

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 624 406**

51 Int. Cl.:

C12N 9/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.12.2013 PCT/EP2013/076147**

87 Fecha y número de publicación internacional: **19.06.2014 WO14090836**

96 Fecha de presentación y número de la solicitud europea: **11.12.2013 E 13818699 (4)**

97 Fecha y número de publicación de la concesión europea: **08.03.2017 EP 2931884**

54 Título: **ADN polimerasas con actividad mejorada**

30 Prioridad:

13.12.2012 US 201261736737 P

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

14.07.2017

73 Titular/es:

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

72 Inventor/es:

SUKO, SHAWN

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 624 406 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 con actividad mejorada

5 **Campo de la invención**

La presente invención proporciona ADN polimerasas con actividades mejoradas, incluyendo eficacia de transcriptasa inversa incrementada, así como el uso de dichas polimerasas en diversas aplicaciones, incluyendo la extensión y amplificación de polinucleótidos de ácidos nucleicos.

10

Antecedentes de la invención

Las ADN polimerasas son responsables de la replicación y mantenimiento del genoma, un papel que es fundamental para transmitir con precisión la información genética de generación a generación. Las ADN polimerasas funcionan en las células como las enzimas responsables de la síntesis del ADN. Polimerizan los desoxirribonucleósidos trifosfato en presencia de un activador de metal, tal como Mg^{2+} , en un orden dictado por el molde de ADN o molde de polinucleótidos que se copia. *In vivo*, las ADN polimerasas participan en un espectro de procesos de síntesis del ADN, incluyendo la replicación del ADN, reparación del ADN, recombinación y amplificación génica. Durante cada proceso de síntesis del ADN, el molde de ADN se copia una vez o, como máximo, unas pocas veces para producir réplicas idénticas. Por el contrario, *in vitro*, la replicación del ADN se puede repetir muchas veces, tal como, por ejemplo, durante la reacción en cadena de la polimerasa (véase, por ejemplo, la patente de EE. UU. n.º 4.683.202).

En los estudios iniciales con la reacción en cadena de la polimerasa (PCR), se añadió la ADN polimerasa al inicio de cada tanda de replicación del ADN (véase la patente de EE. UU. n.º 4.683.202, *supra*). Posteriormente, se determinó que se podían obtener ADN polimerasas termoestables a partir de bacterias que crecen a temperaturas elevadas y que únicamente se necesita añadir estas enzimas una vez (véase la patente de EE. UU. n.º 4.889.818 y la patente de EE. UU. n.º 4.965.188). A las temperaturas elevadas usadas durante la PCR, estas enzimas no se inactivan irreversiblemente. Como resultado, se pueden llevar a cabo ciclos repetitivos de reacciones en cadena de la polimerasa sin añadir nuevas enzimas al inicio de cada proceso de adición de síntesis. Las ADN polimerasas, particularmente las polimerasas termoestables, son la clave de un gran número de técnicas en estudios de ADN recombinante y en el diagnóstico médico de la enfermedad. Para aplicaciones diagnósticas, en particular, una secuencia de ácidos nucleicos diana únicamente puede ser una pequeña porción del ADN o ARN en cuestión, de modo que puede ser difícil detectar la presencia de una secuencia de ácidos nucleicos diana sin amplificación.

El patrón de plegamiento global de las ADN polimerasas se asemeja a la mano derecha humana y contiene tres subdominios distintos de la palma, dedos y pulgar. (Véase Beese *et al.*, *Science* 260:352-355, 1993); Patel *et al.*, *Biochemistry* 34:5351-5363, 1995). Mientras que la estructura de los subdominios de dedos y pulgar varían en gran medida entre polimerasas que difieren en tamaño y en funciones celulares, todos los subdominios catalíticos de la palma se pueden superponer. Por ejemplo, el motivo A, que interacciona con el dNTP de entrada y estabiliza el estado de transición durante la catálisis química, se puede superponer con una desviación media de aproximadamente un Å entre las ADN polimerasas de la familia pol α de mamíferos y pol I de procariotas (Wang *et al.*, *Cell* 89:1087-1099, 1997). El motivo A comienza estructuralmente en una cadena β antiparalela que contiene predominantemente residuos hidrófobos y continúa en una hélice α. La secuencia de aminoácidos primaria de los sitios activos de la ADN polimerasa está excepcionalmente conservada. Por ejemplo, en el caso del motivo A, se retiene la secuencia DYSQIELR (SEQ ID NO:22) en polimerasas de organismos separados por muchos millones de años de evolución, incluyendo, por ejemplo, *Thermus aquaticus*, *Chlamydia trachomatis* y *Escherichia coli*.

Además de estar bien conservado, también se ha demostrado que el sitio activo de las ADN polimerasas es relativamente mutable, puede acomodar determinadas sustituciones de aminoácidos sin reducir significativamente la actividad de la ADN polimerasa (véase, por ejemplo, la patente de EE. UU. n.º 6.602.695). Dichas ADN polimerasas mutantes pueden ofrecer diversas ventajas selectivas, por ejemplo, en aplicaciones diagnósticas y de investigación que comprenden reacciones de síntesis de ácidos nucleicos.

Existen al menos dos etapas en el proceso enzimático de la polimerización del ADN; 1) la incorporación del nucleótido de entrada y 2) la extensión del nucleótido recién incorporado. Generalmente se considera la fidelidad global o "fidelidad" de la ADN polimerasa como un conglomerado de estas dos actividades enzimáticas, pero las etapas son distintas. Una ADN polimerasa puede incorporar erróneamente el nucleótido de entrada, pero, si no se extiende eficazmente, la velocidad de extensión disminuirá gravemente y la formación del producto global será mínima. De manera alternativa, es posible tener una ADN polimerasa que incorpore erróneamente el nucleótido de entrada y fácilmente extienda erróneamente el emparejamiento erróneo recién formado. En este caso, la velocidad de extensión global sería alta, pero la fidelidad global sería baja. Un ejemplo de este tipo de enzima sería la ADN polimerasa ES112 (ADN polimerasa E683R Z05, véase el documento US 7.179.590) cuando se usa Mn^{2+} como el activador de ion de metal divalente. La enzima tiene una eficacia muy alta porque, a diferencia de las ADN polimerasas típicas que tienden a vacilar/bloquearse al encontrar un emparejamiento erróneo, la ADN polimerasa ES112 extiende fácilmente el emparejamiento erróneo. El fenotipo mostrado en ES112 es más pronunciado durante la etapa de RT, supuestamente a causa de los efectos estructurales del heterodúplex ARN/ADN frente al

homodúplex ADN/ADN. Un segundo ejemplo sería si la ADN polimerasa fácilmente no incorpora erróneamente (puede ser incluso menos probable incorporar erróneamente), pero sí tiene capacidad incrementada para extender erróneamente un emparejamiento erróneo. En este caso, no se altera significativamente la fidelidad para el producto global. En general, este tipo de enzima es más favorable para las reacciones de extensión que las características de ES112 en Mn^{2+} porque se mejora la fidelidad del producto. Sin embargo, se puede utilizar este atributo para permitir la extensión errónea de un cebador oligonucleotídico emparejado erróneamente, tal como cuando un cebador oligonucleotídico de una secuencia única se hibrida a una diana que tiene heterogeneidad de secuencia (por ejemplo, dianas víricas), pero la velocidad de incorporación errónea normal o más baja permite la finalización de la síntesis del ADN más allá del cebador oligonucleotídico original. Un ejemplo de este tipo de ADN polimerasa es la ADN polimerasa Z05 D580G (véase la publicación de patente de EE. UU. n.º 2009/0148891). Este tipo de actividad se denomina "tolerante al emparejamiento erróneo" a causa de que es más tolerante a los emparejamientos erróneos en el cebador oligonucleotídico. Mientras que los ejemplos anteriores han analizado las reacciones de tipo de extensión del cebador, la actividad puede ser más significativa en reacciones tales como RT-PCR y PCR donde con frecuencia se produce nuevamente la extensión del cebador. Los datos sugieren que mientras las enzimas, tales como Z05 D580G, son más "tolerantes" a los emparejamientos erróneos, también tienen capacidad potenciada para extender los cebadores oligonucleotídicos que contienen bases modificadas (por ejemplo, bases modificadas con t-butilbencilo) o en presencia de tintes de unión a ADN, tales como SYBR Green I (véase la publicación de patente n.º 2009/028053).

La reacción en cadena de la polimerasa con transcripción inversa (RT-PCR) es una técnica usada en muchas aplicaciones para detectar y/o cuantificar dianas de ARN mediante amplificación. A fin de amplificar dianas de ARN mediante PCR, en primer lugar es necesario transcribir de manera inversa el molde de ARN en ADNc. Típicamente, los ensayos por RT-PCR dependen de una transcriptasa inversa no termoestable (ADN polimerasa dependiente de ARN), derivada de un organismo mesófilo, para la etapa de síntesis ADNc inicial (RT). Se requiere una ADN polimerasa termoestable adicional para la amplificación del ADNc para tolerar las temperaturas elevadas requeridas para la desnaturalización del ácido nucleico en la PCR. Existen varios beneficios potenciales de usar ADN polimerasas termoactivas o termoestables generadas para realizar una transcripción inversa más eficaz para los ensayos por RT-PCR. La actividad de transcriptasa inversa incrementada acoplada con la capacidad de usar temperaturas de incubación de transcripción inversa más altas que permiten la relajación de la estructura secundaria del molde de ARN puede dar como resultado sensibilidad del ensayo y eficacia de síntesis del ADNc global más altas. Una temperatura de incubación más alta también podría incrementar la especificidad reduciendo el falso cebado en la etapa de transcripción inversa. Las enzimas con eficacia de transcripción inversa mejorada pueden simplificar el diseño del ensayo permitiendo una concentración enzimática y/o tiempos de incubación de RT reducidos. Cuando se usan dUTP y UNG, los productos de extensión no específicos que contienen dUMP que se forman durante las condiciones establecidas no restrictivas se degradan mediante UNG y no se pueden utilizar ni como cebadores ni como moldes. Cuando se usa una transcriptasa inversa no termoestable (ADN polimerasa dependiente de ARN) derivada de un organismo mesófilo, no es posible utilizar las metodologías de dUTP y UNG. (Myers, T.W. *et al.*, Amplification of RNA: High Temperature Reverse Transcription and DNA Amplification with *Thermus thermophilus* DNA Polymerase, in PCR Strategies, Innis, M.A., Gelfand, D.H., y Sninsky, J.J., Eds., Academic Press, San Diego, CA, 58-68, (1995)). Sin embargo, el uso de una ADN polimerasa termoactiva o termoestable de la invención para la etapa de transcripción inversa posibilita que la reacción sea completamente compatible con la utilización del sistema de prevención de transferencia de dUTP/uracil-N-glicosilasa (UNG) (Longo *et al.*, Use of Uracil DNA Glycosylase to Control Carry-over Contamination in Polymerase Chain Reactions. *Gene* 93:125-128, (1990). Además de proporcionar un control de la contaminación por transferencia, el uso de dUTP y UNG proporciona un "inicio en caliente" para reducir la amplificación no específica (Innis y Gelfand 1999).

Breve resumen de la invención

En el presente documento se proporcionan ADN polimerasas que tienen actividades mejoradas, incluyendo eficacia de transcriptasa inversa incrementada en relación con una polimerasa de control no modificada correspondiente, y procedimientos para preparar y usar dichas ADN polimerasas. En algunos modos de realización, la ADN polimerasa mejorada tiene eficacia de transcriptasa inversa incrementada en comparación con una ADN polimerasa de control. En algunos modos de realización, la ADN polimerasa mejorada tiene la misma actividad de polimerasa dependiente de ADN o sustancialmente similar en comparación con una ADN polimerasa de control.

En un aspecto, la invención se refiere a una ADN polimerasa que tiene eficacia de transcriptasa inversa incrementada en comparación con una ADN polimerasa de control, en la que la ADN polimerasa comprende una secuencia de aminoácidos al menos un 90 % idéntica a la SEQ ID NO: 1 y en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es T, el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es K y el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es G y en la que la ADN polimerasa de control tiene la misma secuencia de aminoácidos que la ADN polimerasa excepto en que el aminoácido de la ADN polimerasa de control correspondiente a la posición 763 de la SEQ ID NO: 1 es M. También se divulga que la ADN polimerasa mejorada comprende una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, al menos aproximadamente un 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 %) a la SEQ ID NO: 1, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es cualquier aminoácido distinto de M o I.

También se divulga que el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es cualquier aminoácido distinto de D. También se divulga que el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 se selecciona del grupo que consiste en L, G, T, Q, A, S, N, R y K. También se divulga que el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I. También se divulga que el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 se selecciona del grupo que consiste en K, R, S, G y A. También se divulga que el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es G y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es K. En algunos modos de realización, la ADN polimerasa tiene la misma actividad de polimerasa dependiente de ADN o sustancialmente similar en comparación con la ADN polimerasa de control. También se divulga que la ADN polimerasa tiene eficacia de transcriptasa inversa incrementada sin una disminución sustancial de la actividad de polimerasa dependiente de ADN en comparación con una ADN polimerasa de control, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es cualquier aminoácido distinto de M o I, y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I y en la que la ADN polimerasa de control tiene la misma secuencia de aminoácidos que la ADN polimerasa excepto en que el aminoácido de la ADN polimerasa de control correspondiente a la posición 763 de la SEQ ID NO: 1 es M o I y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es I.

También se divulga que la ADN polimerasa mejorada comprende una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, al menos aproximadamente un 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 % idéntica) a la SEQ ID NO: 1, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 580 de la SEQ ID NO: 1 es cualquier aminoácido distinto de D o E. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 580 de la SEQ ID NO: 1 es cualquier aminoácido distinto de D. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 580 de la SEQ ID NO: 1 se selecciona del grupo que consiste en L, G, T, Q, A, S, N, R y K. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 580 de la SEQ ID NO: 1 es G.

También se divulga que la ADN polimerasa mejorada comprende una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, al menos aproximadamente un 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 %) a la SEQ ID NO: 1, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 709 de la SEQ ID NO: 1 se selecciona del grupo que consiste en K, R, S, G y A. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 709 de la SEQ ID NO: 1 es K.

También se divulga que la ADN polimerasa mejorada comprende una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, al menos aproximadamente un 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 % idéntica) a la SEQ ID NO: 1, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es cualquier aminoácido distinto de M o I, el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es cualquier aminoácido distinto de D y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I.

También se divulga que la ADN polimerasa mejorada tiene eficacia de transcriptasa inversa incrementada, opcionalmente sin una disminución sustancial en la actividad de polimerasa dependiente de ADN, en comparación con una ADN polimerasa de control, en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es cualquier aminoácido distinto de M o I, y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I, y en la que la ADN polimerasa de control tiene la misma secuencia de aminoácidos que la ADN polimerasa excepto en que el aminoácido de la ADN polimerasa de control correspondiente a la posición 763 de la SEQ ID NO: 1 es M o I, y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es I. De esta manera, el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es T y el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es K. También se divulga que la ADN polimerasa mejorada comprende adicionalmente una sustitución de aminoácido en el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1. De esta manera, el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es cualquier aminoácido distinto de M o I, el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es cualquier aminoácido distinto de I y el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es cualquier aminoácido distinto de D o E. También se divulga que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es T, el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es K y el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es G.

Diversas ADN polimerasas son susceptibles de mutación de acuerdo con la presente invención. Son particularmente

adecuadas las polimerasas termoestables, incluyendo polimerasas termoestables naturales de diversas especies de bacterias termófilas, así como polimerasas termoestables sintéticas derivadas de dichas enzimas naturales mediante delección, inserción o sustitución de aminoácidos u otra modificación. De esta manera, en algunos modos de realización, la polimerasa es una ADN polimerasa termoestable. Las formas no modificadas ejemplares de polimerasa incluyen, por ejemplo, ADN polimerasa Z05, CS6, CS5 o una ADN polimerasa funcional que tenga al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de aminoácidos con la misma. En determinados modos de realización, la ADN polimerasa funcional tiene al menos un 80 %, más preferentemente un 90 %, más preferentemente un 95 % de identidad de secuencia de aminoácidos con la misma. Otras polimerasas no modificadas incluyen, por ejemplo, ADN polimerasas de cualquiera de las siguientes especies de bacterias termófilas (o una ADN polimerasa funcional que tenga al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de aminoácidos con dicha polimerasa): *Thermotoga maritima*; *Thermus aquaticus*; *Thermus thermophilus*; *Thermus flavus*; *Thermus filiformis*; *Thermus sp. sps17*; *Thermus sp. Z05*; *Thermotoga neopolitana*; *Thermosiphon africanus*; *Thermus caldophilus*; *Deinococcus radiodurans*, *Bacillus stearothermophilus* o *Bacillus caldotenax*. En determinados modos de realización, la ADN polimerasa funcional tiene al menos un 80 %, más preferentemente un 90 %, más preferentemente un 95 % de identidad de secuencia de aminoácidos con estas polimerasas. Las polimerasas adecuadas también incluyen las que tienen actividad de transcriptasa inversa (RT) y/o la capacidad de incorporar nucleótidos no convencionales, tales como ribonucleótidos u otros nucleótidos modificados en 2'.

En algunos modos de realización, la ADN polimerasa es una ADN polimerasa termoactiva. Mientras que las ADN polimerasas termoestables que poseen actividad de transcripción inversa eficaz son particularmente adecuadas para realizar RT-PCR, especialmente RT-PCR de enzima única, las ADN polimerasas termoactivas, pero no termoestables, que poseen actividad de transcripción inversa eficaz también son susceptibles de mutación de acuerdo con la presente invención. Por ejemplo, los atributos de eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de RT son útiles para la etapa de RT en una RT-PCR y esta etapa no se necesita realizar a temperaturas que inactivaran una ADN polimerasa termoactiva, pero no termoestable. Tras la etapa de RT, se podría añadir o bien ya podría estar incluida en la mezcla de reacción una ADN polimerasa termoestable para realizar la etapa de amplificación por PCR. Por ejemplo, se puede combinar la ADN polimerasa mejorada descrita en el presente documento con una segunda ADN polimerasa termoestable antes de la etapa de RT en un tampón adecuado para la extensión y amplificación de moldes ADN y ARN, como se describe en los ejemplos. Los ejemplos de ADN polimerasas termoestables adecuadas se describen en la patente de EE. UU. n.º 4.889.818 y en las patentes de EE. UU. n.ºs 5.773.258 y 5.677.152. La segunda ADN polimerasa termoestable puede ser ADN polimerasa de AmpliTaq® (desoxinucleósido trifosfato: ADN desoxinucleotidiltransferasa, E.C.2.7.7.7). La segunda ADN polimerasa termoestable puede ser una polimerasa termoestable inactivada reversiblemente, como se describe a continuación. La polimerasa termoestable inactivada reversiblemente puede ser ADN polimerasa de AmpliTaq Gold® (Roche Applied Science, Indianapolis, IN, EE. UU.). Esta segunda metodología se beneficiaría especialmente usando una ADN polimerasa termoestable modificada químicamente (u otra tecnología HotStart para inactivar la ADN polimerasa termoestable) de modo que no sea completamente activa durante la etapa de RT. Un ejemplo de una ADN polimerasa termoactiva, pero no termoestable, que posee actividad de transcripción inversa eficaz es la ADN polimerasa de *Carboxydotherrmus hydrogenoformans* (Chy; SEQ ID NO:39). Véase, por ejemplo, las patentes de EE. UU. n.ºs 6.468.775 y 6.399.320.

También se divulga que la ADN polimerasa deriva de una especie *Thermus*. De esta manera, la ADN polimerasa tiene al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % (al menos un 80 %, preferentemente al menos un 90 % o preferentemente al menos un 95 %) de identidad de secuencia de aminoácidos con respecto a una polimerasa seleccionada del grupo que consiste en:

- (a) una ADN polimerasa Z05 de *Thermus sp.* (Z05) (SEQ ID NO: 1);
- (b) una ADN polimerasa de *Thermus aquaticus* (Taq) (SEQ ID NO: 2);
- (c) una ADN polimerasa de *Thermus filiformis* (Tfi) (SEQ ID NO: 3);
- (d) una ADN polimerasa de *Thermus flavus* (Tfl) (SEQ ID NO: 4);
- (e) una ADN polimerasa sps17 de *Thermus sp.* (Sps17) (SEQ ID NO: 5);
- (f) una ADN polimerasa de *Thermus thermophilus* (Tth) (SEQ ID NO: 6); y
- (g) una ADN polimerasa de *Thermus caldophilus* (Tca) (SEQ ID NO: 7).

También se divulga que la ADN polimerasa deriva de una especie *Carboxydotherrmus*. De esta manera, la ADN polimerasa tiene al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % (al menos un 80 %, preferentemente al menos un 90 % o

preferentemente al menos un 95 %) de identidad de secuencia de aminoácidos con respecto a una ADN polimerasa de *Carboxydothemus hydrogenoformans*(Chy) (SEQ ID NO:39).

5 También se divulga que la ADN polimerasa es una ADN polimerasa de *Thermotoga* o deriva de una especie *Thermotoga*. Por ejemplo, la ADN polimerasa tiene al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % (al menos un 80 %, preferentemente al menos un 90 % o preferentemente al menos un 95 %) de identidad de secuencia de aminoácidos con respecto a una polimerasa seleccionada del grupo que consiste en:

10 (a) una ADN polimerasa de *Thermotoga maritima* (Tma) (SEQ ID NO: 34);

(b) una ADN polimerasa de *Thermotoga neopolitana* (Tne) (SEQ ID NO: 35);

15 También se divulga que la ADN polimerasa tiene al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % (al menos un 80 %, preferentemente al menos un 90 % o preferentemente al menos un 95 %) de identidad de secuencia de aminoácidos con respecto a la SEQ ID NO:1. También se divulga que la ADN polimerasa es una ADN polimerasa Z05 de *Thermus* sp. (Z05) (es decir, SEQ ID NO: 1), y el aminoácido en la posición 763 es cualquier aminoácido distinto de M. Por ejemplo, el aminoácido en la posición 763 se selecciona de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, I o H. La ADN polimerasa puede ser una ADN polimerasa Z05 y el aminoácido en la posición 763 puede ser T. La ADN polimerasa puede ser una ADN polimerasa Z05 que comprenda adicionalmente una sustitución en la posición 580 y el aminoácido en la posición 580 puede ser cualquier aminoácido distinto de D o E. La ADN polimerasa puede ser una ADN polimerasa Z05 y el aminoácido en la posición 580 puede ser cualquier aminoácido distinto de D. La ADN polimerasa puede ser una ADN polimerasa Z05 y el aminoácido en la posición 580 se puede seleccionar del grupo que consiste en L, G, T, Q, A, S, N, R y K. La ADN polimerasa puede ser una ADN polimerasa Z05 y el aminoácido en la posición 580 puede ser G. La ADN polimerasa puede ser una ADN polimerasa Z05 que comprenda adicionalmente una sustitución en la posición 709 y el aminoácido en la posición 709 puede ser cualquier aminoácido distinto de I. La ADN polimerasa puede ser una ADN polimerasa Z05, y el aminoácido en la posición 709 se puede seleccionar del grupo que consiste en K, R, S, G, y A. La ADN polimerasa puede ser una ADN polimerasa Z05 y el aminoácido en la posición 709 puede ser K.

La ADN polimerasa de control puede ser una polimerasa Z05, Z05 D580G o Z05 D580G 1709K.

35 Las polimerasas mutantes o mejoradas pueden incluir otras modificaciones no sustitutivas. Una de dichas modificaciones es una modificación covalente reversible térmicamente que inactiva la enzima, pero que se invierte para activar la enzima tras la incubación a una temperatura elevada, tal como una temperatura usada típicamente para la extensión de polinucleótidos. Los reactivos ejemplares para dichas modificaciones reversibles térmicamente se describen en las patentes de EE. UU. n.ºs 5.773.258 y 5.677.152.

40 La actividad de transcriptasa inversa se puede determinar realizando amplificación por RT-PCR en tiempo real y detección de un transcrito del virus de la hepatitis C (VHC) generado a partir de las primeras 800 bases de la 5'NTR del genotipo 1b del VHC en pSP64 poli(A) (Promega). Dos o más mezclas de reacción pueden tener números valorados de copias del transcrito del virus de la hepatitis C (VHC) (por ejemplo, valoraciones 1:5, valoraciones 1:10, por ejemplo, 10.000 copias, 1000 copias, 100 copias, 10 copias, 0 copias en varias mezclas de reacción). Se puede comparar la capacidad de transcriptasa inversa de una polimerasa de la invención con la capacidad de transcriptasa inversa de una polimerasa de referencia (por ejemplo, una polimerasa de control, no modificada o natural), durante una unidad de tiempo preseleccionada, como se describe en el presente documento. Las polimerasas con capacidad de transcriptasa inversa mejorada amplifican el transcrito con mayor eficacia o requieren un número más bajo de ciclos de PCR para amplificar el transcrito (es decir, presentan un valor de P_c más bajo, como se calcula en el presente documento), en comparación con una polimerasa natural o no modificada. Además, las polimerasas con función de RT mejorada también tienen replicación mejorada de moldes de ARN largos (por ejemplo, al menos 500 o 1000 o 2000 o 5000 o más nucleótidos de largo). La eficacia de transcriptasa inversa mejorada puede incluir un tiempo de transcripción inversa más corto en comparación con una polimerasa de control. De esta manera, las polimerasas con eficacia de transcriptasa inversa incrementada transcriben de manera inversa un molde de ARN más rápidamente que una polimerasa de control o de referencia.

55 En diversos otros aspectos, la presente invención proporciona un ácido nucleico recombinante que codifica una ADN polimerasa mutante o mejorada de la invención, un vector que comprende el ácido nucleico recombinante y una célula huésped transformada con el vector. En determinados modos de realización, el vector es un vector de expresión. Las células huésped que comprenden dichos vectores de expresión son útiles en los procedimientos de la invención para producir la polimerasa mutante o mejorada cultivando las células huésped en condiciones adecuadas para la expresión del ácido nucleico recombinante. Las polimerasas de la invención pueden estar contenidas en kits y/o mezclas de reacción. Los aspectos de los ácidos nucleicos recombinantes, células huésped, vectores, vectores de expresión, mezclas de reacción y kits son como se describe anteriormente y en el presente documento.

En aún otro aspecto, se proporciona un procedimiento para llevar a cabo la extensión de polinucleótidos. El procedimiento incluye generalmente poner en contacto una ADN polimerasa que tenga eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT de la invención con un cebador, un molde de polinucleótido y nucleósidos trifosfato en condiciones adecuadas para la extensión del cebador, produciendo de este modo un cebador extendido. En determinados modos de realización las condiciones adecuadas para la extensión comprenden Mg^{2+} . Por ejemplo, el molde de polinucleótido puede ser un molde de ARN o ADN. Los nucleótidos trifosfato pueden incluir nucleótidos no convencionales, tales como, por ejemplo, ribonucleótidos y/o nucleótidos marcados. Además, el cebador y/o molde puede incluir uno o más análogos de nucleótidos. En algunas variaciones, el procedimiento de extensión de polinucleótidos es un procedimiento para la amplificación de polinucleótidos que incluye poner en contacto la ADN polimerasa mutante o mejorada con un par de cebadores, el molde de polinucleótido y los nucleósidos trifosfato en condiciones adecuadas para la amplificación del polinucleótido. La reacción de extensión de polinucleótidos puede ser, por ejemplo, PCR, extensión isotérmica o secuenciación (por ejemplo, la reacción de secuenciación 454). El molde de polinucleótido puede ser de cualquier tipo de muestra biológica.

