Ile
 20 25

45 <210> 15
 <211> 26
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermotoga maritime (Tma)

50

ES 2 403 591 T3

<400> 15

```

Tyr Arg Ile Ala Gly Glu Pro Phe Asn Ile Asn Ser Pro Lys Gln Val
 1           5           10           15
Ser Arg Ile Leu Phe Glu Lys Leu Gly Ile
           20           25

```

5 <210> 16
 <211> 26
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermotoga neapolitana (Tne)

15 <400> 16

```

Tyr Gln Ile Ala Gly Glu Pro Phe Asn Ile Asn Ser Pro Lys Gln Val
 1           5           10           15
Ser Asn Ile Leu Phe Glu Lys Leu Gly Ile
           20           25

```

20 <210> 17
 <211> 26
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica Thermosiphon africanus (Taf)

<400> 17

```

Phe Glu Ile Ala Gly Glu Thr Phe Asn Leu Asn Ser Ser Thr Gln Val
 1           5           10           15
Ala Tyr Ile Leu Phe Glu Lys Leu Asn Ile
           20           25

```

30 <210> 18
 <211> 26
 <212> PRT
 <213> Secuencia artificial

35 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable de la bacteria termofílica familia A Hot Spring (HspA)

40 <400> 18

```

Tyr Ala Gln Ala Gly Glu Val Phe Asn Leu Asn Ser Pro Lys Gln Leu
 1           5           10           15
Gly Ser Leu Leu Phe Glu Lys Leu Lys Leu
           20           25

```

45 <210> 19
 <211> 24
 <212> PRT
 <213> Secuencia artificial

50 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable del bacteriófago T7 (T7).

ES 2 403 591 T3

<400> 19

Val Glu His Val Val Phe Asn Pro Ser Ser Arg Asp His Ile Gln Lys
1 5 10 15
Lys Leu Gln Glu Ala Gly Trp Val
20

5

<210> 20

<211> 893

10 <212> PRT

<213> Secuencia artificial

<220>

<223> ADN polimerasa termoestable quimérica CS5

15

<400> 20

Met Lys Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1 5 10 15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
 20 25 30
 Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
 35 40 45
 Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
 65 70 75 80
 Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
 100 105
 Glu Val Pro Gly Phe Glu Ala Asp Val Leu Ala Thr Leu Ala Lys
 115 120 125
 Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
 130 135 140
 Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
 145 150 155 160
 Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
 165 170 175
 Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
 195 200 205
 Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
 225 230 235 240
 Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
 260 265 270
 Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
 275 280 285
 Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
 290 295 300
 Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
 305 310 315 320
 Ala Ile Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
 325 330 335
 Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
 340 345 350
 Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
 355 360 365
 Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
 370 375 380
 Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
 385 390 395 400
 Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
 405 410 415
 Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
 420 425 430
 Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
 435 440 445
 Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
 450 455 460
 Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
 465 470 475 480
 Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
 485 490 495

```

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
      500      505
Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
      515      520      525
Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
      530      535      540
Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545      550      555      560
Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
      565      570      575
Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
      580      585      590
Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
      595      600      605
Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
610      615      620
Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625      630      635      640
Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
      645      650      655
Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
      660      665      670
Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
      675      680      685
Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
690      695      700
Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705      710      715      720
Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
      725      730      735
Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
      740      745      750
Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
755      760      765
Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
770      775      780
Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785      790      795      800
Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
      805      810      815
Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
820      825      830
Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
835      840      845
His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
850      855      860
Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
865      870      875      880
Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
      885      890

```

- 5 <210> 21
- <211> 893
- <212> PRT
- <213> Secuencia artificial

- <220>
- 10 <223> ADN polimerasa termoestable quimérica CS6

<400> 21

Met Lys Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
 1 5 10 15
 Val Asp Gly His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
 20 25 30
 Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
 35 40 45
 Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
 65 70 75 80
 Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
 100 105 110
 Glu Val Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
 115 120 125
 Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
 130 135 140
 Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
 145 150 155 160
 Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
 165 170 175
 Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
 195 200 205
 Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
 225 230 235 240
 Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
 260 265 270
 Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
 275 280 285
 Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
 290 295 300
 Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
 305 310 315 320
 Ala Ile Ala Leu Ala Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
 325 330 335
 Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
 340 345 350
 Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
 355 360 365
 Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
 370 375 380
 Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
 385 390 395 400
 Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
 405 410 415
 Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
 420 425 430
 Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
 435 440 445
 Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
 450 455 460
 Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
 465 470 475 480

Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
 485 490 495
 Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
 500 505 510
 Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
 515 520 525
 Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
 530 535 540
 Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
 545 550 555 560
 Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
 565 570 575
 Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
 580 585 590
 Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
 595 600 605
 Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
 610 615 620
 Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
 625 630 635 640
 Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
 645 650 655
 Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
 660 665 670
 Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
 675 680 685
 Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
 690 695 700
 Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
 705 710 715 720
 Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
 725 730 735
 Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
 740 745 750
 Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
 755 760 765
 Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
 770 775 780
 Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
 785 790 795 800
 Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
 805 810 815
 Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
 820 825 830
 Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
 835 840 845
 His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
 850 855 860
 Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
 865 870 875 880
 Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
 885 890

<210> 22
 <211> 2682
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> ADN polimerasa termoestable quimérica CS5

