

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 569 723**

51 Int. Cl.:

C12N 9/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.12.2012 E 12794881 (8)**

97 Fecha y número de publicación de la concesión europea: **10.02.2016 EP 2788479**

54 Título: **Polimerasas de DNA con actividad mejorada**

30 Prioridad:

08.12.2011 US 201161568375 P

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

12.05.2016

73 Titular/es:

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

72 Inventor/es:

**BAUER, KEITH;
MYERS, THOMAS W. y
SUKO, SHAWN**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 569 723 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

Polimerasas de DNA con actividad mejorada

5 Campo de la Invención

La presente invención proporciona polimerasas de DNA con actividades mejoradas, incluyendo el aumento de la eficiencia de la transcriptasa inversa, y que puede incluir tolerancia al desemparejamiento, velocidad de la extensión y/ o tolerancia a los inhibidores de la transcriptasa inversa (TI) y la polimerasa, así como el uso de tales polimerasas en diversas aplicaciones, incluyendo la extensión y amplificación de ácidos nucleicos polinucleótido.

Antecedentes de la invención

15 Las polimerasas de DNA son responsables de la replicación y el mantenimiento del genoma, un papel que es vital para transmitir con exactitud la información genética de generación en generación. Las polimerasas de DNA funcionan en las células como las enzimas responsables de la síntesis de DNA. Éstas polimerizan desoxirribonucleósidos trifosfato en presencia de un activador metálico, como el Mg^{2+} , en un orden establecido por el molde de DNA o molde de polinucleótidos que se copia. *In vivo*, las polimerasas de DNA participan en un espectro de procesos de síntesis de DNA, incluyendo la replicación del DNA, la reparación del DNA, la recombinación y la amplificación génica. Durante cada proceso de síntesis de DNA, el molde de DNA se copia una vez o a lo sumo un par de veces para producir réplicas idénticas. Por el contrario, *in vitro*, la replicación del DNA se puede repetir muchas veces, tales como, por ejemplo, durante la reacción en cadena de la polimerasa (véase, por ejemplo, la patente estadounidense N° 4.683.202).

25 En los estudios iniciales con la reacción en cadena de la polimerasa (PCR), se añadió la polimerasa de DNA al inicio de cada ronda de replicación del DNA (ver patente estadounidense N° 4.683.202, supra). Posteriormente, se determinó que podían obtenerse polimerasas de DNA termoestables a partir de bacterias que crecen a temperaturas elevadas, y que estas enzimas sólo tienen que añadirse una vez (véase la patente estadounidense N° 4.889.818 y la patente estadounidense N° 4.965.188). A las temperaturas elevadas utilizadas durante la PCR, estas enzimas no se inactivan de forma irreversible. Como resultado, se pueden llevar a cabo ciclos repetitivos de la reacción en cadena de la polimerasa sin adición de enzima fresca al inicio de cada proceso de adición sintético. Las polimerasas de DNA, en particular las polimerasas termoestables, son la clave de un gran número de técnicas en los estudios de DNA recombinante y en el diagnóstico médico de enfermedades. Para las aplicaciones de diagnóstico, en particular, la secuencia de ácido nucleico diana puede ser sólo una pequeña porción de DNA o RNA en cuestión, por lo que puede ser difícil detectar la presencia de una secuencia de ácido nucleico diana sin amplificación.

40 El patrón general de plegamiento de las polimerasas de DNA se asemeja a la mano derecha humana y contiene tres subdominios distintos de palma, dedos y pulgar (ver Beese et al., Science 260:352-355, 1993.); Patel et al., Biochemistry 34:5351-5363, 1995). Si bien la estructura de los subdominios de los dedos y el pulgar varían mucho entre polimerasas que difieren en tamaño y en las funciones celulares, los subdominios de la palma catalíticos son superponibles. Por ejemplo, el motivo A, que interactúa con el dNTP entrante y estabiliza el estado de transición durante la catálisis química, es superponible con una desviación media de alrededor de un A entre las polimerasas de DNA de la familia pol α en mamíferos y pol I en procariontes (Wang et al., Cell 89:1087-1099, 1997). El motivo A comienza estructuralmente en una lámina β antiparalela que contiene residuos predominantemente hidrófobos y continúa hacia una hélice α . La secuencia primaria de aminoácidos de los sitios activos de la polimerasa de DNA se conserva de forma excepcional. En el caso del motivo A, por ejemplo, la secuencia DYSQIELR (Id. de Sec. N° 22) se mantiene en las polimerasas de organismos separados por muchos millones de años de evolución, incluyendo, por ejemplo, *Thermus aquaticus*, *Chlamydia trachomatis* y *Escherichia coli*.

50 Además de estar muy bien conservado, el sitio activo de las polimerasas de DNA también se ha demostrado que es relativamente mutable, siendo capaz de acomodar ciertas sustituciones de aminoácidos sin reducirse la actividad de la polimerasa de DNA de manera significativa (véase, por ejemplo, la Patente de Estados Unidos N° 6.602.695). Tales polimerasas de DNA mutantes pueden ofrecer varias ventajas selectivas en, por ejemplo, aplicaciones de diagnóstico e investigación, lo que comprende las reacciones de síntesis de ácidos nucleicos.

55 Hay por lo menos dos pasos en el proceso enzimático de la polimerización del DNA; 1) la incorporación del nucleótido entrante y 2) la extensión del nucleótido recién incorporado. La exactitud o "fidelidad" global de la polimerasa de DNA generalmente se considera como un conglomerado de estas dos actividades enzimáticas, pero los pasos son distintos. Una polimerasa de DNA puede incorporar de forma incorrecta el nucleótido entrante, pero si no se extiende de manera eficiente la velocidad de extensión se reducirá severamente y la formación de producto global será mínima. Alternativamente, es posible tener una polimerasa de DNA que incorpore de forma errónea el nucleótido entrante y fácilmente extienda de forma incorrecta el desemparejamiento 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 polimerasa de DNA ES112 (polimerasa de DNA Z05 E683R; véase el documento US 7.179.590) cuando se utiliza Mn^{2+} como ion metálico divalente activador. La enzima tiene una eficacia muy alta porque a diferencia de las polimerasas de DNA típicas que tienden a dudar o detenerse cuando se encuentra una discrepancia, la polimerasa

de DNA ES112 extiende fácilmente el desemparejamiento. El fenotipo que presenta la ES112 es más pronunciado durante la etapa de TI, presumiblemente debido a los efectos estructurales del heterodúplex de RNA/ DNA frente al homodúplex DNA/DNA. Un segundo ejemplo sería si la polimerasa de DNA no incorporara de forma errónea fácilmente (puede ser incluso menos probable incorporar de forma errónea), pero sí que tiene una mayor capacidad para extender un desemparejamiento. En este caso, la fidelidad no se altera significativamente en el producto global. En general, este tipo de enzima es más favorable para las reacciones de extensión que las características de la ES112 en Mn^{2+} porque se mejora la fidelidad del producto. Sin embargo, este atributo se puede utilizar para permitir la extensión errónea de un cebador oligonucleótido no coincidente, como cuando un cebador oligonucleótido de secuencia sencilla se hibrida con una diana que muestra heterogeneidad de secuencia (por ejemplo, dianas virales), pero la tasa de incorporación errónea normal o inferior permite la terminación de la síntesis de DNA más allá del cebador de oligonucleótido inicial. Un ejemplo de este tipo de polimerasa de DNA es la polimerasa de DNA D580G Z05 (véase la publicación de patente estadounidense N° 2009/0148891). Este tipo de actividad se denomina "tolerante al desemparejamiento", porque es más tolerante a los desemparejamientos en el cebador oligonucleótido. Si bien los ejemplos anteriores han discutido reacciones de tipo de extensión del cebador, la actividad puede ser más significativa en reacciones como la RT-PCR y PCR en las que la extensión del cebador está ocurriendo de forma recurrente y frecuente. Los datos sugieren que mientras que las enzimas como la Z05 D580G son más "tolerantes" a los desemparejamientos, también se ha mejorado su capacidad de extender cebadores oligonucleótidos que contienen bases modificadas (por ejemplo, bases modificadas t-butil bencilo) o en presencia de colorantes de unión al DNA tales como el verde SYBR I (véase la publicación de patente estadounidense N° 2009/028053).

La reacción en cadena de la polimerasa de transcripción inversa (RT-PCR) es una técnica utilizada en muchas aplicaciones para detectar y/ o cuantificar dianas de RNA mediante amplificación. Con el fin de amplificar las dianas de RNA por PCR, es necesario realizar primero una transcripción reversa del molde de RNA a cDNA. Típicamente, los ensayos de RT-PCR se basan en una transcriptasa inversa no termoestable (RNA polimerasa dependiente de DNA), derivada de un organismo mesófilo, para la etapa de síntesis del cDNA inicial (TI). Se requiere una polimerasa de DNA termoestable adicional para la amplificación del cDNA para que tolere las temperaturas elevadas necesarias para la desnaturalización de los ácidos nucleicos en la PCR. Hay varias ventajas potenciales en la utilización de polimerasas de DNA termoactivas o termoestables diseñadas para llevar a cabo una transcripción inversa más eficiente en los ensayos de RT-PCR. El aumento de la actividad de la transcriptasa inversa junto con la capacidad de utilizar temperaturas superiores en la incubación de transcripción inversa, que permiten la relajación de la estructura secundaria del molde de RNA, puede resultar en una mayor eficiencia general de la síntesis de cDNA y sensibilidad del ensayo. La incubación a temperatura superior también podría aumentar la especificidad mediante la reducción de falsos cebados en la etapa de transcripción inversa. Las enzimas con una mejor eficacia de la transcripción inversa pueden simplificar el diseño de los ensayos al permitir una reducción de los tiempos de incubación de TI y/ o la concentración de la enzima. Cuando se utiliza dUTP y UNG, se forman productos de extensión no específicos que contienen dUMP en condiciones no astringentes de configuración y son degradados por UNG y no pueden utilizarse ni como cebadores ni como moldes. Cuando se utiliza una transcriptasa inversa no termoestable (polimerasa de RNA dependiente de DNA) derivada de un organismo mesófilo, no es posible utilizar las metodologías de dUTP y UNG (Myers, TW et al, Amplification of RNA: High Temperature Reverse Transcription and DNA amplification with *Thermus thermophilus* DNA Polymerase, en PCR strategies, Innis, M.A., Gelfand, D.H., y Sninsky, J.J., Ed., Academic Press, San Diego, CA, 58-68, (1995)). Sin embargo, el uso de una polimerasa de DNA termoactiva o termoestable de la invención para la etapa de transcripción inversa permite que la reacción sea completamente compatible con la utilización del sistema de prevención de la contaminación cruzada del 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 control sobre la contaminación cruzada, el uso de dUTP y UNG proporciona un "inicio en caliente" que reduce la amplificación no específica (Innis y Gelfand 1999).

Breve resumen de la invención

La invención se expone en las reivindicaciones adjuntas a esta descripción. En el presente documento se proporcionan polimerasas de DNA con actividades mejoradas, incluyendo el aumento de la eficiencia de la transcriptasa inversa, y que puede incluir tolerancia al desemparejamiento, velocidad de extensión y / o la tolerancia a los inhibidores de la TI y la polimerasa, con respecto a la correspondiente polimerasa control no modificada, y los métodos de fabricación y uso de tales polimerasas de DNA. En particular, se proporciona una polimerasa de DNA mejorada con una mayor eficiencia de transcriptasa inversa en comparación con una polimerasa de DNA control. La polimerasa de DNA mejorada puede tener la misma actividad o sustancialmente similar a la polimerasa dependiente de DNA, en comparación con la polimerasa de DNA control. Una polimerasa de DNA mejorada puede comprender una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N°: 1, donde el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I. Una polimerasa de DNA puede comprender una secuencia de aminoácidos idéntica en al menos un 90% a la Id. de Sec. N° 1. Una polimerasa de DNA control puede tener la misma secuencia de aminoácidos que la DNA polimerasa, excepto que el aminoácido de la polimerasa de DNA control correspondiente a la posición 640 de la Id. de Sec. N° 1 es I. El aminoácido en la posición correspondiente a la posición 640 de la Id. de Sec. N° 1 de una

polimerasa mejorada puede seleccionarse de entre G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M o H. En la invención, el aminoácido en la posición correspondiente a la posición 640 de la Id. de Sec. N° 1 de la polimerasa mejorada es F.

5 Una polimerasa de DNA mejorada puede comprender una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N° 1, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 580 de la Id. de Sec. N° 1 es cualquier aminoácido distinto de D o E. El aminoácido de la polimerasa de DNA correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de D. El aminoácido de la polimerasa de DNA correspondiente a la posición 580 de la Id. de Sec. N° 1 se puede seleccionar de entre el grupo que consiste en L, G, T, Q, A, S, N, R y K. El aminoácido de la polimerasa de DNA correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser G.

15 La polimerasa de DNA mejorada puede comprender además una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N° 1, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser cualquier aminoácido que no sea I. El aminoácido de la polimerasa de DNA correspondiente a la posición 709 de la Id. de Sec. N° 1 puede seleccionarse de entre el grupo que consiste en K, R, S, G y A. El aminoácido de la polimerasa de DNA correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser K.

20 La polimerasa de DNA mejorada puede comprender una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N° 1, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I, el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de D o E, y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I. La polimerasa de DNA mejorada puede comprender una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N° 1, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I, en el que el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 puede seleccionarse de entre el grupo que consiste en L, G, T, Q, A, S, N, R y K; y en el que el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 se puede seleccionar de entre el grupo que consiste en K, R, S, G y A. El aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser G; y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser K. La polimerasa de DNA mejorada puede comprender una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, idéntica en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%) a la Id. de Sec. N° 1, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser F, el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser G, y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser K. La polimerasa de DNA puede tener la misma actividad o sustancialmente similar a la de la polimerasa dependiente de DNA, en comparación con la polimerasa de DNA control.

45 Una polimerasa de DNA mejorada puede tener una eficiencia aumentada de la transcriptasa inversa sin una disminución sustancial en la actividad de la polimerasa dependiente de DNA en comparación con una polimerasa de DNA control, en la que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser cualquier aminoácido que no sea I, y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I, y en el que la polimerasa de DNA control tiene la misma secuencia de aminoácidos que la polimerasa de DNA, excepto que el aminoácido de la polimerasa de DNA de control correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser I y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser I. El aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser F, y el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser K. La polimerasa de DNA mejorada puede comprender además una sustitución de aminoácido en el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1. El aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I, el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de I, y el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 puede ser cualquier aminoácido distinto de D o E. El aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. N° 1 es F, el aminoácido correspondiente a la posición 709 de la Id. de Sec. N° 1 es K, y el aminoácido correspondiente a la posición 580 de la Id. de Sec. N° 1 es G en la invención.

60 Varias polimerasas de DNA son susceptibles de mutación de acuerdo con la presente invención. Particularmente adecuadas son las polimerasas termoestables, incluyendo las polimerasas termoestables de tipo salvaje o de origen natural de diversas especies de bacterias termófilas, así como las polimerasas termoestables sintéticas derivadas de las enzimas de tipo salvaje o de origen natural mediante la sustitución, inserción o delección de aminoácidos u otra modificación. Las formas sin modificar de la polimerasa incluyen, a modo de ejemplo, las polimerasas de DNA CS5, CS6 o Z05, o una polimerasa de DNA funcional que tiene 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 ciertas realizaciones, la identidad de secuencia de aminoácidos es de al menos un 80%, preferiblemente de al menos un 90% y más preferiblemente de al menos un 95%. Otras polimerasas no modificadas incluyen, por ejemplo, las polimerasas de DNA de cualquiera de las siguientes especies de bacterias termófilas (o una polimerasa de DNA funcional que tiene 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 una polimerasa tal): *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 ciertas realizaciones, la identidad de secuencia de aminoácidos es de al menos un 80%, preferiblemente al menos un 90% y más preferiblemente al menos un 95%. Las polimerasas adecuadas también incluyen las que tienen actividad transcriptasa inversa (TI) y/ o capacidad de incorporar nucleótidos no convencionales, tales como los ribonucleótidos u otros nucleótidos modificados en 2'.

Mientras que las polimerasas de DNA termoestables que poseen una actividad de transcripción reversa eficiente son particularmente adecuadas para llevar a cabo una RT-PCR, especialmente una RT-PCR con una sola enzima, las polimerasas de DNA termoactivas, pero no termoestables, que poseen actividad eficiente de transcripción reversa también son susceptibles de mutación de acuerdo con la presente invención. Por ejemplo, los atributos de eficiencia aumentada de la transcriptasa inversa, tolerancia al desemparejamiento, tasa de extensión y/ o tolerancia a los inhibidores de TI son útiles para la etapa de TI en una RT-PCR, y este paso no debe realizarse a temperaturas que inactivarían una polimerasa de DNA termoactiva pero no termoestable. Después de la etapa de TI, se podría añadir una polimerasa de DNA termoestable o bien ya podría estar incluida en la mezcla de reacción para llevar a cabo la etapa de amplificación de PCR. Por ejemplo, la polimerasa de DNA mejorada descrita en el presente documento se puede combinar con una segunda polimerasa de DNA termoestable previamente a la etapa de TI en un tampón adecuado para la extensión y la amplificación de moldes de RNA y DNA, como se describe en los Ejemplos. Ejemplos de polimerasas de DNA termoestables adecuadas se describen en la patente de Estados Unidos N° 4.889.818, y las patentes de Estados Unidos N° 5.773.258 y 5.677.152. La segunda polimerasa de DNA termoestable puede ser una polimerasa de DNA AmpliTaq (desoxinucleotidiltransferasa desoxinucleósido trifosfato : DNA, E.C.2.7.7.7). La segunda polimerasa de DNA termoestable puede ser una polimerasa termoestable inactivada de forma reversible, tal y como se describe a continuación. En una realización, la polimerasa termoestable inactivada de forma reversible es la polimerasa de DNA AmpliTaq Gold® (Roche Applied Science, Indianapolis, IN, EE.UU.). Esta segunda metodología se beneficiaría especialmente del uso de una polimerasa de DNA termoestable modificada químicamente (u otra tecnología de HotStart para inactivar la polimerasa de DNA termoestable), de modo que no fuera completamente activa durante la etapa de TI. Un ejemplo de una polimerasa de DNA termoactiva pero no termoestable que posee actividad de transcripción reversa eficiente es la polimerasa de DNA de *Carboxydotherrnus hydrogenoformans* (Chy; Id. de Sec. N° 39); véanse, por ejemplo, las patentes estadounidenses N° 6.468.775 y 6.399.320.