Opcionalmente, la reacción de extensión del cebador comprende un inhibidor real o potencial de una polimerasa de referencia o no modificada. El inhibidor puede inhibir la velocidad de extensión de ácidos nucleicos y/o la eficacia de transcripción inversa de una polimerasa de referencia o no modificada (de control). El inhibidor puede ser hemoglobina, o un producto de degradación de la misma. Por ejemplo, el producto de degradación de hemoglobina es un producto de descomposición de hemo, tal como hemina, hematoporfirina o bilirrubina. El inhibidor puede ser un quelante de hierro o un pigmento púrpura. El inhibidor puede ser heparina o melanina. El inhibidor puede ser un tinte intercalante. En el presente documento, el tinte intercalante puede ser [2-[N-bis-(3-dimetilaminopropil)-amino]-4-[2,3-dihidro-3-metil-(benzo-1,3-tiazol-2-il)-metiliden]-1-fenil-quinolinio]⁺. El tinte intercalante puede ser [2-[N-(3-dimetilaminopropil)-N-propilamino]-4-[2,3-dihidro-3-metil-(benzo-1,3-tiazol-2-il)-metiliden]-1-fenil-quinolinio]⁺. El tinte intercalante puede no ser [2-[N-(3-dimetilaminopropil)-N-propilamino]-4-[2,3-dihidro-3-metil-(benzo-1,3-tiazol-2-il)-metiliden]-1-fenil-quinolinio]⁺. Las condiciones adecuadas para la extensión pueden comprender Mg^{++} . Las condiciones adecuadas para la extensión pueden comprender Mn^{++} .

La presente invención también proporciona un kit útil en dicho procedimiento de extensión de polinucleótidos. El kit incluye al menos un recipiente que proporciona la ADN polimerasa mejorada de acuerdo con la invención. En determinados modos de realización, el kit incluye adicionalmente uno o más recipientes adicionales que proporcionan uno o más reactivos adicionales. Por ejemplo, en variaciones específicas, uno o más recipientes adicionales proporcionan nucleósidos trifosfato; un tampón adecuado para la extensión de polinucleótidos; y/o uno o más polinucleótidos de sonda o cebador, hibridables, en condiciones de extensión de polinucleótidos, a un molde de polinucleótido predeterminado. El molde de polinucleótido puede ser de cualquier tipo de muestra biológica.

Se proporcionan adicionalmente mezclas de reacción que comprenden las polimerasas de la invención. Las mezclas de reacción también pueden contener un ácido nucleico molde (ADN y/o ARN), uno o más polinucleótidos de sonda o cebador, nucleósidos trifosfato (incluyendo, por ejemplo, desoxirribonucleósidos trifosfato, ribonucleósidos trifosfato, nucleósidos trifosfato marcados, nucleósidos trifosfato no convencionales), tampones, sales, marcas (por ejemplo, fluoróforos). La mezcla de reacción puede comprender adicionalmente Mg^{2+} . Las mezclas de reacción pueden comprender un quelante de hierro o un tinte púrpura. Las mezclas de reacción pueden comprender hemoglobina o un producto de degradación de hemoglobina. Por ejemplo, los productos de degradación de hemoglobina incluyen productos de descomposición de hemo, tales como hemina, hematina, hematoforina y bilirrubina. Las mezclas de reacción pueden comprender heparina o una sal de la misma. Opcionalmente, la mezcla de reacción puede comprender un tinte intercalante (incluyendo pero no limitado a los descritos anteriormente o en otra parte en el presente documento). La mezcla de reacción puede contener un ácido nucleico molde que se aísla de la sangre. El ácido nucleico molde puede ser ARN y la mezcla de reacción puede comprender heparina o una sal de la misma.

La mezcla de reacción puede comprender dos o más polimerasas. Por ejemplo, la mezcla de reacción puede comprender una ADN polimerasa mejorada que tenga eficacia de transcripción inversa incrementada (por ejemplo, actividad incrementada que extiende un molde de ARN) como se describe en el presente documento, y otra polimerasa que tenga actividad de polimerasa dependiente de ADN. La mezcla de reacción puede comprender una combinación de una ADN polimerasa mejorada que tiene eficacia de transcripción inversa incrementada como se describe en el presente documento y una segunda polimerasa dependiente de ADN termoestable. La segunda polimerasa dependiente de ADN termoestable puede ser una polimerasa modificada reversiblemente como se describe anteriormente, de tal manera que la enzima sea inactiva a temperaturas adecuadas para la etapa de transcripción inversa, pero sea activa en condiciones adecuadas, por ejemplo, a temperaturas elevadas de aproximadamente 90 °C a 100 °C durante un periodo de tiempo de hasta aproximadamente 12 minutos. Las condiciones adecuadas para la activación de una polimerasa termoestable inactivada reversiblemente se proporcionan, por ejemplo, en una reacción de PCR de inicio en caliente, como se describe en los ejemplos. Los ejemplos de segundas polimerasas dependientes de ADN termoestable adecuadas se describen en las patentes de EE. UU. n.ºs 5.773,258 y 5.677.152, *supra*.

Definiciones

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que se entiende comúnmente por un experto en la técnica a la que pertenece la presente invención. Aunque se puede usar esencialmente cualquier procedimiento y material similar a los descritos en el presente documento, en la práctica o pruebas de la presente invención, únicamente se describen procedimientos y materiales ejemplares. Para los propósitos de la presente invención, se definen a continuación los siguientes términos.

Los términos “un”, “una” y “el/la” incluyen referentes al plural a menos que el contexto indique claramente de otro modo.

Un "aminoácido" se refiere a cualquier unidad monomérica que se puede incorporar en un péptido, polipéptido o proteína. Como se usa en el presente documento, el término “aminoácido” incluye los siguientes veinte aminoácidos alfa naturales o genéticamente codificados: 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), 5 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). En los casos en que los residuos “X” no están definidos, estos se deben definir como “cualquier aminoácido”. Las estructuras de estos veinte aminoácidos naturales se muestran, por ejemplo, en Stryer *et al.*, Biochemistry, 5.ª ed., Freeman and Company (2002).

Los aminoácidos adicionales, tales como selenocisteína y pirrolisina, también se pueden codificar genéticamente (Stadtman (1996) “Selenocysteine,” Annu Rev Biochem. 65:83-100 e Ibbá *et al.* (2002) “Genetic code: introducing pyrrolysine,” Curr Biol. 12(13):R464-R466). El término aminoácido también incluye aminoácidos no naturales, aminoácidos modificados (por ejemplo, que tienen cadenas principales y/o cadenas laterales modificadas) y análogos de aminoácidos. Véase, por ejemplo, Zhang *et al.* (2004) “Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells,” Proc. Natl. Acad. Sci. U.S.A. 101(24):8882-8887, Anderson *et al.* (2004) “An expanded genetic code with a functional quadruplet codon” Proc. Natl. Acad. Sci. U.S.A. 101(20):7566-7571, Ikeda *et al.* (2003) “Synthesis of a novel histidine analogue and its efficient incorporation into a protein *in vivo*,” Protein Eng. Pes. Sel. 16(9):699-706, Chin *et al.* (2003) “An Expanded Eukaryotic Genetic Code,” Science 301(5635):964-967, James *et al.* (2001) “Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues,” Protein Eng. Pes. Sel. 14(12):983-991, Kohrer *et al.* (2001) “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, Bacher *et al.* (2001) “Selection and Characterization of *Escherichia coli* Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue,” J. Bacteriol. 183(18):5414-5425, Hamano-Takaku *et al.* (2000) “A Mutant *Escherichia coli* Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine,” J. Biol. Chem. 275(51):40324-40328, y Budisa *et al.* (2001) “Proteins with {beta}-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids,” Protein Sci. 10(7):1281-1292.

Para ilustrar adicionalmente, un aminoácido es típicamente un ácido orgánico que incluye un grupo amino sustituido o no sustituido, un grupo carboxi sustituido o no sustituido y una o más cadenas laterales o grupos, o análogos de cualquiera de estos grupos. Las cadenas laterales ejemplares incluyen, por ejemplo, tiol, seleno, sulfonilo, alquilo, arilo, acilo, ceto, azido, hidroxilo, hidracina, ciano, halo, hidracida, alquenilo, alquinilo, éter, borato, boronato, fosfo, fosfono, fosfina, heterocíclico, enona, imina, aldehído, éster, tioácido, hidroxilamina o cualquier combinación de estos grupos. Otros aminoácidos representativos incluyen, pero no están limitados a, aminoácidos que comprenden reticuladores fotoactivables, aminoácidos de unión a metal, aminoácidos con marcaje de espín, aminoácidos fluorescentes, aminoácidos que contienen metal, aminoácidos con grupos funcionales novedosos, aminoácidos que interaccionan covalente o no covalentemente con otras moléculas, aminoácidos fotoenmascarados y/o fotoisomerizables, aminoácidos radiactivos, aminoácidos que comprenden biotina o un análogo de biotina, aminoácidos glucosilados, otros aminoácidos modificados con hidrato de carbono, aminoácidos que comprenden polietilenglicol o poliéter, aminoácidos sustituidos con átomos pesados, aminoácidos químicamente escindibles y/o fotoescindibles, aminoácidos que contienen azúcar de enlace a carbono, aminoácidos activos en oxidorreducción, aminotioácidos que contienen aminoácidos y aminoácidos que comprenden uno o más restos tóxicos.

El término "muestra biológica" abarca una variedad de tipos de muestras obtenidos de un organismo y se pueden usar en un ensayo de supervisión o diagnóstico. El término abarca orina, sedimento de orina, sangre, saliva y otras muestras líquidas de origen biológico, muestras de tejidos sólidos, tales como una muestra de biopsia o cultivos tisulares o células derivadas de los mismos y la descendencia de los mismos. El término abarca muestras que se han manipulado de cualquier manera después de su adquisición, tal como mediante tratamiento con reactivos, solubilización, sedimentación o enriquecimiento para determinados componentes. El término abarca una muestra clínica y también incluye células en cultivo celular, sobrenadantes celulares, lisados celulares, suero, plasma, fluidos biológicos y muestras de tejidos.

El término "mutante", en el contexto de las ADN polimerasas de la presente invención, significa un polipéptido, típicamente recombinante, que comprende una o más sustituciones de aminoácidos en relación con una ADN

polimerasa funcional correspondiente.

El término "forma no modificada", en el contexto de una polimerasa mutante, es un término usado en el presente documento para los propósitos de definir una ADN polimerasa mutante de la presente invención: el término "forma no modificada" se refiere a una ADN polimerasa funcional que tiene la secuencia de aminoácidos de la polimerasa mutante, excepto en una o más posiciones de aminoácidos especificada(s) como que caracterizan la polimerasa mutante. De esta manera, la referencia a una ADN polimerasa mutante en términos de (a) su forma no modificada y (b) una o más sustituciones de aminoácidos especificadas significa que, con excepción de la(s) sustitución/sustituciones de aminoácidos especificada(s), la polimerasa mutante de otro modo tiene una secuencia de aminoácidos idéntica a la forma no modificada en el motivo especificado. La "polimerasa no modificada" (y, por lo tanto, también la forma modificada que tiene eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT) puede contener mutaciones adicionales para proporcionar la funcionalidad deseada, por ejemplo, incorporación mejorada de dideoxirribonucleótidos, ribonucleótidos, análogos de ribonucleótidos, nucleótidos marcados con tinte, actividad de 5'-nucleasa moduladora, actividad de 3'-nucleasa moduladora (o corrección de errores) o similares. Por consiguiente, al llevar a cabo la presente invención como se describe en el presente documento, se predetermina la forma no modificada de una ADN polimerasa. La forma no modificada de una ADN polimerasa puede ser, por ejemplo, una ADN polimerasa natural o una ADN polimerasa que ya se ha modificado intencionalmente. Una forma no modificada de la polimerasa es preferentemente una ADN polimerasa termoestable, tal como ADN polimerasas de diversas bacterias termófilas, así como variantes funcionales de la misma que tienen identidad de secuencia sustancial con respecto a una polimerasa termoestable natural. Dichas variantes pueden incluir, por ejemplo, ADN polimerasas quiméricas, tales como, por ejemplo, las ADN polimerasas quiméricas descritas en las patentes de EE. UU. n.ºs 6.228.628 y 7.148.049. En determinados modos de realización, la forma no modificada de una polimerasa tiene actividad de transcriptasa inversa (RT).

El término "polimerasa termoestable" se refiere a una enzima que es estable frente al calor, es resistente al calor y retiene suficiente actividad para efectuar reacciones de extensión de polinucleótidos posteriores y no se desnaturaliza irreversiblemente (inactiva) cuando se somete a temperaturas elevadas durante el tiempo necesario para efectuar la desnaturalización de los ácidos nucleicos bicatenarios. Las condiciones de calentamiento necesarias para la desnaturalización de ácidos nucleicos se conocen bien en la técnica y se ejemplifican, por ejemplo, en las patentes de EE. UU. n.ºs 4.683.202, 4.683.195 y 4.965.188. Como se usa en el presente documento, una polimerasa termoestable es adecuada para su uso en una reacción de ciclos de temperatura, tal como la reacción en cadena de la polimerasa ("PCR"). La desnaturalización irreversible para los propósitos del presente documento se refiere a la pérdida permanente y completa de actividad enzimática. Para una polimerasa termoestable, la actividad enzimática se refiere a la catálisis de la combinación de los nucleótidos de manera apropiada para formar productos de extensión de polinucleótidos que sean complementarios a una cadena de ácido nucleico molde. Las ADN polimerasas termoestables de bacterias termófilas incluyen, por ejemplo, ADN polimerasas de *Thermotoga maritima*, *Thermus aquatilis*, *Thermus thermophilus*, *Thermus flavus*, *Thermus filiformis*, especie *Thermus* sps17, especie *Thermus* Z05, *Thermus caldophilus*, *Bacillus caldotenax*, *Thermotoga neopolitana* y *Thermosiphon africanus*.

El término "termoactiva" se refiere a una enzima que mantiene las propiedades catalíticas a temperaturas usadas comúnmente para la transcripción inversa o etapas de hibridación/extensión en reacciones de RT-PCR y/o PCR (es decir, 45-80 °C). Las enzimas termoestables son las que no están irreversiblemente inactivadas o desnaturalizadas cuando se someten a temperaturas elevadas necesarias para la desnaturalización de ácidos nucleicos. Las enzimas termoactivas pueden o no puede ser termoestables. Las ADN polimerasas termoactivas pueden ser ADN o ARN dependiente de especies termófilas o de especies mesófilas incluyendo, pero no limitadas a, *Escherichia coli*, los virus de la leucemia murina de Moloney y los virus de la mieloblastosis aviar.

Como se usa en el presente documento, una proteína "quimérica" se refiere a una proteína cuya secuencia de aminoácidos representa un producto de fusión de subsecuencias de las secuencias de aminoácidos de al menos dos proteínas distintas. Típicamente no se produce una proteína quimérica mediante manipulación directa de secuencias de aminoácidos, sino que, más bien, se expresa a partir de un gen "quimérico" que codifica la secuencia de aminoácidos quimérica. Por ejemplo, una forma no modificada de una ADN polimerasa mutante como se divulga en el presente documento es una proteína quimérica que consiste en una región de extremo amínico (extremo N) derivada de una ADN polimerasa de una especie *Thermus* y una región de extremo carboxílico (extremo C) derivada de ADN polimerasa de Tma. La región de extremo N se refiere a una región que se extiende desde el extremo N (posición de aminoácido 1) a un aminoácido interno. De manera similar, la región de extremo C se refiere a una región que se extiende desde un aminoácido interno al extremo C.

El término "aptámero" se refiere a un ADN monocatenario que reconoce y se une a ADN polimerasa e inhibe eficazmente la actividad de polimerasa como se describe en la patente de EE. UU. n.º 5.693.502. El uso de aptámero y dUTP/UNG en RT-PCR también se analiza, por ejemplo, en Smith, E.S. *et al.*, (Amplification of RNA: High-temperature Reverse Transcription and DNA Amplification with a Magnesium-activated Thermostable DNA Polymerase, in PCR Primer: A Laboratory Manual, 2nd Edition, Dieffenbach, C.W. y Dveksler, G.S., Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 211-219, (2003)).

En el contexto de las ADN polimerasas mutantes, la "correspondencia" con otra secuencia (por ejemplo, regiones, fragmentos, posiciones de aminoácidos o nucleótidos, o similares) se basa en la convención de numeración de acuerdo con el número de posición de aminoácidos o nucleótidos y, a continuación, alineación de la secuencias de una manera que maximice el porcentaje de identidad de secuencia. Un aminoácido "correspondiente a la posición [X] de [secuencia específica]" se refiere a un aminoácido en un polipéptido de interés que se alinea con el aminoácido equivalente de una secuencia especificada. Generalmente, como se describe en el presente documento, se puede determinar el aminoácido correspondiente a una posición de una polimerasa usando un algoritmo de alineación, tal como BLAST como se describe a continuación. Puesto que no todas las posiciones en una "región correspondiente" dada necesitan ser idénticas, las posiciones de no emparejamiento en una región correspondiente se pueden considerar como "posiciones correspondientes". Por consiguiente, como se usa en el presente documento, la referencia a una "posición de aminoácido correspondiente a la posición de aminoácido [X]" de una ADN polimerasa especificada se refiere a posiciones equivalentes, basadas en la alineación, en otras ADN polimerasas y familias y homólogos estructurales. En algunos modos de realización de la presente invención, se determina la "correspondencia" de posiciones de aminoácidos con respecto a una región de la polimerasa que comprende uno o más motivos de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39. Cuando una secuencia de polipéptidos de polimerasa difiere de las SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39 (por ejemplo, por cambios en aminoácidos o adición o delección de aminoácidos), puede ser que una mutación particular asociada con una actividad mejorada como se analiza en el presente documento no esté en el mismo número de posición que está en las SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39. Por ejemplo, esto se ilustra en la tabla 1.

"Recombinante", como se usa en el presente documento, se refiere a una secuencia de aminoácidos o una secuencia de nucleótidos que se ha modificado intencionalmente mediante procedimientos recombinantes. El término "ácido nucleico recombinante" en el presente documento significa un ácido nucleico, formado originalmente *in vitro*, en general, mediante la manipulación de un ácido nucleico por endonucleasas de restricción, en una forma que no se encuentra normalmente en la naturaleza. De esta manera, un ácido nucleico de ADN polimerasa mutante aislado, en una forma lineal, o un vector de expresión formado *in vitro* ligando moléculas de ADN que no están unidas normalmente se consideran ambos recombinantes para los propósitos de la presente invención. Se entiende que una vez que se prepara un ácido nucleico recombinante y se reintroduce en una célula huésped, se replica de manera no recombinante, es decir, usando la maquinaria celular *in vivo* de la célula huésped en lugar de las manipulaciones *in vitro*; sin embargo, dichos ácidos nucleicos, una vez producidos de manera recombinante, aunque se replican de manera no recombinante posteriormente, todavía se consideran recombinantes para los propósitos de la invención. Una "proteína recombinante" es una proteína preparada usando técnicas recombinantes, es decir a través de la expresión de un ácido nucleico recombinante como se representa anteriormente.

Un ácido nucleico se "enlaza de manera funcional" cuando se dispone en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, un promotor o potenciador se enlaza de manera funcional a una secuencia codificante si afecta a la transcripción de la secuencia; o un sitio de unión a ribosoma se enlaza de manera funcional a una secuencia codificante si se posiciona para facilitar la traducción.

El término "célula huésped" se refiere tanto a los organismos eucariotas (por ejemplo, bacterias, levaduras y actinomicetos) como procariontes unicelulares y células sueltas de animales o plantas de orden superior cuando se cultivan en cultivo celular.

El término "vector" se refiere a una porción de ADN, típicamente bicatenario, que puede tener insertada en la misma una porción de ADN exógeno. Por ejemplo, el vector puede ser de origen plasmídico. Los vectores contienen secuencias de polinucleótidos de "replicón" que facilitan la replicación autónoma del vector en una célula huésped. El ADN exógeno se define como ADN heterógeno, que es ADN no encontrado 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 usa para transportar el ADN exógeno o heterógeno en una célula huésped adecuada. Una vez en la célula huésped, el vector se puede replicar independientemente o coincidiendo con el ADN cromosómico del huésped y se pueden generar varias copias del vector y su ADN insertado. Además, el vector también puede contener los elementos necesarios que permiten la transcripción del ADN insertado en una molécula de ARNm o de otro modo provocar la replicación del ADN insertado en múltiples copias de ARN. Algunos vectores de expresión contienen adicionalmente elementos de secuencia adyacentes al ADN insertado que incrementan la semivida del ARNm expresado y/o permiten la traducción del ARNm en una molécula de proteína. De esta manera, se pueden sintetizar rápidamente muchas moléculas de ARNm y polipéptido codificadas por el ADN insertado.

El término "nucleótido", además de hacer referencia a los monómeros de desoxirribonucleótido o ribonucleótido naturales, se entenderá que en el presente documento se refiere a variantes estructurales relacionadas del mismo, incluyendo derivados y análogos, que sean funcionalmente equivalentes con respecto al contexto particular en el que se usa el nucleótido (por ejemplo, hibridación a una base complementaria), a menos que el contexto indique claramente de otro modo.

El término "ácido nucleico" o "polinucleótido" se refiere a un polímero que se puede corresponder a un polímero de

ácido nucleico de ribosa (ARN) o ácido nucleico de desoxirribosa (ADN), o un análogo del mismo. Esto incluye polímeros de nucleótidos, tales como ARN y ADN, así como formas sintéticas, formas modificadas (por ejemplo, modificadas química o bioquímicamente) de los mismos y polímeros mixtos (por ejemplo, que incluyen tanto subunidades de ADN como ARN). Las modificaciones ejemplares incluyen metilación, sustitución de uno o más de los nucleótidos naturales con un análogo, modificaciones internucleotídicas, tales como enlaces no cargados (por ejemplo, metilfosfonatos, fosfotriésteres, fosfoamidatos, carbamatos y similares), restos pendientes (por ejemplo, polipéptidos) intercalantes (por ejemplo, acridina, soraleno y similares), quelantes, alquilantes y enlaces modificados (por ejemplo, ácidos nucleicos anoméricos alfa y similares). También se incluyen moléculas sintéticas que imitan a los polinucleótidos en su capacidad para unirse a una secuencia designada por medio de enlaces de hidrógeno y otras interacciones químicas. Típicamente, los monómeros de nucleótido se enlazan por medio de enlaces fosfodiéster, aunque las formas sintéticas de ácidos nucleicos pueden comprender otros enlaces (por ejemplo, ácidos nucleicos peptídicos como se describe en Nielsen *et al.* (*Science* 254:1497-1500, 1991). Por ejemplo, un ácido nucleico puede ser o puede incluir un segmento cromosómico o de cromosoma, un vector (por ejemplo, un vector de expresión), un casete de expresión, un polímero de ARN o ADN desnudo, el producto de una reacción en cadena de la polimerasa (PCR), un oligonucleótido, una sonda y un cebador. Por ejemplo, un ácido nucleico puede ser monocatenario, bicatenario o tricatenario y no está limitado a ninguna longitud particular. A menos que se indique de otro modo, una secuencia de ácidos nucleicos particular opcionalmente comprende o codifica secuencias complementarias, además de cualquier secuencia indicada explícitamente.

El término "oligonucleótido" se refiere a un ácido nucleico que incluye al menos dos unidades monoméricas de ácido nucleico (por ejemplo, nucleótidos). Un oligonucleótido incluye típicamente desde aproximadamente seis a aproximadamente 175 unidades monoméricas de ácido nucleico, más típicamente desde aproximadamente ocho a aproximadamente 100 unidades monoméricas de ácido nucleico, y todavía más típicamente desde aproximadamente 10 a aproximadamente 50 unidades monoméricas de ácido nucleico (por ejemplo, aproximadamente 15, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 35 o más unidades monoméricas de ácido nucleico). El tamaño exacto de un oligonucleótido depende de muchos factores, incluyendo el uso o función final del oligonucleótido. Los oligonucleótidos se preparan opcionalmente mediante cualquier procedimiento adecuado, incluyendo, pero no limitado a, aislamiento de una secuencia existente o natural, replicación o amplificación del ADN, transcripción inversa, clonación y digestión de restricción de secuencias apropiadas, o síntesis química directa mediante un procedimiento tal como el procedimiento de fosfotriéster de Narang *et al.* (*Meth. Enzymol.* 68:90-99, 1979); el procedimiento de fosfodiéster de Brown *et al.* (*Meth. Enzymol.* 68:109-151, 1979); el procedimiento de dietilfosforamidita de Beaucage *et al.* (*Tetrahedron Lett.* 22:1859-1862, 1981); el procedimiento de triéster de Matteucci *et al.* (*J. Am. Chem. Soc.* 103:3185-3191, 1981); procedimientos de síntesis automatizados; o el procedimiento de soporte sólido de la patente de EE. UU. n.º 4.458.066 u otros procedimientos conocidos por los expertos en la técnica.

El término "cebador" como se usa en el presente documento se refiere a un polinucleótido que puede actuar como un punto de iniciación de la síntesis de ácidos nucleicos dirigida por molde cuando se dispone en condiciones en las que se inicia la extensión de polinucleótidos (por ejemplo, en condiciones que comprenden la presencia de nucleósidos trifosfato requeridos (como se dicta por el molde que se copia) y una polimerasa en un tampón apropiado y a una temperatura adecuada o ciclo(s) de temperatura (por ejemplo, como en una reacción en cadena de la polimerasa)). Para ilustrar adicionalmente, también se pueden usar los cebadores en una variedad de otros procedimientos de síntesis mediados por oligonucleótidos, incluyendo como iniciadores de la síntesis de ARN *de novo* y procedimientos relacionados con la transcripción *in vitro* (por ejemplo, amplificación basada en secuencias de ácidos nucleicos (NASBA) amplificación mediada por transcripción (TMA), etc.). Un cebador es típicamente un oligonucleótido monocatenario (por ejemplo, oligodesoxirribonucleótido). La longitud apropiada de un cebador depende del uso pretendido del cebador, pero típicamente varía desde 6 a 40 nucleótidos, más típicamente desde 15 a 35 nucleótidos. Las moléculas de cebador cortas generalmente requieren temperaturas más frías para formar complejos híbridos suficientemente estables con el molde. Un cebador no necesita reflejar la secuencia exacta del molde pero debe ser suficientemente complementaria para hibridarse con un molde para que se produzca la elongación del cebador. El término "par de cebadores" puede significar un conjunto de cebadores que incluye un cebador en sentido 5' (a veces llamado "directo") que se hibrida con el complemento del extremo 5' de la secuencia de ácidos nucleicos que se va a amplificar y un cebador en antisentido 3' (a veces llamado "inverso") que se hibrida con el extremo 3' de la secuencia que se va a amplificar (por ejemplo, si la secuencia diana se expresa como ARN o es un ARN). Un cebador se puede marcar, si se desea, incorporando una marca detectable mediante medios espectroscópicos, fotoquímicos, bioquímicos, inmunoquímicos o químicos. Por ejemplo, los marcadores útiles incluyen ³²P, tintes fluorescentes, reactivos densos en electrones, enzimas (usadas comúnmente en ensayos ELISA), biotina o haptenos y proteínas para los que están disponibles antisueros o anticuerpos monoclonales.

El término "convencional" o "natural" cuando se hace referencia a bases de ácidos nucleicos, nucleósidos trifosfato o nucleótidos se refiere a los que se producen naturalmente en el polinucleótido que se describe (es decir, para el ADN estos son dATP, dGTP, dCTP y dTTP). Adicionalmente, con frecuencia se utilizan dITP y 7-deaza-dGTP en lugar de dGTP y se puede utilizar 7-deaza-dATP en lugar de dATP en reacciones de síntesis de ADN *in vitro*, tales como secuenciación. Colectivamente, estos se pueden denominar dNTP.

