

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 274**

51 Int. Cl.:

**C12N 9/12** (2006.01)

**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.10.2007 PCT/GB2007/004031**

87 Fecha y número de publicación internacional: **02.05.2008 WO08050104**

96 Fecha de presentación y número de la solicitud europea: **23.10.2007 E 07824279 (9)**

97 Fecha y número de publicación de la concesión europea: **12.04.2017 EP 2076592**

54 Título: **Polimerasa**

30 Prioridad:

**23.10.2006 GB 0621094**

**30.11.2006 GB 0623977**

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

**27.07.2017**

73 Titular/es:

**MEDICAL RESEARCH COUNCIL (100.0%)  
2nd Floor, David Phillips Building, Polaris House,  
North Star Avenue  
Swindon SN2 1FL, GB**

72 Inventor/es:

**HOLLIGER, PHILIPP;  
D'ABBADIE, MARC y  
BAAR, CLAUDIA**

74 Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 627 274 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

Polimerasa

## 5 Campo de la invención

La presente invención se refiere a polimerasas obtenidas por ingeniería genética. En particular, la invención se refiere a polimerasas obtenidas por ingeniería genética que son resistentes a ciertos inhibidores del entorno y biológicos. Se describen asimismo los usos de dichas polimerasas obtenidas por ingeniería genética y métodos para generar dichas polimerasas obtenidas por ingeniería genética.

Antecedentes de la invención

15 Las enzimas polimerasa, tales como ADN polimerasa, ARN polimerasa o transcriptasa inversa, pueden catalizar la formación de polinucleótidos de ADN o ARN utilizando una cadena de ADN o ARN existente como matriz. Las ARN polimerasas o ADN polimerasas pueden catalizar la polimerización de ARN y ADN respectivamente utilizando una matriz de ADN, mientras que la transcriptasa inversa puede catalizar la formación de ADN utilizando una matriz de ARN.

20 Las ADN polimerasas, por ejemplo, son enzimas intracelulares de origen natural y son utilizadas por una célula para replicar una cadena de ácidos nucleicos utilizando una molécula matriz para fabricar una cadena de ácidos nucleicos complementaria. Asimismo, está muy extendido el uso de ADN polimerasas *in vitro* para diversas aplicaciones bioquímicas, incluyendo la síntesis de ADNc y reacciones de secuenciación de ADN, amplificación de ácidos nucleicos a través de métodos como reacción en cadena de la polimerasa (PCR) y para métodos de amplificación mediados por transcripción de ARN.

30 La reacción en cadena de la polimerasa (PCR) es una técnica de uso muy extendido que permite amplificar exponencialmente una región específica de ADN, siempre y cuando al menos parte de la secuencia de nucleótidos sea conocida. Dicha región conocida de la secuencia se utiliza para diseñar oligonucleótidos de ADN sintéticos complementarios para cada cadena de la doble hélice de ADN. Dichos oligonucleótidos sirven como cebadores para la síntesis de ADN *in vitro*, que es catalizada por ADN polimerasa.

35 Lamentablemente, la eficacia de esta técnica en la investigación básica o en aplicaciones forenses o clínicas está limitada por algunos problemas técnicos. Se conoce una serie de sustancias que son potentes inhibidores de la actividad de polimerasa y que limitan el uso de la reacción en cadena de la polimerasa (PCR) en muestras biológicas, cuando están presentes. Entre los ejemplos se incluyen heme (y sus productos de degradación, como bilirrurbina) presentes en la sangre y las heces. Otro potente inhibidor presente en el entorno es ácido húmico.

40 Los ácidos húmicos son una mezcla compleja de ácidos polifenólicos producidos por descomposición de materia orgánica (p.ej. descomposición de vegetación de la corteza terrestre). Los ácidos húmicos son ubicuos en el suelo y el agua y, por tanto, están presentes en cualquier muestra expuesta al entorno. La inhibición de PCR a través de ácidos húmicos es por tanto especialmente pertinente para muestras de interés paleontológico, arqueológico o forense, expuestas al suelo durante extensos periodos de tiempo.

45 Ha habido algunas tentativas para eludir el problema de la contaminación de ácido húmico en las muestras de PCR. Uno de los enfoques para solventar el problema ha sido la purificación o extracción de ADN de las muestras con antelación a la PCR (La Montagne et al (2002) J Microbiol Methods 49:255-64; Howeler et al J Microbiol Methods 2003 54:37-45). Lamentablemente, es posible que la contaminación siga siendo un problema dependiendo del método de extracción utilizado (La Montagne et al 2002). Además, no todos los contaminantes se eliminan de forma completa durante los protocolos de extracción clásicos (como tratamientos con detergente, proteasa y fenol-cloroformo) y se puede producir la pérdida de muestra original. Otro problema asociado a los procedimientos de extracción incluye el uso de materiales caros, tales como columnas de intercambio iónico, extracción con perlas de vidrio, separación inmunomagnética, cromatografía de exclusión por tamaño, resinas de unión aniónica o columnas de spin (Wilson, IG (1997) Appl. Environ. Microbiol. 63:3741-3751). Por otra parte, las etapas adicionales que requiere cada protocolo de PCR aumentan los riesgos de contaminación transversal y los consiguientes resultados falso-positivo.

60 Otro enfoque para abordar el efecto inhibitorio de ácido húmico en la actividad de polimerasa consiste en aumentar la concentración de polimerasa en cada mezcla de reacción (Sutlovic et al Croat Med J 2005 46:556-62). Se han notificado también diversos aditivos como BSA, T4 gp32 o ADN de esperma de salmón para mitigar la inhibición de la actividad de polimerasa, pero es necesario añadirlas en concentraciones sustanciales (normalmente, más de 0,2 mg/ml) (Tebbe et al Appl Environ Microbiol. (1993) 59:2657-65).

65 En todas las tentativas anteriores para evitar la inhibición de PCR a través del ácido húmico, como las técnicas de extracción o la adición de suplementos, hay un consumo de tiempo y gasto asociado al uso de reactivos adicionales o la aplicación de los protocolos.

Por tanto, en la técnica sigue necesitándose contar con un modo más sencillo y eficaz para solucionar el efecto inhibidor de ácido húmico sobre ADN polimerasas, en particular en PCR.

5 Ravensschlag (*Applications*, 2000; No. 20; Eppendorf; XP-002465385) describe la amplificación utilizando *Taq* polimerasas maestras de genes ADN<sub>r</sub> 16S de ADN contaminado con sustancias húmicas extraídas de sedimento marino.