La polimerasa de DNA puede tener al menos una identidad de secuencia de aminoácidos del 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93 %, 94%, 95%, 96%, 97%, 98% o 99% con una polimerasa seleccionada del grupo que consiste en:

- (a) una polimerasa de DNA Z05 de *Thermus sp.* (Z05) (Id. de Sec. N° 1);
- (b) una polimerasa de DNA de *Thermus aquaticus* (Taq) (Id. de Sec. N° 2);
- (c) una polimerasa de DNA de *Thermus filiformis* (TFI) (Id. de Sec. N° 3);
- (d) una polimerasa de DNA de *Thermus flavus* (Tfl) (Id. de Sec. N° 4);
- (e) una polimerasa de DNA sps17 de *Thermus sp.* (Sps17) (Id. de Sec. N° 5);
- (f) una polimerasa de DNA de *Thermus thermophilus* (Tth) (Id. de Sec. N° 6); y
- (g) una polimerasa de DNA de *Thermus caldophilus* (TCA) (Id. de Sec. N° 7)
- (h) la polimerasa de DNA de *Carboxydotherrnus hydrogenoformans* (Chy) (Id. de Sec. N° 39)

La identidad de secuencia de aminoácidos puede ser de al menos un 80%, preferiblemente al menos un 90% y más preferiblemente al menos un 95%.

La polimerasa de DNA puede ser una polimerasa de DNA de *Thermotoga*. Por ejemplo, la polimerasa de DNA 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 una polimerasa seleccionada del grupo que consiste en:

- (a) una polimerasa de DNA de *Thermotoga maritima* (Tma) (Id. de Sec. N° 34);
- (b) una polimerasa de DNA de *Thermotoga neopolitana* (Tne) (Id. de Sec. N° 35).

La identidad de secuencia de aminoácidos puede ser de al menos un 80%, preferiblemente al menos un 90% y más preferiblemente al menos un 95%.

La polimerasa de DNA 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 la Id. de Sec. N°

1. La identidad de secuencia de aminoácidos puede ser de al menos un 80%, preferiblemente al menos un 90% y más preferiblemente al menos un 95%.

La polimerasa de DNA puede ser una polimerasa de DNA (Z05) de *Thermus sup.* (Z05) (es decir, la Id. de Sec. N° 1), y el aminoácido en la posición 640 es cualquier aminoácido distinto de I. Por ejemplo, el aminoácido en la posición 640 puede seleccionarse de entre G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M o H. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 640 es F. La polimerasa de DNA puede ser una polimerasa de DNA Z05 que comprende además una sustitución en la posición 580, y el aminoácido en la posición 580 puede ser cualquier aminoácido distinto de D o E. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 580 puede ser cualquier aminoácido distinto de D. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 580 puede seleccionarse de entre el grupo que consiste en L, G, T, Q, A, S, N, R y K. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 580 puede ser G. La polimerasa de DNA puede ser una polimerasa de DNA Z05 que comprende además una sustitución en la posición 709, y el aminoácido en la posición 709 puede ser cualquier aminoácido distinto de I. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 709 puede seleccionarse del grupo que consiste en K, R, S, G y A. La polimerasa de DNA puede ser una polimerasa de DNA Z05, y el aminoácido en la posición 709 puede ser K.

La polimerasa de DNA control puede ser una polimerasa Z05, Z05 D580G o Z05 D580G I709K.

Las polimerasas mutantes o mejoradas pueden incluir otras modificaciones, sin sustitución. Una de dichas modificaciones es una modificación covalente reversible térmicamente que inactiva la enzima, pero que se revierte para activar la enzima tras la incubación a una temperatura elevada, como la temperatura utilizada típicamente para la extensión de polinucleótidos. Ejemplos de reactivos para tales modificaciones reversibles térmicamente se describen en las patentes estadounidenses N° 5.773.258 y 5.677.152.

La actividad de transcriptasa inversa se puede determinar mediante la realización de la amplificación de RT-PCR en tiempo real y la detección de un transcrito del virus de la hepatitis C (HCV) generado a partir de las primeras 800 bases del genotipo del HCV Ib 5'NTR en poli(A) de pSP64 (Promega). Dos o más mezclas de reacción pueden tener un número de copias titulado del transcrito de la hepatitis C (HCV) (por ejemplo, titulaciones 1:5, titulaciones 1:10, por ejemplo, 10.000 copias, 1000 copias, 100 copias, 10 copias, 1 copia, 0 copias en varias mezclas de reacción). La capacidad de la transcriptasa inversa de una polimerasa de la invención se puede comparar con la capacidad transcriptasa inversa de una polimerasa de referencia (por ejemplo, una polimerasa de origen natural, no modificada, o control), en una unidad de tiempo preseleccionado, tal como se describe en el presente documento. Las polimerasas con capacidad transcriptasa inversa mejorada amplificarán el transcrito con mayor eficiencia, o se requerirá un menor número de ciclos de PCR para amplificar el transcrito (es decir, muestran un valor de Cp inferior, tal como se calcula en el presente documento), comparado con una polimerasa de origen natural o no modificada. Por otra parte, las polimerasas con mejora de la función TI también poseen una mejora en la replicación de moldes de RNA largo (por ejemplo, al menos 500 o 1000 o 2000 o 5000 o más nucleótidos de longitud). La eficiencia mejorada de la transcriptasa inversa puede incluir un tiempo más corto de transcripción inversa comparado con una polimerasa control. Por lo tanto, las polimerasas con una mayor eficiencia de la transcriptasa inversa pueden realizar la transcripción reversa de un molde de RNA más rápidamente que una polimerasa control o de referencia.

En otros aspectos, la presente invención proporciona un ácido nucleico recombinante que codifica una polimerasa de DNA mutante o mejorada de la invención, un vector que comprende el ácido nucleico recombinante, y una célula huésped transformada con el vector. En ciertas realizaciones, el vector es un vector de expresión. Las células huésped que comprenden dichos vectores de expresión son útiles en los métodos de la invención para la producción de la polimerasa mutante o mejorada mediante el cultivo de las células huésped en condiciones adecuadas para la expresión del ácido nucleico recombinante. Las polimerasas de la invención pueden estar contenidas en las mezclas de reacción y/ o en equipos. Los aspectos de los ácidos nucleicos recombinantes, las células huésped, vectores, vectores de expresión, las mezclas de reacción y los equipos son como se han descrito anteriormente y en el presente documento. Se proporciona un método para la realización de una extensión de polinucleótidos. El método generalmente incluye poner en contacto una polimerasa de DNA que tiene una mayor eficiencia de transcriptasa inversa, tolerancia al desemparejamiento, velocidad de extensión y/ o tolerancia a inhibidores de la TI y de la polimerasa como se describe en el presente documento, con un cebador, un molde de polinucleótidos, y nucleósidos trifosfato en las condiciones adecuadas para la extensión del cebador, produciendo de esta manera un cebador extendido. El molde de polinucleótidos puede ser, por ejemplo, un molde de RNA o DNA. En ciertas realizaciones, la extensión del cebador comprende una etapa de transcripción inversa de menos de aproximadamente cinco minutos. Las condiciones adecuadas para la extensión pueden comprender Mg^{2+} . Los nucleótidos trifosfato pueden incluir nucleótidos no convencionales, tales como, por ejemplo, ribonucleótidos y/ o nucleótidos marcados. Además, el cebador y/ o molde pueden incluir uno o más análogos de nucleótidos. En algunas variaciones, el método de extensión del polinucleótido es un método para la amplificación de polinucleótidos que incluye poner en contacto la polimerasa de DNA mutante o mejorada con un par de cebadores, el molde de polinucleótidos y los nucleósidos trifosfato en condiciones adecuadas para la amplificación del polinucleótido. La reacción de extensión del polinucleótido puede ser, por ejemplo, una PCR, extensión isotérmica o secuenciación (por ejemplo, la reacción de

secuenciación 454). En ciertas realizaciones, el método de extensión del cebador comprende una reacción en cadena de la polimerasa (PCR). El molde de polinucleótido puede ser de cualquier tipo de muestra biológica.

5 Opcionalmente, la reacción de extensión del cebador comprende un inhibidor real o potencial de la polimerasa de referencia o no modificada. El inhibidor puede inhibir la tasa de extensión del ácido nucleico y/ o la eficiencia de transcripción inversa de una polimerasa de referencia o sin modificar (control). El inhibidor puede ser hemoglobina, o un producto de degradación de la misma. Por ejemplo, el producto de degradación de la hemoglobina puede ser un producto de degradación del grupo hemo, tal como hemina, hematoporfirina o bilirrubina. El inhibidor puede ser un quelante de hierro o un pigmento de color púrpura. El inhibidor puede ser heparina o melanina. El inhibidor puede ser un colorante intercalante. El colorante 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 colorante 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 colorante 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 puede comprender Mn⁺⁺.

20 La presente invención también proporciona un equipo útil en tal método de extensión de polinucleótidos. El equipo incluye al menos un recipiente que proporciona la polimerasa de DNA mejorada de la invención. El equipo puede incluir además uno o más recipientes que proporcionen uno o más reactivos. Por ejemplo, uno o más recipientes adicionales pueden proporcionar nucleósidos trifosfato, un tampón adecuado para la extensión de polinucleótidos y/ o uno o más polinucleótidos cebadores o sonda, que se pueden hibridar, en condiciones de extensión de polinucleótidos, a un molde de polinucleótidos predeterminado. El molde de polinucleótidos puede ser de cualquier tipo de muestra biológica.

25 Además se proporcionan mezclas de reacción que comprenden las polimerasas de la invención. Las mezclas de reacción pueden contener también un molde de ácido nucleico (DNA y/ o RNA), uno o más polinucleótidos cebadores o sondas, nucleósidos trifosfato (incluyendo, por ejemplo, desoxirribonucleósidos trifosfato, ribonucleósidos trifosfato, nucleósidos trifosfato marcados, nucleósidos trifosfato no convencionales), tampones, sales y marcajes (por ejemplo, fluoróforos). Las mezclas de reacción pueden comprender un quelante de hierro o un colorante de color púrpura. Las mezclas de reacción pueden comprender hemoglobina o un producto de degradación de la hemoglobina. Por ejemplo, los productos de degradación de la hemoglobina pueden incluir productos de degradación del grupo hemo como hemina, hematina, hematoforina y bilirrubina. Las mezclas de reacción pueden comprender heparina o una sal de la misma. La mezcla de reacción puede comprender un colorante intercalante (lo que incluye pero que no se limita a los descritos anteriormente o en otras partes en este documento). La mezcla de reacción puede contener un molde de ácido nucleico que se ha aislado a partir de sangre. El molde de ácido nucleico puede ser RNA y la mezcla de reacción puede comprender heparina o una sal de la misma. La mezcla de reacción puede comprender además Mg²⁺.

40 La mezcla de reacción puede comprender además una segunda polimerasa de DNA termoestable. La mezcla de reacción puede comprender dos o más polimerasas. Por ejemplo, la mezcla de reacción puede comprender una polimerasa de DNA mejorada que tenga mayor eficiencia de transcripción inversa (por ejemplo, aumento de la actividad de extensión de un molde de RNA) como se describe aquí, y otra polimerasa que tenga actividad polimerasa dependiente de DNA. La mezcla de reacción puede comprender una mezcla de una polimerasa de DNA mejorada que tenga mayor eficiencia de transcripción inversa como se describe en el presente documento, y una segunda polimerasa dependiente de DNA termoestable. La segunda polimerasa dependiente de DNA termoestable puede ser una polimerasa modificada de forma reversible como se ha descrito anteriormente, de manera que la enzima está inactiva a las temperaturas adecuadas para la etapa de transcripción inversa, pero se activa en condiciones adecuadas, por ejemplo, a temperaturas elevadas de aproximadamente 90°C a 100°C durante un periodo de tiempo de hasta alrededor de 12 minutos. Se proporcionan las condiciones adecuadas para la activación de una polimerasa termoestable inactivada de forma reversible, por ejemplo, en una reacción de PCR Hot Start, como se describe en los Ejemplos. Ejemplos de segundas polimerasas dependientes de DNA termoestables adecuadas se describen en las patentes estadounidenses N° 5.773.258 y 5.677.152, *supra*.

Definiciones

55 A menos que se defina de otro modo, todos los términos técnicos y científicos usados en este documento tienen el mismo significado que entiende comúnmente un experto en la materia a la que pertenece esta invención. Aunque esencialmente cualesquier método y materiales similares a los descritos en el presente documento se pueden utilizar en la práctica o análisis de la presente invención, sólo se describen los métodos y materiales de ejemplo. Para el propósito de la presente invención, los siguientes términos se definen a continuación.

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

65 Un "aminoácido" se refiere a cualquier unidad monomérica que puede incorporarse en un péptido, polipéptido o proteína. Tal como se utiliza aquí, el término "aminoácido" incluye los siguientes veinte alfa aminoácidos naturales o

codificados genéticamente: alanina (Ala o A), arginina (Arg o R), asparagina (Asn o N), ácido aspártico (Asp o D), cisteína (Cys o C), glutamina (Gln o Q), ácido glutámico (Glu o E), glicina (Gly o G), histidina (His o H), isoleucina (Ile o I), leucina (Leu o L), lisina (Lys o K), metionina (Met o M), fenilalanina (Phe o F), prolina (Pro o P), serina (Ser o S), treonina (Thr o T), triptófano (Trp o W), tirosina (Tyr o Y) y valina (Val o V). En los casos en los que los residuos "X" no están definidos, éstos deben entenderse como "cualquier aminoácido". Las estructuras de estos veinte aminoácidos naturales se muestran en, por ejemplo, Stryer et al., *Biochemistry*, 5ª Ed., Freeman and Company (2002). Aminoácidos adicionales, como la seleniocisteína y pirrolisina, también pueden estar codificados genéticamente (Stadtman (1996) "Selenocysteine," *Annu Rev Biochem.* 65:83-100 y Ibba et al (2002) "Genetic Code: introducing pyrrolysine" *Curr Biol* 12(13):R464-R466). El término "aminoácido" también incluye aminoácidos no naturales, los aminoácidos modificados (por ejemplo, con las cadenas laterales y/ o el esqueleto central modificados), y los análogos de aminoácidos (véase, por ejemplo, Zhang et al. (2004) "Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells", *Proc Natl Acad Sci USA* 101 (24):8882-8887, Anderson et al (2004), "An expanded genetic code with a functional quadruplet codon", *Proc Natl Acad Sci USA* 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 Des 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. Des. 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. USA.* 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, and Budisa et al. (2001) "Proteins with {beta}-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids", *Protein Sci.* 10(7):1281-1292).

Como ilustración adicional, un aminoácido es típicamente un ácido orgánico que incluye un grupo amino sustituido o no sustituido, un grupo carboxi sustituido o no sustituido, y una o más cadenas o grupos laterales, o análogos de cualquiera de estos grupos. Las cadenas laterales de ejemplo incluyen, por ejemplo, tiol, seleno, sulfonilo, alquilo, arilo, acilo, ceto, azido, hidroxilo, hidracina, ciano, halo, hidrazida, alqueno, alquino, éter, borato, boronato, fosfo, fosfona, fosfina, heterociclo, enona, imina, aldehído, éster, tioácido, hidroxilamina o cualquier combinación de estos grupos. Otros aminoácidos representativos incluyen, pero no se limitan a, los aminoácidos que comprenden entrecruzadores fotoactivables, los aminoácidos que unen metales, aminoácidos con spin marcado, aminoácidos fluorescentes, aminoácidos que contienen metales, aminoácidos con grupos funcionales nuevos, aminoácidos que interactúan de forma covalente o no covalente con otras moléculas, aminoácidos fotoanclados y/ o fotoisomerizables, aminoácidos radiactivos, aminoácidos que comprenden biotina o un análogo de biotina, aminoácidos glicosilados, aminoácidos modificados con otros hidratos de carbono, aminoácidos que comprenden polietilenglicol o poliéter, aminoácidos sustituidos con átomos pesados, aminoácidos químicamente escindibles y/ o fotoescindibles, aminoácidos que contienen azúcares enlazado a carbono, aminoácidos con actividad redox, aminoácidos que contienen aminotioácidos, y aminoácidos que comprenden uno o más residuos tóxicos.

El término "muestra biológica" abarca una variedad de tipos de muestras obtenidas de un organismo y se pueden utilizar en un ensayo de diagnóstico o de control. El término abarca la 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 de tejidos o células derivadas de los mismos y la progenie de los mismos. El término abarca las muestras que se han manipulado de cualquier forma después de su obtención, como mediante un tratamiento con reactivos, solubilización, sedimentación o enriquecimiento de ciertos componentes. El término abarca una muestra clínica y también incluye células de cultivo celular, sobrenadantes celulares, lisados celulares, suero, plasma, fluidos biológicos y muestras de tejido.

El término "mutante", en el contexto de las polimerasas de DNA de la presente invención, significa un polipéptido, normalmente recombinante, que comprende una o más sustituciones de aminoácidos con relación a la correspondiente polimerasa de DNA funcional.

El término "forma no modificada", en el contexto de una polimerasa mutante, es un término usado en el presente documento con el propósito de definir una polimerasa de DNA mutante de la presente invención: el término "forma no modificada" se refiere a una polimerasa de DNA funcional que tiene la secuencia de aminoácidos de la polimerasa mutante, con la excepción de una o más posiciones de aminoácidos especificadas y que caracterizan la polimerasa mutante. Por lo tanto, la referencia a una polimerasa de DNA mutante en términos de (a) su forma no modificada y (b) una o más sustituciones de aminoácidos específicas significa que, con la excepción de las sustituciones de aminoácidos especificadas, la polimerasa mutante tiene el resto de 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 con aumento de la eficiencia de la transcriptasa inversa, de la tolerancia al desemparejamiento, velocidad de extensión y/ o tolerancia a los inhibidores de la TI y la polimerasa) puede contener mutaciones adicionales para proporcionar la funcionalidad deseada, por ejemplo, una mejora de la incorporación de didesoxiribonucleótidos, ribonucleótidos, análogos de ribonucleótidos, nucleótidos marcados con colorantes, modulación de la actividad nucleasa en 5', modulación de la actividad nucleasa en 3' (o autocorrección) o similares.