El término "no convencional" o "modificado" cuando se refiere a una base de ácidos nucleicos, nucleósido o

nucleótido incluye modificación, derivaciones o análogos de bases, nucleósidos o nucleótidos convencionales que se producen naturalmente en un polinucleótido particular. Determinados nucleótidos no convencionales se modifican en la posición 2' del azúcar ribosa en comparación con los dNTP convencionales. De esta manera, aunque para el ARN los nucleótidos naturales son ribonucleótidos (es decir, ATP, GTP, CTP, UTP, colectivamente rNTP), puesto que estos nucleótidos tienen un grupo hidroxilo en la posición 2' del azúcar, que, en comparación, está ausente en los dNTP, como se usa en el presente documento, los ribonucleótidos son nucleótidos no convencionales como sustratos para las ADN polimerasas. Como se usa en el presente documento, los nucleótidos no convencionales incluyen, pero no están limitados a, compuestos usados como terminadores para la secuenciación de ácidos nucleicos. Los compuestos terminadores ejemplares incluyen, pero no están limitados a, los compuestos que tienen una estructura 2',3' dideoxi y se denominan dideoxinucleósidos trifosfato. Los dideoxinucleósidos trifosfato ddATP, ddTTP, ddCTP y ddGTP se denominan colectivamente ddNTP. Los ejemplos adicionales de compuestos terminadores incluyen análogos de 2'-PO₄ de ribonucleótidos (véase, por ejemplo la solicitud de EE. UU. n.ºs 2005/0037991 y 2005/0037398). Otros nucleótidos no convencionales incluyen dNTP de fosforioato ([α-S]dNTP), 5'-[α-borano]-dNTP, dNTP de [α]-metilfosfonato y ribonucleósidos trifosfato (rNTP). Las bases no convencionales se pueden marcar con isótopos radiactivos, tales como ³²P, ³³P o ³⁵S; marcadores fluorescentes; marcadores quimioluminiscentes; marcadores bioluminiscentes; marcadores con hapteno, tales como biotina; o marcadores enzimáticos, tales como estreptavidina o avidina. Los marcadores fluorescentes pueden incluir tintes que están cargados negativamente, tales como tintes de la familia de fluoresceína, o tintes que son de carga neutra, tales como tintes de la familia de rodamina o tintes que están cargados positivamente, tales como tintes de la familia de cianina. Por ejemplo, los tintes de la familia de fluoresceína incluyen FAM, HEX, TET, JOE, NAN y ZOE. Los tintes de la familia de rodamina incluyen Texas Red, ROX, R110, R6G y TAMRA. Diversos tintes o nucleótidos marcados con FAM, HEX, TET, JOE, NAN, ZOE, ROX, R110, R6G, Texas Red y TAMRA están comercializados por Perkin-Elmer (Boston, MA), Applied Biosystems (Foster City, CA) o Invitrogen/Molecular Probes (Eugene, OR). Los tintes de la familia de cianina incluyen Cy2, Cy3, Cy5 y Cy7 y están comercializados por GE Healthcare UK Limited (Amersham Place, Little Chalfont, Buckinghamshire, Inglaterra).

Como se usa en el presente documento, se determina el "porcentaje de identidad de secuencia" comparando dos secuencias alineadas óptimamente sobre una ventana de comparación, en la que la porción 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 determinando el número de posiciones en las que se produce el residuo de aminoácidos o base de ácidos nucleicos en ambas secuencias para producir el número de posiciones emparejadas, dividiendo el número de posiciones emparejadas por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para producir el porcentaje de identidad de secuencia.

Los términos "idéntico" o "identidad" en porcentaje en el contexto de dos o más ácidos nucleicos o secuencias de polipéptidos se refieren a dos o más secuencias o subsecuencias que son la misma.

Las secuencias son "sustancialmente idénticas" entre sí si tienen un porcentaje especificado de nucleótidos o residuos de aminoácidos que son iguales (por ejemplo, al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 % o al menos un 95 % de identidad sobre una región especificada), cuando se comparan y alinean para una correspondencia máxima sobre una ventana de comparación o región designada medida usando uno de los siguientes algoritmos de comparación de secuencias o mediante alineación manual e inspección visual. Estas definiciones también se refieren al complemento de una secuencia de prueba. Opcionalmente, la identidad existe sobre una región que es de al menos aproximadamente 50 nucleótidos de longitud, o más típicamente sobre una región que es de 100 a 500 o 1000 o más nucleótidos de longitud.

Los términos "similitud" o "similitud en porcentaje", en el contexto de dos o más secuencias de polipéptidos, se refieren a dos o más secuencias o subsecuencias que tienen un porcentaje especificado de residuos de aminoácidos que son iguales o bien similares como se define por una sustitución de aminoácidos conservadora (por ejemplo, un 60 % de similitud, opcionalmente un 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 % similar sobre una región determinada) cuando se comparan y alinean para una correspondencia máxima sobre una ventana de comparación o región designada medida usando uno de los siguientes algoritmos de comparación de secuencias o mediante alineación manual e inspección visual. Las secuencias también son "sustancialmente similares" entre sí si son al menos un 20 %, al menos un 25 %, al menos un 30 %, al menos un 35 %, al menos un 40 %, al menos un 45 %, al menos un 50 % o al menos un 55 % similares entre sí. Opcionalmente, esto existe de manera similar sobre una región que es de al menos aproximadamente 50 aminoácidos de longitud, o más típicamente sobre una región que es de al menos aproximadamente 100 a 500 o 1000 o más aminoácidos de longitud.

Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la que se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y de referencia se introducen en un ordenador, se designan las coordenadas de subsecuencia, si es necesario, y se designan parámetros de programa de algoritmo de secuencias. Se usan comúnmente parámetros de programa predeterminados o se pueden designar parámetros alternativos. A continuación, el algoritmo de comparación de secuencias calcula las similitudes o identidades de secuencia en porcentaje para las secuencias de

prueba en relación con la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", como se usa en el presente documento, incluye la referencia a un segmento de uno cualquiera del número de posiciones contiguas seleccionadas del grupo que consiste en desde 20 a 600, habitualmente aproximadamente 50 a aproximadamente 200, más habitualmente aproximadamente 100 a aproximadamente 150 en las que se puede comparar una secuencia con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias estén alineadas óptimamente. Los procedimientos de alineación de secuencias para comparación se conocen bien en la técnica. Se puede llevar a cabo la alineación óptima de secuencias para comparación, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman (*Adv. Appl. Math.* 2:482, 1970), mediante el algoritmo de alineación con homología de Needleman y Wunsch (*J. Mol. Biol.* 48:443, 1970), mediante la búsqueda del procedimiento de similitud de Pearson y Lipman (*Proc. Natl. Acad. Sci. USA* 85:2444, 1988), mediante implementaciones computarizadas de estos 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 (véase, por ejemplo, Ausubel *et al.*, *Current Protocols in Molecular Biology* (suplemento de 1995)).

Los ejemplos de un algoritmo que es adecuado para determinar la identidad de secuencia en porcentaje y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (*Nuc. Acids Res.* 25:3389-402, 1977), y Altschul *et al.* (*J. Mol. Biol.* 215:403-10, 1990), respectivamente. El programa informático para realizar análisis por BLAST está disponible públicamente a través del National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica en primer lugar identificar pares de secuencias de puntuación alta (HSP) identificando palabras cortas de longitud W en la secuencia de consulta, que coinciden o bien satisfacen alguna puntuación umbral de valor positivo T cuando están alineadas con una palabra de la misma longitud en una secuencia de base de datos. T se refiere al umbral de puntuación de palabras de vecindad (Altschul *et al.*, *supra*). Estos resultados de palabras de vecindad iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largas que los contienen. Los resultados de palabras se extienden en ambas direcciones a lo largo de cada secuencia tanto como se pueda incrementar la alineación acumulativa. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos emparejados, siempre > 0) y N (puntuación de penalización para los residuos de emparejamiento erróneo, siempre < 0). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los resultados de palabras en cada dirección se detiene cuando: la puntuación de alineación acumulativa cae en la cantidad X desde su valor máximo logrado; la puntuación acumulativa se desplaza a cero o por debajo, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se alcanza el final de cada secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como valores predeterminados una longitud de palabra (W) de 11, una expectativa (E) o 10, M=5, N=-4 y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa como valores por defecto una longitud de palabra de 3, y la expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915, 1989) alineaciones (B) de 50, expectativa (E) de 10, M=5, N=-4, y una comparación de ambas cadenas.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-87, 1993). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que se produciría por casualidad un emparejamiento entre dos secuencias de aminoácidos o nucleótidos. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con respecto al ácido nucleico de referencia es menos de aproximadamente 0,2, típicamente menos de aproximadamente 0,01 y más típicamente menos de aproximadamente 0,001.

El término "eficacia de transcripción inversa" se refiere a la fracción de moléculas de ARN que se transcriben de manera inversa como ADNc en una reacción de transcripción inversa dada. En determinados modos de realización, las ADN polimerasas mutantes de la invención tienen eficacias de transcripción inversa mejoradas en relación con las formas no modificadas de estas ADN polimerasas. Es decir, estas ADN polimerasas mutantes transcriben de manera inversa una fracción más alta de moldes de ARN que sus formas no modificadas en un conjunto particular de condiciones de reacción. Sin estar limitada por la teoría, la capacidad de una ADN polimerasa mutante descrita en el presente documento para transcribir de manera inversa una fracción más alta de moldes de ARN se puede deber a una actividad de transcripción inversa incrementada, por ejemplo, una velocidad de incorporación de nucleótidos incrementada y/o procesividad incrementada de la enzima. Por ejemplo, se puede medir la eficacia de transcripción inversa midiendo el punto de corte (Pc) de una reacción de PCR usando un molde de ARN y comparando el valor de Pc con un valor de Pc de una reacción de control en la que se amplifica un molde de ADN de la misma secuencia (excepto en que los U se reemplazan por T), en la que las amplificaciones de ARN y ADN usan un conjunto de cebadores común y la misma polimerasa, por ejemplo, como se describe en los ejemplos. Una polimerasa de prueba tiene eficacia de RT mejorada cuando la polimerasa de prueba tiene un valor de Pc disminuido en comparación con una polimerasa de control cuando se usa ARN como molde, pero tiene un valor de Pc sustancialmente sin cambios en relación con la polimerasa de control cuando se usa ADN como molde. En algunos modos de realización, una polimerasa de la invención tiene una eficacia de RT mejorada, de tal manera que

el Pc es al menos una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más unidades menos que la polimerasa de control correspondiente en el molde de ARN.

La eficacia de RT mejorada de una polimerasa de prueba se puede medir como se describe en los ejemplos.

El término "tolerancia al emparejamiento erróneo" se refiere a la capacidad de una polimerasa para tolerar una secuencia que contiene emparejamiento erróneo al extender un ácido nucleico (por ejemplo, un cebador u otro oligonucleótido) de una manera dependiente de molde acoplado (por ejemplo, covalentemente) uno o más nucleótidos al ácido nucleico. El término "tolerancia al emparejamiento erróneo en 3'" se refiere a la capacidad de una polimerasa para tolerar una secuencia (prácticamente complementaria) que contiene emparejamiento erróneo donde el ácido nucleico que se va a extender (por ejemplo, un cebador u otro oligonucleótido) tiene un emparejamiento erróneo con su molde en el nucleótido de extremo 3' del cebador. Los emparejamientos erróneos con respecto al molde también se pueden localizar en el penúltimo nucleótido en 3' del cebador, o en otra posición en la secuencia del cebador.

La expresión "discriminación por emparejamiento erróneo" se refiere a la capacidad de una polimerasa para distinguir una secuencia completamente complementaria de una secuencia que contiene emparejamiento erróneo al extender un ácido nucleico (por ejemplo, un cebador u otro oligonucleótido) de una manera dependiente de molde acoplado (por ejemplo, covalentemente) uno o más nucleótidos al ácido nucleico. El término "discriminación por emparejamiento erróneo en 3'" se refiere a la capacidad de una polimerasa para distinguir una secuencia completamente complementaria de una secuencia (prácticamente complementaria) que contiene emparejamiento erróneo donde el ácido nucleico que se va a extender (por ejemplo, un cebador u otro oligonucleótido) tiene un emparejamiento erróneo en el extremo 3' del ácido nucleico en comparación con el molde al que se hibrida el ácido nucleico. El término "emparejamiento erróneo" se refiere a la existencia de uno o más apareamientos erróneos de bases (u "oposiciones de bases no complementarias") en un tramo de, de otro modo, secuencias formadoras de dúplex complementarias (o potencialmente formadoras de dúplex).

El término "valor de Pc" o valor de "punto de corte" se refiere a un valor que permite la cuantificación de ácidos nucleicos diana de entrada. Se puede determinar el valor de Pc de acuerdo con el procedimiento del máximo de la segunda derivada (Van Luu-The, et al., "Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction," *BioTechniques*, vol. 38, n.º 2, febrero de 2005, pp. 287-293). En el procedimiento de la segunda derivada, un Pc corresponde al primer máximo de una segunda curva derivada. Este máximo corresponde al comienzo de una fase semilogarítmica. El procedimiento de la segunda derivada calcula un valor de segunda derivada de la curva de intensidad de fluorescencia en tiempo real, y únicamente se obtiene un valor. El procedimiento de Pc original se basa en una aproximación diferenciable definida localmente de los valores de intensidad, por ejemplo, mediante una función polinómica. A continuación, se computa la tercera derivada. El valor de Pc es la raíz más pequeña de la tercera derivada. También se puede determinar el Pc usando el procedimiento de punto de ajuste, en el que se determina el Pc mediante la intersección de una paralela a la línea umbral en la región semilogarítmica (Van Luu-The, *et al.*, *BioTechniques*, vol. 38, n.º 2, febrero de 2005, pp. 287-293). El valor de Pc proporcionado por el instrumento LightCycler ofrecido por Roche mediante cálculo de acuerdo con el procedimiento del máximo de la segunda derivada.

El término "eficacia de la PCR" se refiere a una indicación de la eficacia de amplificación de ciclo a ciclo. La eficacia de la PCR se calcula para cada condición usando la ecuación: % de eficacia de la PCR = $(10^{-(\text{pendiente})} - 1) \times 100$, en la que la pendiente se calculó mediante regresión lineal con el logaritmo del número de copias representado en el eje y, y el Pc representado en el eje x. Se puede medir la eficacia de la PCR usando un molde de cebador emparejado o emparejado erróneamente.

El término "velocidad de extensión de ácidos nucleicos" se refiere a la velocidad en la que un biocatalizador (por ejemplo, una enzima, tal como una polimerasa, ligasa o similar) extiende un ácido nucleico (por ejemplo, un cebador u otro oligonucleótido) de una manera independiente de molde o dependiente de molde acoplado (por ejemplo, covalentemente) uno o más nucleótidos al ácido nucleico. Para ilustrar, determinadas ADN polimerasas mutantes descritas en el presente documento tienen velocidades de extensión de ácidos nucleicos mejoradas en relación con formas no modificadas de estas ADN polimerasas, de tal manera que pueden extender cebadores a velocidades más altas que estas formas no modificadas en un conjunto dado de condiciones de reacción.

El término "tolerancia de los inhibidores de la polimerasa y RT" se refiere a la capacidad de una polimerasa para mantener la actividad (actividad de polimerasa o de transcripción inversa) en presencia de una cantidad de un inhibidor que inhibiría la actividad de polimerasa o actividad de transcripción inversa de una polimerasa de control. En algunos modos de realización, la polimerasa mejorada puede tener actividad de polimerasa o de transcripción inversa en presencia de una cantidad del inhibidor que eliminaría esencialmente la actividad de polimerasa de control.

El término "sonda de 5'-nucleasa" se refiere a un oligonucleótido que comprende al menos un resto de marcaje emisor de luz y que se usa en una reacción de 5'-nucleasa para efectuar la detección de ácidos nucleicos diana. En algunos modos de realización, por ejemplo, una sonda de 5'-nucleasa incluye únicamente un único resto emisor de

luz (por ejemplo, un tinte fluorescente, etc.). En determinados modos de realización, las sondas de 5'-nucleasa incluyen regiones de autocomplementariedad, de tal manera que las sondas pueden formar estructuras de horquilla en condiciones seleccionadas. Para ilustrar adicionalmente, en algunos modos de realización una sonda de 5'-nucleasa comprende al menos dos restos de marcaje y emite radiación de intensidad incrementada, después de una de las dos marcas se escinde o, de otro modo, se separa del oligonucleótido. En determinados modos de realización, se marca una sonda de 5'-nucleasa con dos tintes fluorescentes diferentes, por ejemplo, un tinte indicador de extremo 5' y el resto o tinte extintor de extremo 3'. En algunos modos de realización, las sondas de 5'-nucleasa se marcan en una o más posiciones distintas de, o además de, posiciones de extremo. Cuando la sonda está intacta, la transferencia de energía se produce típicamente entre los dos fluoróforos, de tal manera que la emisión fluorescente del tinte indicador se extingue al menos en parte. Durante una etapa de extensión de una reacción en cadena de la polimerasa, por ejemplo, se escinde una sonda de 5'-nucleasa unida a un ácido nucleico molde mediante la actividad de nucleasa de 5' a 3' de, por ejemplo, una Taq polimerasa u otra polimerasa que tenga esta actividad, de tal manera que la emisión fluorescente del tinte indicador ya no se extingue. Las sondas de 5'-nucleasa ejemplares también se describen, por ejemplo, en la patente de EE. UU. n.º 5.210.015, la patente de EE. UU. n.º 5.994.056 y la patente de EE. UU. n.º 6.171.785. En otros modos de realización, se puede marcar una sonda de 5'-nucleasa con dos o más tintes indicadores diferentes y un resto o tinte extintor de extremo 3'.

El término "FRET" o "transferencia de energía por resonancia de fluorescencia" o "transferencia de energía por resonancia de Förster" se refiere a una transferencia de energía entre al menos dos cromóforos, un cromóforo donante y un cromóforo aceptor (denominado extintor). El donante transfiere típicamente la energía al aceptor cuando se excita el donante mediante radiación de luz con una longitud de onda adecuada. El aceptor reemite típicamente la energía transferida en forma de radiación de luz con una longitud de onda diferente. Cuando el aceptor es un extintor "de oscuridad", disipa la energía transferida en una forma distinta de la luz. Si un fluoróforo particular actúa como un donante o un aceptor depende de las propiedades del otro miembro del par de FRET. Los pares donante-aceptor usados comúnmente incluyen el par FAM-TAMRA. Los extintores usados comúnmente son DABCYL y TAMRA. Los extintores de oscuridad usados comúnmente incluyen BlackHole Quenchers™ (BHQ), (Biosearch Technologies, Inc., Novato, Cal.), Iowa Black™ (Integrated DNA Tech., Inc., Coralville, Iowa) y BlackBerry™ Quencher 650 (BBQ-650) (Berry & Assoc., Dexter, Mich.).

Breve descripción de los dibujos

La **figura 1** representa una alineación de secuencias de aminoácidos de una región del dominio de polimerasa de ADN polimerasas ejemplares de diversas especies de bacterias: especie *Thermus* Z05 (Z05) (SEQ ID NO: 12), *Thermus aquaticus* (Taq) (SEQ ID NO: 13), *Thermus filiformis* (Tfi) (SEQ ID NO: 14), *Thermus flavus* (Tfl) (SEQ ID NO: 15), especie *Thermus* sps17 (Sps17) (SEQ ID NO: 16), *Thermus thermophilus* (Tth) (SEQ ID NO: 17), *Thermus caldophilus* (Tca) (SEQ ID NO: 18), *Thermotoga maritima* (Tma) (SEQ ID NO: 19), *Thermotoga neopolitana* (Tne) (SEQ ID NO: 20), *Thermosipho africanus* (Taf) (SEQ ID NO: 21), *Deinococcus radiodurans* (Dra) (SEQ ID NO: 23), *Bacillus stearothermophilus* (Bst) (SEQ ID NO: 24) y *Bacillus caldotenax* (Bca) (SEQ ID NO: 25). Además, las regiones de polipéptidos mostradas comprenden el motivo aminoacídico D-X₁-X₂-K-X₃-A-M-X₄-X₅-X₆-X₇-X₈-X₉-L (SEQ ID NO: 26), cuyas posiciones variables se definen adicionalmente en el presente documento. Este motivo está resaltado en negrita para cada secuencia de polimerasa. Las posiciones de aminoácidos susceptibles de mutación se indican con un asterisco (*). Los huecos en las alineaciones se indican con un punto (.).

La **figura 2** proporciona identidades de secuencia entre las siguientes enzimas ADN polimerasa I: ADN polimerasa Z05 de *Thermus* sp. (Z05); ADN polimerasa de *Thermus aquaticus* (Taq); ADN polimerasa de *Thermus filiformis* (Tfi); ADN polimerasa de *Thermus flavus* (Tfl); ADN polimerasa de *Thermus* sp. sps17 (Sps17); ADN polimerasa de *Thermus thermophilus* (Tth); ADN polimerasa de *Thermus caldophilus* (Tca); ADN polimerasa de *Deinococcus radiodurans* (Dra); ADN polimerasa de *Thermotoga maritima* (Tma); ADN polimerasa de *Thermotoga neopolitana* (Tne); ADN polimerasa de *Thermosipho africanus* (Taf); ADN polimerasa de *Bacillus stearothermophilus* (Bst); y ADN polimerasa de *Bacillus caldotenax* (Bca). **(A)** Identidades de secuencia sobre toda la enzima polimerasa I (correspondiente a los aminoácidos 1-834 de Z05); y **(B)** identidades de secuencia sobre el subdominio de polimerasa correspondiente a los aminoácidos 420-834 de Z05.

La **figura 3** proporciona identidades de secuencia entre diversas enzimas ADN polimerasa I de *Thermus* sp: ADN polimerasa Z05 de *Thermus* sp. (Z05); ADN polimerasa de *Thermus aquaticus* (Taq); ADN polimerasa de *Thermus filiformis* (Tfi); ADN polimerasa de *Thermus flavus* (Tfl); ADN polimerasa de *Thermus* sp. sps17 (Sps17); ADN polimerasa de *Thermus thermophilus* (Tth); y ADN polimerasa de *Thermus caldophilus* (Tca). **(A)** Identidades de secuencia sobre toda la enzima polimerasa I (correspondiente a los aminoácidos 1-834 de Z05); y **(B)** identidades de secuencia sobre el subdominio de polimerasa correspondiente a los aminoácidos 420-834 de Z05.

Descripción detallada

La presente invención proporciona ADN polimerasas mejoradas en las que se han mutado uno o más aminoácidos en el dominio de polimerasa en relación con una ADN polimerasa funcional. Las ADN polimerasas de la invención son enzimas activas que tienen eficacia de transcriptasa inversa incrementada (por ejemplo, en presencia de cationes divalentes de Mn²⁺ y Mg²⁺) en relación con la forma no modificada de la polimerasa. También pueden tener

tolerancia al emparejamiento erróneo, velocidad de extensión y tolerancia de los inhibidores de la polimerasa y RT incrementadas. Se pueden usar las ADN polimerasas mutantes en concentraciones más bajas para un rendimiento superior o equivalente como enzimas precursoras. Las ADN polimerasas mutantes pueden tener eficacia de transcriptasa inversa incrementada mientras que retienen sustancialmente la misma actividad de polimerasa dependiente de ADN en relación con una polimerasa no modificada o de control. La invención se expone en las reivindicaciones adjuntas a esta descripción.

Las ADN polimerasas que realizan más eficazmente la transcripción inversa son útiles, por ejemplo, en una variedad de aplicaciones que implican ensayos que emplean RT-PCR para detectar y/o cuantificar dianas de ARN. Por lo tanto, las ADN polimerasas son útiles en una variedad de aplicaciones que implican la extensión de polinucleótidos, así como la transcripción inversa o amplificación de moldes de polinucleótido, que incluyen, por ejemplo, aplicaciones en estudios de ADN recombinante y en el diagnóstico médico de la enfermedad. Las ADN polimerasas mutantes también son particularmente útiles, debido a su tolerancia para los emparejamientos erróneos, para detectar dianas que posiblemente tienen secuencias variables (por ejemplo, dianas víricas, o marcadores genéticos de cáncer y otras enfermedades).

En el presente documento se divulgan ADN polimerasas que se pueden caracterizar por tener el siguiente motivo:

Asp-X1-X2-Lys-X3-Ala-Met-X4-X5-X6-X7-X8-X9-Leu

(también denominado en el presente documento en el código de una letra D-X1-X2-K-X3-A-M-X4-X5-X6-X7-X8-X9-L) (SEQ ID NO:8); en el que:

X1 es Leu (L) o Ile (I);

X2 es cualquier aminoácido distinto de Met (M) o Ile (I);

X3 es Leu (L), Ile (I) o Lys (K);

X4 es Val (V) o Ile (I);

X5 es Lys (K), Arg (R), Glu (E), Asp (D), Asn (N) o Gin (Q);

X6 es Leu (L) o Ile (I);

X7 es Phe (F), Asp (D), His (H), Ser (S) o Asn (N);

X8 es Pro (P), Arg (R), Glu (E), Asn (N), Val (V) o Ala (A);

X9 es His (H), Arg (R), Glu (E) o Gin (Q).

En el presente documento, X₂ se puede seleccionar de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L o H. También se divulgan ADN polimerasas que se pueden caracterizar por tener el siguiente motivo:

Asp-Leu-X2-Lys-Leu-Ala-Met-Val-X5-Leu-Phe-Pro-X9-Leu

(también denominado en el presente documento en el código de una letra D-L-X2-K-L-A-M-V-X5-L-F-P-X9-L) (SEQ ID NO: 9);

en el que:

X2 es cualquier aminoácido distinto de Met (M);

X5 es Lys (K) o Arg (R);

X9 es His (H) o Arg (R).

También se divulgan ADN polimerasas que se pueden caracterizar por tener el siguiente motivo:

Asp-Leu-X²-Lys- Leu-Ala-Met-Val-Lys-Leu-Phe-Pro-His-Leu

(también denominado en el presente documento en el código de una letra D-L-X2-K-L-A-M-V-K-L-F-P-H-L) (SEQ ID NO: 10); en el que: X₂ es cualquier aminoácido distinto de Met (M).

También se divulgan ADN polimerasas que se pueden caracterizar por tener el siguiente motivo:

Asp-Leu-X2-Lys-Leu-Ala-Met-Val-Lys-Leu-Phe-Pro-His-Leu (también denominado en el presente documento en el código de una letra D-L-X2-K-L-A-M-V-K-L-F-P-H-L) (SEQ ID NO: 11); en el que:

X2 es Thr (T).

También se divulgan las ADN polimerasas que se pueden caracterizar por tener los motivos anteriores (por ejemplo, las SEQ ID NO: 8, 9, 10 y 11), opcionalmente en combinación con motivos adicionales descritos a continuación. Por ejemplo, la ADN polimerasa comprende adicionalmente el motivo de la SEQ ID NO: 29 y/o la SEQ ID NO: 38.

Este motivo está presente en el dominio de los "dedos" (L hélice alfa) de muchas ADN polimerasas dependientes de ADN de tipo A de la de la familia, particularmente ADN polimerasas termoestables de bacterias termófilas (Li *et al.*, *EMBO J.* 17:7514-7525, 1998). Por ejemplo, la figura 1 muestra una alineación de secuencias de aminoácidos de una región del dominio de los "dedos" de las ADN polimerasas de diversas especies de bacterias: *Bacillus caldotenax*, *Bacillus stearothermophilus*, *Deinococcus radiodurans*, *Thermosipho africanus*, *Thermotoga maritima*, *Thermotoga neopolitana*, *Thermus aquaticus*, *Thermus caldophilus*, *Thermus filiformis*, *Thermus flavus*, *Thermus sp. Sps17*, *Thermus sp. Z05* y *Thermus thermophilus*. Como se muestra, la secuencia natural correspondiente al motivo anterior está presente en cada una de estas polimerasas, indicando una función conservada de esta región de la

polimerasa. La figura 2 proporciona identidades de secuencia entre estas ADN polimerasas.