<400> 22

```

atgaaagcta  tgttaccatt  attcgaaccc  aaaggccggg  tcctcctggt  ggacggccac  60
cacctggcct  accgcacctt  cttcgccttg  aagggcctca  ccacgagccg  gggcgaaccg  120
gtgcaggcgg  ttacggctt  cgccaagagc  ctctcaagg  ccctgaagga  ggacgggtac  180
aaggccgtct  tcgtggtctt  tgacgccaag  gcccttct  tccgccacga  ggcctacgag  240
gcctacaagg  caggccgcgc  cccgaccccc  gaggacttcc  cccggcagct  cgcctcatc  300
aaggagctgg  tggacctcct  ggggtttact  cgctcgagg  ttccgggctt  tgaggcggac  360
gacgtcctcg  ccacctggc  caagaaggcg  gaaagggagg  ggtacgaggt  gcgcatcctc  420
accgccgacc  gggaccttta  ccagctcgtc  tccgaccgcg  tcgccgtcct  ccaccccag  480
ggccacctca  tcaccccgga  gtggctttgg  gagaagtacg  gccttaagcc  ggagcagtgg  540
gtggacttcc  gcgccctcgt  gggggacccc  tccgacaacc  tccccggggt  caagggcatc  600
ggggagaaga  ccgccctcaa  gtcctcaag  gagtggggaa  gcctggaaaa  tatcctcaag  660
aacctggacc  gggtgaagcc  ggaaagcgtc  cgggaaagya  tcaaggccca  cctggaagac  720
cttaagctct  ccttggagct  ttccggggtg  cgctcggacc  tccccctgga  ggtggacttc  780
gcccgagggc  gggagcctga  ccgggaaggg  cttcgggctt  ttttggagcg  cttggagttc  840
ggcagcctcc  tccacgagtt  cggccttcta  gaggagtccg  aaccggttg  gtaccgtata  900
gttaaagacc  tggttgaatt  tgaaaaactc  atagagaaac  tgagagaatc  tccttcgttc  960
gctatcgatt  tggaaactag  ttccctcgat  cctttcgaact  gcgacattgt  cggtatctct  1020
gtgtctttca  aaccaaagga  agcgtactac  ataccactcc  atcatagaaa  cgcccagaac  1080
ctggacgaaa  aagaggttct  gaaaaagctc  aaagaaattc  tggaggacc  cggagcaaaag  1140
atcgttggtc  agaatttgaa  attcgaattac  aaggtgttga  tgggtgaagg  tgttgaacct  1200
gttcctcctt  acttcgacac  gatgatagcg  gcttaccttc  ttgagccgaa  cgaaaagaag  1260
ttcaatctgg  acgatctcgc  attgaaattt  cttggataca  aaatgacatc  ttaccaagag  1320
ctcatgtcct  tctcttttcc  gctgtttggt  ttcagttttg  ccgatgttcc  tgtagaaaaa  1380
ggcgaact  actcctgtga  agatgcagac  atcacctaca  gactttaca  gaccctgagc  1440
ttaaactcc  acgagcgaga  tctggaaaac  gtgttctaca  agatagaaat  ccccttgtg  1500
aacgtgcttg  cacggatgga  actgaacggt  gtgtatgtgg  acacagagtt  cctgaagaaa  1560
ctctcagaag  agtacggaaa  aaaactcgaa  gaaactggcag  aggaaatata  caggatagct  1620
ggagagccgt  tcaacataaa  ctaccogaag  caggtttcaa  ggatcctttt  tgaaaaactc  1680
ggcataaaac  cacgtggtaa  aacgacgaaa  acgggagact  attcaacacg  catagaagtc  1740
ctcagagAAC  ttgccgggtga  acacgaaatc  attcctctga  ttcttgaata  cagaaagata  1800
cagaaattga  aatcaaccta  catagacgct  cttcccaaga  tggtaacacc  aaagaccgga  1860
aggattcatg  cttctttcaa  tcaaacgggg  actgccactg  gaagacttag  cagcagcgat  1920
cccaatcttc  agaacctccc  gacgaaaagt  gaagagggaa  aagaaatcag  gaaagcgata  1980
gttcctcagg  atccaaactg  gtggatcgtc  agtgccgact  actcccaaat  agaactgagg  2040
atcctcgccc  atctcagtg  tgatgagaat  cttttgagg  cattcgaaga  gggcatcgac  2100
gtccacactc  taacagcttc  cagaatattc  aacgtgaaac  ccgaagaagt  aaccgaagaa  2160
atgcgccgcg  ctggtaaaat  ggttaatttt  tccatcatat  acggtgtaac  acctacggt  2220
ctgtctgtga  ggcttggagt  acctgtgaaa  gaagcagaaa  agatgatcgt  caactacttc  2280
gtcctctacc  caaagggtcg  cgattacatt  cagagggctg  tatcggaaagc  gaaagaaaaa  2340
ggctatgtta  gaacgctgtt  tggaaagaaa  agagacatac  cacagctcat  ggcccgggac  2400
aggaacacac  aggctgaagg  agaacgaatt  gccataaaca  ctccataca  ggggtacagca  2460
gcggatataa  taaagctggc  tatgatagaa  atagacaggg  aactgaaaga  aagaaaaatg  2520
agatcgaaga  tgatcataca  ggtccacgac  gaactggttt  ttgaagtgcc  caatgaggaa  2580
aaggacgcgc  tcgtcgagct  ggtgaaagac  agaatgacga  atgtggtaaa  gctttcagtg  2640
ccgctcgaag  tggatgtaac  catcggcaaa  acatggtcgt  ga  2682

```

5

<210> 23
 <211> 2682
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> ADN polimerasa termoestable quimérica CS6

15

<400> 23

ES 2 403 591 T3

atgaaagcta tgttaccatt attcgaaccc aaaggccggg tcctcctggt ggacggccac 60
cacctggcct accgcacctt cttcgccttg aagggcctca ccacgagccg gggcgaaccg 120
gtgcaggcgg tttacggctt cgccaagagc ctcctcaagg ccctgaagga ggacgggtac 180
aaggccgtct tcgtggtctt tgacgccaaag gcccttctct tccgccacga ggcctacgag 240
gcctacaagc caggccgcgc cccgaccccc gaggacttcc cccggcagct cgcctctcctc 300