En consecuencia, en la realización de la presente invención como se describe aquí, la forma no modificada de una polimerasa de DNA está predeterminada. La forma no modificada de una polimerasa de DNA puede ser, por ejemplo, una polimerasa de DNA de tipo salvaje y/ o de origen natural, o una polimerasa de DNA que ya ha sido modificada intencionadamente. Una forma no modificada de la polimerasa es preferiblemente una polimerasa de DNA termoestable, como las polimerasas de DNA de varias bacterias termófilas, así como las variantes funcionales de las mismas, que tienen una identidad de secuencia sustancial con una polimerasa termoestable de tipo salvaje o de aparición natural. Tales variantes pueden incluir, por ejemplo, las polimerasas de DNA quiméricas, tales como, por ejemplo, las polimerasas de DNA quiméricas que se describen en las patentes estadounidenses N° 6.228.628 y 7.148.049. En ciertas realizaciones, la forma no modificada de una polimerasa tiene actividad de transcriptasa inversa (TI).

El término "polimerasa termoestable" se refiere a una enzima que es estable en el calor, es resistente al calor y retiene actividad suficiente para efectuar posteriores reacciones de extensión de polinucleótidos y no se queda desnaturalizada de forma irreversible (inactivada) cuando se somete a temperaturas elevadas durante el tiempo necesario para efectuar la desnaturalización de los ácidos nucleicos de doble cadena. Las condiciones de calentamiento necesarias para la desnaturalización de ácidos nucleicos son bien conocidas en la materia y se ejemplifican, por ejemplo, en las patentes estadounidenses N° 4.683.202, 4.683.195 y 4.965.188. Como se utiliza aquí, una polimerasa termoestable es adecuada para su uso en una reacción de ciclos de temperatura como la reacción en cadena de la polimerasa ("PCR"). La desnaturalización irreversible para los propósitos que aquí se describen, 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 nucleótidos de manera apropiada para formar productos de extensión de polinucleótidos que son complementarios a una cadena de ácido nucleico molde. Las polimerasas de DNA termoestables de las bacterias termófilas incluyen, por ejemplo, las polimerasas de DNA de *Thermotoga maritima*, *Thermus aquaticus*, *Thermus thermophilus*, *Thermus flavus*, *Thermus filiformis*, *Thermus especie sps17*, *Thermus especie Z05*, *Thermus caldophilus*, *Bacillus caldotenax*, *Thermotoga neopolitana* y *Thermosiphon africanus*.

El término "termoactiva" se refiere a una enzima que mantiene propiedades catalíticas a temperaturas comúnmente usadas para la transcripción inversa o para los pasos de hibridación/ extensión en las reacciones de RT-PCR y/ o PCR (es decir, 45-80°C). Las enzimas termoestables son aquellas que no se inactivan o desnaturalizan de forma irreversible cuando se las somete a las temperaturas elevadas necesarias para la desnaturalización de los ácidos nucleicos. Las enzimas termoactivas pueden o no ser termoestables. Las polimerasas de DNA termoactivas pueden ser dependientes de DNA o RNA, de especies termófilas o de especies mesófilas, lo que incluye, pero no se limita a, *Escherichia coli*, el virus de la leucemia murina de Moloney y el virus de la mioblastosis aviar.

Tal 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. Una proteína quimérica normalmente no se produce por manipulación directa de las secuencias de aminoácidos, sino, más bien, se expresa a partir de un gen "quimérico" que codifica la secuencia de aminoácidos quimérica. Una forma no modificada de una polimerasa de DNA mutante puede ser una proteína quimérica que consista en una región amino-terminal (N-terminal) derivada de una polimerasa de DNA de la especie *Thermus* y una región carboxiterminal (C-terminal) derivada de una polimerasa de DNA Tma. La región N-terminal se refiere a una región que se extiende desde el extremo N-terminal (posición de aminoácido 1) hasta un aminoácido interno. Del mismo modo, la región C-terminal se refiere a una región que se extiende desde un aminoácido interno al extremo C-terminal.

El término "aptámero" se refiere a un DNA monocatenario que reconoce y se une a la polimerasa de DNA, e inhibe de manera eficiente la actividad de la polimerasa como se describe en la patente estadounidense N° 5.693.502. También se discute el uso de aptámeros y dUTP/ UNG en la RT-PCR, 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, en PCR Primer: A Laboratory Manual, 2ª Edición, Dieffenbach, C.W. and Dveksler, G.S., Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 211-219, (2003)).

En el contexto de las polimerasas de DNA mutantes, "correspondencia" con otra secuencia (por ejemplo, posiciones, regiones, fragmentos, nucleótidos, aminoácidos o similares) se basa en la convención de numeración de acuerdo con el número de posición de los aminoácidos o nucleótidos y a continuación, la alineación de las secuencias de manera que se 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. En general, como se describe en el presente documento, el aminoácido correspondiente a una posición de una polimerasa puede determinarse usando un algoritmo de alineación como BLAST, como se describe a continuación. Debido a que no todas las posiciones dentro de una determinada "región correspondiente" deben ser idénticas, las posiciones no coincidentes dentro de una región correspondiente pueden considerarse "posiciones correspondientes". En consecuencia, como se usa en el presente documento, la referencia a una "posición de aminoácidos correspondiente a la posición del aminoácido [X] "de una polimerasa de DNA especificada se refiere a las posiciones equivalentes, en base a la alineación, en otras polimerasas de DNA, homólogos estructurales y familias. En algunas realizaciones de la presente invención, la "correspondencia" de las

posiciones de aminoácido se determinan con respecto a una región de la polimerasa que comprende uno o más motivos de Id. de Sec. Nº 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39. Cuando una secuencia de polipéptido de la polimerasa difiere de Id. de Sec. Nº 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39 (por ejemplo, por cambios de aminoácidos o por adición o deleción de aminoácidos), puede ser que una mutación particular asociada con una actividad mejorada como se describe en el presente documento no esté el mismo número de posición que en los Id. de Sec. Nº 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37 o 39. Esto se ilustra, por ejemplo, en la Tabla 1.

"Recombinante", como se usa aquí, se refiere a una secuencia de aminoácidos o una secuencia de nucleótidos que ha sido modificada intencionadamente mediante métodos recombinantes. Por el término "ácido nucleico recombinante" se entiende en el presente documento un ácido nucleico, obtenido originalmente *in vitro*, en general, mediante la manipulación de un ácido nucleico con endonucleasas de restricción, en una forma que no se encuentra normalmente en la naturaleza. Así, el ácido nucleico aislado de una polimerasa de DNA mutante, en forma lineal, o un vector de expresión obtenido *in vitro* mediante la ligación de moléculas de DNA que normalmente no se unen, se consideran ambos recombinantes para los fines de esta invención. Se entiende que una vez que se obtiene un ácido nucleico recombinante y se vuelve a introducir en una célula huésped, se replicará de forma no recombinante, es decir, usando la maquinaria celular *in vivo* de la célula huésped, en lugar de manipulaciones *in vitro*; sin embargo, tales ácidos nucleicos, una vez producidos de forma recombinante, aunque posteriormente se repliquen de forma no recombinante, todavía se consideran recombinantes para el propósito de la invención. Una "proteína recombinante" es una proteína producida usando técnicas recombinantes, es decir, a través de la expresión de un ácido nucleico recombinante como se ha indicado anteriormente.

Un ácido nucleico está "unido de forma operativa" cuando está colocado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor o potenciador está unido de forma operativa a una secuencia codificante si afecta a la transcripción de la secuencia; o un sitio de unión al ribosoma está unido de forma operativa a una secuencia codificante si está colocado de manera que facilita la traducción.

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

El término "vector" se refiere a un trozo de DNA, normalmente de doble cadena, en el que puede haber insertado un segmento de DNA extraño. El vector puede ser, por ejemplo, de origen plasmídico. Los vectores contienen secuencias de polinucleótidos de tipo "replicón" que facilitan la replicación autónoma del vector en una célula huésped. El DNA foráneo se define como DNA heterólogo, que es DNA no se encuentra naturalmente en la célula huésped, y que por ejemplo, replica la molécula vector, codifica un marcador seleccionable o cribable, o codifica un transgén. El vector se utiliza para transportar el DNA extraño o heterólogo a una célula huésped adecuada. Una vez en la célula huésped, el vector puede replicarse de forma independiente o coincidente con el DNA cromosómico del huésped, y pueden generarse varias copias del vector y de su DNA insertado. Además, el vector también puede contener los elementos necesarios que permiten la transcripción del DNA insertado en una molécula de mRNA o bien provocan la replicación del DNA insertado en múltiples copias de RNA. Algunos vectores de expresión contienen, además, elementos de secuencia contiguos al DNA insertado que aumentan la vida media del mRNA expresado y/ o permiten la traducción del mRNA en una molécula de proteína. De este modo se pueden sintetizar rápidamente muchas moléculas de mRNA y polipéptidos codificados por el DNA insertado.

El término "nucleótido", además de hacer referencia a los monómeros de ribonucleótidos o desoxirribonucleótidos de origen natural, en este documento se entiende que se refiere a las variantes estructurales relacionadas, incluidos los derivados y análogos, que son funcionalmente equivalentes con respecto al contexto particular en el que se está utilizando el nucleótido (por ejemplo, la hibridación a una base complementaria), a menos que el contexto indique claramente lo contrario.

El término "ácido nucleico" o "polinucleótido" se refiere a un polímero que puede corresponderse con un polímero de ácido nucleico de ribosa (RNA) o ácido nucleico de desoxirribosa (DNA), o un análogo de los mismos. Esto incluye polímeros de nucleótidos, tales como RNA y DNA, así como las formas sintéticas, formas modificadas (por ejemplo, modificadas de forma química o bioquímica) de los mismos, y polímeros mixtos (por ejemplo, incluyendo subunidades tanto de RNA como de DNA). Modificaciones de ejemplo incluyen la metilación, sustitución de uno o más de los nucleótidos naturales con un análogo, modificaciones internucleótido tales como enlazantes no cargados (por ejemplo, fosfonatos de metilo, fosfotriésteres, fosfoamidatos, carbamatos y similares), porciones colgantes (por ejemplo, polipéptidos), intercalantes (por ejemplo, acridina, psoraleno y similares), quelantes, alquilantes y enlaces modificados (por ejemplo, ácidos nucleicos alfa anoméricos y similares). También se incluyen las moléculas sintéticas que imitan a los polinucleótidos en su capacidad de unirse a una secuencia designada a través de enlaces de hidrógeno y otras interacciones químicas. Típicamente, los monómeros de nucleótidos están unidos a través de enlaces fosfodiéster, aunque las formas sintéticas de los ácidos nucleicos pueden comprender otros enlaces (por ejemplo, los ácidos nucleicos peptídicos como se describe en Nielsen et al (Science 254:1497-1500, 1991). Un ácido nucleico puede ser o puede incluir, por ejemplo, un cromosoma o segmento de cromosoma, un vector (por ejemplo, un vector de expresión), un casete de expresión, un polímero de DNA o RNA desnudo, el producto de una reacción en cadena de la polimerasa (PCR), un oligonucleótido, una sonda y un cebador. Un ácido nucleico puede ser, por

ejemplo, de cadena sencilla, de cadena doble o de cadena triple y no se limita a cualquier longitud particular. A menos que se indique de otro modo, una secuencia de ácido nucleico particular opcionalmente comprende o codifica las secuencias complementarias, además de cualquier secuencia que se indique explícitamente.

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

20 El término "cebador" tal como se utiliza aquí, se refiere a un polinucleótido capaz de actuar como un punto de iniciación de la síntesis de ácido nucleico dirigida por el molde cuando se coloca bajo condiciones en las que se inicia la extensión de polinucleótidos (por ejemplo, en condiciones que comprenden la presencia de los nucleósidos trifosfato necesarios (según lo dictado por el molde que se copia) y de una polimerasa en un tampón apropiado y a una temperatura adecuada o ciclo(s) de temperaturas (por ejemplo, como en una reacción en cadena de la polimerasa)). Como ilustración adicional, los cebadores también se pueden utilizar en una variedad de procesos de síntesis diferentes mediados por oligonucleótidos, incluyendo como iniciadores de la síntesis de RNA *de novo* y en los procesos relacionados con la transcripción *in vitro* (por ejemplo, la amplificación basada en la secuencia del ácido nucleico (NASBA), la amplificación mediada por transcripción (TMA), etc.). Un cebador es típicamente un oligonucleótido de cadena sencilla (por ejemplo, un oligodesoxirribonucleótido). La longitud apropiada de un cebador depende del uso previsto del cebador pero típicamente varía de 6 a 40 nucleótidos, más típicamente de 15 a 35 nucleótidos. Las moléculas de cebador cortas generalmente requieren temperaturas más bajas para formar complejos híbridos con el molde lo suficientemente estables. Un cebador no necesita reflejar la secuencia exacta del molde pero debe ser lo suficientemente complementario como para hibridar con el molde y para que la elongación del cebador ocurra. 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 ácido nucleico a amplificar y un cebador antisentido 3' (a veces llamado "reverso") que se hibrida con el extremo 3' de la secuencia a amplificar (por ejemplo, si la secuencia diana se expresa como RNA o es un RNA). Un cebador puede marcarse, si se desea, mediante la incorporación de un marcaje detectable por medios espectroscópicos, fotoquímicos, bioquímicos, inmunquímicos o químicos. Por ejemplo, marcadores útiles incluyen el ³²P, colorantes fluorescentes, reactivos densos en electrones, enzimas (como se utilizan comúnmente en los ensayos de ELISA), biotina, o haptenos y proteínas para las que hay disponibles antisueros o anticuerpos monoclonales.

45 El término "convencional" o "natural" cuando se refiere a las bases de ácido nucleico, nucleósidos trifosfato o nucleótidos se refiere a aquellos que se producen naturalmente en el polinucleótido que se describe (es decir, para el DNA son dATP, dGTP, dCTP y dTTP). Además, se utilizan con frecuencia el 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 DNA *in vitro*, como la secuenciación. Colectivamente, estos pueden denominarse dNTP.

50 El término "no convencional" o "modificado" cuando se refiere a una base de ácido nucleico, nucleósidos o nucleótidos, incluye modificaciones, derivaciones, o análogos de las bases, nucleósidos o nucleótidos convencionales que ocurren naturalmente en un polinucleótido particular. Ciertos nucleótidos no convencionales están modificados en la posición 2' del azúcar ribosa en comparación con los dNTPs convencionales. Así, aunque para el RNA los nucleótidos de origen natural son los ribonucleótidos (es decir, ATP, GTP, CTP, UTP, colectivamente rNTP), porque estos nucleótidos tienen un grupo hidroxilo en la posición 2' del azúcar, que, por comparación está ausente en los dNTPs, como se utiliza en este documento, los ribonucleótidos son nucleótidos no convencionales como sustratos para las polimerasas de DNA. Como se utiliza en el presente documento, los nucleótidos no convencionales incluyen, pero no se limitan a, compuestos usados como terminadores para la secuenciación de ácidos nucleicos. Ejemplos de compuestos de terminación incluyen pero no se limitan a aquellos compuestos que tienen una estructura 2',3' didesoxi y se denominan didesoxinucleósidos trifosfatos. Los didesoxinucleósidos trifosfato ddATP, ddTTP, ddCTP y ddGTP se denominan colectivamente ddNTP. Ejemplos adicionales de compuestos de terminación incluyen los análogos 2'-PO₄ de los ribonucleótidos (véanse, por ejemplo, las solicitudes de patente estadounidenses N° 2005/0037991 y 2005/0037398). Otros nucleótidos no convencionales incluyen los dNTP fosforotioato ([α-S]dNTP), 5'-[α-borano]-dNTP, [α]metil-fosfonato-dNTP, y ribonucleósidos trifosfato (rNTP). Se pueden marcar las bases no convencionales con isótopos radiactivos, tales como ³²P, ³³P o ³⁵S, marcadores fluorescentes, marcadores quimioluminiscentes, marcadores bioluminiscentes, marcadores de haptenos

tales como biotina, o marcadores enzimáticos como la estreptavidina o avidina. Los marcadores fluorescentes pueden incluir colorantes que están cargados negativamente, tales como los colorantes de la familia de la fluoresceína, o colorantes que son de carga neutra, tales como los colorantes de la familia de la rodamina o colorantes que están cargados positivamente, tales como los colorantes de la familia de la cianina. Los colorantes de la familia de la fluoresceína incluyen, por ejemplo, FAM, HEX, TET, JOE, NAN y ZOE. Los colorantes de la familia de la rodamina incluyen Texas Red, ROX, R110, R6G y TAMRA. Varios colorantes o nucleótidos marcados con FAM, HEX, TET, JOE, NAN, ZOE, ROX, R110, R6G, Texas Red y TAMRA son comercializados por Perkin-Elmer (Boston, MA), Applied Biosystems (Foster City, CA), o Invitrogen/ Molecular Probes (Eugene, OR). Los colorantes de la familia de la cianina incluyen Cy2, Cy3, Cy5 y Cy7, y son comercializados por GE Healthcare UK Limited (Amersham Place, Little Chalfont, Buckinghamshire, Inglaterra). Como se usa en este documento, el "porcentaje de identidad de secuencia" se determina comparando dos secuencias óptimamente alineadas 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 la base de ácido nucleico o el residuo de aminoácido son idénticos en ambas secuencias para dar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para dar el porcentaje de identidad de secuencia.

Los términos "identidad" o "porcentaje de identidad" en el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son la misma. Las secuencias son "sustancialmente idénticas" entre ellas si tienen un porcentaje especificado de residuos de nucleótidos o aminoácidos que son iguales (por ejemplo, de 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%, al menos un 55%, 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 de al menos un 95% de identidad a lo largo de una región determinada), cuando se comparan y alinean para conseguir una correspondencia máxima en una ventana de comparación o región designada, medido utilizando uno de los siguientes algoritmos de comparación de secuencia o mediante alineamiento manual e inspección visual. Las secuencias son "sustancialmente idénticas" entre sí si son idénticas en 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%. Estas definiciones se refieren también a la complementaria de una secuencia de prueba. Opcionalmente, existe la identidad en una región que es de al menos aproximadamente 50 nucleótidos de longitud, o más típicamente en una región de 100 a 500 o 1000 o más nucleótidos de longitud.