Por consiguiente, la divulgación proporciona una polimerasa que comprende la SEQ ID NO:8, 9, 10 u 11, que tiene actividad y/o características mejoradas descritas en el presente documento, y en la que la ADN polimerasa es de otro modo una ADN polimerasa natural, tal como, por ejemplo, una polimerasa de cualquiera de las especies de bacterias termófilas enumeradas anteriormente o es sustancialmente idéntica a dicha ADN polimerasa natural. Por ejemplo, la polimerasa de la invención puede comprender la SEQ ID NO: 8, 9, 10 u 11 y es al menos un 80 %, 85 %, 90 % o 95 % idéntica a la SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, o 39. En una variación, la forma no modificada de la polimerasa es de una especie del género *Thermus*. También se divulga que la polimerasa no modificada es de una especie termófila distinta de *Thermus*, por ejemplo, *Thermotoga*. La secuencia de aminoácidos y ácidos nucleicos completa para numerosas ADN polimerasas termoestables está disponible. Las secuencias de cada una de *Thermus aquaticus* (Taq) (SEQ ID NO:2), *Thermus thermophilus* (Tth) (SEQ ID NO:6), especie *Thermus* Z05 (SEQ ID NO:1), especie *Thermus* sps17 (SEQ ID NO:5), *Thermotoga maritima* (Tma) (SEQ ID NO:34) y *Thermosiphon africanus* (Taf) (SEQ ID NO:33) se han publicado en la publicación de patente Internacional PCT n.º WO 92/06200. La secuencia para la ADN polimerasa de *Thermus flavus* (SEQ ID NO:4) se ha publicado en Akhmetzjanov y Vakhitov (*Nucleic Acids Research* 20:5839, 1992). La secuencia de la ADN polimerasa termoestable de *Thermus caldophilus* (SEQ ID NO:7) se encuentra en el n.º de acceso a EMBL/GenBank U62584. La secuencia de la ADN polimerasa termoestable de *Thermus filiformis* se puede recuperar del n.º 42380 del depósito de ATCC usando, por ejemplo, los procedimientos proporcionados en la patente de EE. UU. n.º 4.889.818, así como la información de secuencia proporcionada en la tabla 1. La secuencia de la ADN polimerasa de *Thermotoga neapolitana* (SEQ ID NO: 35) es del n.º de acceso a la base de datos de la patente de GeneSeq R98144 y el documento de PCT WO 97/09451. La secuencia de la ADN polimerasa termoestable de *Bacillus caldodenax* (SEQ ID NO: 37) se describe, por ejemplo, en Uemori *et al.* (*J Biochem (Tokyo)* 113(3):401-410, 1993; véase también el n.º de acceso a la base de datos de Swiss-Prot Q04957 y los n.ºs de acceso a GenBank D12982 y BAA02361). También se describen ejemplos de formas no modificadas de ADN polimerasas que se pueden modificar como se describe en el presente documento, por ejemplo, en la patente de EE. UU. n.ºs 6.228.628; 6.346.379; 7.030.220; 6.881.559; 6.794.177; 6.468.775 y la patente de EE. UU. n.ºs 7.148.049; 7.179.590; 7.410.782; 7.378.262. También se proporcionan secuencias de polimerasas de longitud completa en el listado de secuencias.

También son susceptibles de las mutaciones descritas en el presente documento ADN polimerasas funcionales que se han modificado previamente (por ejemplo, mediante delección, adición o sustitución de aminoácidos). Dichas polimerasas modificadas funcionales retienen el motivo aminoacídico de la SEQ ID NO: 8 (o un motivo de la SEQ ID NO: 9, 10 u 11) y opcionalmente el motivo aminoacídico de la SEQ ID NO: 38. De esta manera, las ADN polimerasas no modificadas adecuadas también incluyen variantes funcionales de polimerasas naturales. Dichas variantes tienen típicamente una similitud o identidad de secuencia sustancial con respecto a la polimerasa natural, típicamente al menos un 80 % de identidad de secuencia y más típicamente al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia.

Una polimerasa, además de tener un dominio de polimerasa que comprende las SEQ ID NO: 8, 9, 10 u 11, puede comprender también un dominio de nucleasa (por ejemplo, correspondiente a las posiciones 1 a 291 de Z05).

Una polimerasa puede ser una polimerasa quimérica, es decir, que comprende regiones de polipéptidos de dos o más enzimas. Los ejemplos de dichas ADN polimerasas quiméricas se describen, por ejemplo, en la patente de EE. UU. n.º 6.228.628. Son particularmente adecuadas las ADN polimerasas de la familia CS quiméricas, que incluyen las polimerasas CS5 (SEQ ID NO: 27) y CS6 (SEQ ID NO: 28) y variantes de las mismas que tienen similitud o identidad de secuencia de aminoácidos sustancial con respecto a la SEQ ID NO: 27 o SEQ ID NO: 28 (típicamente al menos un 80 % de identidad de secuencia de aminoácidos y más típicamente al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de aminoácidos) y, de esta manera, se pueden modificar para contener la SEQ ID NO: 8. Las ADN polimerasas CS5 y CS6 son enzimas quiméricas derivadas de las ADN polimerasas de *Thermus* sp. Z05 y *Thermotoga maritima* (Tma). Comprenden el dominio de 5'-nucleasa de extremo N de la enzima de *Thermus* y los dominios de polimerasa y 3'-5' exonucleasa de extremo C de la enzima de *Tma*. Estas enzimas tienen actividad de transcriptasa inversa eficaz, pueden extender cebadores que contienen análogos de nucleótidos y pueden incorporar dNTP de alfa-fosforotioato, dUTP, dITP y también dNTP marcados de la familia de los tintes de fluoresceína y cianina. Las polimerasas CS5 y CS6 también son enzimas de PCR activadas por Mg²⁺ eficaces. Las polimerasas quiméricas CS5 y CS6 se describen adicionalmente, por ejemplo, en la patente de EE. UU. n.º 7.148.049.

Las sustituciones de aminoácidos pueden ser sustituciones de aminoácidos únicos. Las ADN polimerasas proporcionadas en la presente invención pueden comprender una o más sustituciones de aminoácidos en el sitio activo en relación con la polimerasa no modificada. La(s) sustitución/sustituciones de aminoácidos puede(n) comprender al menos la posición X₂ del motivo expuesto en la SEQ ID NO: 8 (o un motivo de la SEQ ID NO: 9, 10 u 11). La sustitución de aminoácidos en esta posición confiere eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT, produciendo una ADN polimerasa mutante con una eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT en relación

con la polimerasa no modificada. Típicamente, el aminoácido en la posición X_2 está sustituido con un aminoácido que no se corresponde con la secuencia natural en el motivo expuesto en la SEQ ID NO: 8 (o un motivo de la SEQ ID NO: 9, 10 u 11). De esta manera, típicamente, el aminoácido en la posición X_2 , si está sustituido, no es Met (M) o Ile (I), ya que se producen M o I en esta posición en polimerasas naturales. Véase, por ejemplo, la figura 1. Las sustituciones de aminoácidos pueden incluir G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L o H en la posición X_2 . Las sustituciones de aminoácidos pueden incluir treonina (T) en la posición X_2 . Se puede(n) determinar otra(s) sustitución/sustituciones de aminoácidos adecuada(s) en uno o más de los sitios identificados usando, por ejemplo, procedimientos conocidos de mutagénesis dirigida a sitio y determinación del rendimiento de extensión de polinucleótidos en ensayos descritos adicionalmente en el presente documento o, de otro modo, conocidos por expertos en la técnica.

La polimerasa divulgada en el presente documento puede comprender la SEQ ID NO: 8, 9, 10 u 11 y puede comprender adicionalmente uno o más cambios de aminoácidos adicionales (por ejemplo, mediante delección, adición o sustitución de aminoácidos) en comparación con una polimerasa natural. Dichas polimerasas retienen el motivo aminoacídico de la SEQ ID NO: 8 (o un motivo de la SEQ ID NO: 9, 10 u 11) y comprenden adicionalmente el motivo aminoacídico de la SEQ ID NO: 38 (correspondiente a la mutación D580X de Z05 (SEQ ID NO:1)) como sigue:

Thr-Gly-Arg-Leu-Ser-Ser- X_7 - X_8 -Pro-Asn-Leu-Gln-Asn
(también denominado en el presente documento en el código de una letra T-G-R-L-S-S- X_7 - X_8 -P-N-L-Q-N) (SEQ ID NO:38); en el que X_7 es Ser (S) o Thr (T); y X_8 es cualquier aminoácido distinto de Asp (D) o Glu (E)

La mutación caracterizada por la SEQ ID NO: 38 se analiza en más detalle, por ejemplo, en la publicación de patente de EE. UU. n.º 2009/0148891. Dichas polimerasas de variantes funcionales tienen típicamente una similitud o identidad de secuencia sustancial con respecto a la polimerasa natural (por ejemplo, la SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39), típicamente al menos un 80 % de identidad de secuencia de aminoácidos y más típicamente al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de aminoácidos.

También se divulgan polimerasas que comprenden la SEQ ID NO: 8, 9, 10 u 11 y que comprenden adicionalmente el motivo aminoacídico de la SEQ ID NO: 29 (correspondiente a la mutación 1709X de Z05 (SEQ ID NO:1)) como sigue:

X_1 - X_2 - X_3 - X_4 - X_5 - X_6 - X_7 - X_8 - X_9 - X_{10} - X_{11} - X_{12} - X_{13} -Gly-Tyr-Val- X_{14} -Thr-Leu
(también denominado en el presente documento en el código de una letra X_1 - X_2 - X_3 - X_4 - X_5 - X_6 - X_7 - X_8 - X_9 - X_{10} - X_{11} - X_{12} - X_{13} -G-Y-V- X_{14} -T-L) (SEQ ID NO:29); en el que X_1 es Ala (A), Asp (D), Ser (S), Glu (E), Arg (R) o Gln (Q); X_2 es Trp (W) o Tyr (Y); X_3 es cualquier aminoácido distinto de Ile (I), Leu (L) o Met (M); X_4 es Glu (E), Ala (A), Gln (Q), Lys (K), Asn (N) o Asp (D); X_5 es Lys (K), Gly (G), Arg (R), Gln (Q), His (H) o Asn (N); X_6 es Thr (T), Val (V), Met (M) o Ile (I); X_7 es Leu (L), Val (V) o Lys (K); X_8 es Glu (E), Ser (S), Ala (A), Asp (D) o Gln (Q); X_9 es Glu (E) o Phe (F); X_{10} es Gly (G) o Ala (A); X_{11} es Arg (R) o Lys (K); X_{12} es Lys (K), Arg (R), Glu (E), Thr (T) o Gln (Q); X_{13} es Arg (R), Lys (K) o His (H); y X_{14} es Glu (E), Arg (R) o Thr (T).

Dichas polimerasas de variantes funcionales tendrán típicamente una similitud o identidad de secuencia sustancial con respecto a la polimerasa natural (por ejemplo, la SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39), típicamente al menos un 80 % de identidad de secuencia de aminoácidos y más típicamente al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de aminoácidos.

También se divulgan ADN polimerasas que comprenden una sustitución de aminoácidos en la posición X_2 (por ejemplo, como en un motivo seleccionado de la SEQ ID NO: 8, 9, 10 u 11) y que comprenden una sustitución de aminoácidos correspondiente a la SEQ ID NO: 38 y SEQ ID NO: 29.

En el presente documento, se puede sustituir el aminoácido en la posición X_2 con un aminoácido como se expone en la SEQ ID NO: 8, 9, 10 u 11, y el aminoácido en la posición X_8 (de la SEQ ID NO: 38) se sustituye con un aminoácido como se expone en la SEQ ID NO: 38. De esta manera, el aminoácido en la posición X_2 puede ser cualquier aminoácido distinto de Met (M) o Ile (I) y el aminoácido en la posición X_8 puede ser cualquier aminoácido

distinto de Asp (D) o Glu (E). Las sustituciones de aminoácidos pueden incluir leucina (L), glicina (G), treonina (T), glutamina (Q), alanina (A), serina (S), asparagina (N), arginina (R) y lisina (K) en la posición X₈ de la SEQ ID NO: 38. Las sustituciones de aminoácidos pueden incluir independientemente treonina (T) en la posición X₂ de la SEQ ID NO: 8, 9, 10 u 11 y glicina (G) en la posición X₈ de la SEQ ID NO: 38.

5 Se puede sustituir el aminoácido en la posición X₂ con un aminoácido como se expone en la SEQ ID NO: 8, 9, 10 u 11, y se puede sustituir el aminoácido en la posición X₃ con un aminoácido como se expone en la SEQ ID NO: 29. De esta manera, el aminoácido en la posición X₂ puede ser cualquier aminoácido distinto de Met (M) o Ile (I) y el aminoácido en la posición X₃ puede ser cualquier aminoácido distinto de Ile (I), Leu (L) o Met (M). Las sustituciones de aminoácidos pueden incluir lisina (K), arginina (R), serina (S), glicina (G) o alanina (A) en la posición X₃ de la SEQ ID NO: 29. En determinados modos de realización, las sustituciones de aminoácidos incluyen independientemente treonina (T) en la posición X₂ de la SEQ ID NO: 8, 9, 10 u 11, y lisina (K) en la posición X₃ de la SEQ ID NO: 29.

15 Se puede(n) determinar otra(s) sustitución/sustituciones de aminoácidos adecuada(s) en uno o más de los sitios identificados usando, por ejemplo, procedimientos conocidos de mutagénesis dirigida a sitio y determinación del rendimiento de extensión de polinucleótidos en ensayos descritos adicionalmente en el presente documento o, de otro modo, conocidos por expertos en la técnica, por ejemplo, sustituciones de aminoácidos descritas en la publicación de solicitud de patente de EE. UU. n.ºs 2009/0148891 y 2009/0280539.

20 Debido a que varía la longitud precisa de las ADN polimerasas, las posiciones de aminoácidos precisas correspondientes a cada una de X₂ (SEQ ID NO:8), X₈ (SEQ ID NO:38) y X₃ (SEQ ID NO:29) pueden variar dependiendo de la polimerasa mutante particular usada. Los programas de alineación de secuencias de aminoácidos y ácidos nucleicos están fácilmente disponibles (véanse, por ejemplo, los que se hace referencia *supra*) y, dados los motivos particulares identificados en el presente documento, sirven para ayudar a la identificación de los aminoácidos exactos (y codones correspondientes) para la modificación de acuerdo con la presente invención. Las posiciones correspondientes a cada una de X₂, X₈ y X₃ se muestran en la tabla 1 para ADN polimerasas termoestables y ADN polimerasas termoestables quiméricas representativas de especies termófilas ejemplares.

30 **Tabla 1: Posiciones de aminoácidos correspondientes a las posiciones del motivo X₂ (por ejemplo, de las SEQ ID NO: 8, 9, 10 y 11), X₈ (de la SEQ ID NO: 38) y X₃ (de la SEQ ID NO: 29) en polimerasas ejemplares.**

| <u>Organismo o secuencia quimérica</u> | <u>Posición de aminoácidos</u> | | |
|--|--------------------------------|-------------------------------------|-------------------------------------|
| | X ₂ | X ₈ (de la SEQ ID NO:38) | X ₃ (de la SEQ ID NO:29) |
| Consenso (SEQ ID NO:) | | | |
| <i>T. thermophilus</i> (6) | 763 | 580 | 709 |
| <i>T. caldophilus</i> (7) | 763 | 580 | 709 |
| <i>T. sp. Z05</i> (1) | 763 | 580 | 709 |
| <i>T. aquaticus</i> (2) | 761 | 578 | 707 |
| <i>T. flavus</i> (4) | 760 | 577 | 706 |
| <i>T. filiformis</i> (3) | 759 | 576 | 705 |
| <i>T. sp. Sps17</i> (5) | 759 | 576 | 705 |
| <i>T. maritima</i> (34) | 824 | 640 | 770 |
| <i>T. neapolitana</i> (35) | 824 | 640 | 770 |
| <i>T. africanus</i> (33) | 823 | 639 | 769 |
| <i>B. caldotenax</i> (37) | 805 | 621 | 751 |
| <i>B. stearothermophilus</i> (36) | 804 | 620 | 750 |
| CS5 (27) | 824 | 640 | 770 |
| CS6 (28) | 824 | 640 | 770 |

35 La ADN polimerasa puede derivar de *Thermus sp.* La ADN polimerasa Z05 (SEQ ID NO: 1) o una variante de la misma (por ejemplo, que porta la mutación D580G o similar). Como se hace referencia anteriormente, en la ADN

5 polimerasa Z05 de *Thermus sp.*, la posición X₂ corresponde a metionina (M) en la posición 763; la posición X₈ corresponde a aspartato (D) en la posición 580 y la posición X₃ corresponde a isoleucina (I) en la posición 709. La polimerasa mutante puede comprender al menos una sustitución de aminoácidos, en relación con una ADN polimerasa Z05 de *Thermus sp.* (o una ADN polimerasa que es sustancialmente idéntica, por ejemplo, al menos aproximadamente un 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 % idéntica a la SEQ ID NO: 1), en la posición M763, la posición D580 y/o la posición 1709. De esta manera, típicamente, el aminoácido en la posición 763 de la SEQ ID NO: 1 no es M. El aminoácido en la posición 763 de la SEQ ID NO: 1, se puede seleccionar de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, I o H. El residuo de aminoácidos en la posición 763 de la SEQ ID NO: 1 es T en la invención. Los residuos de aminoácidos en la posición 580 de la SEQ ID NO: 1 se pueden seleccionar de leucina (L), glicina (G), treonina (T), glutamina (Q), alanina (A), serina (S), asparagina (N), arginina (R) y lisina (K). El residuo de aminoácidos en la posición 580 de la SEQ ID NO: 1 es glicina (G) en la invención. Adicionalmente, el aminoácido en la posición 709 de la SEQ ID NO: 1 no es I. También se divulga que el aminoácido en la posición 709 de la SEQ ID NO: 1 se selecciona de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M o H. En el presente documento, el aminoácido en la posición 709 de la SEQ ID NO: 1 puede ser K, R, S, G o A. El aminoácido en la posición 709 de la SEQ ID NO: 1 es K en la invención.

Los mutantes de ADN polimerasa Z05 de *Thermus sp* ejemplares incluyen los que comprenden la(s) sustitución/sustituciones de aminoácidos M763T, y/o I709K (o I709R, I709S, I709G, I709A), y/o D580G. La ADN polimerasa Z05 de *Thermus sp.* mutante comprende *por ejemplo*, las sustituciones de residuos de aminoácidos M763T y D580G. En algunos modos de realización, la ADN polimerasa Z05 de *Thermus sp.* mutante comprende, por ejemplo, las sustituciones de residuos de aminoácidos M763T y I709K. En algunos modos de realización, la ADN polimerasa Z05 de *Thermus sp.* mutante comprende, por ejemplo, las sustituciones de residuos de aminoácidos M763T, I709K y D580G. La ADN polimerasa Z05 de *Thermus sp.* mutante comprende, por ejemplo las sustituciones de residuos de aminoácidos seleccionadas independientemente de M763T, I709K y/o D580G.

Se ha mostrado previamente que las sustituciones en el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 descritas anteriormente puede dar como resultado ADN polimerasas que tienen eficacia de transcripción inversa mejorada (es decir, incrementada) actividad de RT-PCR incrementada (por ejemplo, amplificación más eficaz de un molde de ARN sin comprometer la eficacia de la PCR en un molde de ADN), eficacia de RT-PCR incrementada en presencia de Mg²⁺, actividad de transcriptasa inversa incrementada en presencia de inhibidores (por ejemplo, productos de descomposición de hemoglobina, tales como hemina y/o heparina), velocidad de extensión incrementada y tolerancia al emparejamiento erróneo en 3' mejorada en comparación con una polimerasa de control. Véase la solicitud de patente de EE. UU. n.º 13/443.721, presentada el 10 de abril de 2012. De esta manera, se espera que las polimerasas mejoradas que comprenden sustituciones en el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 descritas en el presente documento también tengan las propiedades mejoradas descritas anteriormente.

Además de las mutaciones y sustituciones descritas en el presente documento, las ADN polimerasas de la presente invención también pueden incluir otra(s) modificación/modificaciones no sustitutivas. Dichas modificaciones pueden incluir, por ejemplo, modificaciones covalentes conocidas en la técnica para conferir una ventaja adicional en aplicaciones que comprenden extensión de polinucleótidos. Por ejemplo, una de dichas modificaciones es una modificación covalente reversible térmicamente que inactiva la enzima, pero que se invierte para activar la enzima tras la incubación a una temperatura elevada, tal como una temperatura usada típicamente para la extensión de polinucleótidos. Los reactivos ejemplares para dichas modificaciones reversibles térmicamente se describen en las patentes de EE. UU. n.ºs 5.773.258 y 5.677.152.

Las ADN polimerasas de la presente invención se pueden construir mutando las secuencias de ADN que codifican la polimerasa no modificada correspondiente (por ejemplo, una polimerasa natural o una variante correspondiente de la que deriva la polimerasa de la invención), tal como usando técnicas denominadas comúnmente mutagénesis dirigida a sitio. Las moléculas de ácido nucleico que codifican la forma no modificada de la polimerasa se pueden mutar mediante una diversidad de técnicas de reacción en cadena de la polimerasa (PCR) bien conocidas por un experto en la técnica. (Véase, por ejemplo, PCR Strategies (M. A. Innis, D. H. Gelfand, y J. J. Sninsky eds., 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 eds., Academic Press, NY, 1990).

A modo de ejemplo no limitante, se pueden emplear los dos sistemas de cebadores, utilizados en el kit de mutagénesis dirigida a sitio por transformador de Clontech, para introducir mutantes dirigidos a sitio en un polinucleótido que codifica una forma no modificada de la polimerasa. Después de la desnaturalización del plásmido diana en este sistema, se hibridan simultáneamente dos cebadores al plásmido; uno de estos cebadores contiene la mutación dirigida a sitio deseada, el otro contiene una mutación en otro punto en el plásmido que da como resultado la eliminación de un sitio de restricción. A continuación, se lleva a cabo la síntesis de la segunda cadena, enlazando firmemente estas dos mutaciones, y los plásmidos resultantes se transforman en una cepa mutS de *E. Coli*. El ADN plasmídico se aísla de las bacterias transformadas, se restringe con la enzima de restricción pertinente (linealizando de este modo los plásmidos no mutados), y, a continuación, se retransforma en *E. coli*. Este sistema permite la generación de mutaciones directamente en un plásmido de expresión, sin la necesidad de subclonar o generar fagémidos monocatenarios. El firme enlace de las dos mutaciones y la linealización posterior de los plásmidos no

mutados dan como resultado una eficacia de mutación alta y permiten un cribado mínimo. Después de la síntesis del cebador de sitio de restricción inicial, este procedimiento requiere el uso de únicamente un nuevo tipo de cebador por sitio de mutación. En lugar de preparar cada mutante posicional por separado, se puede sintetizar un conjunto de cebadores oligonucleotídicos "degenerados diseñados" a fin de introducir simultáneamente todas las mutaciones deseadas en un sitio dado. Los transformantes se pueden cribar mediante secuenciación del ADN plasmídico a través de la región mutagenizada para identificar y clasificar clones mutantes. A continuación, cada ADN mutante se puede restringir y analizar mediante electroforesis, tal como, por ejemplo, en un gel de potenciación de la detección de mutaciones (Mallinckrodt Baker, Inc., Phillipsburg, NJ) para confirmar que no se han producido otras alteraciones en la secuencia (mediante comparación por desplazamiento de banda con respecto al control no mutado). De manera alternativa, se puede secuenciar toda la región de ADN para confirmar que no se ha producido ningún acontecimiento mutacional adicional fuera de la región diana.

Las ADN polimerasas con más de un aminoácido sustituido se pueden generar de diversas maneras. En el caso de aminoácidos localizados cerca entre sí en la cadena polipeptídica, se pueden mutar simultáneamente usando un oligonucleótido que codifica todas las sustituciones de aminoácidos deseadas. Sin embargo, si los aminoácidos se localizan a cierta distancia entre sí (separados en más de diez aminoácidos, por ejemplo) es más difícil generar un único oligonucleótido que codifique todos los cambios deseados. En su lugar, se puede emplear uno de dos procedimientos alternativos. En el primer procedimiento, se genera un oligonucleótido separado para cada aminoácido que se va a sustituir. A continuación, se hibridan simultáneamente los oligonucleótidos al ADN molde monocatenario y la segunda cadena de ADN que se sintetiza a partir del molde codifica todas las sustituciones de aminoácidos deseadas. Un procedimiento alternativo implica dos o más rondas de mutagénesis para producir el mutante deseado. La primera tanda es como se describe para los mutantes únicos: se usa el ADN que codifica la polimerasa no modificada para el molde, se hibrida un oligonucleótido que codifica la(s) primera(s) sustitución/sustituciones de aminoácidos deseada(s) a este molde, y, a continuación, se genera la molécula de ADN de heterodúplex. La segunda tanda de mutagénesis utiliza el ADN mutado producido en la primera tanda de mutagénesis como molde. De esta manera, este molde ya contiene una o más mutaciones. A continuación, se hibrida el oligonucleótido que codifica la(s) sustitución/sustituciones de aminoácidos deseada(s) adicional(es) a este molde y la cadena de ADN resultante codifica ahora mutaciones tanto de la primera como de la segunda tandas de mutagénesis. Este ADN resultante se puede usar como un molde en una tercera tanda de mutagénesis, y así sucesivamente. De manera alternativa, se puede utilizar el procedimiento de mutagénesis de sitio múltiple de Seyfang & Jin (*Anal. Biochem.* 324:285-291. 2004).

Por consiguiente, también se proporcionan ácidos nucleicos recombinantes que codifican cualquiera de las ADN polimerasas de la presente invención. Usando un ácido nucleico de la presente invención, que codifica una ADN polimerasa, se pueden preparar una variedad de vectores. Se puede usar cualquier vector que contenga replicón y secuencias de control que deriven de una especie compatible con la célula huésped en la práctica de la invención. Generalmente, los vectores de expresión incluyen regiones de ácido nucleico reguladoras transcripcionales y traduccionales enlazadas de manera funcional al ácido nucleico que codifica la ADN polimerasa. El término "secuencias de control" se refiere a secuencias de ADN necesarias para la expresión de una secuencia codificante enlazada de manera funcional en un organismo huésped particular. Las secuencias de control que son adecuadas para procariontes, por ejemplo, incluyen un promotor, opcionalmente una secuencia operadora y un sitio de unión a ribosoma. Además, el vector puede contener un elemento retrorregulador positivo (PRE) para potenciar la semivida del ARNm transcrito (véase la patente de EE. UU. n.º 4.666.848). Las regiones de ácido nucleico reguladoras transcripcionales y traduccionales generalmente son apropiadas para la célula huésped usada para expresar la polimerasa. En la técnica se conocen numerosos tipos de vectores de expresión apropiados y secuencias reguladoras adecuadas para una variedad de células huésped. En general, las secuencias reguladoras transcripcionales y traduccionales pueden incluir, por ejemplo, secuencias promotoras, sitios de unión ribosómicos, secuencias de inicio y parada transcripcionales, secuencias de inicio y parada traduccionales y secuencias potenciadoras o activadoras. En modos de realización típicos, las secuencias reguladoras incluyen un promotor y secuencias de inicio y parada transcripcionales. Los vectores también incluyen típicamente una región de polienlazador que contiene varios sitios de restricción para la inserción de ADN exógeno. En determinados modos de realización, se usan "señales de fusión" para facilitar la purificación y, si se desea, la retirada posterior de la secuencia etiqueta/señal, por ejemplo, la "etiqueta His". Sin embargo, generalmente son innecesarias cuando se purifica una proteína termoactiva y/o termoestable de un huésped mesófilo (por ejemplo, *E. coli*) donde se puede emplear una "etapa de calor". La construcción de vectores adecuados que contienen ADN que codifica secuencias de replicación, secuencias reguladoras, genes de selección de fenotipo y la polimerasa de interés se preparan usando procedimientos de ADN recombinante estándar. Los plásmidos aislados, vectores víricos y fragmentos de ADN se escinden, adaptan y ligan entre sí en un orden específico para generar los vectores deseados, como se conoce bien en la técnica (véase, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, NY, 2.ª ed. 1989)).

El vector de expresión puede contener un gen marcador seleccionable para permitir la selección de células huésped transformadas. Los genes de selección se conocen bien en la técnica y varían con la célula huésped usada. Los genes de selección adecuados pueden incluir, por ejemplo, genes que codifican la resistencia a tetraciclina y/o ampicilina, que posibilita que las células transformadas con estos vectores crezcan en presencia de estos antibióticos.

Se puede introducir un ácido nucleico que codifica una ADN polimerasa en una célula, ya sea solo o en combinación con un vector. Por "introducir en" o equivalentes gramaticales en el presente documento significa que los ácidos nucleicos entran en las células de una manera adecuada para la posterior integración, amplificación y/o expresión del ácido nucleico. El procedimiento de introducción está dictado en gran medida por el tipo de célula seleccionada como diana. Los procedimientos ejemplares incluyen la precipitación con CaPO₄, fusión de liposomas, LIPOFECTIN®, electroporación, infección vírica y similares.