Sumario de la invención

10 La invención se puede resumir en los siguientes párrafos enumerados:

1. Una polimerasa obtenida por ingeniería genética que tiene una secuencia de aminoácidos que comprende SEQ ID NO 2, 4, 6, 8 o 10.
- 15 2. Una molécula de ácido nucleico aislada que codifica una ADN polimerasa obtenida por ingeniería genética de acuerdo con el párrafo 1.
3. Una molécula de ácido nucleico aislada de acuerdo con el párrafo 2 que comprende una secuencia de nucleótidos tal como se expone en cualquiera de las SEQ ID NO. 1, 3, 5, 7 o 9.
4. Uso de una polimerasa obtenida por ingeniería genética de acuerdo con el párrafo 1, para producir productos de extensión de cebador.
- 20 5. Un kit para amplificar ácido nucleico que comprende una polimerasa obtenida por ingeniería genética aislada de acuerdo con el párrafo 1.

25 La presente invención aborda el problema de la inhibición de la actividad ADN polimerasas mediante ácido húmico. Concretamente, la presente invención proporciona una ADN polimerasa que es resistente a los efectos inhibidores de ácido húmico. Cabe destacar que el problema de la intolerancia al ácido húmico al que se enfrenta la ADN polimerasa en reacciones de PCR se resuelve no alterando la cantidad o la potencia del ácido húmico presente en una muestra, como en la técnica anterior, sino a través de cambios en las propiedades de la propia polimerasa.

30 En un primer aspecto, la presente divulgación proporciona una polimerasa obtenida por ingeniería genética en la que la polimerasa presenta una mejor capacidad para procesar ácido nucleico en presencia de ácido húmico en comparación con la polimerasa de tipo silvestre.

35 De acuerdo con el aspecto mencionado de la divulgación, la expresión "polimerasa obtenida por ingeniería genética" se refiere a una polimerasa que tiene una secuencia de ácidos nucleicos que no es idéntica al 100 % a nivel de ácidos nucleicos a la una o más polimerasa(s) o fragmentos de la misma de la que se deriva, y que es sintética. De acuerdo con la presente divulgación, la polimerasa obtenida por ingeniería genética puede derivarse de ADN polimerasa de tipo silvestre por sustitución, delección o inserción de uno o más aminoácidos. La expresión "polimerasa obtenida por ingeniería genética", tal como se define en el presente documento también incluye dentro de su alcance fragmentos, derivados y homólogos de una "polimerasa obtenida por ingeniería genética" tal como se definen en el presente documento, siempre y cuando presenten la propiedad requerida de poseer una mejor capacidad para procesar ácido nucleico en presencia de ácido húmico en comparación con la polimerasa de tipo silvestre.

45 Una polimerasa de "tipo silvestre" es una polimerasa que no ha sido obtenida por ingeniería genética de acuerdo con la presente divulgación. Preferentemente, una polimerasa de tipo silvestre es una polimerasa que se somete al procedimiento de ingeniería reivindicado; por tanto, la polimerasa de tipo silvestre es una forma no modificada de la polimerasa obtenida por ingeniería genética.

50 "Mejor capacidad" ha de interpretarse como un aumento de cualquier función de la polimerasa obtenida por ingeniería genética que le permita procesar ácido nucleico en comparación con las de la polimerasa de tipo silvestre. Esto incluye un aumento de la capacidad de la polimerasa para catalizar la formación de una unión entre el grupo 3' hidroxilo en el extremo creciente del cebador del ácido nucleico y el grupo 50-fosfato del nucleótido trifosfato. Las funciones de ADN polimerasas incluyen también, pero no se limitan a ellas, la incorporación de subunidades de desoxirribonucleótido o derivados de ellas, transferencia de fosforilo, translocación a lo largo de una matriz de ADN, extensión de sustratos de cebador, reconocimiento de matriz y replicación o amplificación de ADN matriz.

60 Una polimerasa obtenida por ingeniería genética de acuerdo con la divulgación puede ser una ADN polimerasa. Las personas especializadas en la técnica conocerán ADN polimerasa, así como sus funciones y propiedades. Una ADN polimerasa obtenida por ingeniería genética tendrá propiedades y características similares a la polimerasa obtenida por ingeniería genética de la descripción en el sentido de que poseerá una mejor capacidad para procesar ácido nucleico en comparación con las ADN polimerasas de tipo silvestre de las que se haya podido derivar.

65 La ADN polimerasa obtenida por ingeniería genética aislada por los autores de la presente invención tiene una mejor capacidad para procesar ácidos nucleicos a concentraciones comprendidas entre 0,1 % y 50 % de ácido húmico. Preferentemente, la ADN polimerasa obtenida por ingeniería genética tiene una mejor capacidad para procesar ácido nucleico a una concentración comprendida entre 1 y 30 % de ácido húmico. Es sobre todo de preferencia que

la ADN polimerasa obtenida por ingeniería genética tenga una mejor capacidad para procesar ácido nucleico a concentraciones comprendidas entre 5 y 20 % de ácido húmico. El ácido húmico se puede derivar de material orgánico descompuesto como por ejemplo suelo de turba. En el presente documento se describen métodos para la derivación de una solución de ácido húmico a concentraciones comprendidas entre 5 y 20 %. Utilizando dichos métodos, las personas especializadas en la técnica podrán determinar otras polimerasas que pueden procesar ácidos nucleicos a diferentes concentraciones de ácido húmico que entran dentro del ámbito de la presente divulgación. Los autores de la presente invención miden la capacidad de la ADN polimerasa obtenida por ingeniería genética para procesar ácido nucleico comparando la actividad de la ADN polimerasa obtenida por ingeniería genética con la ADN polimerasa de tipo silvestre a varias concentraciones de ácido húmico. Se pueden identificar pues ADN polimerasas obtenidas por ingeniería genética que son activas a concentraciones de ácido húmico en las que las ADN polimerasas de tipo silvestre no.

Se ha observado que las ADN polimerasas obtenidas por ingeniería genética de la presente divulgación son activas a concentraciones de ácido húmico al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 veces mayores que las concentraciones a las que la ADN polimerasa de tipo silvestre sigue siendo activa. Por consiguiente, se proporciona una ADN polimerasa obtenida por ingeniería genética con una mejor capacidad para procesar ácido nucleico en presencia de ácido húmico, en la que dicha capacidad se potencia a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 veces más que la de ADN polimerasa de tipo silvestre.