Los términos "similitud" o "porcentaje de similitud", en el contexto de dos o más secuencias de polipéptidos, se refiere a dos o más secuencias o subsecuencias que tienen un porcentaje especificado de residuos de aminoácidos que son o bien el mismo o similar, tal como se definen las sustituciones conservadoras de aminoácidos (por ejemplo, un 60% de similitud, opcionalmente un 65%, 70%, 75%, 80%, 85%, 90%, o 95% de similitud en una región especificada), cuando se comparan y alinean para conseguir una correspondencia máxima en una ventana de comparación, o región designada según se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante el alineamiento manual e inspección visual. Las secuencias son "sustancialmente similares" entre sí si son similares en 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%, la una de la otra. Opcionalmente, esta similitud existe a lo largo de una región que es de al menos aproximadamente 50 aminoácidos de longitud, o más típicamente en 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 secuencia de referencia, con la que las secuencias de prueba se comparan. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y de referencia se introducen en un ordenador, se indican las coordenadas de subsecuencia, si es necesario, y se indican los parámetros del programa de algoritmo de secuencia. Comúnmente se utilizan los parámetros del programa por defecto, o pueden indicarse parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces la identidad o similitud de secuencia porcentual de la secuencia de prueba con respecto a 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 cualquier número de posiciones contiguas seleccionadas del grupo que consiste en de 20 a 600, por lo general de aproximadamente 50 a aproximadamente 200, más habitualmente de aproximadamente 100 a alrededor de 150, en el que la secuencia puede compararse con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alineen de manera óptima. Los métodos de alineación de secuencias para su comparación son bien conocidos en la materia. El alineamiento óptimo de secuencias para la comparación puede llevarse a cabo, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman (Adv Appl Math 2:482, 1970), mediante el algoritmo de alineamiento por homología de Needleman y Wunsch (J. Mol. Biol. 48:443, 1970), por la búsqueda del método de similitud de Pearson y Lipman (Proc Natl Acad Sci USA. 85:2444, 1988), mediante implementaciones informatizadas 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 alineamiento manual e inspección visual (véase, por ejemplo, Ausubel et al., Current Protocols in Molecular Biology (suplemento de 1995)).

Entre los ejemplos de un algoritmo que es adecuado para determinar el porcentaje de identidad de secuencia 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 software para realizar el análisis BLAST está disponible públicamente a través del Centro Nacional de Información Biotecnológica (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica en primer lugar identificar pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema, que coinciden o satisfacen una puntuación umbral T de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T se denomina el umbral de puntuación de palabra vecina (Altschul et al., supra). Estas dianas de palabras vecinas iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que las contengan. Las dianas de palabras se extienden en ambas direcciones a lo largo de cada secuencia tanto como se consiga aumentar la puntuación de alineamiento acumulativa. Las puntuaciones acumulativas se calculan usando, para las secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre >0) y N (puntuación de penalización para residuos no coincidentes; siempre <0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las palabras diana en cada dirección se detiene cuando: la puntuación de alineamiento acumulativa disminuye en una cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de uno o más alineamientos de residuos de puntuación negativa; o se alcanza el final de cualquiera de las secuencias. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una 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 sumatorio más pequeña (P(N)), que proporciona una indicación de la probabilidad por la cual una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de sumatorio más pequeño en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor de aproximadamente 0,2, típicamente menor de aproximadamente 0,01, y más típicamente menor de aproximadamente 0,001.

El término "eficacia de la transcripción inversa" se refiere a la fracción de moléculas de RNA que se transcriben de forma inversa a cDNA en una reacción de transcripción inversa dada. En ciertas realizaciones, las polimerasas de DNA mutantes de la invención poseen una eficiencia de transcripción inversa mejorada en relación a las formas no modificadas de estas polimerasas de DNA. Es decir, estas polimerasas de DNA mutantes inversa transcriben una mayor fracción de los moldes de RNA que sus formas sin modificar bajo unas condiciones particulares de reacción. Sin limitarse a esta teoría, la capacidad de una polimerasa de DNA mutante descrita en el presente documento de realizar la transcripción inversa de una mayor fracción de moldes de RNA puede deberse a un aumento de la actividad de transcripción inversa, por ejemplo, un aumento de la tasa de incorporación de nucleótidos y/ o aumento de la capacidad de procesamiento de la enzima. La eficacia de la transcripción inversa se puede medir, por ejemplo, mediante la medición del punto de corte (Cp) de una reacción de PCR utilizando un molde de RNA, y comparando el valor de Cp con el valor Cp de una reacción control en la que se amplifica un molde de DNA con la misma secuencia (excepto que los U son reemplazados con T), en la que las amplificaciones de RNA y DNA utilizan un cebador común establecido y la misma polimerasa, por ejemplo, como se describe en los ejemplos. Una polimerasa de prueba ha mejorado la eficiencia TI cuando la polimerasa de prueba tiene un valor de Cp disminuido en comparación con una polimerasa control cuando se utiliza RNA como molde, pero tiene un valor de Cp sustancialmente sin cambios en relación a la polimerasa control cuando se utiliza DNA como molde. En algunas formas de realización una polimerasa de la invención tiene una eficiencia de TI mejorada tal, que el Cp es al menos una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más unidades menor que el correspondiente a la polimerasa control en el molde de RNA. La mejora de la eficiencia de TI de una polimerasa de ensayo se puede medir como se describe en los Ejemplos.

El término "tolerancia al desemparejamiento" se refiere a la capacidad de una polimerasa de tolerar una secuencia que contiene desemparejamientos cuando extiende un ácido nucleico (por ejemplo, un cebador o de otro oligonucleótido) de una manera dependiente del molde uniendo (por ejemplo, covalentemente) uno o más nucleótidos en el ácido nucleico. El término "tolerancia al desemparejamiento en 3'" se refiere a la capacidad de una polimerasa de tolerar una secuencia (casi complementaria) que contiene desemparejamientos en la que el ácido nucleico que se extiende (por ejemplo, un cebador u otro oligonucleótido) tiene falta de coincidencia con su molde en el nucleótido del extremo 3' del cebador. Los desemparejamientos en el molde también pueden estar situados en el penúltimo nucleótido del extremo 3' del cebador, o en otra posición dentro de la secuencia del cebador.

El término "discriminación de desemparejamientos" se refiere a la capacidad de una polimerasa para distinguir una secuencia totalmente complementaria de una secuencia que contiene desemparejamientos, cuando se extiende un ácido nucleico (por ejemplo, un cebador u otro oligonucleótido) de una manera dependiente del molde, uniendo (por ejemplo, covalentemente) uno o más nucleótidos en el ácido nucleico. El término "discriminación de desemparejamientos en 3'" se refiere a la capacidad de una polimerasa de distinguir una secuencia totalmente complementaria de una secuencia (casi complementaria) que contiene desemparejamientos en la que el ácido

nucleico que se extiende (por ejemplo, un cebador u otro oligonucleótido) tiene una falta de coincidencia en el extremo 3' del ácido nucleico en comparación con el molde con el que se hibrida el ácido nucleico. El término "desemparejamiento" se refiere a la existencia de uno o más bases desemparejadas (o bases "opuestas" no complementarias) dentro de un tramo de secuencias que forman dúplex (o potencialmente forman dúplex) complementarios.

El término "valor de Cp" o valor "punto de corte" se refiere a un valor que permite la cuantificación de los ácidos nucleicos diana iniciales. El valor de Cp puede determinarse de acuerdo con el método del máximo de la segunda derivada (Van Luu-La, et al., "Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction", *Biotechniques*, Vol. 38, Nº 2, febrero de 2005, págs. 287-293). En el método de la segunda derivada, el Cp corresponde al primer pico en la curva de la segunda derivada. Este pico corresponde al principio de una fase log-lineal. El método de la segunda derivada calcula un valor de la segunda derivada de la curva de intensidad de fluorescencia en tiempo real, y se obtiene un único valor. El método original de Cp se basa en una aproximación definida localmente, diferenciable, de los valores de intensidad, por ejemplo, mediante una función polinómica. A continuación, se calcula la tercera derivada. El valor de Cp es la raíz más pequeña de la tercera derivada. El Cp también puede determinarse utilizando el método de punto de ajuste, en el que el Cp se determina mediante la intersección de una paralela a la línea umbral en la región log-lineal (Van Luu-La, et al., *BioTechniques*, Vol. 38, Nº 2, febrero de 2005, págs. 287-293). El valor de Cp proporcionado por el instrumento LightCycler de Roche se genera mediante el cálculo de acuerdo con el método del máximo de la segunda derivada.

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

El término "tasa de extensión de ácido nucleico" se refiere a la velocidad a la que un biocatalizador (por ejemplo, una enzima, como una polimerasa, ligasa o similares) extiende un ácido nucleico (por ejemplo, un cebador u otro oligonucleótido) de forma dependiente del molde o independiente del molde uniendo (por ejemplo, covalentemente) uno o más nucleótidos en el ácido nucleico. A modo ilustrativo, ciertas polimerasas de DNA mutantes descritos aquí poseen una tasa de extensión de ácido nucleico mejorada en relación con las formas no modificadas de estas polimerasas de DNA, de manera que pueden extender los cebadores a tasas mayores que las formas sin modificar bajo unas condiciones de reacción dadas.

El término "tolerancia a los inhibidores de la TI y la polimerasa" se refiere a la capacidad de una polimerasa para mantener la actividad (actividad polimerasa o de transcripción reversa) en presencia de una cantidad de un inhibidor que inhibiría la actividad polimerasa o actividad transcriptasa reversa de una polimerasa control. En algunas realizaciones, la polimerasa mejorada es capaz de mantener la actividad polimerasa o de transcripción inversa en presencia de una cantidad del inhibidor que eliminaría esencialmente la actividad de la polimerasa control.

El término "sonda de nucleasa en 5'" se refiere a un oligonucleótido que comprende al menos una porción de marcaje emisora de luz y que se utiliza en una reacción de nucleasa en 5' para efectuar la detección de ácido nucleico diana. En algunas realizaciones, por ejemplo, una sonda de nucleasa en 5' incluye solamente un único residuo emisor de luz (por ejemplo, un colorante fluorescente, etc.). En ciertas realizaciones, las sondas de nucleasa en 5' incluyen regiones de autocomplementariedad, de manera que las sondas son capaces de formar estructuras de horquilla en unas condiciones seleccionadas. Para mayor ilustración, en algunas realizaciones una sonda de nucleasa en 5' comprende al menos dos restos de marcaje y emite radiación de mayor intensidad después de que uno de los dos marcajes se escinde o se separa de otra forma del oligonucleótido. En ciertas realizaciones, una sonda de nucleasa en 5' está marcada con dos colorantes fluorescentes diferentes, por ejemplo, un "colorante marcador en el extremo 5'" y un colorante o residuo bloqueador en el extremo 3'. En algunas realizaciones, las sondas de nucleasa en 5' se marcan en una o más posiciones distintas, o además de, las posiciones terminales. Cuando la sonda está intacta, la transferencia de energía se produce típicamente entre los dos fluoróforos de tal manera que la emisión fluorescente desde el colorante marcador se inactiva al menos en parte. Durante la etapa de extensión de una reacción en cadena de la polimerasa, por ejemplo, una sonda de nucleasa en 5' unida a un ácido nucleico molde se escinde por la actividad nucleasa de 5' a 3', por ejemplo, de una polimerasa Taq u otra polimerasa con esta actividad, de forma que la emisión fluorescente del colorante marcador ya no se bloquea. Ejemplos de sondas de nucleasa en 5' también se describen en, por ejemplo, la patente estadounidense Nº 5.210.015, la patente estadounidense Nº 5.994.056, y la patente estadounidense Nº 6.171.785. En otras realizaciones, una sonda de nucleasa en 5' se puede marcar con dos o más colorantes marcadores diferentes y un colorante o residuo bloqueante en el extremo 3'.

El término "FRET" o "transferencia de la energía de resonancia de la fluorescencia" o "transferencia de la energía de resonancia Foerster" se refiere a una transferencia de energía entre al menos dos cromóforos, un cromóforo donador y un cromóforo aceptor (denominado bloqueador). El donante típicamente transfiere la energía al aceptor cuando el donante se excita con la radiación de una luz con longitud de onda adecuada. El aceptor típicamente re-

emite la energía transferida en forma de radiación de una luz con longitud de onda diferente. Cuando el aceptor es un bloqueador "oscuro", éste disipa la energía transferida en una forma distinta a la luz. Que un fluoróforo particular actúe como un donante o un aceptor depende de las propiedades del otro miembro del par de FRET. Los pares donante-aceptor de uso general incluyen el par FAM-TAMRA. Son comúnmente utilizados los bloqueadores DABCYL y TAMRA. Los bloqueadores oscuros de uso general incluyen los 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 la secuencia de aminoácidos de una región del dominio de la polimerasa de las polimerasas de DNA de varias especies de bacterias: *Thermus especies* Z05 (Z05) (Id. de Sec. N° 12), *Thermus aquaticus* (Taq) (Id. de Sec. N° 13), *Thermus filiformis* (Tfi) (Id. de Sec. N° 14), *Thermus flavus* (Tfl) (Id. de Sec. N° 15), *Thermus especie* sps17 (Sps17) (Id. de Sec. N° 16), *Thermus thermophilus* (Tth) (Id. de Sec. N° 17), *Thermus caldophilus* (Tca) (Id. de Sec. N° 18), *Thermotoga maritima* (Tma) (Id. de Sec. N° 19), *Thermotoga neopolitana* (Tne) (Id. de Sec. N° 20), *Thermosiphon africanus* (Taf) (Id. de Sec. N° 21), *Deinococcus radiodurans* (Dra) (Id. de Sec. N° 23), *Bacillus stearothermophilus* (Bst) (Id. de Sec. N° 24), y *Bacillus caldotenax* (Bca) (Id. de Sec. N° 25). Además, las regiones de polipéptidos mostrados comprenden el motivo de aminoácidos X₁-X₂-X₃-F-X₄-X₅-X₆-X₇-D-X₈-HT-X₉-TA-X₁₀-X₁₁ (Id. de Sec. N°: 26), las posiciones variables de la que se definen adicionalmente en este documento. Este motivo se resalta en negrita para cada secuencia de la polimerasa. Las posiciones de los aminoácidos susceptibles de mutación se indican con un asterisco (*). Las huecos en las alineaciones se indican con un punto (.).

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

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

Descripción detallada

La presente invención proporciona polimerasas de DNA mejoradas en el que uno o más aminoácidos en el dominio de la polimerasa han sido mutados en relación con una polimerasa de DNA funcional. Las polimerasa de DNA de la invención son enzimas activas que tienen una mayor eficiencia de transcriptasa inversa (por ejemplo, en presencia de cationes divalentes Mn²⁺ y Mg²⁺) con relación a la forma no modificada de la polimerasa. Pueden tener una mayor tolerancia al desemparejamiento, tasa de extensión y tolerancia de la TI e inhibidores de la polimerasa. Las polimerasas de DNA mutantes se pueden usar a concentraciones más bajas para un rendimiento superior o equivalente, como las enzimas parentales. Las polimerasas de DNA mutantes pueden haber aumentado la eficiencia de la transcriptasa inversa al tiempo que conservan sustancialmente la misma actividad de la polimerasa dependiente de DNA con respecto a una polimerasa no modificada o control. La invención se expone en las reivindicaciones adjuntas a esta descripción. Las polimerasas de DNA que realizan de manera más eficiente la transcripción inversa son útiles, por ejemplo, en una variedad de aplicaciones que incluyen ensayos que emplean RT-PCR para detectar y/o cuantificar dianas de RNA. Las polimerasas de DNA son por lo tanto útiles en una variedad de aplicaciones que implican la extensión de polinucleótidos, así como la transcripción inversa o amplificación de moldes de polinucleótidos, incluyendo, por ejemplo, las aplicaciones en los estudios de DNA recombinante y diagnóstico médico de la enfermedad. Las polimerasas de DNA mutantes también son particularmente útiles, debido a su tolerancia a los desemparejamientos, para la detección de dianas que posiblemente tienen secuencias variables (por ejemplo, dianas virales, o cáncer y otros marcadores genéticos de la enfermedad).

Las polimerasas de DNA se pueden caracterizar por tener el siguiente motivo:

X₁-X₂-X₃-Phe-X₄-X₅-X₆-X₇-Asp-X₈-His-Thr-X₉-Thr-Ala-X₁₀-X₁₁ (también denominado en la presente memoria en el código de una letra como X₁-X₂-X₃-F-X₄-X₅-X₆-X₇-D-X₈-H-T-X₉-TA-X₁₀-X₁₁ (Id. de Sec. N°: 8); en el que:

X₁ es Ile (I), Leu (L), Val (V), Gln (Q) o Met (M);
 X₂ es Arg (R), Lys (K), Gln (Q) o Glu (E);
 X₃ es Val (V) o Ala (A);
 X₄ es Gln (Q), Arg (R), Glu (E), Lys (K) o Val (V);
 X₅ es Glu (E) o Arg (R);
 X₆ es Gly (G) o Asp (D);
 X₇ es Lys (K), Arg (R), Ile (I), Leu (L), Ala (A);
 X₈ es cualquier aminoácido distinto de Ile (I) o Val (V);
 X₉ es Gln (Q), Glu (E), Leu (L), Ile (I), Arg (R) o Lys (K);
 X₁₀ es Ser (S), Ala (A) o Met (M);
 X₁₁ es Trp (W), Arg (R), Lys (K), Gln (Q) o Asp (D).

X₈ puede ser seleccionado de G, A, W, P, S, T, F, Y, C, N, Q, D, E, K, R, L, M, o H.

15 Las polimerasas de DNA se pueden caracterizar por tener el siguiente motivo:

Ile-Arg-Val-Phe-X₄-Glu-Gly-X₇-Asp-X₈-His-Thr-X₉-Thr-Ala-X para-Trp (también denominado en la presente memoria en el código de una letra como I-R-V-F-X₄-E-G-X₇-D-X₈-H-T-X₉-T-A-X₁₀-W (Id. de Sec. N° 9); en el que:

20 X₄ es Gln (Q) o Arg (R);
 X₇ es Lys (K) o Arg (R);
 X₈ es cualquier aminoácido distinto de Ile (I);
 X₉ es Gln (Q) o Glu (E);
 X₁₀ es Ser (S) o Ala (A).

25 Las polimerasas de DNA se pueden caracterizar por tener el siguiente motivo:

Ile-Arg-Val-Phe-Gln-Glu-Gly-Lys-Asp-X₈-His-Thr-Gln-Thr-Ala-Ser-Trp (también denominado en la presente memoria en el código de una letra como I-R-V-F-Q-E-G-K-D-X₈-H-T-Q-T-A-S-W (Id. de Sec. N° 10); en el que:

30 X₈ es cualquier aminoácido distinto de Ile (I).

Las polimerasas de DNA se pueden caracterizar por tener el siguiente motivo:

35 Ile-Arg-Val-Phe-Gln-Glu-Gly-Lys-Asp-X₈-His-Thr-Gln-Thr-Ala-Ser-Trp (también denominado en la presente memoria en el código de una letra como I-R-V-F-Q-E-G-K-D-X₈-H-T-Q-T-A-S-W (Id. de Sec. N° 11); en el que:

X₈ es Phe (F).