aaggagctgg tggacctcct ggggtttact cgcctcgagg ttccgggctt tgaggcggac 360
gacgtcctcg ccaccctggc caagaaggcg gaaagggagg ggtacgaggt gcgcatcctc 420
accgccgacc gggaccttta ccagctcgtc tccgaccgcg tcgccgtcct ccacccccgag 480
ggccacctca tcaccccgga gtggctttgg gagaagtacg gccttaagcc ggagcagtg 540
gtggaattcc ggcacctcgt gggggacccc tccgacaacc tccccggggt caagggcctc 600
ggggagaaga ccgccctcaa gtcctcaag gagggggaa gcctggaaaa tatcctcaag 660
aacctggacc gggggaagcc ggaagcgtc cgggaaagg tcaaggcca cctggaagac 720
cttaagctct ccttgagct tccccgggtg cgtcggacc tccccctgga ggtggacttc 780
gcccggaggc gggagcctga ccgggaaggg cttcgggctt tttggagcg cttggagttc 840
ggcagcctcc tccacgagtt cggccttcta gaggagtccg aaccggttg gtaccgtata 900
gttaaagacc tgggtgaatt tgaaaaactc atagagaaac tgagagaatc tccttcgctc 960
gcgatcgctc ttgcgactag ttccctcgat cctttcgact gcgacattgt cggtatctct 1020
gtgtctttca aaccaaagga agcgtactac ataccactcc atcatagaaa cgcaccagaac 1080
ctggacgaaa aagaggttct gaaaaagctc aaagaaattc tggaggacc cggagcaaac 1140
atcgttggtc agaatttgaa attcgattac aagggtgtga tggggaaggg tgttgaacct 1200
gttcctcctt acttcgacac gatgatagcg gcttaccttc ttgagccgaa cgaaaagaag 1260
ttcaatctgg acgatctcgc attgaaattt cttggataca aatgacatc ttaccaagag 1320
ctcatgtcct tctcttttcc gctgtttggt ttcagttttg ccgatgttcc tgtagaaaaa 1380
gcagcgaact actcctgtga agatgcagac atcacctaca gactttataa gacctgagc 1440
ttaaactcc acgaggcaga tctggaaaaac gtgttctaca agatagaaat gcccttgtg 1500
aacgtcttg cacggatgga actgaacggt gtgtatgtg acacagagt cctgaaagaa 1560
ctctcagaag agtacggaaa aaaactcgaa gaactggcag aggaaatata caggatagct 1620
ggagagccgt tcaacataaa ctaccgaaag caggtttcaa ggatcctttt tgaaaaactc 1680
ggcataaaac cacgtggtaa aacgacgaaa acgggagact attcaacacg catagaagtc 1740
ctcgaggaac ttgccgtgga acacgaaatc attcctctga ttcttgaata cagaaagata 1800
cagaaattga aatcaacctc catagacgct cttcccaaga tggteaacc aaagaccgga 1860
aggattcatg cttctttcaa tcaaaccggg actgccactg gaagacttag cagcagcgat 1920
cccaatcttc agaacctccc gacgaaaagt gaagagggaa aagaaatcag gaaagcgata 1980
gttcctcagg atccaaactg gtggatcgtc agtgccgact actcccaaat agaactgagg 2040
atcctcggcc atctcagtg tgatgagaat cttttgaggg cattcgaaga gggcatcgac 2100
gtccacactc taacagcttc cagaatattc aacgtgaaac ccgaagaagt aaccgaagaa 2160
atgcccgcgg ctggtaaaat ggttaatttt tccatcatat acggtgtaac accttacggt 2220
ctgctctgta ggcttgagtg acctgtgaaa gaagcagaaa agatgatcgt caactacttc 2280
gtcctctacc caaaggtgcg cgattacatt cagagggtcg tatcggaagc gaaagaaaaa 2340
ggctatgtta gaacgctggt tggaaagaaa agagacatac cacagctcat ggcccgggac 2400
aggaacacac aggctgaagg agaacgaatt gccataaaca ctccataca gggtagacga 2460
gcgatataa taaagctggc tatgatagaa atagacaggg aactgaaaga aagaaaaatg 2520
agatcgaaga tgatcataca ggtccacgac gaactggttt ttgaagtgcc caatgaggaa 2580
aaggacgcgc tcgtcgagct ggtgaaagac agaatgacga atgtggtaaa gctttcagtg 2640
ccgctcgaag tggatgtaac catcggcaaa acatggctgt ga 2682

<210> 24
<211> 26
5 <212> PRT
<213> Secuencia artificial

<220>
10 <223> secuencia de consenso de la región del dominio polimerasa de la ADN polimerasa termoestable, secuencia de consenso de motivo de dominio polimerasa

<221> MOD_RES
<222> (1)...(26)

ES 2 403 591 T3

<223> Xaa = cualquier aminoácido

<400> 24

	Xaa	Xaa	Xaa	Ala	Gly	Xaa	Xaa	Phe	Asn	Xaa	Xaa	Ser	Xaa	Xaa	Gln	Xaa
	1				5				10					15		
5	Xaa	Xaa	Xaa	Leu	Phe	Xaa	Xaa	Leu	Xaa	Xaa						
				20				25								

<210> 25

<211> 16

<212> PRT

10 <213> Secuencia artificial

<220>

<223> región del dominio polimerasa de ADN polimerasa termoestable mutante, motivo de dominio polimerasa

15 <221> MOD_RES

<222> (1)...(1)

<223> Xaa = Ile o Leu

20 <221> MOD_RES

<222> (2)...(2)

<223> Xaa = Gln o Leu

25 <221> MOD_RES

<222> (3)...(3)

<223> Xaa = Gln, His o Glu

30 <221> MOD_RES

<222> (4)...(4)

<223> Xaa = Tyr, His o Phe

35 <221> MOD_RES

<222> (6)...(6)

<223> Xaa = Glu, Gln o Lys

40 <221> MOD_RES

<222> (7)...(7)

<223> Xaa = Ile, Leu o Tyr

45 <221> MOD_RES

<222> (8)...(8)

<223> Xaa = Gln, Thr, Met, Gly o Leu

50 <221> MOD RES

<222> (11)...(11)

<223> Xaa = Lys o Gln

55 <221> MOD RES

<222> (12)...(12)

<223> Xaa = Ser o Asn

60 <221> MOD RES

<222> (15)...(15)

<223> Xaa = Ile o Val

65 <221> MOD RES

<222> (16)...(16)

<223> Xaa = Glu o Asp

<400> 25

60

ES 2 403 591 T3

Xaa Xaa Xaa Xaa Arg Xaa Xaa Xaa Lys Leu Xaa Xaa Thr Tyr Xaa Xaa
 1 5 10 15

<210> 26
 <211> 13
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable mutante, motivo de dominio polimerasa

10 <221> MOD_RES
 <222> (7)...(7)
 <223> Xaa = Ser o Thr

15 <221> MOD_RES
 <222> (8)...(8)
 <223> Xaa = Asp, Glu o Asn

20 <400> 26

Thr Gly Arg Leu Ser Ser Xaa Xaa Pro Asn Leu Gln Asn
 1 5 10

<210> 27
 <211> 15
 25 <212> PRT
 <213> Secuencia artificial

<220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable mutante, motivo de dominio polimerasa