Se pueden usar típicamente procariotas como células huésped para las etapas de clonación iniciales. Son particularmente útiles para la producción rápida de grandes cantidades de ADN, para la producción de moldes de ADN monocatenario usados para la mutagénesis dirigida a sitio, para cribar simultáneamente muchos mutantes y para la secuenciación de ADN de los mutantes generados. Las células huésped procariotas adecuadas incluyen la cepa 94 de *E. coli* K12 (ATCC n.º 31,446), la cepa W3110 de *E. coli* (ATCC n.º 27.325), la cepa DG116 de *E. coli* K12 (ATCC n.º 53.606), *E. coli* X1776 (ATCC n.º 31.537) y *E. coli* B; sin embargo, se pueden usar como huéspedes todas de muchas otras cepas de *E. coli*, tales como HB101, JM101, NM522, NM538, NM539 y muchas otras especies y géneros de procariotas, que incluyen bacilos, tales como *Bacillus subtilis*, otras enterobacterias como *Salmonella typhimurium* o *Serratia marcescens* y diversas especies de *Pseudomonas*. Las células huésped procarióticas u otras células huésped con paredes celulares rígidas se transforman típicamente usando el procedimiento de cloruro de calcio como se describe en la sección 1.82 de Sambrook *et al*, *supra*. De manera alternativa, se puede usar electroporación para la transformación de estas células. Las técnicas de transformación de procariota se exponen, por ejemplo, en Dower, in *Genetic Engineering, Principles and Methods* 12:275-296 (Plenum Publishing Corp., 1990); *Hanahan et al.*, *Meth. Enzymol.*, 204:63, 1991. Los plásmidos usados típicamente para la transformación de *E. coli* incluyen pBR322, pUC18, pUC19, pUC118, pUC119 y Bluescript M13, todos los cuales se describen en las secciones 1.12-1.20 de Sambrook *et al*, *supra*. Sin embargo, también están disponibles muchos otros vectores adecuados.

Las ADN polimerasas de la presente invención se producen típicamente cultivando una célula huésped transformada con un vector de expresión que contiene un ácido nucleico que codifica la ADN polimerasa, en las condiciones apropiadas para inducir o provocar la expresión de la ADN polimerasa. Los procedimientos de cultivo de células huésped transformadas en condiciones adecuadas para la expresión de proteínas se conocen bien en la técnica (véase, por ejemplo, Sambrook *et al*, *supra*). Las células huésped adecuadas para la producción de las polimerasas a partir de vectores plasmídicos que contienen promotor pL lambda incluyen la cepa DG116 de *E. coli* (ATCC n.º 53606) (véase la patente de EE. UU. n.º 5.079.352 y Lawyer, F.C. *et al*, PCR Methods and Applications 2:275-87, 1993). Después de la expresión, la polimerasa se puede recoger y aislar. Los procedimientos para purificar la ADN polimerasa termoestable se describen, por ejemplo, en Lawyer *et al*, *supra*. Una vez purificada, se puede someter a prueba la capacidad de las ADN polimerasas para tener eficacia de RT mejorada, tolerancia al emparejamiento erróneo incrementada, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT (por ejemplo, como se describe en los ejemplos).

Se pueden usar las ADN polimerasas mejoradas de la presente invención para cualquier propósito en el que dicha actividad enzimática sea necesaria o deseada. Por consiguiente, en otro aspecto de la invención, se proporcionan procedimientos de extensión de polinucleótidos (por ejemplo, PCR) usando las polimerasas. Las condiciones adecuadas para la extensión de polinucleótidos se conocen en la técnica. (Véase, por ejemplo, Sambrook *et al.*, *supra*. Véase también Ausubel *et al.*, *Short Protocols in Molecular Biology* (4.º ed., John Wiley & Sons 1999). Generalmente, se hibrida un cebador, es decir, se hibrida a un ácido nucleico diana para formar un complejo cebador-molde. El complejo cebador-molde se pone en contacto con la ADN polimerasa y nucleósidos trifosfato en un ambiente adecuado para permitir la adición de uno o más nucleótidos al extremo 3' del cebador, produciendo de este modo un cebador extendido complementario al ácido nucleico diana. El cebador puede incluir, por ejemplo, uno o más análogos de nucleótidos. Además, los nucleósidos trifosfato 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 polinucleótidos comprende la amplificación de un ácido nucleico diana. También son conocidas en la técnica condiciones adecuadas para la amplificación de ácidos nucleicos usando una ADN polimerasa y un par de cebadores (por ejemplo, procedimientos de amplificación por PCR). (Véase, por ejemplo, Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*; PCR Applications: Protocols for Functional Genomics (Innis *et al.* eds., Academic Press 1999). En otros modos de realización no mutuamente exclusivos, la reacción de extensión de polinucleótidos comprende la transcripción inversa de un molde de ARN (por ejemplo, RT-PCR). En algunos modos de realización, las polimerasas mejoradas encuentran uso en secuenciación

Opcionalmente, la reacción de extensión del cebador comprende un inhibidor real o potencial de una polimerasa de referencia o no modificada. Por ejemplo, el inhibidor puede inhibir la velocidad de extensión de ácidos nucleicos y/o la eficacia de transcripción inversa de una polimerasa de referencia o no modificada (de control). En el presente documento, el inhibidor puede ser hemoglobina, o un producto de degradación de la misma. Por ejemplo, el producto de degradación de hemoglobina puede ser un producto de descomposición de hemo, tal como hemeina, hematorporfirina o bilirrubina. El inhibidor puede ser un quelante de hierro o un pigmento púrpura. El inhibidor puede ser heparina. El inhibidor puede ser un tinte intercalante. El inhibidor puede ser melanina, que se ha descrito como un inhibidor de la polimerasa. Véase, por ejemplo, Ekhardt, *et al*, *Biochem Biophys Res Commun.* 271(3):726-30

(2000).

Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de moldes de polinucleótido aislados de muestras que comprenden inhibidores de la polimerasa, por ejemplo, tales como sangre. Por ejemplo, se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de hemoglobina, un componente principal de la sangre, o en presencia de un producto de degradación de hemoglobina. La hemoglobina puede degradarse en diversos productos de descomposición de hemo, tales como hemina, hematina, hematoporfirina y bilirrubina. De esta manera, se pueden usar ADN polimerasas de la presente invención para extender moldes en presencia de productos de degradación de hemoglobina, incluyendo, pero no limitados a, hemina, hematina, hematoporfirina y bilirrubina. En el presente documento, el producto de degradación de hemoglobina puede ser hemina. Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de hemina aproximadamente 0,5 a 20,0 μM , aproximadamente 0,5 a 10,0 μM , aproximadamente 0,5 a 5,0 μM , aproximadamente 1,0 a 10,0 μM , aproximadamente 1,0 a 5,0 μM , aproximadamente 2,0 a 5,0 μM o aproximadamente 2,0 a 3,0 μM . Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de al menos hemina aproximadamente 0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 4,0, 5,0, 10,0, 20,0 o mayor de 20 μM . Los productos de descomposición de hemoglobina incluyen quelantes de hierro y pigmentos púrpura. De esta manera, se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de quelantes de hierro y/o pigmentos púrpura. Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de cantidades de productos de degradación de hemoglobina que inhiben la extensión del mismo molde mediante una ADN polimerasa de referencia o de control.

Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de heparina. La heparina está presente comúnmente como un anticoagulante en muestras aisladas de sangre. Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de aproximadamente 1,0 a 400 ng/ μl , 1,0 a 300 ng/ μl , 1,0 a 200 ng/ μl , 5,0 a 400 ng/ μl , 5,0 a 300 ng/ μl , 5,0 a 200 ng/ μl , 10,0 a 400 ng/ μl , 10,0 a 300 ng/ μl o 10,0 a 200 ng/ μl de heparina. Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400 ng/ μl o mayor de 400 ng/ μl de heparina. Se pueden usar las ADN polimerasas de la presente invención para extender moldes en presencia de cantidades de heparina que inhiben la extensión del mismo molde mediante una ADN polimerasa de referencia o de control.

Se puede usar una polimerasa mejorada de la invención en una reacción de transcripción inversa. La reacción de transcripción inversa se puede llevar a cabo en una mezcla que contiene el molde de ARN, uno o más cebadores y una ADN polimerasa termoestable de la invención. La mezcla de reacción contiene típicamente los cuatro desoxirribonucleósidos trifosfato estándar (dNTP) y un tampón que contiene un catión divalente y un catión monovalente. Los cationes ejemplares incluyen, por ejemplo, Mg^{2+} , aunque otros cationes, tales como Mn^{2+} o Co^{2+} pueden activar ADN polimerasas. La reacción de transcripción inversa se puede llevar a cabo con una ADN polimerasa termoactiva de la invención. La polimerasa mejorada de la invención puede permitir una amplificación más eficaz de moldes de ARN sin comprometer la amplificación eficaz de un molde de ADN en presencia de Mn^{2+} o Mg^{2+} , como se describe en los ejemplos.

La polimerasa mejorada puede tener una eficacia de transcripción inversa incrementada en comparación con una polimerasa de control. No se apreció previamente que las sustituciones en el aminoácido correspondiente a la posición 763 de la SEQ ID NO: 1 pudieran dar como resultado una mayor eficacia de RT. De esta manera, las ADN polimerasas que tienen una sustitución Met (M) a Thr (T) o una sustitución Ile (I) a Thr (T), en el aminoácido correspondiente a la posición 763 de la SEQ ID NO: 1 pueden tener eficacia de RT incrementada. La ADN polimerasa que tiene una eficacia de transcripción inversa incrementada puede comprender una sustitución de Met (M) a Thr (T), o una sustitución de Ile (I) a Thr (T), en el aminoácido correspondiente a la posición 763 de la SEQ ID NO: 1, y puede tener al menos un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % de identidad de secuencia de aminoácidos con respecto a las SEQ ID NO: 1-7, 32-37 o 39.

La polimerasa mejorada puede tener eficacia de transcripción inversa incrementada usando un molde de ARN sin una disminución sustancial de la actividad de polimerasa usando un molde de ADN. De esta manera, la ADN polimerasa mejorada puede tener una eficacia de RT incrementada sin una disminución sustancial de la actividad de polimerasa dependiente de ADN cuando se compara con una polimerasa de control. La ADN polimerasa mejorada descrita en el presente documento puede tener actividad de polimerasa dependiente de ADN que sea sustancialmente la misma que una polimerasa de control. De esta manera, en algunos modos de realización, la ADN polimerasa mejorada descrita en el presente documento tiene actividad de polimerasa dependiente de ADN que es al menos aproximadamente un 90 % de la actividad de una polimerasa de control, por ejemplo, al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 % o más de la actividad de una polimerasa de control. La actividad de polimerasa dependiente de ADN se puede medir, por ejemplo, amplificando un molde de ADN y determinando los valores de P_c como se describe en el presente documento. De esta manera, la ADN polimerasa puede tener una eficacia de RT mejorada medida como un valor de P_c disminuido en comparación con una polimerasa de control cuando se usa ARN como molde, pero puede tener un valor de P_c sustancialmente sin

cambios en relación con la polimerasa de control cuando se usa ADN como molde. Por ejemplo, cuando se amplifica un molde de ADN, la ADN polimerasa mejorada puede tener un valor de P_c que difiere en menos de 1,0, menos de 0,5, menos de 0,4, menos de 0,3, menos de 0,2 o menos de 0,1 en comparación con una polimerasa de control. La actividad de polimerasa dependiente de ADN se puede determinar como se describe en los ejemplos.

Una polimerasa mejorada de la invención puede incrementar la eficacia de transcripción inversa reduciendo el tiempo de reacción requerido para extender un molde de ARN. Por ejemplo, una polimerasa mejorada descrita en el presente documento puede acortar significativamente el tiempo de reacción requerido para transcribir ARN a ADNc en comparación con una polimerasa de control, incrementando de este modo la eficacia de transcriptasa inversa. Sin estar limitada por la teoría, la polimerasa mejorada puede incrementar la eficacia de RT, por ejemplo, incrementando la actividad de la enzima sobre un molde de ARN, tal como incrementando la velocidad de incorporación de nucleótidos y/o incrementando la procesabilidad de la polimerasa, acortando eficazmente de este modo el tiempo de extensión de un molde de ARN o población de moldes de ARN. Los tiempos de reacción para la etapa RT inicial son típicamente del orden de 30 minutos o más a 65 grados C cuando se usa una polimerasa no modificada o de control. De esta manera, la polimerasa mejorada puede transcribir un molde de ARN en ADNc en menos de aproximadamente 30 minutos, menos de aproximadamente 20 minutos, menos de aproximadamente 10 minutos, menos de aproximadamente 8 minutos, menos de aproximadamente 5 minutos, menos de aproximadamente 4 minutos, menos de aproximadamente 3 minutos o menos de aproximadamente 2 minutos a 65 grados C. La polimerasa mejorada puede transcribir un molde de ARN derivado del transcrito JP2-5 del virus de la hepatitis C (VHC), que contiene las primeras 800 bases de la 5'NTR del genotipo 1b del VHC, en ADNc en menos tiempo o más rápido que una polimerasa de control. Por ejemplo, la polimerasa mejorada puede transcribir 240 bases del molde de ARN de JP2-5 del VHC en ADNc de longitud completa en aproximadamente 15 segundos menos, 30 segundos menos, un minuto menos, dos minutos menos, 3 minutos menos, 4 minutos menos, 5 minutos menos o aproximadamente 10 minutos menos que una polimerasa de control en idénticas condiciones de reacción. En algunos modos de realización, la polimerasa mejorada puede transcribir 240 bases del molde de ARN de JP2-5 del VHC en ADNc de longitud completa más rápido que una polimerasa de control, por ejemplo, aproximadamente 5 segundos, 10 segundos, 15 segundos, 30 segundos, 45 segundos o 60 segundos o más más rápido que una polimerasa de control en idénticas condiciones de reacción. Las condiciones de reacción pueden ser las descritas en los ejemplos. Se puede poner en contacto una polimerasa mejorada descrita en el presente documento con un molde de ARN a 65 grados C durante aproximadamente 2 minutos en la mezcla de reacción descrita anteriormente. La etapa de extensión se puede seguir de amplificación por PCR del molde extendido, como se describe en los ejemplos.

La actividad de RT más eficaz en ADN polimerasas termoestables se ha logrado usando Mn^{2+} como el activador de iones de metal divalentes. Sin embargo, se conoce bien que cuando está presente Mn^{2+} en las reacciones la fidelidad de las ADN polimerasas es más baja. A menos se esté tratando de generar mutaciones, generalmente se favorece mantener una fidelidad más alta. Afortunadamente, la mayoría de las aplicaciones convencionales de secuenciación, PCR y RT-PCR, no requieren condiciones de alta fidelidad porque los sistemas de detección generalmente buscan una población de productos.

Con la llegada de la siguiente generación de secuenciación, PCR digital, etc., la fidelidad del producto es más importante y los procedimientos que permiten una síntesis de ADN de fidelidad más alta son críticos. El logro de una actividad de RT eficaz usando Mg^{2+} como activador de iones de metal divalentes es una excelente manera de incrementar sustancialmente la fidelidad de la ADN polimerasa y permitir una copia más fiable de la diana de ácido nucleico. Por consiguiente, la polimerasa mejorada de la invención puede permitir una extensión y/o amplificación eficaces de moldes de ARN usando Mg^{2+} como activador de iones de metal divalentes, como se describe en los ejemplos.

Debido a que las polimerasas descritas en el presente documento también pueden tener tolerancia al emparejamiento erróneo incrementada, las polimerasas encuentran uso en procedimientos en los que es probable la variación del molde diana y, no obstante, se desea que el molde aún se amplifique independientemente de la variación en el molde diana. Un ejemplo de dichos moldes puede incluir, por ejemplo, secuencias víricas, bacterianas o patógenas. Puede ser deseable determinar simplemente si un sujeto (humano o animal no humano) tiene una infección vírica u otra infección, independientemente de la variante vírica precisa que ha infectado al sujeto. Como ejemplo, se puede usar un par de cebadores para amplificar el VHC usando una polimerasa de la invención y detectando la presencia del VHC incluso si el virus particular que infecta al sujeto tiene una mutación que da como resultado un emparejamiento erróneo en el sitio de hibridación del cebador.

Los ácidos nucleicos diana pueden proceder de una fuente biológica o sintética. La diana puede ser, por ejemplo, ADN o ARN. Generalmente, cuando se generan amplicones, los amplicones se componen de ADN, aunque también se pueden incorporar ribonucleótidos o nucleótidos sintéticos en el amplicón. Cuando se desea detectar un ARN, el procedimiento de amplificación implica típicamente el uso de transcripción inversa, incluyendo, por ejemplo, PCR de transcripción inversa (RT-PCR).

Las secuencias diana específicas pueden incluir, por ejemplo, ácidos nucleicos víricos (por ejemplo, virus de la inmunodeficiencia humana (VIH), virus de la hepatitis B (VHB), (citomegalovirus (CMV), parvovirus B19, virus de

Epstein-Barr, virus de la hepatitis C (VHC), virus del papiloma humano (VPH), virus de la encefalitis japonesa (VEJ), virus del Nilo Occidental (VNO), virus de la encefalitis de San Luis (VESL), virus de la encefalitis del valle de Murray y el virus Kunjin), ácidos nucleicos bacterianos (por ejemplo, *S. aureus*, *Neisseria meningitidis*, *Plasmodium falciparum*, *Chlamydia muridarum*, *Chlamydia trachomatis*), micobacterias, ácidos nucleicos fúngicos o ácidos nucleicos de animales o plantas. Los ácidos nucleicos diana pueden ser ácidos nucleicos de animales (por ejemplo, humanos) o pueden derivar de una muestra animal (por ejemplo, humana) (es decir, pueden estar presentes ácidos nucleicos de organismos víricos u otros patógenos una muestra de una biopsia animal, muestra de sangre, muestra de orina, muestra fecal, saliva, etc.). Los ácidos nucleicos diana, por ejemplo, son regiones genéticas humanas que pueden incluir variantes asociadas con enfermedad (por ejemplo, cáncer, diabetes, etc.). Debido a que las polimerasas de la invención pueden tener tolerancia al emparejamiento erróneo, dichas enzimas pueden ser particularmente útiles, por ejemplo, donde una diversidad de secuencias relacionadas puede estar en una secuencia diana. Como ejemplo, se puede usar la invención para detectar patógenos víricos, donde los patógenos víricos tienen suficiente variación en sus genomas para hacer difícil o imposible diseñar un conjunto único o pequeño de cebadores que amplifiquen la mayoría o todos los genomas víricos posibles o en cáncer u otros marcadores genéticos de enfermedades en los que se sabe o es probable que se produzca una variación en la secuencia.

Otros procedimientos para detectar productos de extensión o productos de amplificación que usan las polimerasas mejoradas descritas en la presente invención incluyen el uso de tintes de unión a nucleótidos bicatenarios fluorescentes o tintes intercalantes de nucleótidos bicatenarios fluorescentes. Los ejemplos de tintes de unión a ADN bicatenario fluorescentes incluyen SYBR-verde (Molecular Probes). Se pueden usar los tintes de unión a ADN bicatenario junto con el análisis de curvas de fusión para medir productos de extensión del cebador y/o productos de amplificación. El análisis de curvas de fusión se puede realizar en un instrumento de PCR en tiempo real, tal como el instrumento ABI 5700/7000 (formato de 96 pocillos) o ABI 7900 (384 pocillos) con programa informático integrado (SDS 2.1). De manera alternativa, el análisis de curvas de fusión se puede realizar como un análisis de criterio de valoración. Los procedimientos ejemplares de análisis de puntos de fusión se describen en la publicación de patente de EE. UU. n.º 2006/0172324.

En otro aspecto de la presente invención, se proporcionan kits para su uso en procedimientos de extensión del cebador descritos en el presente documento. En el presente documento, el kit puede compartimentarse para facilitar su uso y contiene al menos un recipiente que proporciona una ADN polimerasa mejorada de acuerdo con la presente invención. También se pueden incluir uno o más recipientes adicionales que proporcionan un reactivo(s) adicional(es). El kit también puede incluir un tubo de extracción sanguínea, recipiente o unidad que comprenda heparina o una sal de la misma, o libera heparina en solución. La unidad de extracción sanguínea puede ser un tubo heparinizado. Dichos recipientes adicionales pueden incluir cualquier reactivo u otros elementos reconocidos por el experto en la técnica para su uso en procedimientos de extensión del cebador de acuerdo con los procedimientos descritos anteriormente, incluyendo reactivos para su uso, por ejemplo, en procedimientos de amplificación de ácidos nucleicos (por ejemplo, PCR, RT-PCR), procedimientos de secuenciación de ADN o procedimientos de marcaje de ADN. Por ejemplo, el kit puede incluir adicionalmente un recipiente que proporciona un cebador en sentido 5' hibridable, en condiciones de extensión del cebador, a un molde de polinucleótido predeterminado, o un par de cebadores que comprende el cebador en sentido 5' y un cebador antisentido en 3' correspondiente. El kit puede incluir uno o más recipientes que proporcionan nucleósidos trifosfato (convencionales y/o no convencionales). Más específicamente, el kit puede incluir dNTP de alfa-fosforotioato, dUTP, dITP, y/o dNTP marcados, tales como, por ejemplo, dNTP de la familia de los tintes de fluoresceína y cianina. El kit también puede incluir uno o más recipientes que proporcionan un tampón adecuado para una reacción de extensión del cebador.

En otro aspecto de la presente invención, se proporcionan mezclas de reacción que comprenden las polimerasas con eficacia de transcriptasa inversa incrementada, tolerancia al emparejamiento erróneo, velocidad de extensión y/o tolerancia de los inhibidores de la polimerasa y RT como se describe en el presente documento. Las mezclas de reacción pueden comprender adicionalmente reactivos para su uso, por ejemplo, en procedimientos de amplificación de ácidos nucleicos (por ejemplo, PCR, RT-PCR), procedimientos de secuenciación de ADN o procedimientos de marcaje de ADN. Por ejemplo, las mezclas de reacción pueden comprender un tampón adecuado para una reacción de extensión del cebador. Las mezclas de reacción también pueden contener un ácido nucleico molde (ADN y/o ARN), uno o más polinucleótidos de sonda o cebadores, nucleósidos trifosfato (incluyendo, por ejemplo, desoxirribonucleótidos, ribonucleótidos, nucleótidos marcados, nucleótidos no convencionales), sales (por ejemplo,), marcas (por ejemplo, fluoróforos). Las mezclas de reacción pueden contener un cebador en sentido 5' hibridable, en condiciones de extensión del cebador, a un molde de polinucleótido predeterminado, o un par de cebadores que comprende el cebador en sentido 5' y un cebador antisentido en 3' correspondiente. Las mezclas de reacción pueden contener dNTP de alfa-fosforotioato, dUTP, dITP, y/o dNTP marcados, tales como, por ejemplo, dNTP de la familia de los tintes de fluoresceína y cianina. Las mezclas de reacción pueden comprender un quelante de hierro o un tinte púrpura. Las mezclas de reacción pueden comprender hemoglobina o un producto de degradación de hemoglobina. Por ejemplo, los productos de degradación de hemoglobina pueden incluir productos de descomposición de hemo, tales como hemina, hematina, hematófisis y bilirrubina. Las mezclas de reacción pueden comprender heparina o una sal de la misma. La mezcla de reacción puede contener un ácido nucleico molde que se aísla de la sangre. El ácido nucleico molde puede ser ARN y la mezcla de reacción puede comprender heparina o una sal de la misma.

La mezcla de reacción puede comprender dos o más polimerasas. Por ejemplo, la mezcla de reacción puede comprender una primera ADN polimerasa que tenga eficacia de transcriptasa inversa incrementada en comparación con una polimerasa de control y una segunda ADN polimerasa que tenga actividad de polimerasa dependiente de ADN. La segunda ADN polimerasa puede ser una polimerasa natural o no modificada, o puede ser una polimerasa mejorada que tenga actividad de polimerasa dependiente de ADN incrementada. Dichas mezclas de reacción son útiles para la amplificación de moldes de ARN (por ejemplo, RT-PCR) proporcionando tanto una polimerasa que tiene actividad de transcriptasa inversa incrementada como un polímero que tiene actividad de polimerasa dependiente de ADN.

10 Ejemplos

Se ofrecen los siguientes ejemplos para ilustrar, pero no para limitar la invención reivindicada.

15 **Ejemplo 1: generación de genotecas**

En resumen, las etapas en este procedimiento de cribado incluyen la generación de genotecas, expresión y purificación parcial de las enzimas mutantes, cribado de las enzimas frente a las propiedades deseadas, secuenciación de ADN, purificación de clones y caracterización adicional de mutantes candidatos seleccionados. Cada una de estas etapas se describe más adelante.

Generación de genotecas de clones: Se sometió un ácido nucleico que codificaba el dominio de polimerasa de la ADN polimerasa D580G1709K Z05 a PCR propensa a errores (mutagénica) entre los sitios de restricción Bsp I y Bgl II de un plásmido que incluía esta secuencia de ácido nucleico. Los cebadores usados para esto se dan a continuación:

25 Cebador directo: 5'- CTACCTCCTGGACCCCTCCAA-3' (SEQ ID NO:30); y cebador inverso: 5'-ATAACCAACTGGTAGTGGCGTGTA-3' (SEQ ID NO:31)

30 Se realizó la PCR usando una concentración de Mg^{2+} de 1,8 mM a fin de generar una genoteca con una velocidad de mutación deseada. Las condiciones del tampón fueron bicina 50 mM pH 8,2, KOAc 115 mM, glicerol al 8 % p/v y 0,2 mM de cada dNTP. Se usó una enzima de PCR de inicio en caliente GeneAmp® AccuRT a 0,15 U/ μ l. Comenzando con 5×10^5 copias de ADN plasmídico D580G1709K Z05 linealizado por volumen de reacción de 50 μ l, se desnaturalizaron las reacciones usando una temperatura de 94 °C durante 60 segundos, a continuación, se realizaron 30 ciclos de amplificación, usando una temperatura de desnaturalización de 94 °C durante 15 segundos, una temperatura de hibridación de 60 °C durante 15 segundos, una temperatura de extensión de 72 °C durante 120 segundos y seguido de una extensión final a una temperatura de 72 °C durante 5 minutos.

40 El amplicón resultante se purificó con un kit de purificación de PCR QIAquick (Qiagen, Inc., Valencia, CA, EE. UU.) y se cortó con Bsp I y Bgl II, y, a continuación, se repurificó con un kit de purificación de PCR QIAquick. Se preparó un plásmido de vector D580G1709K Z05 cortando con las mismas dos enzimas de restricción y tratando con fosfatasa alcalina, recombinante (RAS, n.º cat 03359123001) y se purificó con un kit de purificación de PCR QIAquick. El vector de corte y el inserto mutado se mezclaron en una proporción de 1:3 y se trataron con ADN ligasa T4 durante 5 minutos a temperatura ambiente (NEB Quick Ligation™ Kit). Las ligaciones se purificaron con un kit de purificación de PCR QIAquick y se transformaron en una cepa huésped de *E. coli* mediante electroporación.

45 Se sembraron en placas alícuotas de los cultivos expresados en medio selectivo de ampicilina a fin de determinar el número de transformantes únicos en cada transformación. Las transformaciones se agruparon y almacenaron a -70 °C hasta -80 °C en presencia de glicerol como crioprotector.

50 A continuación, la genoteca se extendió en placas de agar selectivas a ampicilina de gran formato. Las colonias individuales se transfirieron a placas de 384 pocillos que contenían 2X de caldo Luria con ampicilina y glicerol al 10 % p/v usando un colector automático de colonias (QPix2, Genetix Ltd). Estas placas se incubaron durante la noche a 30 °C para permitir que los cultivos crecieran y, a continuación, se almacenaron a -70 °C a -80 °C. El glicerol añadido al 2X de caldo Luria era lo suficientemente bajo para permitir el crecimiento del cultivo y aún lo suficientemente alto para proporcionar crioprotección. Se prepararon diversos miles de colonias de esta manera para su uso posterior.