40 Las polimerasas de DNA se pueden caracterizar por tener los motivos anteriores (por ejemplo, Id. de Sec. N° 8, 9, 10, y 11), opcionalmente en combinación con motivos adicionales que se describen a continuación. Por ejemplo, la polimerasa de DNA puede comprender además el motivo de Id. de Sec. N° 29 y/o Id. de Sec. N° 38.

45 Este motivo está presente en el dominio de "dedos" (N hélice alfa) de muchas polimerasas de DNA dependientes de DNA de la familia de tipo A, en particular las polimerasas de DNA termoestables a partir de bacterias termófilas (Li et al., EMBO J. 17: 7514-7525, 1998). Por ejemplo, la Figura 1 muestra un alineamiento de la secuencia de aminoácidos de una región del dominio de "dedos" de las polimerasas de DNA de varias especies de bacterias: *Bacillus caldotenax*, *Bacillus stearothermophilus*, *Deinococcus radiodurans*, *Thermosiphon africanus*, *Thermotoga maritima*, *Thermotoga neopolitana*, *Thermus aquaticus*, *Thermus caldophilus*, *Thermus filiformis*, *Thermus flavus*,
 50 *Thermus sp. sps17*, *Thermus sp. Z05*, y *Thermus thermophilus*. Como se muestra, la secuencia nativa correspondiente al motivo anterior está presente en cada una de estas polimerasas, que indica una función conservada de esta región de la polimerasa. La Figura 2 proporciona identidades de secuencia entre estas polimerasas de DNA.

55 Por consiguiente, la descripción proporciona una polimerasa que comprende el Id. de Sec. N°: 8, 9, 10, o 11, que tiene la actividad y/o características descritas en este documento mejoradas, y en el que la polimerasa de DNA es por lo demás una polimerasa de DNA de tipo salvaje o producida de forma natural, tal como, por ejemplo, una polimerasa de cualquiera de las especies de bacterias termófilas enumeradas anteriormente, o es sustancialmente idéntica a una polimerasa de DNA de tipo salvaje tal o de origen natural. La polimerasa puede comprender el Id. de
 60 Sec. N° 8, 9, 10, o 11 y puede ser al menos un 80%, 85%, 90%, o 95% idéntica a Id. de Sec. N° 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, o 39. la forma no modificada de la polimerasa puede ser de una especie del género *Thermus*. La polimerasa no modificada puede ser de una especie termofílica distinta de *Thermus*, por ejemplo, *Thermotoga*. La secuencia de ácidos nucleicos y aminoácidos completa para numerosas polimerasas de DNA termoestables está disponible. Las secuencias de cada uno de las polimerasas de *Thermus aquaticus* (Taq) (Id. de Sec. N° 2), *Thermus thermophilus* (Tth) (Id. de Sec. N° 6), *Thermus especie Z05*- (Id. de Sec. N° 1), *Thermus especie sps17* (Id. de Sec. N° 5), *Thermotoga maritima* (Tma) (Id. de Sec. N° 34), y *Thermosiphon africanus* (Taf) (Id. de Sec. N° 33) se han
 65

publicado en la publicación de patente internacional PCT nº WO 92/06200. La secuencia para la polimerasa de DNA de *Thermus flavus* (Id. de Sec. Nº 4) ha sido publicada en Akhmetzjanov y Vakhitov (Nucleic Acids Research 20: 5839, 1992). La secuencia de la polimerasa de DNA termoestable de *Thermus caldophilus* (Id. de Sec. Nº 7) se encuentra en EMBL/GenBank Nº de acceso U62584. La secuencia de la polimerasa de DNA termoestable de *Thermus filiformis* se puede recuperar de depósito de la ATCC Nº 42380 usando, por ejemplo, los métodos proporcionados en la patente de EE.UU. Nº 4.889.818, así como la información de secuencia proporcionada en la Tabla 1. La secuencia de la polimerasa de DNA de *Thermotoga neapolitana* (Id. de Sec. Nº 35) es de la base de datos de Patentes de GeneSeq nº de acceso R98144 y PCT WO 97/09451. La secuencia de la polimerasa de DNA termoestable de *Bacillus caldotenax* (Id. de Sec. Nº 37) se describe en, por ejemplo, Uemori et al. (J Biochem (Tokyo) 113 (3): 401-410, 1993; véase también, la base de datos Swiss-Prot Nº de acceso Q04957 y GenBank números de acceso D12982 y BAA02361). Ejemplos de formas no modificadas de las polimerasas de DNA que pueden ser modificadas como se describe en el presente documento también se describen en, por ejemplo, las patentes de Estados Unidos Nº 6.228.628; 6.346.379; 7.030.220; 6.881.559; 6.468.775; y las patentes de Estados Unidos nº 7.148.049; 7.179.590; 7.410.782; 7378262. Las secuencias representativas de la polimerasa de longitud completa también están dentro de la lista de secuencias.

También son susceptibles a las mutaciones descritas en este documento las polimerasas de DNA funcionales que se han modificado previamente (por ejemplo, por sustitución, adición, o supresión de aminoácidos). Tales polimerasas modificadas funcionales pueden retener el motivo de aminoácidos de Id. de Sec. Nº 8 (o un motivo de la Id. de Sec. Nº 9, 10 o 11), y, opcionalmente, el motivo de aminoácidos de Id. de Sec. Nº 38. Por lo tanto, las polimerasas de DNA no modificadas adecuadas incluyen también variantes funcionales de tipo salvaje o las polimerasas de origen natural. Tales variantes tendrán normalmente una identidad o similitud de secuencia sustancial con la polimerasa de tipo salvaje o de origen 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, que posee un dominio de la polimerasa que comprende los Id. de Sec. Nº: 8, 9, 10, o 11 también puede comprender un dominio de nucleasa (por ejemplo, que corresponde a las posiciones 1 a 291 de Z05)

Una polimerasa puede ser una polimerasa quimérica, es decir, que comprende las regiones de polipéptidos de dos o más enzimas. Ejemplos de tales polimerasas de DNA quiméricas se describen en, por ejemplo, la patente de EE.UU. Nº 6.228.628. Particularmente adecuadas son las polimerasas de DNA quiméricas de la familia CS, que incluyen las polimerasas CS5 (Id. de Sec. Nº 27) y CS6 (Id. de Sec. Nº 28) y variantes de las mismas que tienen una identidad sustancial de la secuencia de aminoácidos o similar a Id. de Sec. Nº 27 o Id. de Sec. 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 ácido amino) y por lo tanto pueden ser modificadas para contener el Id. de Sec. Nº 8. Las polimerasas de DNA CS5 y CS6 son enzimas quiméricas derivadas de las polimerasas de DNA de *Thermus sp.* Z05 y *Thermotoga maritima* (Tma). Estas comprenden el dominio N-terminal de la nucleasa 5' de la enzima *Thermus* y los dominios exonucleasa 3'-5' C-terminal y polimerasa de la enzima Tma. Estas enzimas tienen actividad de transcriptasa inversa eficiente, pueden extender cebadores que contienen análogos de nucleótidos, y pueden incorporar dNTP alfa-fosforotioato, dUTP, dITP, y también dNTP marcados con los marcadores de la familia fluoresceína y cianina. Las polimerasas CS5 y CS6 son también enzimas de PCR activados de forma eficiente por Mg²⁺. Las polimerasas quiméricas CS5 y CS6 se describen adicionalmente en, por ejemplo, la patente de EE.UU. Nº 7.148.049.

Las sustituciones de aminoácidos pueden ser sustituciones individuales de aminoácidos. Las polimerasas de DNA proporcionadas en el presente documento pueden comprender una o más sustituciones de aminoácidos en el sitio activo con respecto a la polimerasa no modificada. La sustitución de aminoácidos puede comprender al menos la posición X₈ del motivo expuesto en Id. de Sec. Nº 8 (o un motivo de Id. de Sec. Nº 9, 10 o 11). La sustitución de aminoácidos en esta posición confiere una mayor eficiencia de la transcriptasa inversa, la tolerancia a los desemparejamientos, velocidad de extensión y/o la tolerancia de los inhibidores de TI y de la polimerasa, produciendo una polimerasa de DNA mutante con una mayor eficiencia de la transcriptasa inversa, la tolerancia a los desemparejamientos, velocidad de extensión y/o la tolerancia de los inhibidores de TI y de la polimerasa con respecto a la polimerasa no modificada. Típicamente, el aminoácido en la posición X₈ está sustituido con un aminoácido que no se corresponde con la secuencia nativa en el motivo expuesto en Id. de Sec. Nº 8 (o un motivo de Id. de Sec. Nº 9, 10 o 11). Por lo tanto, típicamente, el aminoácido en la posición X₈, si está sustituido, no es Ile (I) o Val (V), ya que I o V aparecen en esta posición en las polimerasas de origen natural (véase, por ejemplo, la Figura 1). En ciertas realizaciones, las sustituciones de aminoácidos incluyen G, A, W, P, S, T, F, Y, C, N, Q, D, E, K, R, L, M, o H en la posición X₈. Las sustituciones de aminoácidos pueden incluir fenilalanina (F) en la posición X₈. Las sustituciones de otros aminoácidos adecuados en uno o más de los sitios identificados se puede determinar usando, por ejemplo, los métodos conocidos de mutagénesis y determinación del rendimiento de la extensión de polinucleótidos en ensayos de localización dirigida descritos en más detalle en el presente documento o conocidos de otra manera por los expertos en la materia.

Una polimerasa puede comprender el Id. de Sec. Nº 8, 9, 10, o 11 y comprende además uno o más cambios de aminoácidos (por ejemplo, por sustitución, adición, o supresión de aminoácidos) en comparación con una polimerasa

nativa. Una polimerasa puede retener el motivo de aminoácidos de Id. de Sec. N° 8 (o un motivo de Id. de Sec. N° 9, 10 o 11), y comprenden además el motivo de aminoácidos de Id. de Sec. N° 38 (correspondiente a la mutación D580X de Z05 (Id. de Sec. N° 1)) como sigue:

- 5 Thr-Gly-Arg-Leu-Ser-Ser-X_{b7}-X_{b8}-Pro-Asn-Leu-Gln-Asn (también denominado en este documento en el código de una letra como T-G-R-L-S-S-X_{b7}-X_{b8}-P-N-L-Q-N) (Id. de Sec. N° 38); en el que

X_{b7} es Ser (S) o Thr (T); y

X_{b8} es cualquier aminoácido distinto de Asp (D) o Glu (E)

- 10 La mutación caracterizada por EL Id. de Sec. N° 38 se discute con más detalle en, por ejemplo, la Publicación de Patente de Estados Unidos N° 2009/0148891. Tales variantes funcionales de polimerasas típicamente tendrán una identidad o similitud de secuencia sustancial con la de la polimerasa de tipo salvaje o de origen natural (por ejemplo, Id. de Sec. N° 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.

- Una polimerasa puede comprender el Id. de Sec. N° 8, 9, 10, o 11 y además comprender el motivo de aminoácidos de Id. de Sec. N° 29 (correspondiente a la mutación I709X de Z05 (Id. de Sec. N° 1)) como sigue:

- 20 X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈-X₉-X₁₀-X₁₁-X₁₂-X₁₃-Gly-Tyr-Val-X₁₄-Thr-Leu (también denominado en este documento en el código de una letra como X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈-X₉-X₁₀-X₁₁-X₁₂-X₁₃-X₁₄-G-Y-V-T-L) (Id. de Sec. N° 29); en el que

- 25 X₁ es Ala (A), Asp (D), Ser (S), Glu (E), Arg (R) o Gln (Q);
 X₂ es Trp (W) o Tyr (y);
 X₃ es cualquier aminoácido distinto de Ile (I), Leu (L) o Met (M);
 X₄ es Glu (E), Ala (A), Gln (Q), Lys (K), Asn (N) o Asp (D);
 X₅ es Lys (K), Gly (G), Arg (R), Gln (Q), Su (H) o Asn (N);
 X₆ es Thr (T), Val (V), Met (M) o Ile (I);
 30 X₇ es Leu (L), Val (V) o Lys (K);
 X₈ es Glu (E), Ser (S), Ala (A), Asp (D) o Gln (Q);
 X₉ es Glu (E) o Phe (F);
 X₁₀ es Gly (G) o Ala (A);
 X₁₁ es Arg (R) o Lys (K);
 35 X₁₂ es Lys (K), Arg (R), Glu (E), Thr (T) o Gln (Q);
 X₁₃ es Arg (R), Lys (K) o His (H), y
 X₁₄ es Glu (e), Arg (R) o Thr (T).

- 40 Tales variantes funcionales de polimerasas pueden tener típicamente una identidad o similitud de secuencia sustancial con la de la polimerasa de tipo salvaje o de origen natural (por ejemplo, Id. de Sec. N° 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.

- 45 Una polimerasa de DNA puede comprender una sustitución de aminoácido en la posición X₈ (por ejemplo, como en un motivo seleccionado de Id. de Sec. N° 8, 9, 10 o 11) y comprende una sustitución de aminoácidos correspondiente a Id. de Sec. N° 38 e Id. de Sec. N° 29.

- 50 Un aminoácido en la posición X₈ sustituido con un aminoácido como se expone en Id. de Sec. N° 8, 9, 10 o 11, y el aminoácido en la posición X_{b8} sustituido con un aminoácido como se expone en Id. de Sec. N° 38. Por lo tanto, el aminoácido en la posición X₈ es cualquier aminoácido que no sea Ile (I) y el aminoácido en la posición X_{b8} es 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 X_{b8} de Id. de Sec. N° 38. Las sustituciones de aminoácidos pueden incluir independientemente metionina (M) en la posición X₈ de Id. de Sec. N° 8, 9, 10 o 11, y glicina (G) en la posición X_{b8} de Id. de Sec. N° 38.

- El aminoácido en la posición X₈ puede estar sustituido con un aminoácido como se expone en Id. de Sec. N° 8, 9, 10 o 11, y el aminoácido en la posición X₃ (de Id. de Sec. N° 29) puede estar sustituido con un aminoácido como se expone en Id. de Sec. N° 29. Por lo tanto, el aminoácido en la posición X₈ puede ser cualquier aminoácido distinto de Ile (I) y el aminoácido en la posición X₃ es 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 Id. de Sec. N° 29. Las sustituciones de aminoácidos pueden incluir independientemente metionina (M) en la posición X₈ de Id. de Sec. N° 8, 9, 10 o 11, y la lisina (K) en la posición X₃ de Id. de Sec. N° 29. Otra sustitución de aminoácido adecuado en uno o más de los sitios identificados se puede determinar usando, por ejemplo, los métodos conocidos de mutagénesis y determinación del rendimiento de la extensión de polinucleótidos en ensayos

de localización dirigida descritos en más detalle en el presente documento o conocido de otra manera por los expertos en la materia, por ejemplo, las sustituciones de aminoácidos descritas en la patente de los Estados Unidos. solicitud de publicación N° 2009/0148891 y 2009/0280539.

- 5 Debido a que la longitud precisa de las polimerasas de DNA varían, las posiciones de los aminoácidos precisos que corresponden a cada uno de X_8 (Id. de Sec. N° 8), X_{b8} (Id. de Sec. N° 38) y X_3 (Id. de Sec. N° 29) pueden variar dependiendo en la polimerasa mutante particular utilizada. Los programas de alineación de secuencias de ácidos nucleicos y aminoácidos se encuentran fácilmente disponibles (véase, por ejemplo, los mencionados más arriba) y, dados los motivos particulares identificados en este documento, sirven para ayudar en la identificación de los aminoácidos exactos (y los codones correspondientes) para la modificación de acuerdo con la presente invención. Las posiciones correspondientes a cada uno de X_8 , X_{b8} y X_3 se muestran en la Tabla 1 para las polimerasas de DNA termoestables químicas representativas y las polimerasas de DNA termoestables a partir de ejemplos de especies termófilas.
- 10
- 15 Tabla 1. Posiciones de aminoácidos correspondientes a las posiciones del motivo X_8 (por ejemplo, de Id. de Sec. N°: 8, 9, 10, y 11), X_{b8} (de Id. de Sec. N° 38) y X_3 (de Id. de Sec. N° 29) en ejemplos de polimerasas.

| Organismo o secuencia química | posición de aminoácidos | | |
|-------------------------------|-------------------------|---------------------------------|------------------------------|
| | X_8 | X_{b8} (de Id. de Sec. N° 38) | X_3 (de Id. de Sec. N° 29) |
| T. thermophilus (6) | 640 | 580 | 709 |
| T. caldophilus (7) | 640 | 580 | 709 |
| T. sp. Z05 (1) | 640 | 580 | 709 |
| T. aquaticus (2) | 638 | 578 | 707 |
| T. flavus (4) | 637 | 577 | 706 |
| T. filiformis (3) | 636 | 576 | 705 |
| T. sp. sps17 (5) | 636 | 576 | 705 |
| T. maritima (34) | 701 | 640 | 770 |
| T. neapolitana (35) | 701 | 640 | 770 |
| T. africanus (33) | 700 | 639 | 769 |
| B. caldotenax (37) | 682 | 621 | 751 |
| B. stearothermophilus (36) | 681 | 620 | 750 |
| CS5 (27) | 701 | 640 | 770 |
| CS6 (28) | 701 | 640 | 770 |