30 <221> MOD_RES
 <222> (1)...(1)
 <223> Xaa = Gly, Asn o Asp

35 <221> MOD_RES
 <222> (2)...(2)
 <223> Xaa = Trp o His

40 <221> MOD_RES
 <222> (3)...(3)
 <223> Xaa = Trp, Ala, Leu o Val

45 <221> MOD_RES
 <222> (4)...(4)
 <223> Xaa = Ile o Leu

50 <221> MOD_RES
 <222> (5)...(5)
 <223> Xaa = Val, Phe o Leu

55 <221> MOD_RES
 <222> (6)...(6)
 <223> Xaa = Ser, Ala, Val o Gly

60 <221> MOD RES
 <222> (7)...(7)
 <223> Xaa = Ala o Leu

<400> 27

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Tyr Ser Gln Ile Glu Leu Arg
 1 5 10 15

ES 2 403 591 T3

- <210> 28
<211> 30
<212> ADN
5 <213> Secuencia artificial
- <220>
<223> cebador de oligonucleótido molde para cribado de ensayo de extensión de M13mp18 (ADN de M13)
- 10 <400> 28
gggaagggcg atcgggtagcgc gcctcttcgc 30
- <210> 29
<211> 21
15 <212> PRT
<213> Secuencia artificial
- <220>
<223> región del dominio polimerasa de forma no modificada de ADN polimerasa termoestable, motivo de dominio polimerasa
- 20 <221> MOD_RES
<222> (3)...(3)
<223> Xaa = His, Glu o Gln
- 25 <221> MOD_RES
<222> (4)...(4)
<223> Xaa = cualquier aminoácido
- 30 <221> MOD_RES
<222> (6)...(6)
<223> Xaa = Asn o His
- 35 <221> MOD_RES
<222> (7)...(7)
<223> Xaa = Leu o Ile
- <221> MOD_RES
<222> (8)...(8)
40 <223> Xaa = cualquier aminoácido
- <221> MOD_RES
<222> (10)...(10)
<223> Xaa = cualquier aminoácido
- 45 <221> MOD_RES
<222> (11)...(11)
<223> Xaa = Asp, Lys o Thr
- 50 <221> MOD_RES
<222> (13)...(13)
<223> Xaa = Leu o Val
- <221> MOD_RES
<222> (14)...(15)
<223> Xaa = cualquier aminoácido
- 55 <221> MOD_RES
<222> (16)...(16)
<223> Xaa = Val, Ile o Leu
- 60 <221> MOD_RES
<222> (18)...(18)
<223> Xaa = Phe o Tyr

ES 2 403 591 T3

<221> MOD_RES
 <222> (19)...(19)
 <223> Xaa = Asp o Glu
 5
 <221> MOD_RES
 <222> (20)...(20)
 <223> Xaa = Lys o Glu
 10
 <221> MOD_RES
 <222> (21)...(21)
 <223> Xaa = Leu or Gln
 <400> 29
 15
 Ala Gly Xaa Xaa Phe Xaa Xaa Xaa Ser Xaa Xaa Gln Xaa Xaa Xaa Xaa
 1 5 10 15
 Leu Xaa Xaa Xaa Xaa
 20
 <210> 30
 <211> 8
 20 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> motivo A conservado de sitio activo de ADN polimerasa
 25
 <400> 30
 Asp Tyr Ser Gln Ile Glu Leu Arg
 1 5
 30 <210> 31
 <211> 26
 <212> PRT
 <213> Secuencia artificial
 35 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable quimérica CS5
 <400> 31
 Tyr Arg Ile Ala Gly Glu Pro Phe Asn Ile Asn Ser Pro Lys Gln Val
 1 5 10 15
 Ser Arg Ile Leu Phe Glu Lys Leu Gly Ile
 40 20 25
 <210> 32
 <211> 26
 <212> PRT
 45 <213> Secuencia artificial
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable quimérica CS6
 50 <400> 32
 Tyr Arg Ile Ala Gly Glu Pro Phe Asn Ile Asn Ser Pro Lys Gln Val
 1 5 10 15
 Ser Arg Ile Leu Phe Glu Lys Leu Gly Ile
 20 25

<210> 33
 <211> 21
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> región del dominio polimerasa de ADN polimerasa termoestable mejorada modificada, motivo de dominio polimerasa
 10 <221> MOD_RES
 <222> (3)...(3)
 <223> Xaa = His, Glu o Gln
 15 <221> MOD_RES
 <222> (4)...(4)
 <223> Xaa = cualquier aminoácido
 20 <221> MOD_RES
 <222> (6)...(6)
 <223> Xaa = Asn o His
 25 <221> MOD_RES
 <222> (7)...(7)
 <223> Xaa = Leu o Ile
 30 <221> MOD_RES
 <222> (8)...(8)
 <223> Xaa = cualquier aminoácido
 35 <221> MOD_RES
 <222> (10)...(10)
 <223> Xaa = cualquier aminoácido
 40 <221> MOD_RES
 <222> (11)...(11)
 <223> Xaa = Asp, Lys o Thr
 45 <221> MOD_RES
 <222> (13)...(13)
 <223> Xaa = Leu o Val
 50 <221> MOD_RES
 <222> (14)...(15)
 <223> Xaa = cualquier aminoácido
 55 <221> MOD_RES
 <222> (16)...(16)
 <223> Xaa = Val, Ile o Leu
 60 <221> MOD_RES
 <222> (18)...(18)
 <223> Xaa = Phe o Tyr
 65 <221> MOD_RES
 <222> (20)...(20)
 <223> Xaa = Lys o Glu
 70 <221> MOD_RES
 <222> (21)...(21)
 <223> Xaa = Leu o Gln
 <400> 33

ES 2 403 591 T3

Ala Gly Xaa Xaa Phe Xaa Xaa Xaa Ser Xaa Xaa Gln Xaa Xaa Xaa Xaa
1 5 10 15
Leu Xaa Gly Xaa Xaa
20

REIVINDICACIONES

1. ADN polimerasa, que comprende A-G-X₁-X₂-F-X₃-X₄-X₅-S-X₆-X₇-Q-X₈- X₉- X₁₀-X₁₁-L-X₁₂-X₁₃-X₁₄X₁₅, en la que:

- 5 X₁ es E,
X₂ es P,
X₃ es N,
X₄ es I,
X₅ es N,
- 10 X₆ es P,
X₇ es K,
X₈ es V,
X₉ es S,
X₁₀ es R,
- 15 X₁₁ es I,
X₁₂ es F,
X₁₃ es G,
X₁₄ es K, y
X₁₅ es L,

20 en la que la polimerasa presenta una tasa mejorada de extensión de ácidos nucleicos en comparación con una ADN polimerasa de otro modo idéntica con la excepción de que X₁₃ es E.