Las ADN polimerasas obtenidas por ingeniería genética de la presente divulgación se pueden utilizar en cualquier reacción *in vitro* en la que se considere beneficiosa la propiedad de resistencia al ácido húmico. Las ADN polimerasas se utilizan en muchas de las aplicaciones de biología molecular *in vitro*, entre las que se incluyen mutagénesis, genotecas de ADNc, secuenciado y reacción en cadena de la polimerasa (PCR). Se sabe que la inhibición de la actividad de ADN polimerasa de tipo silvestre mediante ácido húmico puede inhibir o dañar una reacción en cadena de la polimerasa (PCR). En esta técnica es esencial la actividad de ADN polimerasa, que participa en la replicación del ADN matriz en sitios marcados por cebadores, incorporando subunidades desoxirribonucleico para sintetizar una nueva cadena de ADN. Las muestras biológicas de interés paleontológico, arqueológico o forense que contienen ADN matriz pueden quedar expuestas al suelo durante periodos de tiempo prolongados y pueden contener ácido húmico. Las ADN polimerasas obtenidas por ingeniería química de la presente divulgación son particularmente adecuadas para su uso en reacciones de PCR realizadas sobre estas muestras. Por lo tanto, preferentemente, la ADN polimerasa obtenida por ingeniería genética de la presente divulgación tiene una mejor capacidad para procesar ácido nucleico dentro de una reacción en cadena de la polimerasa.

Las moléculas de ácido nucleico que codifican ADN polimerasas obtenidas por ingeniería genética de la presente divulgación se pueden obtener fácilmente de varias maneras, sin limitación, síntesis química, rastreo de genoteca genómica o ADNc, rastreo de genoteca de expresión y/o amplificación PCR de ADNc.

Los autores de la presente invención proporcionan también métodos para introducir mutaciones en ácidos nucleicos y para generar genotecas de los mismos. Las personas especializadas en la técnica estarán enteradas de las distintas técnicas para generar diversidad dentro de un gen y dentro de ácido nucleico. Las moléculas de ácido nucleico que codifican variantes se pueden producir utilizando mutagénesis dirigida a sitio, amplificación por PCR propensa a error, amplificación por PCR, u otros métodos apropiados, en los que el (los) cebador(es) tienen las mutaciones puntuales deseadas (véase Sambrook et al., supra, and Ausubel et al., supra, para las descripciones de técnicas de mutaciones). Se puede recurrir también a la síntesis química utilizando los métodos descritos por Engels et al., para preparar dichas variantes. Se pueden utilizar también otros métodos conocidos entre las personas especializadas en la técnica.

Los autores de la presente invención describen la generación de una genoteca de variantes génicas de polimerasa química que se pueden derivar por técnicas de barajado genético como "proceso de extensión escalonado" (StEP) (Zhao et al Biotechnol (1998) 16:258-261). Esta técnica permite que dos o más genes de interés de diferentes especies se recombinen aleatoriamente para producir quimeras, cuya secuencia contiene partes de los genes parentales de entrada originales.

Por consiguiente, se proporciona en la presente divulgación una ADN polimerasa obtenida por ingeniería que comprende una ADN polimerasa que se genera a partir de una genoteca derivada por recombinación de los genes de ADN polimerasa de tipo silvestre relacionados. Ventajosamente, la ADN polimerasa obtenida por ingeniería genética con resistencia a ácido húmico de acuerdo con la divulgación se deriva de una ADN polimerasa de la familia pol A. Preferentemente, la polimerasa de tipo silvestre se selecciona del grupo que consiste en Taq, T8 (una variante de Taq 11 veces más termoestable seleccionada previamente; Ghadessey et al. 2001), TTh (*Thermus thermophilus*) y Ttl (*Thermus flavus*).

Se proporciona asimismo en la presente divulgación una ADN polimerasa obtenida por ingeniería genética generada a partir de una genoteca de ácidos nucleicos derivados por mutagénesis y/o recombinación por reacción en cadena de la polimerasa propensa a error de genes de ADN polimerasa de tipo silvestre relacionados.

En un segundo aspecto de la divulgación, se proporcionan métodos para la generación de ADN polimerasas obtenidas por ingeniería genética que son resistentes a ácido húmico. Por consiguiente, se proporciona un método para producir una ADN polimerasa de la presente divulgación que comprende:

- 5 (a) preparación de una molécula de ácido nucleico que codifica ADN polimerasa;
- (b) introducción de una mutación en la molécula de ácido nucleico que codifica la polimerasa de la etapa (a) de manera que uno o más nucleótidos en una o más regiones no son idénticas a la ADN polimerasa de la que se derivan;
- 10 (c) selección de ADN polimerasas modificadas expresadas por molécula de ácido nucleico mutada; y
- (d) aislamiento y purificación de ADN polimerasa.

Un método preferente sobre todo para generar las ADN polimerasas por ingeniería genética de la presente divulgación es por evolución dirigida. Las técnicas de evolución dirigida y la autorreplicación compartimentalizada se explican con detalle en GB 97143002 y GB 98063936 y GB 01275643, a nombre de los autores de la presente invención.

Los autores de la presente invención han modificado los métodos de autorreplicación compartimentalizada y han generado de manera sorprendente ADN polimerasas que presentan resistencia a ácido húmico. Por consiguiente, en otro aspecto más de la invención, se proporciona un método para generar una ADN polimerasa por ingeniería genética que comprende las etapas de:

- a) proporcionar una agrupación de ácidos nucleicos que comprende miembros que codifican cada uno de ellos individualmente una ADN polimerasa obtenida por ingeniería genética;
- 25 b) proporcionar ácido húmico;
- c) subdividir la agrupación de ácidos nucleicos en compartimentos, de manera que cada compartimento comprende esencialmente un miembro ácido nucleico de la agrupación junto con la ADN polimerasas obtenida por ingeniería genética o variante codificada por el miembro de ácido nucleico, y ácido húmico;
- d) dejar que tenga lugar el procesamiento del miembro de ácido nucleico, y
- 30 e) detectar el procesamiento del miembro de ácido nucleico mediante la ADN polimerasa obtenida por ingeniería genética; y
- f) opcionalmente, repetir la serie de las etapas (a) a (f) una o más veces.

Preferentemente, el procesamiento de dicho miembro de ácido nucleico forma parte de una reacción en cadena de polimerasa.