- 20 Una polimerasa de DNA puede derivar de la polimerasa de DNA de *Thermus sp.* Z05 (Id. de Sec. N° 1) o una variante de la misma (por ejemplo, llevando la mutación D580G o similares). Como se menciona anteriormente, en la polimerasa de DNA de *Thermus sp.* Z05, la posición X_8 corresponde a isoleucina (I) en la posición 640; posición X_{b8} corresponde a aspartato (D) en la posición 580, y la posición X_3 corresponde a isoleucina (I) en la posición 709. La polimerasa mutante puede comprender al menos una sustitución de aminoácido, con relación a la polimerasa de DNA de *Thermus sp.* Z05 (o una polimerasa de DNA que es sustancialmente idéntica (por ejemplo, al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90%, o 95% idéntica) a Id. de Sec. N° 1), en 1640, D580 y/o 1709. Por lo tanto, típicamente, el aminoácido en la posición 640 de Id. de Sec. N°: 1 no es I. El aminoácido en la posición 640 de Id. de Sec. N°: 1 puede ser seleccionado de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M, o H. El residuo de aminoácido en la posición 640 de Id. de Sec. N° 1 es F en la invención. Los residuos de aminoácidos en la posición D580 de Id. de Sec. N° 1 se pueden seleccionar de leucina (L), glicina (G), treonina (T), glutamina (Q), alanina (A), serina (S), asparagina (N), arginina (R) y lisina (K). Por lo tanto, en algunas realizaciones, el residuo de aminoácido en la posición 580 de Id. de Sec. N° 1 es la glicina (G). Además, el aminoácido en la posición 709 de Id. de Sec. N° 1 no puede ser I. El aminoácido en la posición 709 de Id. de Sec. N° 1 puede ser seleccionado de G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M, o H. El aminoácido en la posición 709 de la Id. de Sec. N° 1 pueden ser K, R, S, G o A. El aminoácido en la posición 709 de Id. de Sec. N° 1 puede ser K.
- 25
- 30
- 35

- Los mutantes de la polimerasa de DNA de *Thermus sp.* Z05 pueden incluir los que comprenden las sustituciones de aminoácidos I640F, y/o I709K (o I709R, I709S, I709G, I709A), y/o D580G. La polimerasa de DNA mutante de *Thermus sp.* Z05 puede comprender, por ejemplo, las sustituciones de residuos de aminoácidos I640F y D580G. La polimerasa de DNA mutante de *Thermus sp.* Z05 puede comprender, por ejemplo, las sustituciones de residuos de aminoácidos I640F y I709K. La polimerasa de DNA mutante de *Thermus sp.* Z05 puede comprender, por ejemplo, sustituciones de residuos de aminoácido I640F, I709K, y D580G. La polimerasa de DNA mutante de *Thermus sp.* Z05 puede comprender, por ejemplo, las sustituciones de residuos de aminoácidos independientemente seleccionados de I640F, I709K, y/o D580G.
- 40

- 45 El aminoácido correspondiente a la posición 324 de Id. de Sec. N° 1 puede ser Lys (K). El aminoácido correspondiente a la posición 324 de Id. de Sec. N°: 1 no puede ser Met (M). El aminoácido correspondiente a la posición 461 de Id. de Sec. N° 1 puede ser Leu (L). El aminoácido correspondiente a la posición 461 de Id. de Sec. N° 1 no puede ser Met (M). El aminoácido correspondiente a la posición 517 de Id. de Sec. N° 1 puede ser Ser (S). El aminoácido correspondiente a la posición 517 de Id. de Sec. N° 1 no puede ser Arg (R). El aminoácido

correspondiente a la posición 741 de Id. de Sec. N° 1 puede ser Ser (S). El aminoácido correspondiente a la posición 741 de Id. de Sec. N° 1 no puede ser Gly (G). El aminoácido correspondiente a la posición 775 de Id. de Sec. N° 1 puede ser Arg (R). El aminoácido correspondiente a la posición 775 de Id. de Sec. N° 1 no puede ser Gly (G). El aminoácido correspondiente a la posición 791 de Id. de Sec. N° 1 puede ser Leu (L). El aminoácido correspondiente a la posición 789 de Id. de Sec. N° 1 no puede ser Phe (F).

Las sustituciones en el aminoácido correspondiente a la posición 709 de Id. de Sec. N° 1 descrito anteriormente puede resultar en polimerasas de DNA que tienen mejorada (es decir, aumentada) la eficacia de la transcripción inversa, aumentada la actividad RT-PCR (por ejemplo, amplificación más eficiente de un molde de RNA sin comprometer la eficiencia de la PCR en un molde de DNA), aumentada la eficiencia de RT-PCR en presencia de Mg^{2+} , aumentada la actividad de la transcriptasa inversa en presencia de inhibidores (por ejemplo, productos de degradación de la hemoglobina, tales como hemina, y/o heparina), aumentada la tasa de extensión y una mejor tolerancia al desemparejamiento en 3' en comparación con una polimerasa de control (véase la solicitud de patente de Estados Unidos No. 61/474.160). Por lo tanto, se espera que las polimerasas mejoradas que comprenden sustituciones en el aminoácido correspondiente a la posición 709 de Id. de Sec. N° 1 descritas en este documento también tendrán las propiedades mejoradas descritas anteriormente.

Además de las mutaciones y sustituciones descritas en el presente documento, las polimerasas de DNA también pueden incluir otras modificaciones, no sustitutivas. Tales modificaciones pueden incluir, por ejemplo, las modificaciones covalentes conocidas en la técnica para conferir una ventaja adicional en las aplicaciones que comprenden la extensión de polinucleótidos. Por ejemplo, una de tales modificaciones es una modificación covalente térmicamente reversible que inactiva la enzima, pero que se invierte para activar la enzima tras la incubación a una temperatura elevada, tal como la temperatura usada típicamente para la extensión de polinucleótidos. Ejemplos de reactivos para tales modificaciones térmicamente reversibles se describen en la patente de EE.UU. N° 5.773.258 y 5.677.152.

Las polimerasas de DNA pueden construirse mediante la mutación de las secuencias de DNA que codifican la correspondiente polimerasa no modificada (por ejemplo, una polimerasa de tipo salvaje o una variante correspondiente a partir de la cual deriva la polimerasa), tal como mediante el uso de técnicas comúnmente referidas como mutagénesis dirigida al sitio. Las moléculas de ácido nucleico que codifican la forma no modificada de la polimerasa pueden mutarse mediante una serie 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, Nueva York, 1990)).

A modo de ejemplo no limitativo, el sistema de dos cebadores, utilizado en el equipo Transformer Site-Directed Mutagenesis de Clontech, se puede emplear para la introducción de mutantes dirigidos al sitio en un polinucleótido que codifica una forma no modificada de la polimerasa. Después de la desnaturalización del plásmido diana en este sistema, dos cebadores se hibridan simultáneamente al plásmido; uno de estos cebadores contiene la mutación dirigida al sitio deseado, el otro contiene una mutación en otro punto del plásmido que resulta en la eliminación de un sitio de restricción. La síntesis de la segunda cadena se lleva a cabo a continuación, que une fuertemente estas dos mutaciones, y los plásmidos resultantes se transforma en una cepa mutS de E. coli. El DNA del plásmido se aísla de las bacterias transformadas, cortado con la enzima de restricción relevante (linearizando así los plásmidos no mutados) y, a continuación retransformado en E. coli. Este sistema permite la generación de mutaciones directamente en un plásmido de expresión, sin la necesidad de subclonación o la generación de fagémidos de cadena simple. La unión fuerte de las dos mutaciones y la posterior linearización de los plásmidos no mutados dan como resultado una alta eficiencia de mutación y permiten un cribado mínimo. Después de la síntesis del cebador inicial del sitio de restricción, este método requiere el uso de sólo un nuevo tipo de cebador por sitio de mutación. En lugar de preparar cada mutante de posición por separado, un conjunto de cebadores de oligonucleótidos "degenerados" se pueden sintetizar con el fin de introducir todas las mutaciones deseadas en un sitio determinado de forma simultánea. Los transformantes se pueden seleccionar mediante la secuenciación del DNA de plásmido a través de la región mutagenizada para identificar y escoger los clones mutantes. Cada DNA mutante puede cortarse y analizarse por electroforesis, tal como, por ejemplo, en un gel de aumento de detección de mutaciones (Mallinckrodt Baker, Inc., Phillipsburg, NJ) para confirmar que no se han producido otras alteraciones en la secuencia (por comparación de desplazamiento de banda respecto al control no mutado). Alternativamente, toda la región de DNA puede secuenciarse para confirmar que no han ocurrido eventos mutacionales adicionales fuera de la región diana.

Las polimerasas de DNA con más de un aminoácido sustituido pueden generarse de diversas maneras. En el caso de aminoácidos localizados juntos en la cadena polipeptídica, pueden mutarse simultáneamente utilizando un oligonucleótido que codifique para todas las sustituciones de aminoácidos deseadas. Sin embargo, si los aminoácidos están situados a cierta distancia unos de otros (separados por más de diez aminoácidos, por ejemplo) es más difícil generar un único oligonucleótido que codifique todos los cambios deseados. En lugar de ello, puede emplearse uno de los dos métodos alternativos. En el primer método, se genera un oligonucleótido separado para cada aminoácido a ser sustituido. Los oligonucleótidos se hibridan a continuación con el molde de DNA de una sola hebra de forma simultánea, y la segunda cadena de DNA que se sintetiza a partir del molde codificará todas las

sustituciones de aminoácidos deseadas. Un método alternativo implica dos o más rondas de mutagénesis para producir el mutante deseado. La primera ronda es como se describe para los mutantes individuales: el DNA que codifica la polimerasa no modificada se utiliza para el molde, un oligonucleótido que codifica la primera sustitución de aminoácidos deseada se hibrida a este molde y se genera entonces la molécula de DNA heterodúplex. La segunda ronda de mutagénesis utiliza el DNA mutado producido en la primera ronda de mutagénesis como molde. Por lo tanto, este molde ya contiene una o más mutaciones. El oligonucleótido que codifica las sustituciones adicionales deseadas de aminoácidos se hibrida a continuación con este molde, y la cadena de DNA resultante codifica ahora las mutaciones de la primera y segunda ronda de mutagénesis. Este DNA resultante se puede utilizar como un molde en una tercera ronda de mutagénesis, y así sucesivamente. Alternativamente, se puede utilizar el método de mutagénesis multisitio de Seyfang y Jin (Anal. Biochem. 324: 285-291 2004).

También se proporcionan ácidos nucleicos recombinantes que codifican cualquiera de las polimerasas de DNA de la presente invención. El uso de un ácido nucleico de la presente invención que codifica una polimerasa de DNA, puede proporcionar varios vectores. Puede utilizarse en la práctica de la invención cualquier secuencia de replicón y de control que contenga el vector que derivan de una especie compatible con la célula huésped. Generalmente, los vectores de expresión incluyen regiones de ácido nucleico reguladoras de la transcripción y de la traducción unidas operativamente al ácido nucleico que codifica la polimerasa de DNA. El término "secuencias de control" se refiere a secuencias de DNA necesarias para la expresión de una secuencia codificante unida operativamente en un organismo huésped particular. Las secuencias de control que son adecuadas para procariotas, por ejemplo, incluyen un promotor, opcionalmente una secuencia operadora, y un sitio de unión al ribosoma. Además, el vector puede contener un elemento retroregulador positivo (PRE) para mejorar la vida media del RNAm transcrito (véase la patente de EE.UU. n° 4.666.848). Las regiones de ácido nucleico reguladoras de la transcripción y de la traducción generalmente serán apropiadas para la célula huésped utilizada para expresar la polimerasa. Numerosos tipos de vectores de expresión apropiados, y secuencias reguladoras adecuadas son conocidos en la técnica 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 finalización de la transcripción, secuencias de inicio y finalización de la traducción, y secuencias potenciadoras o activadoras. En realizaciones típicas, las secuencias reguladoras incluyen un promotor y secuencias de inicio y finalización de la transcripción. Los vectores también incluyen típicamente una región de polienlazador que contiene varios sitios de restricción para la inserción de DNA extraño. En ciertas realizaciones, se utilizan "señales de fusión" para facilitar la purificación y, si se desea, la retirada posterior de la secuencia marcadora/señal, por ejemplo, "His-Tag". Sin embargo, éstos son generalmente innecesarios cuando se purifica una proteína termoactiva y/o termoestable de un huésped mesófilo (por ejemplo, *E. coli*) en el que se puede emplear un paso "heat-step". La construcción de vectores adecuados que contienen secuencias de replicación de codificación de DNA, secuencias reguladoras, genes de selección fenotípica, y la polimerasa de interés se preparan usando procedimientos de DNA recombinante estándar. Los plásmidos aislados, vectores virales, y fragmentos de DNA se escinden, se adaptan y se ligan juntos en un orden específico para generar los vectores deseados, como es bien conocido en la técnica (véase, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Nueva York, NY, segunda ed. 1989)).

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

Un ácido nucleico que codifica una polimerasa de DNA puede introducirse en una célula, ya sea solo o en combinación con un vector. Por "introducido en" o equivalentes gramaticales en el presente documento se entiende que los ácidos nucleicos entran en las células de una manera adecuada para la integración, amplificación y/o expresión posterior del ácido nucleico. El método de introducción está dictado en gran medida por el tipo de célula diana. Los métodos ejemplares incluyen la precipitación con CaPO_4 , fusión de liposomas, LIPOFECTINA®, electroporación, infección viral, y similares.

Los procariotas pueden ser utilizados típicamente como células huésped para los pasos iniciales de clonación. Son particularmente útiles para la producción rápida de grandes cantidades de DNA, para la producción de moldes de DNA de cadena sencilla utilizadas para la mutagénesis dirigida al sitio, para el cribado simultáneo de muchos mutantes, y para la secuenciación del DNA de los mutantes generados. Las células huésped procariotas adecuadas incluyen *E. coli* K12 cepa 94 (ATCC N° 31446), *E. coli* cepa W3110 (ATCC N° 27325), *E. coli* K12 cepa DG116 (ATCC N° 53606), *E. coli* X1776 (ATCC N° 31.537), y *E. coli* B; sin embargo muchas otras cepas de *E. coli*, tales como HB101, JM101, NM522, NM538, NM539, y muchas otras especies y géneros de procariotas incluyendo bacilos tales como *Bacillus subtilis*, otras enterobacteriáceas tales como *Salmonella typhimurium* o *Serratia marcesans*, y diversas especies de *Pseudomonas*, pueden utilizarse como huéspedes. Las células huésped procariotas u otras células huésped con paredes celulares rígidas se transforman típicamente usando el método del cloruro de calcio como se describe en la sección 1.82 de Sambrook et al., supra. Alternativamente, la electroporación puede utilizarse para la transformación de estas células. Las técnicas de transformación de procariotas se exponen en, por ejemplo Dower, en *Genetic Engineering, Principles and Methods* 12: 275-296 (Plenum Publishing Corp., 1990); Hanahan et al., *Meth. Enzymol*, 204:63, 1991. Los plásmidos usados típicamente para la transformación de *E. coli* incluyen

pBR322, pUC18, pUC19, pUC118, pUC119 y Bluescript M13, todos ellos se describen en las secciones 1.12-1.20 de Sambrook et al, supra . Sin embargo, también están disponibles muchos otros vectores adecuados.

Las polimerasas de DNA de la presente invención se producen típicamente mediante el cultivo de una célula huésped transformada con un vector de expresión que contiene un ácido nucleico que codifica la polimerasa de DNA, en las condiciones apropiadas para inducir o causar la expresión de la polimerasa de DNA. Los métodos de cultivo de células huésped transformadas en condiciones adecuadas para la expresión de proteínas son bien conocidos en la técnica (véase, por ejemplo, Sambrook et al., Supra). Las células huésped adecuadas para la producción de las polimerasas de vectores plasmídicos que contienen el promotor pL lambda incluyen la cepa de *E. coli* DG116 (ATCC N° 53606) (véase la patente de EE.UU. N° 5.079.352 y Lawyer, F.C. et al., PCR Methods and Applications 2: 275-87, 1993). Después de la expresión, la polimerasa puede recogerse y aislarse. Los métodos para purificar la polimerasa de DNA termoestable se describen en, por ejemplo, Lawyer et al., Supra. Una vez purificada, puede analizarse la capacidad de mejora de las polimerasas de DNA de la eficiencia de TI, aumento de la tolerancia al desemparejamiento, velocidad de extensión y/o la tolerancia a los inhibidores de TI y de la polimerasa (por ejemplo, como se describe en los ejemplos).

Las polimerasas de DNA mejoradas de la presente invención se pueden usar para cualquier propósito en el que tal actividad enzimática sea necesaria o deseada. Por consiguiente, se proporcionan métodos de extensión de polinucleótidos (por ejemplo, PCR), utilizando las polimerasas. Las condiciones adecuadas para la extensión de polinucleótidos son conocidos en la técnica (véase, por ejemplo, Sambrook et al., supra; véase también Ausubel et al., Short Protocols in Molecular Biology (4ª ed, John Wiley & Sons 1999). En general, un cebador se anilla, 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 polimerasa de DNA y trifosfatos de nucleósidos en un ambiente adecuado para permitir la adición de uno o más nucleótidos al 3' del cebador, produciendo de esta manera un cebador extendido complementario al ácido nucleico diana. El cebador puede incluir, por ejemplo, uno o más análogo(s) de nucleótidos. Además, los trifosfatos de nucleósidos pueden ser nucleótidos convencionales, nucleótidos no convencionales (por ejemplo, ribonucleótidos o nucleótidos marcados), o una mezcla de los mismos. La reacción de extensión de polinucleótidos puede comprender la amplificación de un ácido nucleico diana. Las condiciones adecuadas para la amplificación de ácidos nucleicos utilizando una polimerasa de DNA y un par de cebadores también son conocidos en la técnica (por ejemplo, PCR amplification methods) (véase, por ejemplo, Sambrook et al., supra; Ausubel et al., Supra; PCR Applications: Protocols for Functional Genomics (Innis et al. eds., Academic Press 1999). La reacción de extensión del polinucleótido puede comprender la transcripción inversa de un molde de RNA (por ejemplo, RT-PCR). La polimerasa mejorada puede encontrar uso en la secuenciación 454 (Margulies, M et al., 2005, Nature, 437, 376-380).

Opcionalmente, la reacción de extensión de cebador comprende un inhibidor real o potencial de una polimerasa de referencia o no modificada. El inhibidor puede inhibir, por ejemplo, la tasa de extensión de ácido nucleico y/o la eficiencia de transcripción inversa de una polimerasa de referencia o sin modificar (control). El inhibidor puede ser hemoglobina, o un producto de degradación de la misma. Por ejemplo, el producto de degradación de la hemoglobina puede ser un producto de degradación hemo, tales como hemina, hematoporfirina, o bilirrubina. El inhibidor puede ser un quelante de hierro o un pigmento púrpura. El inhibidor puede ser heparina. El inhibidor puede ser un marcador intercalante. El inhibidor puede ser melanina, que se ha descrito como un inhibidor de la polimerasa (ver, por ejemplo, Ekhardt, et al., Biochem Biophys Res Commun. 271 (3): 726-30 (2000)).