2. ADN polimerasa según la reivindicación 1, en la que la polimerasa comprende una polimerasa quimérica.

25 3. ADN polimerasa según la reivindicación 2, en la que la polimerasa quimérica presenta una identidad de secuencia de por lo menos 90% respecto a una ADN polimerasa CS5 (SEC ID n° 20) o la ADN polimerasa CS6 (SEC ID n° 21).

30 4. ADN polimerasa según la reivindicación 4, en la que la polimerasa quimérica comprende SEC ID n° 20 ó SEC ID n° 21 ó una o más sustituciones de aminoácidos respecto a SEC ID n° 20 ó SEC ID n° 21, las cuales se seleccionan de entre el grupo que consiste de: G46E, L329A y E678G, e incluye un cambio E558G respecto a SEC ID n° 20 ó SEC ID n° 21.

35 5. ADN polimerasa según la reivindicación 1, que comprende además una modificación covalente térmicamente reversible.

6. Ácido nucleico recombinante que codifica la ADN polimerasa según la reivindicación 1.

40 7. Vector de expresión que comprende el ácido nucleico recombinante según la reivindicación 6.

8. Célula huésped que comprende el vector de expresión según la reivindicación 7.

9. Método para producir una ADN polimerasa, comprendiendo dicho método:

45 cultivar la célula huésped según la reivindicación 8 bajo condiciones adecuadas para la expresión del ácido nucleico codificante de la ADN polimerasa mutante.

50 10. Método para llevar a cabo la extensión de cebadores, que comprende:

poner en contacto una ADN polimerasa según la reivindicación 1 con un cebador, un polinucleótido molde y nucleótidos libres bajo condiciones adecuadas para la extensión del cebador, produciendo de esta manera un cebador extendido.

55 11. Método según la reivindicación 10, en el que el polinucleótido molde es un ARN o un ADN.

12. Método según la reivindicación 10, en el que los nucleótidos libres comprenden nucleótidos no convencionales, comprendiendo estos últimos ribonucleótidos o nucleótidos marcados.

60 13. Método según la reivindicación 10, que comprende poner en contacto la ADN polimerasa con una pareja de cebadores, el polinucleótido molde y los nucleótidos libres bajo condiciones adecuadas para la amplificación del polinucleótido.

14. Kit para producir un cebador extendido, que comprende:

por lo menos un recipiente que proporciona una ADN polimerasa según la reivindicación 1 y uno o más recipientes adicionales seleccionados de entre el grupo que consiste de:

5

(a) un recipiente que proporciona un cebador hibridable, bajo condiciones de extensión de cebadores, con un polinucleótido molde predeterminado,

(b) un recipiente que proporciona nucleótidos libres, y

(c) un recipiente que proporciona un tampón adecuado para la extensión de cebadores.

10

*		
Tth	FRLAGHPFNLNSRDQLERVLFDELRL	(SEC ID n° 3)
Tca	FRLAGHPFNLNSRDQLERVLFDELRL	(SEC ID n° 4)
Z05	FRLAGHPFNLNSRDQLERVLFDELRL	(SEC ID n° 5)
Tag	FRLAGHPFNLNSRDQLERVLFDELGL	(SEC ID n° 6)
Tfl	FRLAGHPFNLNSRDQLERVLFDELGL	(SEC ID n° 7)
Tfi	HRLAGHPFNLNSRDQLERVLFDELGL	(SEC ID n° 8)
Sps17	HRLAGHPFNLNSRDQLERVLFDELGL	(SEC ID n° 9)
Dra	HEYAGEEFHIRSPKQLETVLYDKLEL	(SEC ID n° 10)
HspB7	YTLAGEAFNIGSPKQLGAILFEKLGL	(SEC ID n° 11)
Bst	YELAGQEFNINSPKQLGTVLFDKLQL	(SEC ID n° 12)
Bca	YELAGQEFNINSPKQLGVILFEKLQL	(SEC ID n° 13)
Eco	HEIAGEEFNLSSTKQLQITILFEKQGI	(SEC ID n° 14)
Tma	YRIAGEPFNINSPKQVSRILFEKLG I	(SEC ID n° 15)
Tne	YQIAGEPFNINSPKQVSNILFEKLG I	(SEC ID n° 16)
Taf	FEIAGETFNLSSTQVAYILFEKLNI	(SEC ID n° 17)
HspA	YAQAGEVFNLNSPKQLGSLLFEKLKL	(SEC ID n° 18)
CS5	YRIAGEPFNINSPKQVSRILFEKLG I	(SEC ID n° 31)
CS6	YRIAGEPFNINSPKQVSRILFEKLG I	(SEC ID n° 32)
T7	VEHVV. . FNPSSRDHIQKKLQEAGWV	(SEC ID n° 19)
Cons	---AG--FN--S--Q----LF--L--	(SEC ID n° 24)

Figura 1

1 MKAMPLPLFEP KGRVLLVDGH HLAYRTFFAL KGLTTSRGEP VQAVYGFSAKS
 51 LLKALKEDGY KAVFVVFDAK APSFRHEAYE AYKAGRAPT EDPFRQLALI
 101 KELVDLLGFT RLEVPGFEAD DVLATLAKKA EREGYEVRI TADRDLYQLV
 151 SDRVAVLHPE GHLITPEWLW EKYGLKPEQW VDFRALVGDV SDNLPGVKGI
 201 GEKTALKLLK EWGSLENILK NLDKVKPESV RERIKAHLED LKLSLELSRV
 251 RSDLPLEVDF ARREPDREG LRAFLERLEF GSSLHEFGLL ESEEPVGYRI
 301 VKDLVEFEKL IEKLRESPTF AIDLETSSLD PFDKDIVGIS VSEKPKAEAYY
 351 IPLHHRNAQN LDEKEVLKKL KEILEDPGAK IVGQNLKFDY KVLNVKGVPEP
 401 VPPYFDTMIA AYLLEPNEKK FNLDLALKE LGYKMTSYQE LMSFSFPLFG
 451 FSEADVPVEK AANYSCEDAD ITYRLYKTLK LKLHEADLEN VFYKIEMPLV
 501 NVLARMELNG VYVDTEFLKK LSEYGGKLE ELAEEIYRIA GEPFNINSPK
 551 QVSRIKFEKL GIKFRGKTTK TGDYSTRIEV LEELAGEHEI IPLILEYRKI
 601 QKPKSTYIDA LPKMNVNPKTG RIHASFNQTG TATGRLSSSD PNLQNLPTKS
 651 EEGKEIRKAI VPQDPNWWIV SADYSQIELR ILAHLSGDEN LLRAFEEGID
 701 VHTLTASRIF NVKPEEVTEE MRRAGKMNVE SIIYGVTPYG LSVRLGVVVK
 751 EAEKMIVNYF VLYPKVRDYI QRVVSEAKEK GYVRTLFGRK RDIPQLMARD
 801 RNTQAEGERI AINTPIQGTG ADIIKLAMIE IDRELKERKM RSKMIIQVHD
 851 ELVFEVPNEE KDALVELVKD RMTNVVKLSV PLEVDTVIGK TWS