60 **Preparación de la genoteca de extractos parte 1—fermentación:** A partir de las genotecas de clones descritas anteriormente, se preparó una genoteca correspondiente de extractos parcialmente purificados adecuados para propósitos de cribado. La primera etapa de este procedimiento fue preparar cultivos de expresión a pequeña escala de cada clon. Estos cultivos se cultivaron en un formato de 96 pocillos; por lo tanto, había 4 placas de cultivo de expresión para cada placa de la genoteca de 384 pocillos. Se transfirieron 0,5 μ l desde cada pocillo de la placa de la genoteca de clones a un pocillo de una placa de siembra de 96 pocillos, que contenía 1150 μ l de medio A (véase la tabla 3 a continuación). Esta placa de siembra se agitó durante la noche a 1150 rpm a 30 °C, en un agitador/incubador de placas iEMS (ThermoElectron). A continuación, estos cultivos de siembra se usaron para

65

5 inocular el mismo medio, esta vez inoculando 20 µl en 250 µl de medio A en placas de 96 pocillos de formato grande (Nunc n.º 267334). Estas placas se incubaron durante la noche a 37 °C con agitación. El plásmido de expresión contenía elementos de control transcripcionales, que permiten la expresión a 37 °C, pero no a 30 °C. Después de la incubación durante la noche, los cultivos expresaron la proteína del clon en típicamente un 1-10 % de la proteína celular total. Las células de estos cultivos se recogieron mediante centrifugación. Estas células se congelaron (-20 °C) o bien se procesaron inmediatamente, como se describe a continuación.

Tabla 2: Medio A (esterilizado con filtro antes de su uso)

| 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 |
| Ácidos casamino | 2,5 g/l |
| Glucosa | 2 g/l |
| Tiamina·HCl | 10 mg/l |
| Ampicilina | 100 mg/l |

10 **Preparación de la genoteca de extractos parte 2—extracción:** Los sedimentos celulares de la etapa de fermentación se resuspendieron en 25 µl de tampón de lisis (tabla 3 a continuación) y se transfirieron a placas de termociclador de 384 pocillos y se sellaron. Obsérvese que el tampón contenía lisozima para ayudar en la lisis celular y DNasa para retirar el ADN del extracto. Para lisis de las células, las placas se incubaron a 37 °C durante 15 minutos, se congelaron durante la noche a -20 °C y se incubaron de nuevo a 37 °C durante 15 minutos. Se añadió sulfato de amonio (1,5 µL de una solución 2 M) y las placas se incubaron a 75 °C durante 15 minutos a fin de precipitar e inactivar proteínas contaminantes, incluyendo las nucleasas añadidas de manera exógena. Las placas se centrifugaron a 3000 x g durante 15 minutos a 4 °C y los sobrenadantes se transfirieron a una placa de termociclador de 384 pocillos nueva. Estas placas de extracto se congelaron a -20 °C para su uso posterior en cribados. Cada pocillo contenía aproximadamente 0,5-3 µM de la enzima polimerasa de la genoteca mutante.

Tabla 3. Tampón de lisis

| Componente | Concentración o porcentaje |
|---------------------|----------------------------|
| Tris pH 7.5 | 50 mM |
| EDTA | 1 mM |
| MgCl ₂ | 6 mM |
| Tween 20 | 0,5 % v/v |
| Lisozima (de polvo) | 1 mg/ml |
| DNasa I | 0,05 unidades/µl |

25 **Ejemplo 2: Identificación de ADN polimerasas mutantes con eficacia de transcripción inversa mejorada**

Cribado de genotecas de extractos para una eficacia de transcripción inversa mejorada: La genoteca de extractos se cribó comparando los valores de P_c (punto de corte) de las curvas de crecimiento generadas mediante la actividad de 5'-nucleasa fluorescente (TaqMan) de extractos enzimáticos en bruto en un sistema de RT-PCR a

partir de la amplificación de un amplicón de 240 pares de bases del transcrito JP2-5 del virus de la hepatitis C (VHC), que contenía las primeras 800 bases de la 5'NTR del genotipo 1b del VHC en pSP64 poli(A) (Promega).

Las reacciones se llevaron a cabo en el termociclador cinético LC 480 de Roche en formato de 384 pocillos, conteniendo cada pocillo 3 µl de un extracto enzimático único diluido 10 veces con tampón que contenía Tris-HCl 20 mM, pH 8, KCl 100 mM, EDTA 0,1 mM y Tween-20 al 0,1 % añadido a 12 µl de mezcla madre de RT-PCR descrita en la tabla 4). Las condiciones de termociclado fueron: 2 minutos a 50 °C (etapa de "UNG"); 2 minutos a 65 °C (etapa de "RT"); 5 ciclos de 94 °C durante 15 segundos seguido de 62 °C durante 30 segundos; y 45 ciclos de 91 °C durante 15 segundos seguido de 62 °C durante 30 segundos.

Tabla 4

| <u>Componente</u> | <u>Concentración</u> |
|----------------------|----------------------|
| Tricina pH 8.3 | 50 mM |
| KOAc | 60 mM |
| Glicerol | 5 % (v/v) |
| DMSO | 2 % (v/v) |
| Cebador 1 | 200 nM |
| Cebador 2 | 200 nM |
| Sonda de TaqMan | 100 nM |
| Aptámero | 200 nM |
| dATP | 200 µM |
| dCTP | 200 µM |
| dGTP | 200 µM |
| dUTP | 400 µM |
| UNG | 2 unidades/µl |
| ARN diana | 6666 copias/µL |
| Mg(OAc) ₂ | 2 mM |

Aproximadamente 5000 clones se cribaron usando el protocolo anterior. Se seleccionaron 40 clones del grupo original para cribar de nuevo basándose en los valores de punto de corte (Pc) más tempranos y valores de meseta fluorescente por encima de un corte arbitrario calculado mediante el procedimiento de cuantificación absoluta/máximo de la segunda derivada. Los pocillos de cultivo correspondientes a los extractos superiores se muestrearon en medio de cultivo reciente y se cultivaron nuevamente para producir nuevas placas de cultivo que contenían los mejores mutantes, así como una serie de cultivos de D580G_I709K Z05 precursores que se usarían para controles comparativos. A continuación, se usaron estas placas de cultivo para preparar nuevos extractos en bruto que se cribaron de nuevo con el mismo ARN diana y condiciones como se describe previamente para el cribado original. La tabla 5 muestra los valores de Pc promedio obtenidos a partir del incremento de señales fluorescentes debido a la hidrólisis en 5' de una sonda marcada con FAM. Los resultados muestran que el clon 0691-D20 amplifica la diana de ARN con una eficacia más alta a la del precursor D580G_I709K Z05.

Tabla 5

| <u>Clon</u> | <u>Pc promedio</u> |
|-------------|--------------------|
| 0691-D20 | 20,9 |

| | |
|-----------------|------|
| D580G_I709K Z05 | 28,0 |
|-----------------|------|

5 La secuencia de ADN de la región mutada del gen de polimerasa se secuenció para determinar la(s) mutación/mutaciones de cambio de aminoácidos que estaba(n) presente(s) en cualquier clon único. Los resultados de secuenciación revelaron que la polimerasa expresada por el clon 0691-D20 llevaba la mutación M763T, además de las mutaciones D580G e I709K precursoras. El clon 0691-D20 se eligió para pruebas adicionales, de modo que la proteína polimerasa mutante se expresó en cultivo en matraz, se purificó hasta la homogeneidad y se cuantificó.

10 **Uso del mutante D580G_I709K Z05 en RT-PCR basada en Mg²⁺:** El mutante purificado D580G_I709K_M763T Z05 se comparó con el D580G_I709K Z05 precursor en RT-PCR basada en Mg²⁺ TaqMan. La transcripción inversa y las eficacias de PCR se midieron comparando los valores de Pc de las amplificaciones del transcrito de ARN JP2-5 y del plásmido lineal de ADN pJP2-5 digeridos con la endonucleasa de restricción *EcoRI*. Los oligonucleótidos y las condiciones de mezcla madre (tabla 4) fueron los mismos que se usaron en el cribado original. Cada reacción tenía 100.000 copias de ARN de transcrito JP2-5, 100.000 copias de ADN plasmídico lineal de pJP2-5 o bien 1000 copias de ADN plasmídico lineal de pJP2-5. Se amplificaron todas las dianas con el cebador 1 y cebador 2, como se describe anteriormente, en reacciones por duplicado para generar un amplicón de 240 pares de bases. Todas las reacciones se realizaron en el termociclador Light Cycler 480 de Roche con un volumen de reacción de 15 µl. Se calcularon los puntos de corte (Pc) mediante el procedimiento de cuantificación absoluta/máximo de la segunda derivada y se promediaron. Las amplificaciones se llevaron a cabo usando un intervalo de concentraciones de ADN polimerasa desde 5 nM-40 nM. Las condiciones de termociclado fueron: 2 minutos a 50 °C (etapa de "UNG"); 2 minutos a 65 °C (etapa de "RT"); 5 ciclos de 94 °C durante 15 segundos seguido de 62 °C durante 30 segundos; y 45 ciclos de 91 °C durante 15 segundos seguido de 62 °C durante 30 segundos. La tabla 6 muestra los valores de Pc obtenidos a partir del incremento de señales fluorescentes debido a la escisión de la sonda TaqMan en una condición enzimática a 20 nM.

Tabla 6

| Enzima | Pc de 10 ⁵ copias de ARN | Pc de 10 ⁵ copias de ADN | Pc de 10 ³ copias de ADN |
|-----------------------|--|--|--|
| D580G_I709K Z05 | 28,7 | 18,8 | 25,7 |
| D580G_I709K_M763T Z05 | 20,6 | 18,9 | 25,8 |

25 Los resultados indican que D580G_I709K_M763T Z05 mutante permite una amplificación más eficaz de la diana de ARN sin comprometer la eficacia de la PCR en una diana de ADN, en comparación con la enzima precursora D580G_I709K.

30 Se entiende que los ejemplos y modos de realización descritos en el presente documento son únicamente para propósitos ilustrativos y que se sugerirán diversas modificaciones o cambios a la luz de los mismos a expertos en la técnica.

LISTADO DE SECUENCIAS

<110> F. Hoffmann-La Roche AG Roche Diagnostics GmbH

5 <120> ADN polimerasas con actividad mejorada

<130> 31275 WO-HS

<150> US 61/736.737

10 <151> 13-12-2012

<160> 39

<170> FastSEQ para Windows versión 4.0

<210> 1

15 <211> 834

<212> PRT

<213> Thermus sp.

<220>

20 <223> ADN polimerasa Z05 de Thermus sp. (Z05)

<400> 1

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 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 | | | | | 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 |

ES 2 624 406 T3

Lys Leu Phe Pro His Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
 770 775 780
 Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
 785 790 795 800
 Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
 805 810 815
 Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala
 820 825 830
 Lys Gly

<210> 2
 <211> 832
 5 <212> PRT
 <213> Thermus aquaticus

<220>
 <223> ADN polimerasa de Thermus aquaticus (Taq)

10 <400> 2

Met Arg Gly 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 His 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 Asp Ala Val Ile Val
 50 55 60
 Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
 65 70 75 80
 Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
 85 90 95
 Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
 100 105 110
 Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
 115 120 125
 Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
 130 135 140
 Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
 145 150 155 160
 Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
 165 170 175
 Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
 180 185 190
 Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
 195 200 205
 Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
 210 215 220
 Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
 225 230 235 240
 Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
 245 250 255
 Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
 260 265 270
 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
 275 280 285
 Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
 290 295 300
 Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
 305 310 315 320
 Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
 325 330 335
 Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu

ES 2 624 406 T3

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | 340 | | | | | 345 | | | | 350 | | | |
| Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | Pro |
| | | 355 | | | | | | 360 | | | | 365 | | | |
| Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | Asn |
| | | 370 | | | | 375 | | | | | | 380 | | | |
| Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | Glu |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | Leu |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Trp | Gly | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | Glu |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | Gly |
| | | 435 | | | | | 440 | | | | | 445 | | | |
| Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | Ala |
| | | 450 | | | | 455 | | | | | 460 | | | | |
| Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | His |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | Asp |
| | | | | 485 | | | | 490 | | | | | | 495 | |
| Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys | Arg |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | Ile |
| | | 515 | | | | | 520 | | | | | 525 | | | |
| Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | Thr |
| | | 530 | | | | 535 | | | | | 540 | | | | |
| Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | Leu |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | Ser |
| | | | | 565 | | | | | 570 | | | | | | 575 |
| Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | Gln |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | Ala |
| | | | | 595 | | | 600 | | | | | 605 | | | |
| Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | Gly |
| | | | | | | 615 | | | | | 620 | | | | |
| Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | Thr |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | Pro |
| | | | | 645 | | | | | 650 | | | | | 655 | |
| Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | Gly |
| | | | | 660 | | | | 665 | | | | | 670 | | |
| Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | Arg |
| | | | | | 690 | | 695 | | | | 700 | | | | |
| Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr | Val |
| 705 | | | | | 710 | | | | 715 | | | | | | 720 |
| Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala | Arg |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | Pro |
| | | | | 740 | | | | 745 | | | | | 750 | | |
| Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | Leu |
| | | 755 | | | | | 760 | | | | | 765 | | | |
| Phe | Pro | Arg | Leu | Glu | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val | His |
| | | 770 | | | | 775 | | | | | 780 | | | | |
| Asp | Glu | Leu | Val | Leu | Glu | Ala | Pro | Lys | Glu | Arg | Ala | Glu | Ala | Val | Ala |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 |
| Arg | Leu | Ala | Lys | Glu | Val | Met | Glu | Gly | Val | Tyr | Pro | Leu | Ala | Val | Pro |
| | | | | 805 | | | | | 810 | | | | | 815 | |
| Leu | Glu | Val | Glu | Val | Gly | Ile | Gly | Glu | Asp | Trp | Leu | Ser | Ala | Lys | Glu |
| | | | 820 | | | | | 825 | | | | | 830 | | |

<210> 3
 <211> 830
 <212> PRT
 <213> Thermus filiformis

5

<220>
 <223> ADN polimerasa de Thermus filiformis (Tfi)

<400> 3

10

```

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

ES 2 624 406 T3

Arg Pro Leu Ala Gln Val Leu Ala His Met Glu Ala Thr Gly Val Arg
 435 440 445
 Leu Asp Val Pro Tyr Leu Glu Ala Leu Ser Gln Glu Val Ala Phe Glu
 450 455 460
 Leu Glu Arg Leu Glu Ala Glu Val His Arg Leu Ala Gly His Pro Phe
 465 470 475 480
 Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
 485 490 495
 Gly Leu Pro Pro Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
 500 505 510
 Ser Ala Ala Val Leu Glu Leu Leu Arg Glu Ala His Pro Ile Val Gly
 515 520 525
 Arg Ile Leu Glu Tyr Arg Glu Leu Met Lys Leu Lys Ser Thr Tyr Ile
 530 535 540
 Asp Pro Leu Pro Arg Leu Val His Pro Lys Thr Gly Arg Leu His Thr
 545 550 555 560
 Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
 565 570 575
 Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
 580 585 590
 Arg Lys Ala Phe Ile Ala Glu Glu Gly His Leu Leu Val Ala Leu Asp
 595 600 605
 Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
 610 615 620
 Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
 625 630 635 640
 Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
 645 650 655
 Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
 660 665 670
 Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
 675 680 685
 Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
 690 695 700
 Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
 705 710 715 720
 Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys
 725 730 735
 Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
 740 745 750
 Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
 755 760 765
 Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu
 770 775 780
 Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu
 785 790 795 800
 Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu
 805 810 815
 Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Glu
 820 825 830

<210> 4

<211> 831

<212> PRT

5 <213> Thermus flavus

<220>

<223> ADN polimerasa de Thermus flavus (Tfl)

10 <400> 4

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

ES 2 624 406 T3

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
 20 25 30
 Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
 35 40 45
 Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Val Val Val Val
 50 55 60
 Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr
 65 70 75 80
 Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
 85 90 95
 Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val
 100 105 110
 Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Arg Ala
 115 120 125
 Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
 130 135 140
 Tyr Gln Leu Leu Ser Glu Arg Ile Ala Ile Leu His Pro Glu Gly Tyr
 145 150 155 160
 Leu Ile Thr Pro Ala Trp Leu Tyr Glu Lys Tyr Gly Leu Arg Pro Glu
 165 170 175
 Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile
 180 185 190
 Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Gln Arg Leu Ile Arg
 195 200 205
 Glu Trp Gly Ser Leu Glu Asn Leu Phe Gln His Leu Asp Gln Val Lys
 210 215 220
 Pro Ser Leu Arg Glu Lys Leu Gln Ala Gly Met Glu Ala Leu Ala Leu
 225 230 235 240
 Ser Arg Lys Leu Ser Gln Val His Thr Asp Leu Pro Leu Glu Val Asp
 245 250 255
 Phe Gly Arg Arg Arg Thr Pro Asn Leu Glu Gly Leu Arg Ala Phe Leu
 260 265 270
 Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu
 275 280 285
 Gly Pro Lys Ala Ala Glu Glu Ala Pro Trp Pro Pro Glu Gly Ala
 290 295 300
 Phe Leu Gly Phe Ser Phe Ser Arg Pro Glu Pro Met Trp Ala Glu Leu
 305 310 315 320
 Leu Ala Leu Ala Gly Ala Trp Glu Gly Arg Leu His Arg Ala Gln Asp
 325 330 335
 Pro Leu Arg Gly Leu Arg Asp Leu Lys Gly Val Arg Gly Ile Leu Ala
 340 345 350
 Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Phe Pro
 355 360 365
 Glu Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr
 370 375 380
 Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp
 385 390 395 400
 Ala Gly Glu Arg Ala Leu Leu Ala Glu Arg Leu Phe Gln Thr Leu Lys
 405 410 415
 Glu Arg Leu Lys Gly Glu Glu Arg Leu Leu Trp Leu Tyr Glu Glu Val
 420 425 430
 Glu Lys Pro Leu Ser Arg Val Leu Ala Arg Met Glu Ala Thr Gly Val
 435 440 445
 Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Val Glu Ala
 450 455 460
 Glu Val Arg Gln Leu Glu Glu Val Phe Arg Leu Ala Gly His Pro
 465 470 475 480
 Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu
 485 490 495
 Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser
 500 505 510
 Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val

ES 2 624 406 T3

```

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

```

<210> 5

<211> 830

<212> PRT

5 <213> Thermus sp.

<220>

<223> ADN polimerasa de Thermus sp. sps17 (Sps17)

10 <400> 5