Preferentemente, el ácido húmico se proporciona a una concentración que inhibe la actividad de ADN polimerasa de tipo silvestre. Ventajosamente, se añade ácido húmico a una concentración suficiente como para proporcionar una presión de selección, pero no hasta el punto de inhibir toda la actividad de polimerasa. Al aplicar el método mencionado para generar por ingeniería genética ADN polimerasa, solamente las ADN polimerasas que son resistentes para una cantidad de ácido húmico dada serán capaces de procesar ácido nucleico y, posteriormente, ser detectadas.

En otro aspecto del método mencionado, el miembro comprende una célula bacteriana que expresa la ADN polimerasa obtenida por ingeniería genética de acuerdo con la presente invención. Preferentemente, la célula bacteriana es *E. Coli*.

En el método para generar ADN polimerasa por ingeniería genética mencionado, solamente las ADN polimerasas que presentan al menos alguna resistencia a ácido húmico serán capaces de procesar ácido nucleico y, posteriormente, ser detectadas. Por consiguiente, el número de copias post-amplificación del miembro de ácido nucleico que codifica la ADN polimerasa obtenida por ingeniería genética de acuerdo con la divulgación es esencialmente proporcional a la actividad de la ADN polimerasa. Preferentemente, el procesamiento de ácido nucleico se detecta determinando el número de copias del miembro de ácido nucleico.

En un a realización preferente, los compartimentos consisten en el componente acuoso encapsulado de emulsión agua en aceite. La emulsión agua en aceite se produce preferentemente emulsionando una fase acuosa con una fase oleosa en presencia de un tensioactivo que comprende 4,5 % v/v Span 80, 0,4 % v/v Tween 80 y 0,1 % v/v Triton X 100, o un tensioactivo que comprende Span 80, Tween 80 y Triton X 100 esencialmente en las mismas proporciones. Preferentemente, la relación de agua: fase oleosa es 1:2, que supone un tamaño de gota adecuado. Dichas emulsiones tienen una mayor estabilidad térmica que las emulsiones más ricas en aceite.

En otro aspecto más de la divulgación, se proporciona una ADN polimerasa obtenida por ingeniería genética caracterizada por que la secuencia de aminoácidos de la polimerasa comprende, preferentemente consiste en, la secuencia de aminoácidos designada en el presente documento como SEQ ID NO: 2.

Se proporciona asimismo una molécula de ácido nucleico aislada que codifica un polipéptido de ADN polimerasa obtenida por ingeniería genética que comprende una secuencia de aminoácidos con al menos un 80 % de identidad

con cualquiera de las SEQ ID NO 2, 4 o 6 y en la que dicho polipéptido tiene una actividad de ADN polimerasa en presencia de 5 a 20 % ácido húmico. Preferentemente, dicho polipéptido tiene al menos 90 % de identidad con restos amino de SEQ ID NO: 2, 4 o 6. Más preferentemente, dicho polipéptido tiene al menos 95 % de identidad con restos amino de SEQ ID NO: 2, 4 o 6. De forma sobre todo preferente, dicho polipéptido tiene al menos 99 % de identidad con los restos de SEQ ID NO: 2, 4 o 6.

En otra realización más de la divulgación, se proporciona una molécula de ácido nucleico aislada que codifica ADN polimerasa obtenida por ingeniería genética de acuerdo con la presente divulgación que comprende una secuencia de nucleótidos tal como se expone en las SEQ ID NO.1, 3 o 5. Preferentemente, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos que tiene al menos 90 % de identidad de secuencia con las SEQ ID NO. 1, 3 o 5.

Se proporciona asimismo en la presente divulgación, una ADN polimerasa obtenida por ingeniería genética en la que dicha ADN polimerasa obtenida por ingeniería genética tiene al menos 80 % de identidad con los restos amino de la polimerasa de tipo silvestre. Preferentemente, dicha ADN polimerasa obtenida por ingeniería genética tiene al menos 90 % de identidad con los restos amino de la polimerasa de tipo silvestre. De forma sobre todo preferente, dicha ADN polimerasa obtenida por ingeniería genética tiene al menos 95 % de identidad con los restos amino de la polimerasa de tipo silvestre. Preferentemente, dicha ADN polimerasa de tipo silvestre es una ADN polimerasa de la familia Pol A. Ventajosamente dicha ADN polimerasa de tipo silvestre se selecciona del grupo que comprende *Taq*, T8, TTh y Ttl ADN polimerasas.

Preferentemente, una ADN polimerasa obtenida por ingeniería genética de la presente divulgación tiene al menos 95 % de homología con la secuencia de aminoácidos y al menos 95 % de la capacidad de corrección de errores y termoestabilidad de la ADN polimerasa de tipo silvestre aislada de *Thermus aquaticus*, *Thermus thermophilus* o *Thermus flavus*.

En un aspecto más de la divulgación, se proporciona, una secuencia de nucleótidos que codifica los polipéptidos antes descritos. Se proporciona asimismo, una molécula de ácido nucleico recombinante que comprende una secuencia promotora operativamente unida a una molécula de ácido nucleico en la que dicha secuencia promotora puede tener una función constitutiva, inducible o específica de tejido. Se proporciona además una célula transformada con dicha molécula de ácido nucleico recombinante. Las células huésped pueden ser células huésped procariontas (como *E.coli*) o células huésped eucariontas (como células de levaduras, insectos o vertebrados). Preferentemente, la célula huésped es una célula huésped bacteriana. De forma sobre todo preferente la célula huésped es *E.coli*.

Ventajosamente, el polipéptido antes descrito se utiliza para producir productos de extensión de cebador. Preferentemente, la ADN polimerasa obtenida por ingeniería genética de la presente divulgación se utiliza en una reacción en cadena de la polimerasa.

En un aspecto más, se proporciona un kit para amplificar ADN que comprende la ADN polimerasa obtenida por ingeniería genética de la presente divulgación aislada.

#### Descripción detallada

En la puesta en práctica de la presente invención se emplearán, a no ser que se indique de otro modo, técnicas convencionales de química, biología molecular, microbiología, ADN recombinante e inmunología, asequibles para las personas especializadas en la técnica habitual. Dichas técnicas se explican en la bibliografía. Véase, p.ej., J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual, Second Edition, Books 1-3*, Cold Spring Harbor Laboratory Press; B. Roe, J. Crabtree, y A. Kahn, 1996, *DNA Isolation and Sequencing: Essential Techniques*, John Wiley & Sons; J. M. Polak and James O'D. McGee, 1990, *In Situ Hybridization: Principles and Practice*; Oxford University Press; M. J. Gait (Editor), 1984, *Oligonucleotide Synthesis: A Practical Approach*, Irl Press; and, D. M. J. Lilley and J. E. Dahlberg, 1992, *Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology*, Academic Press.