Figura 2A

ES 2 403 591 T3

1 ATGAAAGCTA TGTTACCATT ATTCGAACCC AAAGGCCGGG TCCTCCTGGT
51 GGACGGCCAC CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA
101 CCACGAGCCG GGGCGAACCG GTGCAGGCGG TTTACGGCTT CGCCAAGAGC
151 CTCCTCAAGG CCCTGAAGGA GGACGGGTAC AAGGCCGTCT TCGTGGTCTT
201 TGACGCCAAG GCCCCTTCTT TCCGCCACGA GGCCTACGAG GCCTACAAGG
251 CAGGCCCGCG CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC
301 AAGGAGCTGG TGGACCTCCT GGGGTTTACT CGCCTCGAGG TTCCGGGCTT
351 TGAGGCGGAC GACGTCTCTG CCACCCTGGC CAAGAAGGCG GAAAGGGAGG
401 GGTACGAGGT GCGCATCTTC ACCGCCGACC GGGACCTTTA CCAGCTCGTC
451 TCCGACCGCG TCGCCGTCTT CCACCCCGAG GGCCACCTCA TCACCCCGGA
501 GTGGCTTTGG GAGAAGTACG GCCTTAAGCC GGAGCAGTGG GTGGACTTCC
551 GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC
601 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGAAAA
651 TATCCTCAAG AACCTGGACC GGGTGAAGCC GGAAAGCGTC CGGGAAAGGA
701 TCAAGGCCCA CCTGGAAGAC CTTAAGCTCT CCTTGGAGCT TTCCCCGGTG
751 CGCTCGGACC TCCCCCTGGA GGTGGACTTC GCCCGGAGGC GGGAGCCTGA
801 CCGGAAGGG CTTCGGCCTT TTTTGGAGCG CTTGGAGTTC GGCAGCCTCC
851 TCCACGAGTT CGGCCTTCTA GAGGAGTCCG AACCCGTTGG GTACCGTATA
901 GTTAAAGACC TGGTTGAATT TGAAAACTC ATAGAGAAAC TGAGAGAATC
951 TCCTTCGTTT GCTATCGATT TGGAAACTAG TTCCTCGAT CCTTTCGACT
1001 GCGACATTGT CGGTATCTCT GTGTCTTTCA AACCAAAGGA AGCGTACTAC
1051 ATACCACTCC ATCATAGAAA CGCCCAGAAC CTGGACGAAA AAGAGGTTCT
1101 GAAAAAGCTC AAAGAAATTC TGGAGGACCC CGGAGCAAAG ATCGTTGGTC
1151 AGAATTTGAA ATTCGATTAC AAGGTGTTGA TGGTGAAGGG TGTTGAACCT
1201 GTTCCTCCTT ACTTCGACAC GATGATAGCG GCTTACCTTC TTGAGCCGAA
1251 CGAAAAGAAG TTCAATCTGG ACGATCTCGC ATTGAAATTT CTTGGATACA
1301 AAATGACATC TTACCAAGAG CTCATGTCTT TCTCTTTTCC GCTGTTTGGT

Figura 2B-1

ES 2 403 591 T3

1351 TTCAGTTTTG CCGATGTTCC TGTAGAAAAA GCAGCGAACT ACTCCTGTGA
1401 AGATGCAGAC ATCACCTACA GACTTTACAA GACCCTGAGC TTAAAACCTCC
1451 ACGAGGCAGA TCTGGAAAAC GTGTTCTACA AGATAGAAAT GCCCCTGTG
1501 AACGTGCTTG CACGGATGGA ACTGAACGGT GTGTATGTGG ACACAGAGTT
1551 CCTGAAGAAA CTCTCAGAAG AGTACGGAAA AAAACTCGAA GAACTGGCAG
1601 AGGAAATATA CAGGATAGCT GGAGAGCCGT TCAACATAAA CTCACCGAAG
1651 CAGGTTTCAA GGATCCTTTT TGAAAACTC GGCATAAABC CACGTGGTAA
1701 AACGACGAAA ACGGGAGACT ATTCAACACG CATAGAAGTC CTCGAGGAAC
1751 TTGCCGGTGA ACACGAAATC ATTCCTCTGA TTCTTGAATA CAGAAAGATA
1801 CAGAAATTGA AATCAACCTA CATAGACGCT CTTCCCAAGA TGGTCAACCC
1851 AAAGACCGGA AGGATTCATG CTTCTTTCAA TCAAACGGGG ACTGCCACTG
1901 GAAGACTTAG CAGCAGCGAT CCCAATCTTC AGAACCTCCC GACGAAAAGT
1951 GAAGAGGGAA AAGAAATCAG GAAAGCGATA GTTCCTCAGG ATCCAAACTG
2001 GTGGATCGTC AGTGCCGACT ACTCCCAAAT AGAACTGAGG ATCCTCGCCC
2051 ATCTCAGTGG TGATGAGAAAT CTTTTGAGGG CATTGGAAGA GGGCATCGAC
2101 GTCCACACTC TAACAGCTTC CAGAATATTC AACGTGAAAC CCGAAGAAGT
2151 AACCGAAGAA ATGCGCCGCG CTGGTAAAAT GGTTAATTTT TCCATCATAT
2201 ACGGTGTAAC ACCTTACGGT CTGTCTGTGA GGCTTGGAGT ACCTGTGAAA
2251 GAAGCAGAAA AGATGATCGT CAACTACTTC GTCCTCTACC CAAAGGTGCG
2301 CGATTACATT CAGAGGGTCG TATCGGAAGC GAAAGAAAAA GGCTATGTTA
2351 GAACGCTGTT TGGAAGAAAA AGAGACATAC CACAGCTCAT GGCCCGGGAC
2401 AGGAACACAC AGGCTGAAGG AGAACGAATT GCCATAAACA CTCCATACA
2451 GGGTACAGCA GCGGATATAA TAAAGCTGGC TATGATAGAA ATAGACAGGG
2501 AACTGAAGA AAGAAAATG AGATCGAAGA TGATCATACA GGTCCACGAC
2551 GAACTGGTTT TTGAAGTGCC CAATGAGGAA AAGGACGCGC TCGTCGAGCT
2601 GGTGAAAGAC AGAATGACGA ATGTGGTAAA GCTTTCAGTG CCGCTCGAAG
2651 TGGATGTAAC CATCGGCAAA ACATGGTCGT GA