```

Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
 1                5                10                15
His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu Thr Thr
          20                25                30
Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
          35                40                45
Leu Lys Ala Leu Lys Glu Asp Gly Glu Val Ala Ile Val Val Phe Asp
   50                55                60
Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr Lys Ala
65                70                75                80
Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile
          85                90                95
Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val Pro Gly

```


ES 2 624 406 T3

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
 610 615 620
 Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
 625 630 635 640
 Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
 645 650 655
 Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
 660 665 670
 Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
 675 680 685
 Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
 690 695 700
 Ile Ala Lys Thr Leu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
 705 710 715 720
 Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys
 725 730 735
 Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
 740 745 750
 Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
 755 760 765
 Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu
 770 775 780
 Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu
 785 790 795 800
 Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu
 805 810 815
 Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Ala
 820 825 830

<210> 6

<211> 834

<212> PRT

5 <213> Thermus thermophilus

<220>

<223> ADN polimerasa de Thermus thermophilus (Tth)

10 <400> 6

Met Glu 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 Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys 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 Arg
 165 170 175
 Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
 180 185 190

ES 2 624 406 T3

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 Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
 225 230 235 240
 Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Leu Ala Gln Gly 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 Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
 290 295 300
 Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
 305 310 315 320
 Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg
 325 330 335
 Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
 340 345 350
 Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
 355 360 365
 Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
 370 375 380
 Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
 385 390 395 400
 Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
 405 410 415
 Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
 450 455 460
 Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
 515 520 525
 Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
 530 535 540
 Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
 595 600 605
 Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
 645 650 655
 Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
 675 680 685
 Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys

690 695 700
 Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
 725 730 735
 Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
 770 775 780
 Val His Asp Glu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
 785 790 795 800
 Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
 805 810 815
 Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
 820 825 830
 Lys Gly

<210> 7

<211> 834

5 <212> PRT

<213> Thermus caldophilus

<220>

10 <223> ADN polimerasa de Thermus caldophilus (Tca)

<400> 7

Met Glu 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
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
 115 120 125
 Asn Pro Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
 130 135 140
 Asp Leu Asp 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 Gln 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 Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
 225 230 235 240
 Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg
 260 265 270

ES 2 624 406 T3

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
 275 280 285
 Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
 290 295 300
 Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
 305 310 315 320
 Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg
 325 330 335
 Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
 340 345 350
 Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
 355 360 365
 Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
 370 375 380
 Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
 385 390 395 400
 Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
 405 410 415
 Asn Leu Leu Lys Arg Leu Gln Gly Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
 450 455 460
 Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
 515 520 525
 Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
 530 535 540
 Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Asn Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
 595 600 605
 Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
 645 650 655
 Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
 675 680 685
 Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
 690 695 700
 Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
 725 730 735
 Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln

ES 2 624 406 T3

| | | | | | | | | | | | | | | | | | |
|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| | 770 | | | | | 775 | | | | | | 780 | | | | | |
| | Val | His | Asp | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Gly | Ala | Glu | Glu | |
| | 785 | | | | | | 790 | | | | | 795 | | | | 800 | |
| | Val | Ala | Ala | Leu | Ala | Lys | Glu | Ala | Met | Glu | Lys | Ala | Tyr | Pro | Leu | Ala | |
| | | | | | 805 | | | | | 810 | | | | | 815 | | |
| | Val | Pro | Leu | Glu | Val | Glu | Val | Gly | Met | Gly | Glu | Asp | Trp | Leu | Ser | Ala | |
| | | | | 820 | | | | | 825 | | | | | 830 | | | |

Lys Gly

<210> 8

<211> 14

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> motivo de ADN polimerasa mejorada sintética

<220>

<221> VARIANTE

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

15 <223> Xaa = Leu o Ile

<220>

<221> VARIANTE

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

20 <223> Xaa = cualquier aminoácido distinto de Met o Ile

<220>

<221> VARIANTE

<222> (5)...(5)

25 <223> Xaa = Leu, Ile o Lys

<220>

<221> VARIANTE

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

30 <223> Xaa = Val o Ile

<220>

<221> VARIANTE

<222> (9)...(9)

35 <223> Xaa = Lys, Arg, Glu, Asp, Asn o Gln

<220>

<221> VARIANTE

<222> (10)...(10)

40 <223> Xaa= Leu o Ile

<220>

<221> VARIANTE

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

45 <223> Xaa = Phe, Asp, His, Ser o Asn

<220>

<221> VARIANTE

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

50 <223> Xaa = Pro, Arg, Glu, Asn, Val o Ala

<220>

<221> VARIANTE

<222> (13)...(13)

55 <223> Xaa = His, Arg, Glu o Gln

<400> 8

Asp Xaa Xaa Lys Xaa Ala Met Xaa Xaa Xaa Xaa Xaa Xaa Leu

1

5

10

<210> 9

<211> 14
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> motivo de ADN polimerasa mejorada sintética

<220>
 <221> VARIANTE
 10 <222> (3)...(3)
 <223> Xaa = cualquier aminoácido distinto de Met

<220>
 <221> VARIANTE
 15 <222> (9)...(9)
 <223> Xaa = Lys o Arg

<220>
 <221> VARIANTE
 20 <222> (13)...(13)
 <223> Xaa = His o Arg

<400> 9

Asp Leu Xaa Lys Leu Ala Met Val Xaa Leu Phe Pro Xaa Leu
 25 1 5 10

<210> 10
 <211> 14
 <212> PRT
 30 <213> Secuencia artificial

<220>
 <223> motivo de ADN polimerasa mejorada sintética

35 <220>
 <221> VARIANTE
 <222> (3)...(3)
 <223> Xaa = cualquier aminoácido distinto de Met

40 <400> 10

Asp Leu Xaa Lys Leu Ala Met Val Lys Leu Phe Pro His Leu
 1 5 10

<210> 11
 <211> 14
 <212> PRT
 45 <213> Secuencia artificial

<220>
 50 <223> motivo de ADN polimerasa mejorada sintética

<400> 11

Asp Leu Thr Lys Leu Ala Met Val Lys Leu Phe Pro His Leu
 55 1 5 10

<210> 12
 <211> 35
 <212> PRT
 <213> Secuencia artificial

60 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa Z05 de Thermus sp. (Z05)

ES 2 624 406 T3

<400> 12
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
 1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro His Leu Arg Glu Met Gly Ala
 20 25 30
Arg Met Leu
 35
 <210> 13
 <211> 35
 5 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de *Thermus aquaticus* (Taq)
 10
 <400> 13
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
 1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala
 20 25 30
Arg Met Leu
 35
 <210> 14
 <211> 35
 15 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de *Thermus filiformis* (Tfi)
 20
 <400> 14
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
 1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Pro Leu Gly Val
 20 25 30
Arg Ile Leu
 35
 <210> 15
 <211> 35
 25 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de *Thermus flavus* (Tfl)
 30
 <400> 15
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
 1 5 10 15
Lys Leu Ala Met Val Arg Leu Phe Pro Arg Leu Gln Glu Leu Gly Ala
 20 25 30
Arg Met Leu
 35
 35 <210> 16
 <211> 35
 <212> PRT
 <213> Secuencia artificial

 40 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de *Thermus sp. sps17* (Sps17)

 45

ES 2 624 406 T3

<400> 16
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala
20 25 30
Arg Met Leu
35

5 <210> 17
 <211> 35
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Thermus thermophilus (Tth)

<400> 17
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala
20 25 30
Arg Met Leu
35

15 <210> 18
 <211> 35
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Thermus caldophilus (Tca)

<400> 18
Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met
1 5 10 15
Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala
20 25 30
Arg Met Leu
35

25 <210> 19
 <211> 37
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Thermotoga maritima (Tma)

<400> 19
Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp Ile Ile
1 5 10 15
Lys Leu Ala Met Ile Glu Ile Asp Arg Glu Leu Lys Glu Arg Lys Met
20 25 30
Arg Ser Lys Met Ile
35

35 <210> 20
 <211> 37
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Thermotoga neopolitana (Tne)

45

ES 2 624 406 T3

<400> 20
Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp Ile Ile
 1 5 10 15
Lys Leu Ala Met Ile Asp Ile Asp Glu Glu Leu Arg Lys Arg Asn Met
 20 25 30
Lys Ser Arg Met Ile
 35
 <210> 21
 <211> 37
 5 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Thermosipho africanus (Taf)
 10
 <400> 21
Arg Ile Ala Val Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp Ile Ile
 1 5 10 15
Lys Ile Ala Met Ile Asn Ile His Asn Arg Leu Lys Lys Glu Asn Leu
 20 25 30
Arg Ser Lys Met Ile
 35
 <210> 22
 <211> 8
 15 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> motivo A de sitio activo de ADN polimerasa conservado sintético
 20
 <400> 22
Asp Tyr Ser Gln Ile Glu Leu Arg
 1 5
 <210> 23
 <211> 35
 25 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Deinococcus radiodurans (Dra)
 30
 <400> 23
Arg Leu Ala Tyr Asn Met Pro Ile Gln Gly Thr Ala Ala Asp Ile Met
 1 5 10 15
Lys Leu Ala Met Val Gln Leu Asp Pro Gln Leu Asp Ala Ile Gly Ala
 20 25 30
Arg Met Leu
 35
 <210> 24
 <211> 37
 35 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> región de dominio de polimerasa sintética de ADN polimerasa de Bacillus stearothermophilus (Bst)
 40
 <400> 24
Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala Ala Asp Ile Ile
 1 5 10 15
Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg Glu Glu Arg Leu
 20 25 30
Gln Ala Arg Leu Leu
 35
 <210> 25
 45 <211> 37
 <212> PRT

ES 2 624 406 T3

<213> Secuencia artificial

<220>

<223> región de dominio de polimerasa sintética de ADN polimerasa de Bacillus caldotenax (Bca)

5

<400> 25
Arg Met Ala Met Asn Thr Pro Ile Gln Gly Ser Ala Ala Asp Ile Ile
1 5 10 15
Lys Lys Ala Met Ile Asp Leu Asn Ala Arg Leu Lys Glu Glu Arg Leu
20 25 30
Gln Ala Arg Leu Leu
35

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

10

<220>

<223> motivo de consenso de ADN polimerasa de región de dominio de polimerasa sintética

15

<220>

<221> VARIANTE

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

<223> Xaa = Leu o Ile

20

<220>

<221> VARIANTE

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

<223> Xaa = Met o Ile

25

<220>

<221> VARIANTE

<222> (5)...(5)

<223> Xaa = Leu, Ile o Lys

30

<220>

<221> VARIANTE

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

<223> Xaa = Val o Ile

35

<220>

<221> VARIANTE

<222> (9)...(9)

<223> Xaa = Lys, Arg, Glu, Asp, Asn o Gln

40

<220>

<221> VARIANTE

<222> (10)...(10)

<223> Xaa= Leu o Ile

45

<220>

<221> VARIANTE

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

<223> Xaa = Phe, Asp, His, Ser o Asn

50

<220>

<221> VARIANTE

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

<223> Xaa = Pro, Arg, Glu, Asn, Val o Ala

55

<220>

<221> VARIANTE

<222> (13)...(13)

<223> Xaa = His, Arg, Glu o Gln

60

<400> 26

ES 2 624 406 T3

Asp Xaa Xaa Lys Xaa Ala Met Xaa Xaa Xaa Xaa Xaa Xaa Leu
1 5 10

<210> 27

<211> 893

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> ADN polimerasa CS5 quimérica sintética derivada de dominio de 5'-nucleasa de extremo N de Thermus sp. Z05 y 3'-5' exonucleasa de extremo C y dominios de polimerasa de ADN polimerasas de Thermotoga maritima

<400> 27

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
100 105 110
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

ES 2 624 406 T3

<400> 28
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
100 105 110
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

- 5
 <220>
 <221> VARIANTE
 <222> (2)...(2)
 <223> Xaa = Trp o Tyr
- 10
 <220>
 <221> VARIANTE
 <222> (3)...(3)
 <223> Xaa = cualquier aminoácido distinto de Ile, Leu o Met
- 15
 <220>
 <221> VARIANTE
 <222> (4)...(4)
 <223> Xaa = Glu, Ala, Gln, Lys, Asn o Asp
- 20
 <220>
 <221> VARIANTE
 <222> (5)...(5)
 <223> Xaa = Lys, Gly, Arg, Gln, His o Asn
- 25
 <220>
 <221> VARIANTE
 <222> (6)...(6)
 <223> Xaa = Thr, Val, Met o Ile
- 30
 <220>
 <221> VARIANTE
 <222> (7)...(7)
 <223> Xaa = Leu, Val o Lys
- 35
 <220>
 <221> VARIANTE
 <222> (8)...(8)
 <223> Xaa = Glu, Ser, Ala, Asp o Gln
- 40
 <220>
 <221> VARIANTE
 <222> (9)...(9)
 <223> Xaa = Glu o Phe
- 45
 <220>
 <221> VARIANTE
 <222> (10)...(10)
 <223> Xaa = Gly o Ala
- 50
 <220>
 <221> VARIANTE
 <222> (11)...(11)
 <223> Xaa = Arg o Lys
- 55
 <220>
 <221> VARIANTE
 <222> (12)...(12)
 <223> Xaa = Lys, Arg, Glu, Thr o Gln
- 60
 <220>
 <221> VARIANTE
 <222> (13)...(13)
 <223> Xaa = Arg, Lys o His
- 60
 <220>
 <221> VARIANTE
 <222> (17)...(17)
 <223> Xaa = Glu, Arg o Thr

ES 2 624 406 T3

<400> 29
Xaa Xaa Gly Tyr Val
1 5 10 15
Xaa Thr Leu

5
 <210> 30
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10
 <220>
 <223> cebador directo de amplificación por PCR propensa a errores (mutagénica) sintético

15
 <400> 30
 ctacctctg gaccctcca a 21

20
 <210> 31
 <211> 25
 <212> ADN
 <213> Secuencia artificial

25
 <220>
 <223> cebador inverso de amplificación por PCR propensa a errores (mutagénica) sintético

30
 <400> 31
 ataaccaact gtagtggcg tgtaa 25

35
 <210> 32
 <211> 921
 <212> PRT
 <213> Deinococcus radiodurans

35
 <220>
 <223> ADN polimerasa de Deinococcus radiodurans (Dra)

<400> 32
Met Ala Asp Ala Ser Pro Asp Pro Ser Lys Pro Asp Ala Leu Val Leu
1 5 10 15
Ile Asp Gly His Ala Leu Ala Phe Arg Ser Tyr Phe Ala Leu Pro Pro
20 25 30
Leu Asn Asn Ser Lys Gly Glu Met Thr Asp Ala Ile Val Gly Phe Met
35 40 45
Lys Leu Leu Leu Arg Leu Ala Arg Gln Lys Ser Asn Gln Val Ile Val
50 55 60
Val Phe Asp Pro Pro Val Lys Thr Leu Arg His Glu Gln Tyr Glu Gly
65 70 75 80
Tyr Lys Ser Gly Arg Ala Gln Thr Pro Glu Asp Leu Arg Gly Gln Ile
85 90 95
Asn Arg Ile Arg Ala Leu Val Asp Ala Leu Gly Phe Pro Arg Leu Glu
100 105 110
Glu Pro Gly Tyr Glu Ala Asp Asp Val Ile Ala Ser Leu Thr Arg Met
115 120 125

ES 2 624 406 T3

Ala Glu Gly Lys Gly Tyr Glu Val Arg Ile Val Thr Ser Asp Arg Asp
130 135 140
Ala Tyr Gln Leu Leu Asp Glu His Val Lys Val Ile Ala Asn Asp Phe
145 150 155 160
Ser Leu Ile Gly Pro Ala Gln Val Glu Glu Lys Tyr Gly Val Thr Val
165 170 175
Arg Gln Trp Val Asp Tyr Arg Ala Leu Thr Gly Asp Ala Ser Asp Asn
180 185 190
Ile Pro Gly Ala Lys Gly Ile Gly Pro Lys Thr Ala Ala Lys Leu Leu
195 200 205
Gln Glu Tyr Gly Thr Leu Glu Lys Val Tyr Glu Ala Ala His Ala Gly
210 215 220
Thr Leu Lys Pro Asp Gly Thr Arg Lys Lys Leu Leu Asp Ser Glu Glu
225 230 235 240
Asn Val Lys Phe Ser His Asp Leu Ser Cys Met Val Thr Asp Leu Pro
245 250 255
Leu Asp Ile Glu Phe Gly Val Arg Arg Leu Pro Asp Asn Pro Leu Val
260 265 270
Thr Glu Asp Leu Leu Thr Glu Leu Glu Leu His Ser Leu Arg Pro Met
275 280 285
Ile Leu Gly Leu Asn Gly Pro Glu Gln Asp Gly His Ala Pro Asp Asp
290 295 300
Leu Leu Glu Arg Glu His Ala Gln Thr Pro Glu Glu Asp Glu Ala Ala
305 310 315 320
Ala Leu Pro Ala Phe Ser Ala Pro Glu Leu Ala Glu Trp Gln Thr Pro
325 330 335
Ala Glu Gly Ala Val Trp Gly Tyr Val Leu Ser Arg Glu Asp Asp Leu
340 345 350
Thr Ala Ala Leu Leu Ala Ala Ala Thr Phe Glu Asp Gly Val Ala Arg
355 360 365
Pro Ala Arg Val Ser Glu Pro Asp Glu Trp Ala Gln Ala Glu Ala Pro
370 375 380
Glu Asn Leu Phe Gly Glu Leu Leu Pro Ser Asp Lys Pro Leu Thr Lys
385 390 395 400
Lys Glu Gln Lys Ala Leu Glu Lys Ala Gln Lys Asp Ala Glu Lys Ala
405 410 415
Arg Ala Lys Leu Arg Glu Gln Phe Pro Ala Thr Val Asp Glu Ala Glu
420 425 430
Phe Val Gly Gln Arg Thr Val Thr Ala Ala Ala Ala Lys Ala Leu Ala
435 440 445
Ala His Leu Ser Val Arg Gly Thr Val Val Glu Pro Gly Asp Asp Pro
450 455 460
Leu Leu Tyr Ala Tyr Leu Leu Asp Pro Ala Asn Thr Asn Met Pro Val
465 470 475 480
Val Ala Lys Arg Tyr Leu Asp Arg Glu Trp Pro Ala Asp Ala Pro Thr
485 490 495
Arg Ala Ala Ile Thr Gly His Leu Val Arg Glu Leu Pro Pro Leu Leu
500 505 510
Asp Asp Ala Arg Arg Lys Met Tyr Asp Glu Met Glu Lys Pro Leu Ser
515 520 525
Gly Val Leu Gly Arg Met Glu Val Arg Gly Val Gln Val Asp Ser Asp
530 535 540
Phe Leu Gln Thr Leu Ser Ile Gln Ala Gly Val Arg Leu Ala Asp Leu
545 550 555 560
Glu Ser Gln Ile His Glu Tyr Ala Gly Glu Glu Phe His Ile Arg Ser
565 570 575
Pro Lys Gln Leu Glu Thr Val Leu Tyr Asp Lys Leu Glu Leu Ala Ser
580 585 590
Ser Lys Lys Thr Lys Leu Thr Gly Gln Arg Ser Thr Ala Val Ser Ala
595 600 605
Leu Glu Pro Leu Arg Asp Ala His Pro Ile Ile Pro Leu Val Leu Glu
610 615 620
Phe Arg Glu Leu Asp Lys Leu Arg Gly Thr Tyr Leu Asp Pro Ile Pro

ES 2 624 406 T3

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Phe | Glu | Lys | Val | Asn | Ile | Ile | Thr | Gly | Asp | Lys | Asp | Leu | Leu | Gln | Leu |
| | 130 | | | | | | 135 | | | | | 140 | | | |
| Val | Ser | Asp | Lys | Val | Phe | Val | Trp | Arg | Val | Glu | Arg | Gly | Ile | Thr | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Leu | Val | Leu | Tyr | Asp | Arg | Asn | Lys | Val | Ile | Glu | Lys | Tyr | Gly | Ile | Tyr |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Pro | Glu | Gln | Phe | Lys | Asp | Tyr | Leu | Ser | Leu | Val | Gly | Asp | Gln | Ile | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Asn | Ile | Pro | Gly | Val | Lys | Gly | Ile | Gly | Lys | Lys | Thr | Ala | Val | Ser | Leu |
| | 195 | | | | | | 200 | | | | | 205 | | | |
| Leu | Lys | Lys | Tyr | Asn | Ser | Leu | Glu | Asn | Val | Leu | Lys | Asn | Ile | Asn | Leu |
| | 210 | | | | | 215 | | | | | | 220 | | | |
| Leu | Thr | Glu | Lys | Leu | Arg | Arg | Leu | Leu | Glu | Asp | Ser | Lys | Glu | Asp | Leu |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Gln | Lys | Ser | Ile | Glu | Leu | Val | Glu | Leu | Ile | Tyr | Asp | Val | Pro | Met | Asp |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Val | Glu | Lys | Asp | Glu | Ile | Ile | Tyr | Arg | Gly | Tyr | Asn | Pro | Asp | Lys | Leu |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Leu | Lys | Val | Leu | Lys | Lys | Tyr | Glu | Phe | Ser | Ser | Ile | Ile | Lys | Glu | Leu |
| | 275 | | | | | | 280 | | | | | 285 | | | |
| Asn | Leu | Gln | Glu | Lys | Leu | Glu | Lys | Glu | Tyr | Ile | Leu | Val | Asp | Asn | Glu |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Asp | Lys | Leu | Lys | Lys | Leu | Ala | Glu | Glu | Ile | Glu | Lys | Tyr | Lys | Thr | Phe |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Ser | Ile | Asp | Thr | Glu | Thr | Thr | Ser | Leu | Asp | Pro | Phe | Glu | Ala | Lys | Leu |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Val | Gly | Ile | Ser | Ile | Ser | Thr | Met | Glu | Gly | Lys | Ala | Tyr | Tyr | Ile | Pro |
| | | | 340 | | | | | 345 | | | | | 350 | | |
| Val | Ser | His | Phe | Gly | Ala | Lys | Asn | Ile | Ser | Lys | Ser | Leu | Ile | Asp | Lys |
| | 355 | | | | | | 360 | | | | | 365 | | | |
| Phe | Leu | Lys | Gln | Ile | Leu | Gln | Glu | Lys | Asp | Tyr | Asn | Ile | Val | Gly | Gln |
| | 370 | | | | | 375 | | | | | 380 | | | | |
| Asn | Leu | Lys | Phe | Asp | Tyr | Glu | Ile | Phe | Lys | Ser | Met | Gly | Phe | Ser | Pro |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Asn | Val | Pro | His | Phe | Asp | Thr | Met | Ile | Ala | Ala | Tyr | Leu | Leu | Asn | Pro |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Asp | Glu | Lys | Arg | Phe | Asn | Leu | Glu | Glu | Leu | Ser | Leu | Lys | Tyr | Leu | Gly |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Tyr | Lys | Met | Ile | Ser | Phe | Asp | Glu | Leu | Val | Asn | Glu | Asn | Val | Pro | Leu |
| | 435 | | | | | | 440 | | | | | 445 | | | |
| Phe | Gly | Asn | Asp | Phe | Ser | Tyr | Val | Pro | Leu | Glu | Arg | Ala | Val | Glu | Tyr |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Ser | Cys | Glu | Asp | Ala | Asp | Val | Thr | Tyr | Arg | Ile | Phe | Arg | Lys | Leu | Gly |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| Arg | Lys | Ile | Tyr | Glu | Asn | Glu | Met | Glu | Lys | Leu | Phe | Tyr | Glu | Ile | Glu |
| | | | | 485 | | | | | 490 | | | | | 495 | |
| Met | Pro | Leu | Ile | Asp | Val | Leu | Ser | Glu | Met | Glu | Leu | Asn | Gly | Val | Tyr |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Phe | Asp | Glu | Glu | Tyr | Leu | Lys | Glu | Leu | Ser | Lys | Lys | Tyr | Gln | Glu | Lys |
| | 515 | | | | | | 520 | | | | | | 525 | | |
| Met | Asp | Gly | Ile | Lys | Glu | Lys | Val | Phe | Glu | Ile | Ala | Gly | Glu | Thr | Phe |
| | 530 | | | | | 535 | | | | | 540 | | | | |
| Asn | Leu | Asn | Ser | Ser | Thr | Gln | Val | Ala | Tyr | Ile | Leu | Phe | Glu | Lys | Leu |
| 545 | | | | | 550 | | | | | | 555 | | | | 560 |
| Asn | Ile | Ala | Pro | Tyr | Lys | Lys | Thr | Ala | Thr | Gly | Lys | Phe | Ser | Thr | Asn |
| | | | | 565 | | | | | 570 | | | | | 575 | |
| Ala | Glu | Val | Leu | Glu | Glu | Leu | Ser | Lys | Glu | His | Glu | Ile | Ala | Lys | Leu |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Leu | Leu | Glu | Tyr | Arg | Lys | Tyr | Gln | Lys | Leu | Lys | Ser | Thr | Tyr | Ile | Asp |
| | 595 | | | | | | 600 | | | | | | 605 | | |
| Ser | Ile | Pro | Leu | Ser | Ile | Asn | Arg | Lys | Thr | Asn | Arg | Val | His | Thr | Thr |
| | 610 | | | | | 615 | | | | | | | 620 | | |

ES 2 624 406 T3

Phe His Gln Thr Gly Thr Ser Thr Gly Arg Leu Ser Ser Ser Asn Pro
625 630 635 640
Asn Leu Gln Asn Leu Pro Thr Arg Ser Glu Glu Gly Lys Glu Ile Arg
645 650 655
Lys Ala Val Arg Pro Gln Arg Gln Asp Trp Trp Ile Leu Gly Ala Asp
660 665 670
Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Val Ser Lys Asp Glu
675 680 685
Asn Leu Leu Lys Ala Phe Lys Glu Asp Leu Asp Ile His Thr Ile Thr
690 695 700
Ala Ala Lys Ile Phe Gly Val Ser Glu Met Phe Val Ser Glu Gln Met
705 710 715 720
Arg Arg Val Gly Lys Met Val Asn Phe Ala Ile Ile Tyr Gly Val Ser
725 730 735
Pro Tyr Gly Leu Ser Lys Arg Ile Gly Leu Ser Val Ser Glu Thr Lys
740 745 750
Lys Ile Ile Asp Asn Tyr Phe Arg Tyr Tyr Lys Gly Val Phe Glu Tyr
755 760 765
Leu Lys Arg Met Lys Asp Glu Ala Arg Lys Lys Gly Tyr Val Thr Thr
770 775 780
Leu Phe Gly Arg Arg Arg Tyr Ile Pro Gln Leu Arg Ser Lys Asn Gly
785 790 795 800
Asn Arg Val Gln Glu Gly Glu Arg Ile Ala Val Asn Thr Pro Ile Gln
805 810 815
Gly Thr Ala Ala Asp Ile Ile Lys Ile Ala Met Ile Asn Ile His Asn
820 825 830
Arg Leu Lys Lys Glu Asn Leu Arg Ser Lys Met Ile Leu Gln Val His
835 840 845
Asp Glu Leu Val Phe Glu Val Pro Asp Asn Glu Leu Glu Ile Val Lys
850 855 860
Asp Leu Val Arg Asp Glu Met Glu Asn Ala Val Lys Leu Asp Val Pro
865 870 875 880
Leu Lys Val Asp Val Tyr Tyr Gly Lys Glu Trp Glu
885 890

<210> 34

<211> 893

<212> PRT

5 <213> Thermotoga maritima

<220>

<223> ADN polimerasa de Thermotoga maritima (Tma)

10 <400> 34

Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala
1 5 10 15
Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr
20 25 30
Asn Ala Thr Tyr Gly Val Ala Arg Met Leu Val Arg Phe Ile Lys Asp
35 40 45
His Ile Ile Val Gly Lys Asp Tyr Val Ala Val Ala Phe Asp Lys Lys
50 55 60
Ala Ala Thr Phe Arg His Lys Leu Leu Glu Thr Tyr Lys Ala Gln Arg
65 70 75 80
Pro Lys Thr Pro Asp Leu Leu Ile Gln Gln Leu Pro Tyr Ile Lys Lys
85 90 95
Leu Val Glu Ala Leu Gly Met Lys Val Leu Glu Val Glu Gly Tyr Glu
100 105 110
Ala Asp Asp Ile Ile Ala Thr Leu Ala Val Lys Gly Leu Pro Leu Phe
115 120 125
Asp Glu Ile Phe Ile Val Thr Gly Asp Lys Asp Met Leu Gln Leu Val
130 135 140

ES 2 624 406 T3

Asn Glu Lys Ile Lys Val Trp Arg Ile Val Lys Gly Ile Ser Asp Leu
 145 150 155 160
 Glu Leu Tyr Asp Ala Gln Lys Val Lys Glu Lys Tyr Gly Val Glu Pro
 165 170 175
 Gln Gln Ile Pro Asp Leu Leu Ala Leu Thr Gly Asp Glu Ile Asp Asn
 180 185 190
 Ile Pro Gly Val Thr Gly Ile Gly Glu Lys Thr Ala Val Gln Leu Leu
 195 200 205
 Glu Lys Tyr Lys Asp Leu Glu Asp Ile Leu Asn His Val Arg Glu Leu
 210 215 220
 Pro Gln Lys Val Arg Lys Ala Leu Leu Arg Asp Arg Glu Asn Ala Ile
 225 230 235 240
 Leu Ser Lys Lys Leu Ala Ile Leu Glu Thr Asn Val Pro Ile Glu Ile
 245 250 255
 Asn Trp Glu Glu Leu Arg Tyr Gln Gly Tyr Asp Arg Glu Lys Leu Leu
 260 265 270
 Pro Leu Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln
 275 280 285
 Leu Tyr 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
 500 505 510
 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> 35

<211> 893

<212> PRT

5 <213> Thermotoga neopolitana

<220>

<223> ADN polimerasa de Thermotoga neopolitana (Tne)

10 <400> 35

Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala
 1 5 10 15
 Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr
 20 25 30