#### Ácido húmico

El ácido húmico es una sustancia compleja que se halla en ciertos depósitos de materia orgánica parcialmente descompuesta, en particular, plantas muertas (Hertkorn N. et al. 2002). Dichos depósitos existen especialmente en áreas anteriormente densamente forestadas en condiciones de humedad de tipo pantanoso. Los depósitos representan una etapa entre la vegetación en vías de descomposición (humus/humatos/turba) y la posible formación de carbón y petróleo final.

El término "ácido húmico" se refiere a cualquiera de los diversos ácidos orgánicos que se obtienen del humus, siendo el humus una materia orgánica parcialmente descompuesta. Otros términos que se utilizan para ácido húmico incluyen, pero sin limitarse sólo a ellos, humina, sustancia húmica, materia orgánica natural, ácido fúlvico, landa, ulmina, geína, ácido úlmico o géico. Las sustancias húmicas están dotadas de grupos funcionales ácidos,

principalmente ácido carboxílico, que confieren a estas moléculas la capacidad de quelar cationes multivalentes como  $Mg^{2+}$ ,  $Ca^{2+}$ , y  $Fe^{2+}$ . El ácido húmico contiene un conjunto de entidades de peso molecular relativamente bajo que incluye metales, ácidos alifáticos, éteres, ésteres, alcoholes, fenoles (ácidos carboxílicos), compuestos fenólicos, fragmentos derivados de lignina aromáticos, polisacáridos y polipéptidos (Simpson A J et al. 2002). Otras referencias para su consulta incluyen Flaig, Soil Components pp. 1-219 (Gieseking Ed., Springer, Berlin 1975) y Humic Substances II Hays et al. Ed., Wiley Interscience, John Wiley, Nueva York (1989), así como Humus Chemistry, Genesis Composition Reactions, autor F. J. Stevenson, John Wiley & Sons, Nueva York (1994).

Las muestras del entorno en las que está presente ácido húmico incluyen, pero sin limitarse a ellas, suelo, sedimento, barros, materia biológica en descomposición, restos arqueológicos, ciénagas de turba, compost y agua, de origen terrestre o subterráneo. La ADN polimerasa obtenida por ingeniería genética de la presente invención puede ser particularmente útil para reacciones de replicación o amplificación como PCR, en las que el ácido nucleico que se ha de replicar o amplificar está comprendido o ha sido expuesto a dicha muestra del entorno. Entre los usos se incluyen por ejemplo, analíticas, clonación, diagnóstico y reacciones de detección en los campos de agricultura, horticultura, forestación, forense, investigación biológica, así como en la identificación de organismos y composiciones demuestra.

#### Aislamiento y uso de ácido húmico

El ácido húmico se extrae normalmente de humus en virtud de su solubilidad en álcali fuerte y la posterior precipitación de ácido fuerte (Swift, R.S. in "Methods of soil analysis. Parte 3: Chemical methods", Sparks, D. L. (Ed.), Soil Sci. Soc. Am., Madison, 1996, pp.1018-1020). El material solubilizado que queda es en cierto modo una versión refinada de ácido húmico, denominado ácido fúlvico. Las preparaciones solubles de ácidos húmicos también están disponibles en el mercado, especialmente como suplementos de nutrientes para las plantas. Se puede obtener ácido húmico de calidad técnica, por ejemplo de Sigma-Aldrich Company Ltd, Gillingham, RU, número de producto 53680, número CAS 1415-93-6. Tal como se ilustra en el Ejemplo 1, más adelante, las soluciones de ácido húmico se pueden preparar y utilizar para analizar la ADN polimerasa obtenida por ingeniería genética candidata en cuanto a su resistencia al ácido húmico. Las ADN polimerasas candidatas se pueden someter a ensayo en cualquier reacción de replicación o amplificación, por ejemplo, PCR. Preferentemente, las ADN polimerasas candidatas se seleccionan por evolución dirigida de ADN polimerasas en presencia de ácido húmico como presión de selección. Se puede añadir el ácido húmico a cada uno de los compartimentos o microcápsulas durante la auto-replicación compartimentalizada, por ejemplo, o en cualquier otro método de evolución dirigida. Se puede emplear la adición de ácido húmico a cada compartimento para seleccionar ADN polimerasas que tengan actividad en dichas condiciones.

La resistencia o una mejor capacidad para procesar ácido nucleico se expresa convenientemente en lo que se refiere a la concentración de ácido húmico que según las observaciones inhibe la actividad de la ADN polimerasa seleccionada obtenida por ingeniería genética, en comparación con la concentración que según las observaciones inhibe la enzima ADN polimerasa de tipo silvestre. Por tanto, las ADN polimerasas obtenidas por ingeniería genética seleccionadas de acuerdo con la divulgación pueden tener 2X, 4X, 6X, 8X, 10X, 12X, 14X, 15X, 16X, 18X, 20X, 22X, 25X, 30X, o más resistencia o mejor capacidad para procesar ácido nucleico, en comparación con la enzima ADN polimerasa de tipo silvestre. De forma sobre todo preferente, las ADN polimerasas obtenidas por ingeniería genética de la presente divulgación tienen capacidad para procesar ácido nucleico 16x mejor cuando se comparan de esta forma. Las ADN polimerasas obtenidas por ingeniería genética seleccionadas tienen preferentemente 50 % o más, 60 % o más, 70 % o más, 80 % o más, 90 % o más, o incluso 100 % de actividad a la concentración del factor de inhibición.

#### Compuestos fenólicos

El ácido húmico consiste en una mezcla de macromoléculas complejas que tienen estructuras fenólicas poliméricas (Índice Merck 13ª edición).

El término "compuestos fenólicos" o polifenoles se refiere a una gama de sustancias que poseen un anillo aromático que lleva una o más sustituciones hidroxilo. Compuestos fenólicos son productos del metabolismo secundario de las plantas y están muy extendidos en el reino vegetal. Las principales clases de compuestos fenólicos de las plantas incluyen fenoles simples, ácidos fenólicos, ácidos fenilacéticos, cumarinas, naftoquinonas, estilbenos/antraquinonas, flavonoides/isoflavonoides y ligninas (para mayor detalle, véase Harbone JB 1980, "Plant Phenolics in Encyclopedia of Plant Physiology, volumen 8, páginas 329-395, editado por Bell EA and Charlwood BV, publicado por Springer-Verlag, Berlín Heidelberg Nueva York). La detección y extracción de compuestos fenólicos del suelo y otras muestras biológicas es muy conocida entre las personas especializadas en la técnica (Mahugo Santana et al Anal Bioanal Chem (2005) 382(1):125-33, Shin et al J Biotechnol (2005) 119(1):36-43).