Figura 2B-2

1 MKAMPLPFEP KGRVLLVDGH HLAYRTFFAL KGLTTSRGEP VQAVYGFAXS
 51 LLKALKEDGY KAVFVVFDAK APSFRHEAYE AYKAGRAPTP EDFPRQLALI
 101 KELVDLLGFT RLEVPGFEAD DVLATLAKKA EREGYEVRII TADRDLYQLV
 151 SDRVAVLHPE GHLITPEWLW EKYGLKPEQW VDFRALVQDP SDNLPGVKGI
 201 GEKTKALKLK EWGSLENILK NLDVRKPEV RERIKAHLED LKLSLELSRV
 251 RSDLPLEVDF ARRREPDREG LRAFLELRF GSKLHEFGLL EEESEVGYRI
 301 VKDLVEFEKL IEKLRESFSF AIALATSSLD PFDCCIVGIS VSEKPKAEAY
 351 IPLHHRNAQN LDEKEVLKKL KEILEDPAK IVGQNLKFDY KVLNVKGVPEP
 401 VPPYFDTMIA AYLLEPNEKK FNLDLALKF LGYKMTSYQE LMSFSFPLFG
 451 FSFADVPVEK AANYSCEDAD ITYRLYKTLK LKLHEADLEN VFYKIEMPLV
 501 NVLARMELNG VYVDTEFLKK LSEYGGKLE ELAEEIYRIA GEPFNINSPK
 551 QVSRILFEKL GIKPRGKTK TGDYSTRIEV LEELAGEHEI IPLILEYRKI
 601 QKLKSTYIDA LPKMVNPKTG RIHASFNQTG TATGRLSSSD PNLQNLPTKS
 651 EEGKEIRKAI VPQDPNWWIV SADYSQIELR ILAHLSDEN LLRAFEEGID
 701 VHTLTASRIF NVKPEEVTEE MRRAGKMNK SIIYGVTPYG LSVRLGVPVK
 751 EAEKMIVNYF VLYPKVRDYI QRVVSEAKEK GYVRTLFGRK RDIPQLMARD
 801 RNTQAEGERI AINTPIQTA ADIKLAMIE IDRELKERKM RSKMIIQVHD
 852 ELVFEVPNEE KDALVELVKD RMTNVVKLSV PLEVDTIGK TWS

Figura 3A

1 ATGAAAGCTA TGTTACCATT ATTCGAACCC AAAGGCCGGG TCCTCCTGGT
 51 GGACGGCCAC CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA
 101 CCACGAGCCG GGGCGAACCG GTGCAGGCGG TTTACGGCTT CGCCAAGAGC
 151 CTCCTCAAGG COCTGAAGGA GGACGGGTAC AAGGCCGTCT TCGTGGTCTT
 201 TGACGCCAAG GCCCCTTCCT TCCGCCACGA GGCCTACGAG GCCTACAAGG
 251 CAGGCCGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC
 301 AAGGAGCTGG TGGACCTCCT GGGGTTTACT CGCCTCGAGG TTCGGGGCTT
 351 TGAGGCGGAC GACGTCTCTG CCACCCTGGC CAAGAAGGCG GAAAGGGAGG
 401 GGTACGAGGT GGCATCCTC ACCGCCGACC GGGACCTTTA CCAGCTCGTC
 451 TCCGACCGCG TCGCGTCTCT CCACCCGAG GGCCACCTCA TCACCCCGGA
 501 GTGGCTTTGG GAGAAGTACG GCCTTAAGCC GGAGCAGTGG GTGGACTTCC
 551 GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCC0GGGGT CAAGGGCATC
 601 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGAAAA
 651 TATCTCAAG AACCTGGACC GGGTGAAGCC GGAAAGCGTC CGGGAAAGGA
 701 TCAAGGCCCA CCTGGAAGAC CTTAAGCTCT CCTTGGAGCT TTC0CGGGTG
 751 CGCTCGGACC TCCCCCTGGA GGTGGACTTC GCCCGGAGGC GGGAGCCTGA
 801 CCGGAAGGG CTTCGGGCTT TTTTGGAGCG CTTGGAGTTC GGCAGCCTCC
 851 TCCACGAGTT CGGCCTTCTA GAGGAGTCCG AACCCGTGG GTACCGTATA
 901 GTTAAAGACC TGGTTGAATT TGAAAACTC ATAGAGAAAC TGAGAGAATC
 951 TCCTTCGTTT GCGATCGCTC TTGCGACTAG TTCCCTCGAT CCTTTCGACT
 1001 GCGACATTGT CGGTATCFCT GTGTCTTTCA AACCAAAGGA AGCGTACTAC
 1051 ATACCACTCC ATCATAGAAA CGCCGAGAAC CTGGACGAAA AAGAGGTTCT
 1101 GAAAAAGCTC AAAGAAATTC TGGAGGACCC CGGAGCAAAG ATCGTTGGTC
 1151 AGAATTTGAA ATTCGATTAC AAGGTGTTGA TGGTGAAGGG TGTGAACCT
 1201 GTTCCTCCTT ACTTCGACAC GATGATAGCG GCTTACCTTC TTGAGCCGAA
 1251 CGAAAAGAAG TTCAATCTGG ACGATCTCGC ATTGAAATTT CTGGATACA
 1301 AAATGACATC TTACCAAGAG CTCATGTCCT TCTCTTTTCC GCTGTTTGGT