Las polimerasas de DNA de la presente invención pueden utilizarse para extender los moldes en presencia de moldes de polinucleótidos aislados de muestras que comprenden inhibidores de la polimerasa, por ejemplo, como la sangre. Por ejemplo, las polimerasas de DNA de la presente invención pueden utilizarse para ampliar los moldes en presencia de hemoglobina, un componente principal de la sangre, o en presencia de un producto de degradación de la hemoglobina. La hemoglobina puede ser degradada en varios productos de degradación hemo, tales como hemina, hematina, hematoporfirina, y bilirrubina. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en la presencia de productos de degradación de la hemoglobina, que incluye pero no se limita a, hemina, hematina, hematoporfirina, y bilirrubina. El producto de degradación de la hemoglobina puede ser hemina. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de aproximadamente 0,5 a 20,0 μm , aproximadamente de 0,5 a 10,0 μm , aproximadamente de 0,5 a 5,0 μm , aproximadamente de 1,0 a 10,0 μm , aproximadamente de 1,0 a 5,0 μM , aproximadamente de 2,0 a 5,0 μM , o aproximadamente de 2,0 a 3,0 μM de hemina. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de al menos aproximadamente 0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 4,0, 5,0, 10,0, 20,0, o más de 20 μM de hemina. Los productos de degradación de la hemoglobina incluyen quelantes de hierro y pigmentos de color púrpura. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de quelantes de hierro y/o pigmentos de color púrpura. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de cantidades de productos de degradación de la hemoglobina que inhibirían la extensión del mismo molde por una polimerasa de DNA de referencia o control.

Las polimerasas de DNA de la presente invención pueden utilizarse para extender los moldes en presencia de heparina. La heparina está comúnmente presente como anticoagulante en muestras aisladas de sangre. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de

aproximadamente 1,0 a 400 ng/μl, 1,0 a 300 ng/μl, 1,0 a 200 ng/ml, 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. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los 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 más de 400 ng/μl de heparina. Las polimerasas de DNA de la presente invención se pueden utilizar para extender los moldes en presencia de cantidades de heparina que inhibirían la extensión del mismo molde por una polimerasa de DNA de referencia o control.

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

Una polimerasa mejorada de la invención puede aumentar la eficiencia de la transcripción inversa mediante la reducción del tiempo de reacción requerido para la ampliación de un molde de RNA. Por ejemplo, una polimerasa mejorada descrita en este documento puede reducir significativamente el tiempo de reacción requerido para transcribir RNA a cDNA en comparación con una polimerasa control, aumentando así la eficiencia de la transcriptasa inversa. Sin estar limitados por la teoría, la polimerasa mejorada puede aumentar la eficiencia de la TI, por ejemplo, aumentando la actividad de la enzima sobre un molde de RNA, aumentando la velocidad de incorporación de nucleótidos y/o aumentando la capacidad de procesamiento de la polimerasa, acortando de ese modo de forma eficaz el tiempo de extensión de un molde de RNA o población de moldes de RNA. Los tiempos de reacción para la etapa inicial de TI son típicamente del orden de 30 minutos o más a 65 grados C cuando se utiliza una polimerasa no modificada o control. Una polimerasa mejorada puede ser capaz de transcribir un molde de RNA en DNAC en 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. Una polimerasa mejorada puede ser capaz de transcribir un molde de RNA derivado del transcrito JP2-5 de la hepatitis C (VHC), que contiene las primeras 800 bases del VHC genotipo 1b NTR 5', en DNAC en menos tiempo o más rápido que una polimerasa control. Una polimerasa mejorada puede ser capaz de transcribir 240 bases del molde de RNA JP2-5 de VHC en DNAC de longitud completa en unos 15 segundos menos, 30 segundos menos, un minuto menos, dos minutos menos, 3 minutos menos, 4 minutos menos, 5 minutos menos, o unos 10 minutos de menos que una polimerasa control en condiciones de reacción idénticas. Una polimerasa mejorada puede ser capaz de transcribir 240 bases del molde de RNA JP2-5 de VHC en DNAC de longitud completa más rápido que una polimerasa control, por ejemplo, aproximadamente 5 segundos, 10 segundos, 15 segundos, 30 segundos, 45 segundos, o 60 segundos o más rápido que una polimerasa de control en condiciones de reacción idénticas. Las condiciones de reacción pueden ser las descritos en los Ejemplos. Una polimerasa mejorada descrita en el presente documento puede ponerse en contacto con un molde de RNA a 65 grados C durante 2 minutos en la mezcla de reacción descrita anteriormente. La etapa de extensión puede continuar con la amplificación por PCR del molde extendido, como se describe en los ejemplos.

La actividad TI más eficiente en las polimerasas de DNA termoestables se ha logrado usando Mn^{2+} como el activador de ion metálico divalente. Sin embargo, es bien sabido que cuando Mn^{2+} está presente en las reacciones, la fidelidad de las polimerasas de DNA es menor. A menos que se esté tratando de generar mutaciones, por lo general se intenta mantener una mayor fidelidad. Afortunadamente, la mayoría de secuenciación convencional, aplicaciones PCR y RT-PCR no requieren condiciones de alta fidelidad, porque los sistemas de detección generalmente están buscando en una población de productos. Con el advenimiento de la secuenciación de siguiente generación, PCR digital, etc., la fidelidad del producto es más importante y los métodos que permiten la síntesis de DNA de mayor fidelidad son críticos. Conseguir una actividad TI eficiente usando Mg^{2+} como el activador de ion metálico divalente es una excelente manera de aumentar sustancialmente la fidelidad de la polimerasa de DNA y permitir una copia más fiable del ácido nucleico diana. En consecuencia, una polimerasa mejorada puede permitir la extensión y/o la amplificación de moldes de RNA utilizando Mg^{2+} como el activador de ion metálico divalente eficiente, como se describe en los ejemplos.

Debido a que las polimerasas descritas en este documento también pueden tener una mayor tolerancia a los desemparejamientos, las polimerasas encuentran uso en procedimientos en los que es probable la variación del molde diana y, aún así, se desea que el molde se amplifique independientemente de la variación en el molde diana. Un ejemplo de tales moldes puede incluir, por ejemplo, secuencias virales, bacterianas, o de otros patógenos. Puede ser deseable determinar simplemente si un individuo (humano o animal no humano) tiene una infección viral o de otro tipo, independientemente de la variante viral que ha infectado el individuo. Como ejemplo, se puede utilizar un par de cebadores para amplificar el VHC usando una polimerasa de la invención y detectar la presencia del VHC incluso si el virus particular que infecta el individuo tiene una mutación que resulta en una falta de coincidencia en el sitio de hibridación del cebador.

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

Las secuencias diana específicas pueden incluir, por ejemplo, los ácidos nucleicos virales (por ejemplo, virus de la inmunodeficiencia humana (VIH), virus de la hepatitis B (VHB), citomegalovirus (CMV), parvovirus B19, virus Epstein-Barr, virus de la hepatitis C (VHC), virus del papiloma humano (VPH), el virus de la encefalitis japonesa (VEJ), el virus del Nilo Occidental (VNO), virus de la encefalitis St. Louis (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 origen animal (por ejemplo, humano) o derivan de una muestra de un animal (por ejemplo, humano) (es decir, los ácidos nucleicos virales o de un organismo patógeno puede estar presente en una muestra de una biopsia animal, muestra de sangre, muestra de orina, muestra fecal, saliva, etc.). Los ácidos nucleicos diana pueden ser, por ejemplo, regiones genéticas humanas que pueden incluir variantes asociadas con la enfermedad (por ejemplo, cáncer, diabetes, etc.). Debido a que las polimerasas de la invención pueden tener tolerancia al desemparejamiento, tales enzimas pueden ser particularmente útiles, por ejemplo, cuando la diversidad de secuencias relacionadas podría estar en una secuencia diana. La descripción se puede utilizar para detectar patógenos virales, en la que los patógenos virales tienen suficiente variación en sus genomas para que sea difícil o imposible diseñar un solo cebador o un pequeño conjunto de cebadores que amplifiquen la mayoría o todos los posibles genomas virales o en el cáncer u otra enfermedad los marcadores genéticos en que la variación en la secuencia sea conocida o que pueda producirse.

Otros métodos para detectar productos de extensión o productos de amplificación utilizando las polimerasas mejoradas descritas en este documento incluyen el uso de colorantes de unión a nucleótidos de doble cadena fluorescentes o colorantes intercalantes de nucleótidos de doble cadena fluorescentes. Ejemplos de colorantes de unión al DNA de doble cadena fluorescentes incluyen SYBR-green (Molecular Probes). Los colorantes de unión al DNA de doble cadena se pueden utilizar en conjunción con el análisis de curva de fusión para medir los productos de extensión de cebadores y/o productos de amplificación. El análisis de la curva de fusión se puede realizar en un instrumento de PCR en tiempo real, tales como el instrumento ABI 5700/7000 (formato de 96 pocillos) o el instrumento ABI 7900 (formato de 384 pocillos) con el programa incorporado (SDS 2.1). Alternativamente, el análisis de la curva de fusión se puede realizar como un análisis de punto final. Ejemplos de métodos para el análisis de punto de fusión se describen en la publicación de patente de EE.UU. n° 2006/0172324.

Los equipos se proporcionan para su uso en los métodos de extensión de cebadores descritos en el presente documento. El equipo puede estar compartimentado para facilitar su uso y contiene al menos un recipiente que proporciona una mejora de la polimerasa de DNA de acuerdo con la presente invención. También puede incluirse uno o más contenedores adicionales que proporcionan reactivos adicionales. El equipo también puede incluir un tubo de recogida de sangre, contenedor o unidad que comprende heparina o una sal de la misma, o libera la heparina en la solución. La unidad de recogida de sangre puede ser un tubo heparinizado. Tales recipientes adicionales pueden incluir cualquier reactivo u otros elementos reconocidos por el experto en la técnica para su uso en los procedimientos de extensión de cebadores de acuerdo con los métodos descritos anteriormente, incluyendo reactivos para su uso en, por ejemplo, procedimientos de amplificación de ácidos nucleicos (por ejemplo, PCR, RT-PCR), procedimientos de secuenciación de DNA, o los procedimientos de marcaje de DNA. Por ejemplo, en ciertas realizaciones, el equipo incluye además un recipiente que proporciona un cebador sentido 5' hibridable, en condiciones de extensión del cebador, a un molde de polinucleótidos predeterminado, o un par de cebadores que comprende el cebador sentido 5' y un cebador antisentido 3' correspondiente. El equipo puede incluir uno o más recipientes que proporcionan trifosfatos de nucleósidos (convencionales y/o no convencionales). El equipo puede incluir dNTP alfafosforotioato, dUTP, dTTP, y/o dNTPs marcados tales como, por ejemplo, dNTP de la familia de colorantes fluoresceína o cianina. El equipo puede incluir uno o más recipientes que proporcionan un tampón adecuado para una reacción de extensión del cebador.

Las mezclas de reacción pueden proporcionarse comprendiendo las polimerasas con una eficiencia de la transcriptasa inversa, tolerancia al desemparejamiento, velocidad de extensión y/o la tolerancia de los inhibidores de TI y de la polimerasa aumentados tal como se describe en el presente documento. Las mezclas de reacción pueden comprender además reactivos para su uso en, por ejemplo, procedimientos de amplificación de ácidos nucleicos (por ejemplo, PCR, RT-PCR), procedimientos de secuenciación de DNA, o de marcaje de DNA. Por ejemplo, las mezclas de reacción pueden comprender un tampón adecuado para una reacción de extensión del cebador. Las mezclas de reacción pueden contener también un molde de ácido nucleico (DNA y/o RNA), uno o más cebadores o sondas de polinucleótidos, trifosfatos de nucleósidos (incluyendo, por ejemplo, desoxirribonucleótidos, ribonucleótidos, nucleótidos marcados, nucleótidos no convencionales), sales (por ejemplo, Mn^{2+} , Mg^{2+}), marcajes (por ejemplo, fluoróforos). Las mezclas de reacción pueden contener un cebador sentido 5' hibridable, en condiciones de extensión del cebador, a un molde de polinucleótido predeterminado, o un par de cebadores que comprende el cebador sentido 5' y un cebador antisentido 3' correspondiente. Las mezclas de reacción pueden contener alfafosforotioato dNTP, dUTP, dTTP, y/o dNTP marcados tales como, por ejemplo, dNTP de la familia de colorantes

fluoresceína o cianina. Las mezclas de reacción pueden comprender un quelante de hierro o un colorante de color púrpura. Las mezclas de reacción pueden comprender la hemoglobina, o un producto de degradación de la hemoglobina. Los productos de degradación de la hemoglobina pueden incluir productos de degradación hemo como hemina, hematina, hematoforina y bilirrubina. Las mezclas de reacción pueden comprender heparina o una sal de la misma. La mezcla de reacción puede contener un molde de ácido nucleico que está aislado de la sangre. El molde de ácido nucleico puede ser RNA 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 polimerasa de DNA que tiene una mayor eficiencia de la transcriptasa inversa en comparación con una polimerasa control, y una segunda polimerasa de DNA que tiene actividad de la polimerasa dependiente de DNA. La segunda polimerasa de DNA puede ser una de tipo salvaje o no modificada, o puede ser una polimerasa mejorada que tiene una mayor actividad de la polimerasa dependiente de DNA. Tales mezclas de reacción son útiles para la amplificación de moldes de RNA (por ejemplo, RT-PCR), proporcionando tanto una polimerasa que tiene actividad aumentada de transcriptasa inversa y una polimerasa que tiene actividad de la polimerasa dependiente de DNA.

Ejemplos

Los siguientes ejemplos se ofrecen para ilustrar, pero no para limitar la invención reivindicada.

Ejemplo 1: generación de bibliotecas

En resumen, los pasos de este proceso de selección incluyen la generación de bibliotecas, expresión y purificación parcial de las enzimas mutantes, la detección de las enzimas para las propiedades deseadas, secuenciación de DNA, purificación clonal, y una mayor caracterización de los mutantes candidatos seleccionados. Cada uno de estos pasos se describe más adelante.

Generación de biblioteca clonal: Un ácido nucleico que codifica el dominio de la polimerasa D580G_I709K de la polimerasa de DNA Z05 se sometió a PCR propensa a errores (mutagénica) entre los sitios de restricción Bsp I y Bgl II de un plásmido que incluye esta secuencia de ácido nucleico. Los cebadores utilizados para esto se proporcionan a continuación:

Cebador directo: 5'CTACCTCCTGGACCCCTCCAA-3' (Id. de Sec. N° 30); y,
Cebador reverso: 5'-ATAACCAACTGGTAGTGGCGTGTA-3' (Id. de Sec. N° 31)

La PCR se realizó utilizando una concentración de Mg^{2+} de 1,8 mM, con el fin de generar una biblioteca con una tasa de mutación deseada. Las condiciones del tampón fueron bicina 50 mM pH 8,2, KOAc 115 mM, 8% p/v de glicerol, y 0,2 mM de cada dNTP. Una enzima de inicio caliente de PCR GeneAmp® AccuRT se utilizó a 0,15 U/mL. A partir de 5×10^5 copias de DNA de plásmido linearizado Z05 D580G_I709K por volumen de reacción de 50 μ l, las reacciones fueron desnaturalizadas utilizando una temperatura de 94 °C durante 60 segundos, a continuación, se realizaron 30 ciclos de amplificación, utilizando 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 por una extensión final a una temperatura de 72 °C durante 5 minutos.

El amplicón resultante se purificó con un equipo de purificación de PCR QIAquick (Qiagen, Inc., Valencia, CA, EE.UU.) y se cortó con Bsp I y Bgl II, y luego se volvió a purificar con un equipo de purificación de PCR QIAquick. Un plásmido vector Z05 D580G_I709K se preparó cortando con las mismas dos enzimas de restricción y tratando con fosfatasa alcalina recombinante (RAS, n° cat 03359123001) y se purificó con un equipo de purificación QIAquick PCR. El vector cortado y el inserto mutado se mezclaron en una proporción de 1:3 y se trataron con DNA ligasa de T4 durante 5 minutos a temperatura ambiente (Equipo NEB Quick Ligation™). Las ligaciones se purificaron con un equipo de purificación QIAquick PCR y se transforman en una cepa huésped de *E. coli* por electroporación.

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

La biblioteca se extendió luego sobre grandes placas de agar selectivas con ampicilina. Las colonias individuales fueron transferidas a placas de 384 pocillos que contienen caldo Luria 2X con ampicilina y 10% p/v de glicerol utilizando un selector automático de colonias (QPix2, Genetix Ltd). Estas placas se incubaron durante la noche a 30 °C para permitir el crecimiento de los cultivos y luego se almacenaron entre -70 °C y -80 °C. El glicerol añadido al caldo Luria 2X fue lo suficientemente bajo como para permitir el crecimiento del cultivo y sin embargo lo suficientemente alto como para proporcionar protección contra la congelación. Varios miles de colonias se prepararon de esta manera para su uso posterior.

Preparación de la biblioteca de extractos Parte 1 - Fermentación: A partir de las bibliotecas clonales descritas anteriormente, se preparó la correspondiente biblioteca de extractos parcialmente purificados adecuados para los

propósitos de cribaje. El primer paso de este proceso era hacer cultivos de expresión a pequeña escala de cada clon. Estos cultivos se cultivaron en formato de 96 pocillos; por lo tanto, había 4 placas de cultivo de expresión para cada placa de biblioteca de 384 pocillos. Se transfirió 0,5 µl de cada pocillo de la placa de biblioteca clonal a un pocillo de una placa de siembra de 96 pocillos, que contiene 150 µl 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 una placa incubadora/agitadora iEMS (ThermoElectron). Estos cultivos de siembra se utilizaron luego para inocular el mismo medio, pero esta vez la inoculación fue de 20 µl en 250 µl de medio A en placas de gran formato de 96 pocillos (Nunc N° 267334). Estas placas se incubaron durante la noche a 37 °C con agitación. El plásmido de expresión contenía elementos de control transcripcional, que permiten la expresión a 37 °C pero no a 30 °C. Después de la incubación durante la noche, los cultivos expresaban la proteína del clon en típicamente 1-10% de la proteína celular total. Las células de estos cultivos se recogieron por centrifugación. Estas células se congelaron (-20 °C) o se procesaron inmediatamente, como se describe a continuación.