 Asn Ala Val Tyr Gly Val Ala Arg Met Leu Val Lys Phe Ile Lys Glu
 35 40 45
 His Ile Ile Pro Glu Lys Asp Tyr Ala Ala Val Ala Phe Asp Lys Lys
 50 55 60
 Ala Ala Thr Phe Arg His Lys Leu Leu Val Ser Asp Lys Ala Gln Arg
 65 70 75 80
 Pro Lys Thr Pro Ala Leu Leu Val Gln Gln Leu Pro Tyr Ile Lys Arg
 85 90 95
 Leu Ile Glu Ala Leu Gly Phe Lys Val Leu Glu Leu Glu Gly Tyr Glu
 100 105 110
 Ala Asp Asp Ile Ile Ala Thr Leu Ala Val Arg Ala Ala Arg Phe Leu
 115 120 125
 Met Arg Phe Ser Leu Ile Thr Gly Asp Lys Asp Met Leu Gln Leu Val
 130 135 140
 Asn Glu Lys Ile Lys Val Trp Arg Ile Val Lys Gly Ile Ser Asp Leu
 145 150 155 160
 Glu Leu Tyr Asp Ser Lys Lys Val Lys Glu Arg Tyr Gly Val Glu Pro

ES 2 624 406 T3

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
 675 680 685
 Glu Asn Leu Val Lys Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
 690 695 700
 Thr Ala Ser Arg Ile Tyr Asn Val Lys Pro Glu Glu Val Asn Glu Glu
 705 710 715 720
 Met Arg Arg Val 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 Ile Pro Val Lys Glu Ala
 740 745 750
 Glu Lys Met Ile Ile Ser Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser
 755 760 765
 Tyr Ile Gln Gln Val Val Ala 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
 Lys Asn Thr Gln Ser 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 Asp Ile Asp
 820 825 830
 Glu Glu Leu Arg Lys Arg Asn Met Lys Ser Arg Met Ile Ile Gln Val
 835 840 845
 His Asp Glu Leu Val Phe Glu Val Pro Asp Glu Glu Lys Glu Glu Leu
 850 855 860
 Val Asp Leu Val Lys Asn Lys Met Thr Asn Val Val Lys Leu Ser Val
 865 870 875 880
 Pro Leu Glu Val Asp Ile Ser Ile Gly Lys Ser Trp Ser
 885 890

<210> 36

<211> 876

<212> PRT

5 <213> Bacillus stearothermophilus

<220>

<223> ADN polimerasa de Bacillus stearothermophilus (Bst)

10 <400> 36

Met Lys Asn Lys Leu Val Leu Ile Asp Gly Asn Ser Val Ala Tyr Arg
 1 5 10 15
 Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr
 20 25 30
 Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu
 35 40 45
 Glu Gln Pro Thr His Ile Leu Val Ala Phe Asp Ala Gly Lys Thr Thr
 50 55 60
 Phe Arg His Glu Thr Phe Gln Asp Tyr Lys Gly Gly Arg Gln Gln Thr
 65 70 75 80
 Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Lys
 85 90 95
 Ala Tyr Arg Ile Pro Ala Tyr Glu Leu Asp His Tyr Glu Ala Asp Asp
 100 105 110
 Ile Ile Gly Thr Met Ala Ala Arg Ala Glu Arg Glu Gly Phe Ala Val
 115 120 125
 Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro Gln
 130 135 140
 Val Thr Val Glu Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Ser Tyr
 145 150 155 160
 Thr Pro Glu Thr Val Val Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile
 165 170 175
 Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly
 180 185 190

ES 2 624 406 T3

Val Pro Gly Ile Gly Glu Lys Thr Ala Val Lys Leu Leu Lys Gln Phe
 195 200 205
 Gly Thr Val Glu Asn Val Leu Ala Ser Ile Asp Glu Ile Lys Gly Glu
 210 215 220
 Lys Leu Lys Glu Asn Leu Arg Gln Tyr Arg Asp Leu Ala Leu Leu Ser
 225 230 235 240
 Lys Gln Leu Ala Ala Ile Cys Arg Asp Ala Pro Val Glu Leu Thr Leu
 245 250 255
 Asp Asp Ile Val Tyr Lys Gly Glu Asp Arg Glu Lys Val Val Ala Leu
 260 265 270
 Phe Gln Glu Leu Gly Phe Gln Ser Phe Leu Asp Lys Met Ala Val Gln
 275 280 285
 Thr Asp Glu Gly Glu Lys Pro Leu Ala Gly Met Asp Phe Ala Ile Ala
 290 295 300
 Asp Ser Val Thr Asp Glu Met Leu Ala Asp Lys Ala Ala Leu Val Val
 305 310 315 320
 Glu Val Val Gly Asp Asn Tyr His His Ala Pro Ile Val Gly Ile Ala
 325 330 335
 Leu Ala Asn Glu Arg Gly Arg Phe Phe Leu Arg Pro Glu Thr Ala Leu
 340 345 350
 Ala Asp Pro Lys Phe Leu Ala Trp Leu Gly Asp Glu Thr Lys Lys Lys
 355 360 365
 Thr Met Phe Asp Ser Lys Arg Ala Ala Val Ala Leu Lys Trp Lys Gly
 370 375 380
 Ile Glu Leu Arg Gly Val Val Phe Asp Leu Leu Ala Ala Tyr Leu
 385 390 395 400
 Leu Asp Pro Ala Gln Ala Ala Gly Asp Val Ala Ala Val Ala Lys Met
 405 410 415
 His Gln Tyr Glu Ala Val Arg Ser Asp Glu Ala Val Tyr Gly Lys Gly
 420 425 430
 Ala Lys Arg Thr Val Pro Asp Glu Pro Thr Leu Ala Glu His Leu Ala
 435 440 445
 Arg Lys Ala Ala Ala Ile Trp Ala Leu Glu Glu Pro Leu Met Asp Glu
 450 455 460
 Leu Arg Arg Asn Glu Gln Asp Arg Leu Leu Thr Glu Leu Glu Gln Pro
 465 470 475 480
 Leu Ala Gly Ile Leu Ala Asn Met Glu Phe Thr Gly Val Lys Val Asp
 485 490 495
 Thr Lys Arg Leu Glu Gln Met Gly Ala Glu Leu Thr Glu Gln Leu Gln
 500 505 510
 Ala Val Glu Arg Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile
 515 520 525
 Asn Ser Pro Lys Gln Leu Gly Thr Val Leu Phe Asp Lys Leu Gln Leu
 530 535 540
 Pro Val Leu Lys Lys Thr Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val
 545 550 555 560
 Leu Glu Lys Leu Ala Pro His His Glu Ile Val Glu His Ile Leu His
 565 570 575
 Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu Leu
 580 585 590
 Lys Val Val His Pro Val Thr Gly Lys Val His Thr Met Phe Asn Gln
 595 600 605
 Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Val Glu Pro Asn Leu Gln
 610 615 620
 Asn Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala Phe
 625 630 635 640
 Val Pro Ser Glu Pro Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser Gln
 645 650 655
 Ile Glu Leu Arg Val Leu Ala His Ile Ala Glu Asp Asp Asn Leu Ile
 660 665 670
 Glu Ala Phe Arg Arg Gly Leu Asp Ile His Thr Lys Thr Ala Met Asp
 675 680 685
 Ile Phe His Val Ser Glu Glu Asp Val Thr Ala Asn Met Arg Arg Gln

ES 2 624 406 T3

| | | | | | |
|-----|---|-----|---|-----|---|
| 690 | Ala Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly | 695 | Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly | 700 | Gly Ile Ser Asp Tyr Gly |
| 705 | Leu Ala Gln Asn Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile | 710 | Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile | 715 | Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile |
| | | 725 | | 730 | |
| | Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Asp Asn | 740 | Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Asp Asn | 745 | Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Asp Asn |
| | Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His | 755 | Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His | 760 | Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His |
| | Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg | 770 | Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg | 775 | Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg |
| | Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala | 785 | Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala | 790 | Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala |
| | Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg | 805 | Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg | 810 | Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg |
| | Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu | 820 | Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu | 825 | Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu |
| | Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Arg Leu Val | 835 | Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Arg Leu Val | 840 | Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Arg Leu Val |
| | Pro Glu Val Met Glu Gln Ala Val Ala Leu Arg Val Pro Leu Lys Val | 850 | Pro Glu Val Met Glu Gln Ala Val Ala Leu Arg Val Pro Leu Lys Val | 855 | Pro Glu Val Met Glu Gln Ala Val Ala Leu Arg Val Pro Leu Lys Val |
| | Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys | 865 | Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys | 870 | Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys |
| | | | | 875 | |

<210> 37

<211> 877

5

<212> PRT

<213> Bacillus caldotenax

<220>

<223> ADN polimerasa de Bacillus caldotenax (Bca)

10

<400> 37

| | | | | |
|---|-----|-----|-----|-----|
| Met Lys Lys Lys Leu Val Leu Ile Asp Gly Ser Ser Val Ala Tyr Arg | 1 | 5 | 10 | 15 |
| Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr | 20 | 25 | 30 | |
| Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu | 35 | 40 | 45 | |
| Glu Glu Pro Thr His Met Leu Val Ala Phe Asp Ala Gly Lys Thr Thr | 50 | 55 | 60 | |
| Phe Arg His Glu Ala Phe Gln Glu Tyr Lys Gly Gly Arg Gln Gln Thr | 65 | 70 | 75 | 80 |
| Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Arg | 85 | 90 | 95 | |
| Ala Tyr Arg Ile Pro Ala Tyr Glu Leu Glu Asn Tyr Glu Ala Asp Asp | 100 | 105 | 110 | |
| Ile Ile Gly Thr Leu Ala Ala Arg Ala Glu Gln Glu Gly Phe Glu Val | 115 | 120 | 125 | |
| Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro His | 130 | 135 | 140 | |
| Val Thr Val Asp Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Pro Tyr | 145 | 150 | 155 | 160 |
| Thr Pro Glu Ala Val Arg Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile | 165 | 170 | 175 | |
| Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly | 180 | 185 | 190 | |
| Val Pro Gly Ile Gly Glu Lys Thr Ala Val Lys Leu Leu Arg Gln Phe | 195 | 200 | 205 | |
| Gly Thr Val Glu Asn Val Leu Ala Ser Ile Asp Glu Ile Lys Gly Glu | 210 | 215 | 220 | |
| Lys Leu Lys Glu Thr Leu Arg Gln His Arg Glu Met Ala Leu Leu Ser | | | | |

ES 2 624 406 T3

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| Lys | Lys | Leu | Ala | Ala | Ile | Arg | Arg | Asp | Ala | Pro | Val | Glu | Leu | Ser | Leu |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Asp | Asp | Ile | Ala | Tyr | Gln | Gly | Glu | Asp | Arg | Glu | Lys | Val | Val | Ala | Leu |
| | | | 260 | | | | | 265 | | | | | | 270 | |
| Phe | Lys | Glu | Leu | Gly | Phe | Gln | Ser | Phe | Leu | Glu | Lys | Met | Glu | Ser | Pro |
| | | | 275 | | | | | 280 | | | | | 285 | | |
| Ser | Ser | Glu | Glu | Glu | Lys | Pro | Leu | Ala | Lys | Met | Ala | Phe | Thr | Leu | Ala |
| | | | | | | 295 | | | | | | 300 | | | |
| Asp | Arg | Val | Thr | Glu | Glu | Met | Leu | Ala | Asp | Lys | Ala | Ala | Leu | Val | Val |
| 305 | | | | | | 310 | | | | | | 315 | | | 320 |
| Glu | Val | Val | Glu | Glu | Asn | Tyr | His | Asp | Ala | Pro | Ile | Val | Gly | Ile | Ala |
| | | | | | 325 | | | | 330 | | | | | | 335 |
| Val | Val | Asn | Glu | His | Gly | Arg | Phe | Phe | Leu | Arg | Pro | Glu | Thr | Ala | Leu |
| | | | 340 | | | | | 345 | | | | | | 350 | |
| Ala | Asp | Pro | Gln | Phe | Val | Ala | Trp | Leu | Gly | Asp | Glu | Thr | Lys | Lys | Lys |
| | | | 355 | | | | | 360 | | | | | 365 | | |
| Ser | Met | Phe | Asp | Ser | Lys | Arg | Ala | Ala | Val | Ala | Leu | Lys | Trp | Lys | Gly |
| | | | 370 | | | 375 | | | | | | 380 | | | |
| Ile | Glu | Leu | Cys | Gly | Val | Ser | Phe | Asp | Leu | Leu | Leu | Ala | Ala | Tyr | Leu |
| 385 | | | | | | 390 | | | | | | 395 | | | 400 |
| Leu | Asp | Pro | Ala | Gln | Gly | Val | Asp | Asp | Val | Ala | Ala | Ala | Ala | Lys | Met |
| | | | | 405 | | | | | 410 | | | | | | 415 |
| Lys | Gln | Tyr | Glu | Ala | Val | Arg | Pro | Asp | Glu | Ala | Val | Tyr | Gly | Lys | Gly |
| | | | 420 | | | | | 425 | | | | | | 430 | |
| Ala | Lys | Arg | Ala | Val | Pro | Asp | Glu | Pro | Val | Leu | Ala | Glu | His | Leu | Val |
| | | | 435 | | | | | 440 | | | | | 445 | | |
| Arg | Lys | Ala | Ala | Ala | Ile | Trp | Ala | Leu | Glu | Arg | Pro | Phe | Leu | Asp | Glu |
| | | | | | | 455 | | | | | | 460 | | | |
| Leu | Arg | Arg | Asn | Glu | Gln | Asp | Arg | Leu | Leu | Val | Glu | Leu | Glu | Gln | Pro |
| 465 | | | | | | 470 | | | | | | 475 | | | 480 |
| Leu | Ser | Ser | Ile | Leu | Ala | Glu | Met | Glu | Phe | Ala | Gly | Val | Lys | Val | Asp |
| | | | | 485 | | | | | | 490 | | | | | 495 |
| Thr | Lys | Arg | Leu | Glu | Gln | Met | Gly | Glu | Glu | Leu | Ala | Glu | Gln | Leu | Arg |
| | | | 500 | | | | | 505 | | | | | | 510 | |
| Thr | Val | Glu | Gln | Arg | Ile | Tyr | Glu | Leu | Ala | Gly | Gln | Glu | Phe | Asn | Ile |
| | | | 515 | | | | | 520 | | | | | | 525 | |
| Asn | Ser | Pro | Lys | Gln | Leu | Gly | Val | Ile | Leu | Phe | Glu | Lys | Leu | Gln | Leu |
| | | | | | | 535 | | | | | 540 | | | | |
| Pro | Val | Leu | Lys | Lys | Ser | Lys | Thr | Gly | Tyr | Ser | Thr | Ser | Ala | Asp | Val |
| 545 | | | | | | 550 | | | | | 555 | | | | 560 |
| Leu | Glu | Lys | Leu | Ala | Pro | Tyr | His | Glu | Ile | Val | Glu | Asn | Ile | Leu | Gln |
| | | | | 565 | | | | | | 570 | | | | | 575 |
| His | Tyr | Arg | Gln | Leu | Gly | Lys | Leu | Gln | Ser | Thr | Tyr | Ile | Glu | Gly | Leu |
| | | | 580 | | | | | 585 | | | | | | 590 | |
| Leu | Lys | Val | Val | Arg | Pro | Asp | Thr | Lys | Lys | Val | His | Thr | Ile | Phe | Asn |
| | | | 595 | | | | | 600 | | | | | 605 | | |
| Gln | Ala | Leu | Thr | Gln | Thr | Gly | Arg | Leu | Ser | Ser | Thr | Glu | Pro | Asn | Leu |
| | | | 610 | | | 615 | | | | | | 620 | | | |
| Gln | Asn | Ile | Pro | Ile | Arg | Leu | Glu | Glu | Gly | Arg | Lys | Ile | Arg | Gln | Ala |
| 625 | | | | | | 630 | | | | | 635 | | | | 640 |
| Phe | Val | Pro | Ser | Glu | Ser | Asp | Trp | Leu | Ile | Phe | Ala | Ala | Asp | Tyr | Ser |
| | | | | 645 | | | | | | 650 | | | | | 655 |
| Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Ile | Ala | Glu | Asp | Asp | Asn | Leu |
| | | | 660 | | | | | 665 | | | | | | 670 | |
| Met | Glu | Ala | Phe | Arg | Arg | Asp | Leu | Asp | Ile | His | Thr | Lys | Thr | Ala | Met |
| | | | 675 | | | | | 680 | | | | | 685 | | |
| Asp | Ile | Phe | Gln | Val | Ser | Glu | Asp | Glu | Val | Thr | Pro | Asn | Met | Arg | Arg |
| | | | | | | 695 | | | | | | | 700 | | |
| Gln | Ala | Lys | Ala | Val | Asn | Phe | Gly | Ile | Val | Tyr | Gly | Ile | Ser | Asp | Tyr |
| 705 | | | | | | 710 | | | | | 715 | | | | 720 |
| Gly | Leu | Ala | Gln | Asn | Leu | Asn | Ile | Ser | Arg | Lys | Glu | Ala | Ala | Glu | Phe |
| | | | | 725 | | | | | | | 730 | | | | 735 |

ES 2 624 406 T3

Ile Glu Arg Tyr Phe Glu Ser Phe Pro Gly Val Lys Arg Tyr Met Glu
 740 745 750
 Asn Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu
 755 760 765
 His Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val
 770 775 780
 Arg Ser Phe Ala Glu Arg Met Ala Met Asn Thr Pro Ile Gln Gly Ser
 785 790 795 800
 Ala Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Asn Ala Arg Leu
 805 810 815
 Lys Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu
 820 825 830
 Leu Ile Leu Glu Ala Pro Lys Glu Glu Met Glu Arg Leu Cys Arg Leu
 835 840 845
 Val Pro Glu Val Met Glu Gln Ala Val Thr Leu Arg Val Pro Leu Lys
 850 855 860
 Val Asp Tyr His Tyr Gly Ser Thr Trp Tyr Asp Ala Lys
 865 870 875

<210> 38
 <211> 13
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> motivo de ADN polimerasa sintética correspondiente a la mutación D580X de Z05

<220>
 <221> VARIANTE
 <222> (7)...(7)
 <223> Xaa = Ser o Thr

<220>
 <221> VARIANTE
 <222> (8)...(8)
 <223> Xaa = cualquier aminoácido distinto de Asp o Glu

<400> 38
 Thr Gly Arg Leu Ser Ser Xaa Xaa Pro Asn Leu Gln Asn
 1 5 10

<210> 39
 <211> 831
 <212> PRT
 <213> Carboxydotherrnus hydrogenoformans

<220>
 <223> ADN polimerasa de Carboxydotherrnus hydrogenoformans (Chy)

<400> 39
 Met Gly Lys Val Val Leu Val Asp Gly Asn Ser Leu Leu His Arg Ala
 1 5 10 15
 Phe Phe Ala Leu Pro Pro Leu Lys Thr Thr Lys Gly Glu Pro Thr Gly
 20 25 30
 Ala Val Tyr Glu Phe Leu Thr Met Leu Phe Arg Val Ile Lys Asp Glu
 35 40 45
 Lys Pro Glu Tyr Leu Ala Val Ala Phe Asp Ile Ser Arg Lys Thr Phe
 50 55 60
 Arg Thr Glu Gln Phe Thr Ala Tyr Lys Gly His Arg Lys Glu Ala Pro
 65 70 75 80

ES 2 624 406 T3

Asp Glu Leu Val Pro Gln Phe Ala Leu Val Arg Glu Val Leu Lys Val
 85 90 95
 Leu Asn Val Pro Tyr Ile Glu Leu Asp Gly Tyr Glu Ala Asp Asp Ile
 100 105 110
 Ile Gly His Leu Ser Arg Ala Phe Ala Gly Gln Gly His Glu Val Val
 115 120 125
 Ile Tyr Thr Ala Asp Arg Asp Met Leu Gln Leu Val Asp Glu Lys Thr
 130 135 140
 Val Val Tyr Leu Thr Lys Lys Gly Ile Thr Glu Leu Val Lys Met Asp
 145 150 155 160
 Leu Ala Ala Ile Leu Glu Asn Tyr Gly Leu Lys Pro Lys Gln Leu Val
 165 170 175
 Asp Val Lys Gly Leu Met Gly Asp Pro Ser Asp Asn Ile Pro Gly Val
 180 185 190
 Pro Gly Ile Gly Glu Lys Thr Ala Leu Asp Leu Ile Lys Thr Tyr Gly
 195 200 205
 Ser Val Glu Glu Val Leu Ala Arg Lys Asp Glu Leu Lys Pro Lys Leu
 210 215 220
 Arg Glu Lys Leu Ala Glu His Glu Asn Leu Ala Lys Ile Ser Lys Gln
 225 230 235 240
 Leu Ala Thr Ile Leu Arg Glu Ile Pro Leu Glu Ile Ser Leu Glu Asp
 245 250 255
 Leu Lys Val Lys Glu Pro Asn Tyr Glu Glu Val Ala Lys Leu Phe Leu
 260 265 270
 His Leu Glu Phe Lys Ser Phe Leu Lys Glu Ile Glu Pro Lys Ile Lys
 275 280 285
 Lys Glu Tyr Gln Glu Gly Lys Asp Leu Val Gln Val Glu Thr Val Glu
 290 295 300
 Thr Glu Gly Gln Ile Ala Val Val Phe Ser Asp Gly Phe Tyr Val Asp
 305 310 315 320
 Asp Gly Glu Lys Thr Lys Phe Tyr Ser Leu Asp Arg Leu Asn Glu Ile
 325 330 335
 Glu Glu Ile Phe Arg Asn Lys Lys Ile Ile Thr Asp Asp Ala Lys Gly
 340 345 350
 Ile Tyr His Val Cys Leu Glu Lys Gly Leu Thr Phe Pro Glu Val Cys
 355 360 365
 Phe Asp Ala Arg Ile Ala Ala Tyr Val Leu Asn Pro Ala Asp Gln Asn
 370 375 380
 Pro Gly Leu Lys Gly Leu Tyr Leu Lys Tyr Asp Leu Pro Val Tyr Glu
 385 390 395 400
 Asp Val Ser Leu Asn Ile Arg Gly Leu Phe Tyr Leu Lys Lys Glu Met
 405 410 415
 Met Arg Lys Ile Phe Glu Gln Glu Gln Glu Arg Leu Phe Tyr Glu Ile
 420 425 430
 Glu Leu Pro Leu Thr Pro Val Leu Ala Gln Met Glu His Thr Gly Ile
 435 440 445
 Gln Val Asp Arg Glu Ala Leu Lys Glu Met Ser Leu Glu Leu Gly Glu
 450 455 460
 Gln Ile Glu Glu Leu Ile Arg Glu Ile Tyr Val Leu Ala Gly Glu Glu
 465 470 475 480
 Phe Asn Leu Asn Ser Pro Arg Gln Leu Gly Val Ile Leu Phe Glu Lys
 485 490 495
 Leu Gly Leu Pro Val Ile Lys Lys Thr Lys Thr Gly Tyr Ser Thr Asp
 500 505 510
 Ala Glu Val Leu Glu Glu Leu Leu Pro Phe His Glu Ile Ile Gly Lys
 515 520 525
 Ile Leu Asn Tyr Arg Gln Leu Met Lys Leu Lys Ser Thr Tyr Thr Asp
 530 535 540
 Gly Leu Met Pro Leu Ile Asn Glu Arg Thr Gly Lys Leu His Thr Thr
 545 550 555 560
 Phe Asn Gln Thr Gly Thr Leu Thr Gly Arg Leu Ala Ser Ser Glu Pro
 565 570 575
 Asn Leu Gln Asn Ile Pro Ile Arg Leu Glu Leu Gly Arg Lys Leu Arg

REIVINDICACIONES

- 5 1. Una ADN polimerasa que tiene eficacia de transcriptasa inversa incrementada en comparación con una ADN polimerasa de control, en la que la ADN polimerasa comprende una secuencia de aminoácidos al menos un 90 % idéntica a la SEQ ID NO: 1 y en la que el aminoácido de la ADN polimerasa correspondiente a la posición 763 de la SEQ ID NO: 1 es T, el aminoácido correspondiente a la posición 709 de la SEQ ID NO: 1 es K y el aminoácido correspondiente a la posición 580 de la SEQ ID NO: 1 es G, y en la que la ADN polimerasa de control tiene la misma secuencia de aminoácidos que la ADN polimerasa excepto en que el aminoácido de la ADN polimerasa de control correspondiente a la posición 763 de la SEQ ID NO: 1 es M.
- 10
2. La ADN polimerasa de la reivindicación 1, en la que la ADN polimerasa tiene la misma actividad de polimerasa dependiente de ADN o sustancialmente similar en comparación con la ADN polimerasa de control.
- 15 3. Un ácido nucleico recombinante que codifica la ADN polimerasa de acuerdo con una cualquiera de las reivindicaciones 1 o 2.
4. Un vector de expresión que comprende el ácido nucleico recombinante de la reivindicación 3.
- 20 5. Una célula huésped que comprende el vector de expresión de la reivindicación 4.
6. Un procedimiento de producción de una ADN polimerasa, comprendiendo dicho procedimiento:
- 25 cultivar la célula huésped de la reivindicación 5 en condiciones adecuadas para la expresión del ácido nucleico que codifica la ADN polimerasa.
7. Un procedimiento para llevar a cabo la extensión del cebador, que comprende:
- 30 poner en contacto una ADN polimerasa como en una de las reivindicaciones 1 o 2 con un cebador, un molde de polinucleótido y nucleósidos trifosfato en condiciones adecuadas para la extensión del cebador, produciendo de este modo un cebador extendido.
8. Un kit para producir un cebador extendido, que comprende:
- 35 al menos un recipiente que proporciona una ADN polimerasa como en una de las reivindicaciones 1 o 2.
9. El kit de acuerdo con la reivindicación 8, que comprende adicionalmente uno o más recipientes adicionales seleccionados del grupo que consiste en:
- 40 (a) un recipiente que proporciona un cebador hibridable, en condiciones de extensión del cebador, a un molde de polinucleótido predeterminado;
- (b) un recipiente que proporciona nucleósidos trifosfato; y
- 45 (c) un recipiente que proporciona un tampón adecuado para la extensión del cebador.
10. Una mezcla de reacción que comprende una ADN polimerasa como en una de las reivindicaciones 1 o 2, al menos un cebador, un molde de polinucleótido y nucleósidos trifosfato.
- 50 11. La mezcla de reacción de la reivindicación 10, que comprende adicionalmente una segunda ADN polimerasa termoestable.

Figura 1

| | | | | | | |
|-------|--------------------------------------|-------|--|--------|------------|----------------|
| Z05 | RMAFNMPVQGTAA | D L M | KL AM V K L | FP H L | REM..GARML | (SEQ ID NO:12) |
| Taq | RMAFNMPVQGTAA | D L M | KL AM V K L | FP R L | EEM..GARML | (SEQ ID NO:13) |
| Tfi | RMAFNMPVQGTAA | D L M | KL AM V K L | FP R L | RPL..GVRIL | (SEQ ID NO:14) |
| Tfl | RMAFNMPVQGTAA | D L M | KL AM V R L | FP R L | QELGAR..ML | (SEQ ID NO:15) |
| Sps17 | RMAFNMPVQGTAA | D L M | KL AM V K L | FP R L | RPL..GVRIL | (SEQ ID NO:16) |
| Tth | RMAFNMPVQGTAA | D L M | KL AM V K L | FP R L | REM..GARML | (SEQ ID NO:17) |
| Tca | RMAFNMPVQGTAA | D L M | KL AM V K L | FP R L | REMGAR..ML | (SEQ ID NO:18) |
| Tma | RIANTPIQGTAA | D I I | KL AM I E I | DR E L | KERKMRSKMI | (SEQ ID NO:19) |
| Tne | RIANTPIQGTAA | D I I | KL AM I D I | DE E L | RKRNMKSRMI | (SEQ ID NO:20) |
| Taf | RIAVNTPIQGTAA | D I I | KI AM I N I | HN R L | KKENLRSKMI | (SEQ ID NO:21) |
| Dra | RLAYNMPIQGTAA | D I M | KL AM V Q L | DP Q L | DAIGAR..ML | (SEQ ID NO:23) |
| Bst | RTAMNTPIQGSAA | D I I | KK AM I D L | SV R L | REERLQARLL | (SEQ ID NO:24) |
| Bca | RMAMNTPIQGSAA | D I I | KK AM I D L | NA R L | KEERLQARLL | (SEQ ID NO:25) |
| | -----D X ₁ X ₂ | | AM X ₄ X ₅ X ₆ X ₇ X ₈ X ₉ L | | ----- | (SEQ ID NO:26) |

FIGURA 2

| A. Identidades de secuencia sobre toda la enzima polimerasa I (correspondiente a los aminoácidos 1-834 de Z05) | | | | | | | | | | | | | | |
|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Nombre | Z05 | Taq | Tfi | Tfi | Tfi | Sps17 | Tth | Tca | Dra | Tma | Tnc | Taf | Bst | Bca |
| Z05 | | 0,864 | 0,833 | 0,859 | 0,839 | 0,839 | 0,962 | 0,958 | 0,459 | 0,374 | 0,368 | 0,359 | 0,407 | 0,408 |
| Taq | 0,864 | | 0,831 | 0,854 | 0,836 | 0,836 | 0,872 | 0,864 | 0,468 | 0,382 | 0,368 | 0,351 | 0,397 | 0,397 |
| Tfi | 0,833 | 0,831 | | 0,82 | 0,991 | 0,829 | 0,829 | 0,824 | 0,45 | 0,371 | 0,375 | 0,353 | 0,405 | 0,397 |
| Tfi | 0,859 | 0,854 | 0,82 | | 0,824 | 0,853 | 0,848 | 0,848 | 0,462 | 0,381 | 0,374 | 0,356 | 0,397 | 0,398 |
| Sps17 | 0,839 | 0,836 | 0,991 | 0,824 | | 0,835 | 0,83 | 0,83 | 0,452 | 0,375 | 0,377 | 0,355 | 0,407 | 0,399 |
| Tth | 0,962 | 0,872 | 0,829 | 0,853 | 0,835 | | 0,989 | 0,989 | 0,463 | 0,373 | 0,367 | 0,358 | 0,406 | 0,406 |
| Tca | 0,958 | 0,864 | 0,824 | 0,848 | 0,83 | 0,989 | | | 0,46 | 0,371 | 0,365 | 0,356 | 0,404 | 0,404 |
| Dra | 0,459 | 0,468 | 0,45 | 0,462 | 0,452 | 0,463 | 0,46 | 0,46 | | 0,334 | 0,325 | 0,314 | 0,338 | 0,339 |
| Tma | 0,374 | 0,382 | 0,371 | 0,381 | 0,375 | 0,373 | 0,371 | 0,371 | 0,334 | | 0,854 | 0,567 | 0,37 | 0,377 |
| Tnc | 0,368 | 0,368 | 0,375 | 0,374 | 0,377 | 0,367 | 0,365 | 0,365 | 0,325 | 0,854 | | 0,558 | 0,377 | 0,376 |
| Taf | 0,359 | 0,351 | 0,353 | 0,356 | 0,355 | 0,358 | 0,356 | 0,356 | 0,314 | 0,567 | 0,558 | | 0,356 | 0,364 |
| Bst | 0,407 | 0,397 | 0,405 | 0,397 | 0,407 | 0,406 | 0,404 | 0,404 | 0,338 | 0,37 | 0,377 | 0,356 | | 0,881 |
| Bca | 0,408 | 0,397 | 0,397 | 0,398 | 0,399 | 0,406 | 0,404 | 0,404 | 0,339 | 0,377 | 0,376 | 0,364 | 0,881 | |

| B. Identidades de secuencia sobre el subdominio de polimerasa únicamente (correspondiente a los aminoácidos 420-834 de Z05) | | | | | | | | | | | | | | |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Nombre | Z05 | Taq | Tfi | Tfi | Tfi | Sps17 | Tth | Tca | Dra | Tma | Tnc | Taf | Bst | Bca |
| Z05 | | 0,901 | 0,845 | 0,891 | 0,845 | 0,845 | 0,975 | 0,973 | 0,563 | 0,483 | 0,478 | 0,44 | 0,498 | 0,49 |
| Taq | 0,901 | | 0,879 | 0,901 | 0,877 | 0,877 | 0,906 | 0,901 | 0,561 | 0,488 | 0,473 | 0,44 | 0,503 | 0,495 |
| Tfi | 0,845 | 0,879 | | 0,857 | 0,997 | 0,853 | 0,853 | 0,853 | 0,566 | 0,495 | 0,49 | 0,449 | 0,512 | 0,49 |
| Tfi | 0,891 | 0,901 | 0,857 | | 0,855 | 0,889 | 0,889 | 0,889 | 0,571 | 0,492 | 0,48 | 0,444 | 0,494 | 0,485 |
| Sps17 | 0,845 | 0,877 | 0,997 | 0,855 | | 0,853 | 0,853 | 0,853 | 0,566 | 0,495 | 0,49 | 0,449 | 0,512 | 0,49 |
| Tth | 0,975 | 0,906 | 0,853 | 0,889 | 0,853 | | | 0,99 | 0,563 | 0,478 | 0,473 | 0,437 | 0,496 | 0,488 |
| Tca | 0,973 | 0,901 | 0,853 | 0,889 | 0,853 | 0,99 | | | 0,563 | 0,478 | 0,473 | 0,437 | 0,496 | 0,488 |
| Dra | 0,563 | 0,561 | 0,566 | 0,571 | 0,566 | 0,563 | 0,563 | 0,563 | | 0,45 | 0,448 | 0,426 | 0,474 | 0,454 |
| Tma | 0,483 | 0,488 | 0,495 | 0,492 | 0,495 | 0,478 | 0,478 | 0,478 | 0,45 | | 0,883 | 0,622 | 0,474 | 0,475 |
| Tnc | 0,478 | 0,473 | 0,49 | 0,48 | 0,49 | 0,473 | 0,473 | 0,473 | 0,448 | 0,883 | | 0,615 | 0,476 | 0,473 |
| Taf | 0,44 | 0,44 | 0,449 | 0,444 | 0,449 | 0,437 | 0,437 | 0,437 | 0,426 | 0,622 | 0,615 | | 0,46 | 0,473 |
| Bst | 0,498 | 0,503 | 0,512 | 0,494 | 0,512 | 0,496 | 0,496 | 0,496 | 0,474 | 0,474 | 0,476 | 0,46 | | 0,898 |
| Bca | 0,49 | 0,495 | 0,49 | 0,485 | 0,49 | 0,488 | 0,488 | 0,488 | 0,454 | 0,475 | 0,473 | 0,473 | 0,898 | |

FIGURA 3

| A. Identidades de secuencia sobre toda la enzima polimerasa I (correspondiente a los aminoácidos 1-834 de Z05) | | | | | | | | | |
|--|-------|-------|-------|-------|-------|-------|-------|--|--|
| Nombre | Z05 | Tth | Tfi | Tfl | Tca | Taq | Sps17 | | |
| Z05 | | 0,962 | 0,833 | 0,859 | 0,958 | 0,864 | 0,839 | | |
| Tth | 0,962 | | 0,829 | 0,853 | 0,989 | 0,872 | 0,835 | | |
| Tfi | 0,833 | 0,829 | | 0,82 | 0,824 | 0,831 | 0,991 | | |
| Tfl | 0,859 | 0,853 | 0,82 | | 0,848 | 0,854 | 0,824 | | |
| Tca | 0,958 | 0,989 | 0,824 | 0,848 | | 0,864 | 0,83 | | |
| Taq | 0,864 | 0,872 | 0,831 | 0,854 | 0,864 | | 0,836 | | |
| Sps17 | 0,839 | 0,835 | 0,991 | 0,824 | 0,83 | 0,836 | | | |
| | | | | | | | | | |
| B. Identidades de secuencia sobre el subdominio de polimerasa únicamente (correspondiente a los aminoácidos 420-834 de Z05) | | | | | | | | | |
| Nombre | Z05 | Tth | Tfi | Tfl | Tca | Taq | Sps17 | | |
| Z05 | | 0,975 | 0,845 | 0,891 | 0,973 | 0,901 | 0,845 | | |
| Tth | 0,975 | | 0,853 | 0,889 | 0,99 | 0,906 | 0,853 | | |
| Tfi | 0,845 | 0,853 | | 0,857 | 0,853 | 0,879 | 0,997 | | |
| Tfl | 0,891 | 0,889 | 0,857 | | 0,889 | 0,901 | 0,855 | | |
| Tca | 0,973 | 0,99 | 0,853 | 0,889 | | 0,901 | 0,853 | | |
| Taq | 0,901 | 0,906 | 0,879 | 0,901 | 0,901 | | 0,877 | | |
| Sps17 | 0,845 | 0,853 | 0,997 | 0,855 | 0,853 | 0,877 | | | |