El término "ácidos fenólicos" se refiere a derivados ácidos de fenol, incluyendo, pero sin limitarse ellos, ácido caféico, vanilina, ácido ferúlico, ácido gálico, ácido elágico y ácido cumárico. Los ácidos fenólicos forman un grupo diverso que incluye dos categorías principales: los ácidos hidroxibenzoicos y los ácidos hidroxicinámicos (King and Young, J Am Diet Assoc. (1999) 99(2):213-8). "Ácidos fitofenólicos" se refiere a ácidos fenólicos derivados de material vegetal. Los ácidos fenólicos pueden darse en las plantas en forma de ésteres o glucósidos conjugados con otros

compuestos naturales como flavonoides, alcoholes, ácidos grasos hidroxilados, esteroides y glucósidos. Los compuestos de ácido cinámico hidroxilados se dan con más frecuencia sobre todo como ésteres simples con ácidos hidroxil carboxílicos o glucosa. Los compuestos de ácido hidroxibenzoico están presentes principalmente en forma de glucósidos. Los métodos para la extracción de ácidos fenólicos a partir de muestras biológicas se explican en Luthria and Mukhopadhyay (2006) J. Agric. Food Chem 54:41-47.

### Polimerasas

Las enzimas polimerasas son capaces de catalizar la producción de ADN o ARN nuevos a partir de una matriz de ADN o ARN existente, un proceso denominado polimerización. Existen muchos tipos diferentes de polimerasas incluyendo ADN polimerasas, ARN polimerasas y transcriptasas inversas. Los métodos descritos en la presente divulgación pueden utilizarse para generar polimerasas por ingeniería genética, incluyendo ARN polimerasas, ADN polimerasas o transcriptasas inversas que son resistentes a ácido hímico o compuestos fenólicos.

### ARN polimerasas

Las ARN polimerasas (ARNP) catalizan la polimerización de una cadena de ARN a partir de una matriz de ADN en un proceso de transcripción. ARNP puede iniciar la transcripción en secuencias de ADN específicas conocidas como promotores. A continuación, produce una cadena de ADN que es complementaria para la cadena de ADN utilizada como matriz. El proceso de adición de nucleótidos a la cadena de ARN se conoce como elongación. Al contrario que las ADN polimerasas, ARNP incluye una actividad helicasa, de manera que no se necesita una enzima por separado para desenrollar ADN. No obstante, las RNP sí que funcionan en asociación con una serie de factores accesorios. Dichos factores pueden controlar diversos procesos relacionados con polimerasas como temporalización y especificidad de la expresión génica. (Kaiser et al Trends Biochem Sci. (1996) 21(9):325-6) o reparación acoplada a transcripción (Lane TF, Cancer Biol Ther. (2004) 3(6):528-33).

En eucariotas, la transcripción de genes que codifican núcleo es llevada a cabo por tres ARN polimerasas diferenciadas denominadas I, II y III (Archambault J et al; Microbiol Rev. (1993) 57(3):703-24). ARN polimerasa I participa en la síntesis de ARN ribosómico, ARN polimerasa II participa en la síntesis de precursores de ARNm y ARNs y ARN polimerasa III sintetiza ARNt y otros ARNs pequeños. En las bacterias, la misma enzima cataliza la síntesis de tres tipos de ARN: ARNm, ARNr y ARNt. La enzima nuclear tiene 5 subunidades (dos subunidades  $\alpha$ ,  $\beta$ ,  $\beta^1$  y  $\omega$ ) entre las cuales, la subunidad  $\beta$  cataliza la síntesis de RNA. En Mooney et al Cell, (1999), Vol. 98:687-690 and Cramer P (2002) Curr Opin Struct Biol 12(1):89-97, se proporciona además una explicación de la estructura, la función y la regulación de otras ARN polimerasas en eucariotas y procariontes.

### ADN polimerasas

Las ADN polimerasas obtenidas por ingeniería genética de acuerdo con la presente divulgación presentan al menos alguna resistencia a ácido hímico o compuestos fenólicos.

Las enzimas ADN polimerasas son enzimas intracelulares de origen natural y son utilizadas por una célula para replicar cadenas de ácido nucleico. Durante el proceso de replicación, se copia una secuencia de nucleótidos de una cadena de ADN por apareamiento de bases complementario en una secuencia de ácidos nucleicos complementaria. Cada nucleótido en la cadena de ADN es reconocido por un nucleótido complementario sin polimerizar y requiere que las dos cadenas de la hélice de ADN se separen, al menos transitoriamente, para que los grupos donador y aceptor de enlace de hidrogeno de cada base quede expuesto para el apareamiento de bases. Los nucleótidos sueltos de entrada apropiados se alinean así para su polimerización catalizada por enzima en una nueva cadena de ácidos nucleicos.

Las enzimas que tienen actividad ADN polimerasa catalizan la formación de un enlace entre el grupo 3'-hidroxilo en el extremo en crecimiento de una secuencia de cebador de ácido nucleico y el grupo 5'-fosfato del nucleótido trifosfato. Estos nucleótidos trifosfato se seleccionan normalmente entre desoxiadenosina trifosfato (A), desoxitimidina trifosfato (T), desoxicitidina trifosfato (C) y dexosiguanosina trifosfato (G). Sin embargo, las ADN polimerasas pueden incorporar versiones modificadas o alteradas de estos nucleótidos. El orden en el que se añadan los nucleótidos está dictado por el apareamiento de bases para una cadena de la matriz de ADN; dicho apareamiento de bases se lleva a cabo a través de la unión hidrógeno "canónica" (unión de hidrógeno entre nucleótidos A y T y entre nucleótidos G y C de cadenas de ADN opuestas), si bien en la técnica se conoce el apareamiento de bases no canónico, como por ejemplo apareamiento de bases G:U. Véase, p.ej., Adams et al., The Biochemistry of the Nucleic Acids 14-32 (11th ed. 1992). El uso *in vitro*, de enzimas que tienen actividad ADN polimerasa se está haciendo habitual en los últimos años en diversas aplicaciones bioquímicas, incluyendo la síntesis de ADNc y reacciones de secuenciación de ADN (véase, Sambrook et al., (2ª ed. Cold Spring Harbor Laboratory Press, 1989)), y la amplificación de ácidos nucleicos a través de métodos como la reacción en cadena de la polimerasa (PCR) (Mullis et al., Patentes estadounidenses Nos. 4.683.195, 4.683.202, y 4.800.159) y métodos de amplificación mediada por transcripción de ARN (p.ej., Kacian et al., Publicación PCT No. WO91/01384).