Figura 3B-1

1351 TTCAGTTTTG CCGATGTTCC TGTAGAAAAA GCAGCGAACT ACTCCTGTGA
 1401 AGATGCAGAC ATCACCTACA GACTTTACAA GACCCTGAGC TTAAAACTCC
 1451 ACGAGGCAGA TCTGGAAAAC GTGTCTTACA AGATAGAAAT GCCCCTTGTG
 1501 AACGTGCTTG CACGGATGGA ACTGAACGGT GTGTATGTGG ACACAGAGTT
 1551 CCTGAAGAAA CTCTCAGAAG AGTACGGAAA AAAACTCGAA GAACTGGCAG
 1601 AGGAAATATA CAGGATAGCT GGAGAGCCGT TCAACATAAA CTCACCGAAG
 1651 CAGGTTTCAA GGATCCTTTT TGAAAACTC GGCATAAAAC CACGTGGTAA
 1701 AACGACGAAA ACGGGAGACT ATTCAACACG CATAGAAGTC CTCGAGGAAC
 1751 TTGCCGGTGA ACACGAAATC ATTCCTCTGA TTCTTGAATA CAGAAAGATA
 1801 CAGAAATTGA AATCAACCTA CATAGACGCT CTTCCCAAGA TGGTCAACCC
 1851 AAAGACCGGA AGGATTCATG CTTCTTTCAA TCAAACGGGG ACTGCCACTG
 1901 GAAGACTTAG CAGCAGCGAT CCAATCTTC AGAACCTCCC GACGAAAAGT
 1951 GAAGAGGGAA AAGAAATCAG GAAAGCGATA GTTCCTCAGG ATCCAAACTG
 2001 GTGGATCGTC AGTGCCGACT ACTCCCAAAT AGAACTGAGG ATCCTCGCCC
 2051 ATCTCAGTGG TGATGAGAAT CTTTGAGGG CATTCGAAGA GGCATCGAC
 2101 GTCCACACTC TAACAGCTTC CAGAATATTC AACGTGAAAC CCGAAGAAGT
 2151 AACCGAAGAA ATGCGCCGCG CTGGTAAAAT GGTAAATTTT TCCATCATAT
 2201 ACGGTGTAAC ACCTFACGGT CTGTCTGTGA GGCTTGAGT ACCTGTGAAA
 2251 GAAGCAGAAA AGATGATCGT CAACTACTTC GTCCTCTACC CAAAGGTGCG
 2301 CGATTACATT CAGAGGGTCG TATCGGAAGC GAAAGAAAAA GGCTATGTTA
 2351 GAACGCTGTT TGGAAGAAAA AGAGACATAC CACAGCTCAT GGCCCGGGAC
 2401 AGGAACACAC AGGCTGAAGG AGAACGAATT GCCATAACA CTCCCATACA
 2451 GGGTACAGCA GCGGATATAA TAAAGCTGGC TATGATAGAA ATAGACAGGG
 2501 AACTGAAAGA AAGAAAAATG AGATCGAAGA TGATCATACA GGTCCACGAC
 2551 GAACTGGTTT TTGAAGTGCC CAATGAGGAA AAGGACGCGC TCGTCGAGCT
 2601 GGTGAAAGAC AGAATGACGA ATGTGGTAAA GCTTTCAGTG CCGCTCGAAG
 2651 TGGATGTAAC CATCGGCAAA ACATGGTCGT GA

Figura 3B-2

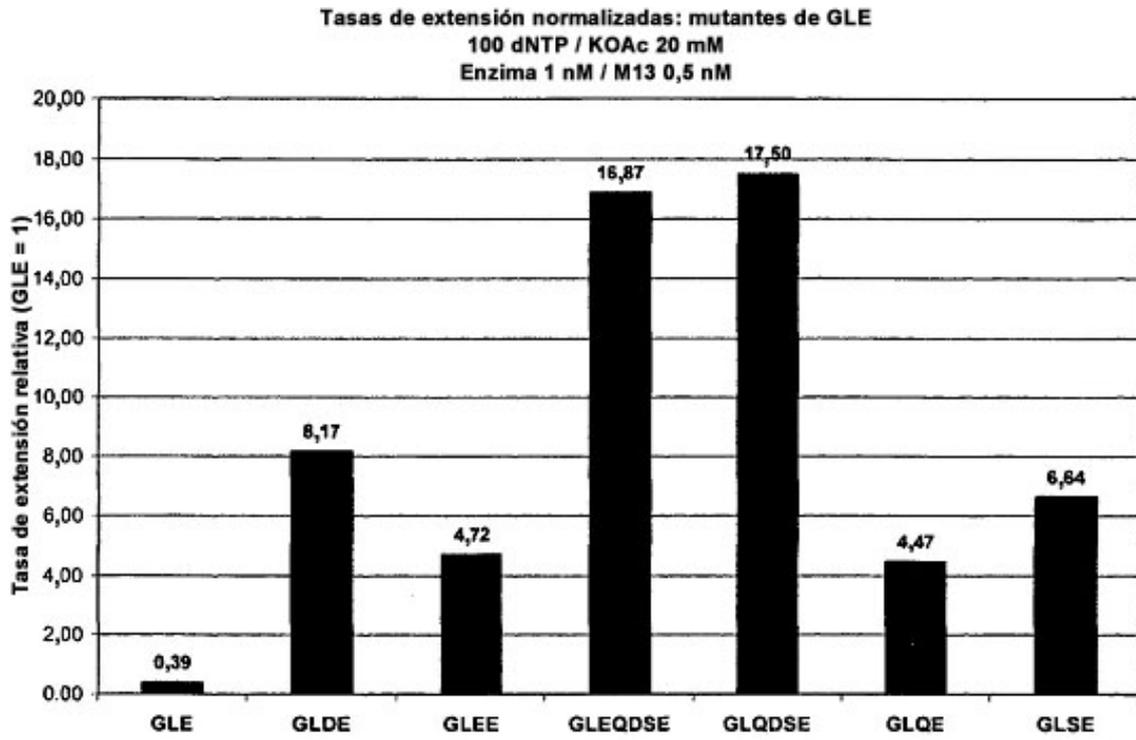


Figura 4