Tabla 2. Medio A (esterilizada por filtración 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 |
| Casaminoácidos | 2,5 g/L |
| glucosa | 2 g/L |
| tiamina.HCl | 10 mg/L |
| ampicilina | 100 mg/L |

Preparación de la biblioteca de extractos Parte 2 - Extracción: Los sedimentos celulares del paso 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. Hay que tener en cuenta que el tampón contenía lisozima para ayudar en la lisis celular, y DNasa para eliminar el DNA del extracto. Para lisar las células de 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 con el fin de precipitar e inactivar las proteínas contaminantes, incluyendo las nucleasas añadidas exógenamente. Las placas se centrifugaron a 3000 x g durante 15 minutos a 4 °C y los sobrenadantes se transfirieron a una nueva placa de termociclador de 384 pocillos. Estas placas de extracto se congelaron a -20 °C para su uso posterior para cribaje. Cada pocillo contenía aproximadamente de 0,5 a 3 µM de la enzima polimerasa de la biblioteca de mutantes.

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/mL |

Ejemplo 2: Identificación de las polimerasas de DNA mutantes con eficiencia mejorada de la transcripción inversa

Cribaje de las bibliotecas de extracto de proteína bruta para la eficiencia mejorada de la transcripción reversa: La biblioteca de extractos se cribó mediante la comparación de los valores de Cp (Puntos de cruce) a partir de las curvas de crecimiento generadas mediante la actividad fluorescente de la nucleasa 5' (TaqMan) de extractos de enzima bruta en un sistema de RT-PCR para la amplificación de un amplicón de 240 pares de bases a partir del transcrito JP2-5 del virus de la hepatitis C (VHC), que contiene las primeras 800 bases del NTR 5' de VHC genotipo 1b en pSP64 poli(A) (Promega).

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

Tabla 4

| Componente | Concentración |
|----------------------|-----------------------|
| 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 TaqMan | 100 nM |
| Aptámero | 200 nM |
| dATP | 200 μ M |
| dCTP | 200 μ M |
| dGTP | 200 μ M |
| dUTP | 400 μ M |
| UNG | 0,2 Unidades/ μ l |
| RNA diana | 6666 copias/ μ l |
| Mg(OAc) ₂ | 2 mM |

Aproximadamente 5.000 clones fueron seleccionados utilizando el protocolo anterior. Cuarenta clones fueron seleccionados del conjunto original para volver a cribar en base a los valores de punto de cruce (Cp) y los valores de meseta fluorescente por encima de un corte arbitrario según los cálculos por el método de cuantificación absoluta/valor máximo de la 2ª derivada. Los pocillos de cultivo correspondientes a los mejores extractos se sembraron en medio de crecimiento fresco y se cultivaron para producir placas de cultivo nuevas que contienen los mejores mutantes, así como un número de cultivos parentales Z05 D580G_I709K que se utilizarán para los controles de comparación. Estas placas de cultivo se usaron entonces para proporcionar los extractos crudos frescos que fueron cribados de nuevo con el mismo RNA diana y las mismas condiciones que las descritas anteriormente para el cribaje original. La Tabla 5 muestra los valores medios de Cp obtenidos a partir del aumento de la señal fluorescente debido a la hidrólisis 5' de una sonda marcada con FAM. Los resultados muestran que el clon 0691-L24 amplifica la diana de RNA con mayor eficiencia que la Z05_D580G_I709K parental.

Tabla 5

| Clon | Cp media |
|-----------------|----------|
| 0691-L24 | 19,1 |
| Z05 D580G_I709K | 28,0 |

La secuencia de DNA de la región mutada del gen de la polimerasa se secuenció para determinar la(s) mutación(es) que estaban presentes en un clon único. El clon 0691-L24 fue elegida para análisis adicionales, por lo que la proteína polimerasa mutante se expresó en un frasco de cultivo, se purificó a homogeneidad, y se cuantificó.

Uso de mutante 0691-L24 totalmente purificado en RT-PCR basado en Mg²⁺: El mutante 0691-L24 purificado y cuantificado se comparó con Z05_D580G_I709K parental en RT-PCR TaqMan basado en Mg²⁺. Las eficiencias de transcripción inversa y PCR se midieron por comparación de los valores Cp de amplificaciones del transcrito de RNA JP2-5 y el plásmido lineal de DNA pJP2-5 digerido con la endonucleasa de restricción EcoRI. Los oligonucleótidos y las condiciones de la mezcla maestra (Tabla 4) fueron los mismos que los utilizados en el cribaje original. Cada reacción tuvo 100.000 copias de transcrito de RNA JP2-5, 100.000 copias de plásmido lineal de DNA pJP2-5 o 1000 copias de plásmido lineal de DNA pJP2-5. Todas las dianas se amplificaron con el cebador 1 y el cebador 2, como se describe anteriormente, en reacciones duplicadas 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 μ L. El punto de corte (Cp) se calculó por el método de cuantificación absoluta/valor máximo de la 2ª derivada y se calculó la media. Las condiciones de la mezcla maestra fueron las mismas que las descritas anteriormente en la Tabla 4, excepto que las reacciones se llevaron a cabo usando un intervalo de concentraciones de polimerasa de DNA a partir de 5 nM- 40 nM. Las condiciones de termociclado fueron: 2 minutos a 50 °C (paso "UNG"); 2 minutos a 65 °C (paso "RT"); 5 ciclos de 94 °C durante 15 segundos, seguido por 62 °C durante 30 segundos; y 45 ciclos de 91 °C durante 15 segundos, seguido por 62 °C durante 30 segundos. La Tabla 6 muestra los valores Cp obtenidos a partir del aumento de la señal fluorescente debido a la escisión de la sonda TaqMan en condiciones de enzima 20 nM.

Tabla 6

| Enzima | RNA 10 ⁵ copias Cp | DNA 10 ⁵ copias Cp | DNA 10 ³ copias Cp |
|-----------------|-------------------------------|-------------------------------|-------------------------------|
| Z05 D580G_I709K | 28,8 | 17,5 | 24,4 |
| 0691-L24 | 18,9 | 17,3 | 24,0 |

Los resultados indican que la polimerasa de DNA mutante 0691-L24 permite la amplificación más eficiente de dianas de RNA sin comprometer la eficiencia de la PCR sobre dianas de DNA, en comparación con el Z05 D580G_I709K parental.

5 Determinación de mutación(es) que confieren el fenotipo: Los resultados de la secuenciación reveló que la polimerasa expresada por el clon 0691-L24 lleva mutaciones N629D y I640F, además de las mutaciones D580G e I709K parentales. Un mutante Z05 D580G_I709K_I640F se construyó por subclonación, se purificó, cuantificó, y se comparó con 0691-L24 (Z05 D580G_I709K_N629D_I640F) y Z05 D580G_I709K parental en RT-PCR TaqMan activado con Mg²⁺ variando las concentraciones de KOAc desde 40 mM a 110 mM y purificada enzima 25 nM. Las
 10 condiciones de la mezcla maestra eran las mismas que las descritos anteriormente en la Tabla 4. Cada reacción tuvo 100.000 copias de transcrito de RNA JP2-5, 100.000 copias de plásmido lineal de DNA pJP2-5 o 1000 copias de plásmido lineal de DNA pJP2-5. Todas las dianas se ampliaron con el mismo conjunto de cebadores en reacciones duplicadas para generar un amplicón de 240 pares de bases. Las eficiencias de PCR y RT-PCR se determinaron por comparación de los valores Cp entre el DNA y el RNA. Todas las reacciones se realizaron en el
 15 termociclador Light Cycler 480 de Roche con un volumen de reacción de 15 µL. Se calcularon los puntos de cruce (Cp) por el método método de cuantificación absoluta/ valor máximo de la 2ª derivada y se calculó la media de Cp. Las condiciones de termociclado fueron: 2 minutos a 50 °C (paso "UNG"); 65 °C durante 2 minutos (paso "RT"); 5 ciclos de 94 °C durante 15 segundos, seguido por 62 °C durante 30 segundos; y 45 ciclos de 91 °C durante 15 segundos, seguido por 62 °C durante 30 segundos. La Tabla 7 muestra los valores Cp obtenidos del aumento de la
 20 señal fluorescente debido a la escisión de la sonda TaqMan en la condición de KOAc 60 mM.

Tabla 7

| Enzima | RNA 10 ⁵ copias Cp | DNA 10 ⁵ copias Cp | DNA 10 ³ copias Cp |
|-----------------------|-------------------------------|-------------------------------|-------------------------------|
| Z05 D580G_I709K | 28,1 | 17,2 | 24,1 |
| 0691-L24 | 18,8 | 17,2 | 24,3 |
| Z05 D580G_I709K_I640F | 19,6 | 17,0 | 24,2 |

25 0691-L24 (Z05 D580G_I709K_N629D_I640F) y Z05 D580G_I709K_I640F tienen valores similares de Cp en ambas dianas de RNA y DNA, lo que demuestra que la mutación I640F confiere la mejora observada en el rendimiento de RT-PCR.

30 Se entiende que los ejemplos y realizaciones descritos en el presente documento son sólo para fines ilustrativos y que varias modificaciones o cambios en la luz de la misma serán sugeridos por los expertos en la técnica.

REIVINDICACIONES

1. Un polimerasa de DNA que tiene al menos un 80% de identidad de secuencia con Id. de Sec. Nº 1 y que tiene aumentada la eficiencia de la transcriptasa inversa en comparación con una polimerasa de DNA control, en el que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de Id. de Sec. Nº 1 es F, y en el que la polimerasa de DNA control tiene la misma secuencia de aminoácidos que la polimerasa de DNA, excepto que el aminoácido de la polimerasa de DNA control correspondiente a la posición 640 de Id. de Sec. Nº 1 es I.
2. La polimerasa de DNA de la reivindicación 1, en el que la polimerasa de DNA comprende una secuencia de aminoácidos al menos un 90% idéntica a Id. de Sec. Nº 1.
3. La polimerasa de DNA de cualquiera de las reivindicaciones 1 o 2, en el que el aminoácido correspondiente a la posición 580 de Id. de Sec. Nº 1 se selecciona del grupo que consiste en L, G, T, Q, A, S, N, R y K.
4. La polimerasa de DNA de cualquiera de las reivindicaciones 1 a 3, en el que el aminoácido correspondiente a la posición 709 de Id. de Sec. Nº 1 se selecciona de entre el grupo que consiste en K, R, S, G, y A:
5. La polimerasa de DNA de cualquiera de las reivindicaciones 1 a 4, en el que el aminoácido correspondiente a la posición 709 de Id. de Sec. Nº 1 es K.
6. Un polimerasa de DNA que tiene al menos un 80% de identidad de secuencia con Id. de Sec. Nº 1 y que tiene aumentada la eficiencia de la transcriptasa inversa sin una disminución sustancial en la actividad de la polimerasa dependiente de DNA en comparación con una polimerasa de DNA control, en el que el aminoácido de la polimerasa de DNA correspondiente a la posición 640 de la Id. de Sec. Nº 1 es F, y el aminoácido correspondiente a la posición 709 de Id. de Sec. Nº 1 es K, y en el que la polimerasa de DNA control tiene la misma secuencia de aminoácidos que la polimerasa de DNA, excepto que el aminoácido de la polimerasa de DNA control correspondiente a la posición 640 de la Id. de Sec. Nº 1 es I y el aminoácido correspondiente a la posición 709 de Id. de Sec. Nº 1 es I.
7. La polimerasa de DNA de la reivindicación 6, en el que el aminoácido de la polimerasa de DNA correspondiente a la posición 580 de Id. de Sec. Nº 1 es G.
8. Un ácido nucleico recombinante que codifica la polimerasa de DNA de acuerdo con cualquiera de las reivindicaciones 1 a 7.
9. Un vector de expresión que comprende el ácido nucleico recombinante de la reivindicación 8.
10. Una célula huésped que comprende el vector de expresión de la reivindicación 9.
11. Un método de producción de una polimerasa de DNA, comprendiendo dicho método:
cultivar la célula huésped de la reivindicación 10 en condiciones adecuadas para la expresión del ácido nucleico que codifica la polimerasa de DNA.
12. Un método para la realización de la extensión del cebador, que comprende:
poner en contacto una polimerasa de DNA como en una de las reivindicaciones 1 a 7 con un cebador, un molde de polinucleótidos, y trifosfatos de nucleósidos en condiciones adecuadas para la extensión del cebador, produciendo de esta manera un cebador extendido.
13. Un equipo para producir un cebador extendido, que comprende:
al menos un recipiente que proporciona una polimerasa de DNA como en una de las reivindicaciones 1 a 7.
14. El equipo de acuerdo con la reivindicación 13, que comprende además uno o más contenedores seleccionados de entre el grupo que consiste en:
(a) un contenedor que proporciona un cebador hibridable, en condiciones de extensión del cebador, a un molde de polinucleótido predeterminado;
(b) un contenedor que proporciona trifosfatos de nucleósidos; y
(c) un contenedor que proporciona un tampón adecuado para la extensión del cebador.
15. Una mezcla de reacción que comprende una polimerasa de DNA como en una de las reivindicaciones 1 a 7, al menos un cebador, un molde de polinucleótidos, y trifosfatos de nucleósidos.

Figura 1

| | | | | | | | | | | | |
|-------|---|--|-----------------|------------------|-------------------|-----------------|-------|-----|---|-------|----------------------|
| Z05 | AHLSGDENLI | RV | FQ | EG | K | DI | HTQ | TAS | W | MFGVS | (Id. de Sec. N°: 12) |
| Taq | AHLSGDENLI | RV | FQ | EG | R | DI | HTE | TAS | W | MFGVP | (Id. de Sec. N°: 13) |
| Tfi | AHLSGDENLI | RV | FR | EG | K | DI | HTE | TAA | W | MFGVP | (Id. de Sec. N°: 14) |
| Tfl | AHLSGDENLI | RV | FQ | EG | R | DI | HTQ | TAS | W | MFGVS | (Id. de Sec. N°: 15) |
| Sps17 | AHLSGDENLI | RV | FR | EG | K | DI | HTE | TAA | W | MFGVP | (Id. de Sec. N°: 16) |
| Tth | AHLSGDENLI | RV | FQ | EG | K | DI | HTQ | TAS | W | MFGVP | (Id. de Sec. N°: 17) |
| Tca | AHLSGDENLI | RV | FQ | EG | K | DI | HTQ | TAS | W | MFGVP | (Id. de Sec. N°: 18) |
| Tma | AHLSGDENLL | RA | FE | EG | I | DV | HTL | TAS | R | IFNVK | (Id. de Sec. N°: 19) |
| Tne | AHLSGDENLV | KA | FE | EG | I | DV | HTL | TAS | R | IYNVK | (Id. de Sec. N°: 20) |
| Taf | AHVSKDENLL | KA | FK | ED | L | DI | HTI | TAA | K | IFGVS | (Id. de Sec. N°: 21) |
| Dra | AHIAADDPLMQ | QA | FV | EG | A | DI | HRR | TAA | Q | VLGLD | (Id. de Sec. N°: 23) |
| Bst | AHIAEDDNLJ | EA | FR | RG | L | DI | HTK | TAM | D | IFHVS | (Id. de Sec. N°: 24) |
| Bca | AHIAEDDNLJM | EA | FR | RD | L | DI | HTK | TAM | D | IFQVS | (Id. de Sec. N°: 25) |
| | -----X ₁ X ₂ X ₃ FX ₄ | X ₅ X ₆ X ₇ | DX ₈ | HTX ₉ | TAX ₁₀ | X ₁₁ | ----- | | | | (Id. de Sec. N°: 26) |

Figura 2

| A. Identities de secuencia sobre la enzima polimerasa I completa (corresponde a los aa 1-834 de Z05) | | | | | | | | | | | | | |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Nombre | Z05 | Taq | Tfi | Tfl | Sps17 | Tth | Tca | Dra | Tma | Tne | Taf | Bst | Bca |
| Z05 | | 0.864 | 0.833 | 0.859 | 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 | 0.872 | 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 | 0.824 | 0.45 | 0.371 | 0.375 | 0.353 | 0.405 | 0.397 |
| Tfl | 0.859 | 0.854 | 0.82 | | 0.824 | 0.853 | 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 | 0.452 | 0.375 | 0.377 | 0.355 | 0.407 | 0.399 |
| Tth | 0.962 | 0.872 | 0.829 | 0.853 | 0.835 | | 0.989 | 0.463 | 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 | | 0.334 | 0.325 | 0.314 | 0.338 | 0.339 |
| Tma | 0.374 | 0.382 | 0.371 | 0.381 | 0.375 | 0.373 | 0.371 | 0.334 | | 0.854 | 0.567 | 0.37 | 0.377 |
| Tne | 0.368 | 0.368 | 0.375 | 0.374 | 0.377 | 0.367 | 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 | 0.314 | 0.567 | 0.558 | | 0.356 | 0.364 |
| Bst | 0.407 | 0.397 | 0.405 | 0.397 | 0.407 | 0.406 | 0.404 | 0.338 | 0.37 | 0.377 | 0.356 | | 0.881 |
| Bca | 0.408 | 0.397 | 0.397 | 0.398 | 0.399 | 0.406 | 0.404 | 0.339 | 0.377 | 0.376 | 0.364 | 0.881 | |
| | | | | | | | | | | | | | |
| B. Identities de secuencia solo sobre el subdominio de la polimerasa (corresponde a los aa 420-834 de Z05) | | | | | | | | | | | | | |
| Nombre | Z05 | Taq | Tfi | Tfl | Sps17 | Tth | Tca | Dra | Tma | Tne | Taf | Bst | Bca |
| Z05 | | 0.901 | 0.845 | 0.891 | 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 | 0.906 | 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 | 0.566 | 0.495 | 0.49 | 0.449 | 0.512 | 0.49 |
| Tfl | 0.891 | 0.901 | 0.857 | | 0.855 | 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 | 0.566 | 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 | | 0.45 | 0.448 | 0.426 | 0.474 | 0.454 |
| Tma | 0.483 | 0.488 | 0.495 | 0.492 | 0.495 | 0.478 | 0.478 | 0.45 | | 0.883 | 0.622 | 0.474 | 0.475 |
| Tne | 0.478 | 0.473 | 0.49 | 0.48 | 0.49 | 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 | 0.426 | 0.622 | 0.615 | | 0.46 | 0.473 |
| Bst | 0.498 | 0.503 | 0.512 | 0.494 | 0.512 | 0.496 | 0.496 | 0.474 | 0.474 | 0.476 | 0.46 | | 0.898 |
| Bca | 0.49 | 0.495 | 0.49 | 0.485 | 0.49 | 0.488 | 0.488 | 0.454 | 0.475 | 0.473 | 0.473 | 0.898 | |

Figura 3

| A. Identidades de secuencia sobre la enzima polimerasa I completa (corresponde a los aa 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 solo sobre el subdominio de la polimerasa (corresponde a los aa 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 | |