Los métodos como PCR hacen uso de ciclos de la extensión del cebador aprovechando la actividad de ADN polimerasa, seguido de desnaturalización térmica del ácido nucleico de doble cadena resultante para proporcionar una nueva matriz para otra ronda de hibridación y extensión de cebador. Dado que las altas temperaturas necesarias para la desnaturalización de cadena tienen como resultado inactivaciones irreversibles de muchas ADN polimerasas, el descubrimiento y uso de ADN polimerasas capaces de seguir activas a temperaturas por encima de aproximadamente 37 °C a 42 °C (enzimas ADN polimerasa termoestables) proporciona ventajas desde el punto de vista de los costes y la eficiencia del trabajo. Se han descubierto ADN polimerasas termoestables en una serie de organismos termófilos entre los que se incluyen, sin limitarse sólo a ellos, *Thermus aquaticus*, *Thermus thermophilus*, y especies de los géneros *Bacillus*, *Thermococcus*, *Sulfolobus*, *Pyrococcus*. Las ADN polimerasas se pueden purificar directamente a partir de estos organismos termófilos. No obstante, se pueden conseguir sustanciales aumentos del rendimiento de ADN polimerasa clonando primero el gen que codifica la enzima en un vector de expresión multicopia a través de métodos de tecnología de ADN recombinante, insertando el vector en una cepa de célula huésped capaz de expresar la enzima, desarrollando el cultivo de las células huésped que contienen el vector, y extrayendo después la ADN polimerasa desde la cepa de la célula huésped que ha expresado la enzima.

Preferentemente, la DNA polimerasa de la presente divulgación es una polimerasa termoestable. Una ADN polimerasa "termoestable" tal como se utiliza en el presente documento es una polimerasa que presenta una significativa resistencia a la desnaturalización térmica a temperaturas elevadas, normalmente por encima de la temperatura corporal (37 °C). Preferentemente, dicha temperatura se encuentra en el intervalo de 42 °C a 160 °C, más preferentemente, entre 60 y 100 °C, de forma sobre todo preferente, por encima de 90 °C. En comparación con una polimerasa no termoestable, la polimerasa termoestable presenta un significativo aumento de su vida media (tiempo de incubación a temperatura elevada que tiene como resultado un 50 % de pérdida de la actividad). Preferentemente, la polimerasa termoestable retiene un 30 % o más de su actividad tras la incubación a una temperatura elevada, más preferentemente, 40 %, 50 %, 60 %, 70 % o 80 % o más de su actividad. Con mayor preferencia aún, la polimerasa retiene un 80 % de la actividad. De forma sobre todo preferente, se retiene 90 %, 95 % o más, incluso 100 % de la actividad. Las polimerasas no termoestables podrían presentar una escasa retención, o ninguna retención de la actividad tras incubaciones similares a temperatura elevada.

Las ADN polimerasas bacterianas que se han caracterizado hasta la fecha presentan ciertos patrones de similitudes y diferencias que han llevado a algunos a dividir estas enzimas en dos grupos: aquellas cuyos genes contienen intrones/inteínas (ADN polimerasas clase B) y aquellas cuyos genes ADN polimerasa son prácticamente similares a ADN polimerasa I de *E. coli* y no contienen intrones (ADN polimerasas Clase A). Se han clonado y expresado varias ADN polimerasas termoestables Clase A y Clase B derivadas de organismos termófilos. Entre las enzimas de clase A: Lawyer, et al., J. Biol. Chem. 264:6427-6437 (1989) y Gelfund et al, patente estadounidense No. 5.079.352, notifican la clonación y expresión de ADN polimerasas termoestables de longitud completa derivadas de *Thermus aquaticus* (Taq). Lawyer et al., en PCR Methods and Applications, 2:275-287 (1993), y Barnes, publicación PCT No. WO92/06188 (1992), divulgan la clonación y expresión de versiones truncadas de la misma ADN polimerasa, mientras que Sullivan, publicación EPO No. 0482714A1 (1992), notifica la clonación de una versión mutada de la Taq ADN polimerasa. Asakura et al., J. Ferment. Bioeng. (Japón), 74:265-269 (1993), aparentemente, han clonado y expresado ADN polimerasa a partir de *Thermus thermophilus*. Gelfund et al., publicación PCT No. WO92/06202 (1992), han divulgado una ADN polimerasa termoestable purificada a partir de *Thermosipho africanus*. Akhmetzjanov y Vakhitov, Nucleic Acids Res., 20:5839 (1992) han notificado una ADN polimerasa termoestable a partir de *Thermus flavus*. Uemori et al., J. Biochem. 113:401-410 (1993) y la publicación EPO No. 0517418A2 (1992) han notificado la clonación y expresión de ADN polimerasas a partir de la bacteria termófila *Bacillus caldotenax*. Ishino et al., solicitud de patente japonesa No. HEI 4[1992]-131400 (fecha de publicación Nov. 19, 1993) notifican la clonación de una ADN polimerasa a partir de *Bacillus stearothermophilus*. Entre las enzimas de Clase B: Comb et al., publicación EPO No. 0 455 430 A3 (1991), Comb et al., publicación EPO No. 0547920A2 (1993), y Perler et al., Proc. Natl. Acad. Sci. (Estados Unidos), 89:5577-5581 (1992), notifican una ADN polimerasa termoestable recombinante a partir de *Thermococcus litoralis*. En Pisani et al., Nucleic Acids Res. 20:2711-2716 (1992) y en la publicación PCT WO93/25691 (1993) se divulga una ADN polimerasa termoestable clonada a partir de *Sulfolobus solofatarius*. En Uemori et al., Nucleic Acids Res., 21:259-265 (1993), se describe una enzima termoestable de *Pyrococcus furiosus*, mientras que en Comb et al., publicación EPO No. 0547359A1 (1993), se divulga una enzima termoestable recombinante de *Pyrococcus* sp.

Muchas ADN polimerasas termoestables poseen actividades adicionales aparte de la actividad ADN polimerasa; entre ellas se incluyen actividad 5'-3'-exonucleasa y/o actividad 3'-5' exonucleasa. Las actividades 5'-3'-exonucleasa y 3'-5' exonucleasa son conocidas entre las personas especializadas en la técnica. La actividad 3'-5' exonucleasa mejora la precisión de la cadena recién sintetizada eliminando las bases incorrectas que puedan incorporarse; las ADN polimerasas en las que dicha actividad es baja o está ausente, que incluyen aparentemente Taq ADN polimerasa (véase Lawyer et al., J. Biol Chem. 264:6427-6437), tienen tasas de error elevadas en la incorporación de restos de nucleótidos en la cadena de extensión de cebador. En aplicaciones como los procedimientos de amplificación de ácido nucleico en los que la replicación de ADN suele ser geométrica en relación con el número de ciclos de extensión de cebador, dichos errores pueden conducir a serios problemas artificiales como la heterogeneidad de la secuencia del producto de amplificación de ácido nucleico (amplicón). Por tanto la actividad 3'-5' exonucleasa es una característica deseable de una ADN polimerasa termoestable utilizada para estos propósitos.

En contraposición, la actividad 5'-3' exonucleasa presente a menudo en enzimas ADN polimerasa suele ser indeseable en una aplicación en particular, ya que puede digerir ácidos nucleicos, incluyendo cebadores, que tienen un extremo 5' desprotegido. Por lo tanto, una ADN polimerasa termoestable con una actividad 5'-3' exonucleasa atenuada, o en la que dicha actividad está ausente, también es una característica deseable de una enzima para aplicaciones bioquímicas. Se han descrito varias enzimas ADN polimerasa en las que se ha introducido una modificación en una ADN polimerasa, que cumplen con este objetivo. Por ejemplo, se puede producir el fragmento de Klenow de ADN polimerasa I de *E. coli* como fragmento proteolítico de la holoenzima en la que se ha eliminado el dominio de la proteína que controla la actividad 5'-3' exonucleasa. El fragmento de Klenow sigue reteniendo la actividad polimerasa y la actividad 3'-5' exonucleasa. Barnes, supra, y Gelfund et al., patente estadounidense. No. 5.079.352 han producido *Taq* ADN polimerasas recombinantes deficientes en 5'-3' exonucleasa. Ishino et al., publicación EPO No. 0517418A2, han producido ADN polimerasa deficiente en 5'-3' exonucleasa derivada de *Bacillus caldotenax*. Por otra parte, las polimerasas que carecen del dominio 5'-3' exonucleasa tienen una menor procesabilidad.

## 15 Reacción en cadena de la polimerasa y otras técnicas de amplificación

### Amplificación

Los métodos para generar ADN polimerasa por ingeniería genética de acuerdo con la divulgación implican la amplificación por matriz de los ácidos nucleicos deseados. "Amplificación" se refiere al aumento del número de copias de un fragmento de ácido nucleico en particular (o una porción de él) que resulta de una reacción en cadena enzimática (como la reacción en cadena de la polimerasa), o de la replicación de todo o parte del vector en el que ha sido clonado. Preferentemente, la amplificación de acuerdo con la divulgación es una amplificación exponencial, tal como presenta por ejemplo la reacción en cadena de la polimerasa.

En la bibliografía se han descrito muchas dianas y métodos de amplificación de señal, por ejemplo, las revisiones generales de estos métodos en Landegren, U., et al., *Science* 242:229-237 (1988) y Lewis, R., *Genetic Engineering News* 10:1, 54-55 (1990).

### 30 Reacción en cadena de la polimerasa (PCR)

La PCR es un método de amplificación de ácidos nucleicos descrito, entre otros, en las patentes estadounidenses Nos. 4.683.195 y 4.683.202. PCR consiste en ciclos repetidos de reacciones de extensión de cebador generadas por ADN polimerasa. Se desnaturaliza por calor el ADN diana y se hibridan dos oligonucleótidos, que soportan la secuencia diana en cadenas opuestas del ADN que se va a amplificar. Estos oligonucleótidos se convierten en cebadores para su uso con la ADN polimerasa. El cebador puede ser químicamente igual o diferente de la secuencia extendida (por ejemplo, se sabe que ADN polimerasa de mamífero extiende una secuencia de ADN desde un cebador de ARN). El ADN es copiado por la extensión de cebador para hacer una segunda copia de ambas cadenas. Al repetir el ciclo de desnaturalización térmica, hibridación de cebador y extensión, se puede amplificar el ADN diana un millón de veces o más en aproximadamente dos a cuatro horas. La PCR es una herramienta de la biología molecular que debe ser utilizada en combinación con una técnica de detección para determinar los resultados de la amplificación. Una ventaja de la PCR es que aumenta la sensibilidad ampliando la cantidad de ADN diana a entre 1 millón y mil millones de veces más en aproximadamente 4 horas.

La reacción en cadena de la polimerasa se puede utilizar en los métodos de selección de la divulgación del siguiente modo. Por ejemplo, se puede utilizar PCR para seleccionar variantes de *Taq* polimerasa que tiene actividad polimerasa. Tal como se ha descrito con mayor detalle anteriormente, se genera una genoteca de ácidos nucleicos que codifican individualmente ADN polimerasa o una variante de ADN polimerasa, por ejemplo *Taq* polimerasa, y se subdivide en compartimentos. Cada compartimento comprende esencialmente un miembro de la genoteca junto con ADN polimerasa o una variante codificada por ese miembro.

La ADN polimerasa o variante puede expresarse *in vivo* dentro de la bacteria transformada o cualquier otro huésped de expresión adecuado, por ejemplo, células de levadura, insectos o mamíferos, y se puede encapsular el huésped de expresión dentro del compartimento. Se aplica calor u otro medio adecuado para interrumpir el huésped y liberar la variante de polimerasa y su ácido nucleico codificador dentro del compartimento. En el caso de un huésped bacteriano, se puede emplear la expresión temporalizada de una proteína lítica, por ejemplo proteína E de  $\Phi$ X174, o el uso de un lisógeno  $\lambda$  inducible, para interrumpir la bacteria.

Está suficientemente claro que ADN polimerasa no tiene por qué ser una proteína heteróloga expresada en ese huésped (p.ej. un plásmido), sino que se puede expresar a partir de un gen que forma parte del genoma huésped. Por tanto, la polimerasa puede ser por ejemplo una polimerasa bacteriana endógena o nativa. Así pues, los métodos de selección de acuerdo con la divulgación se pueden emplear para la clonación funcional directa de ADN polimerasas desde diversas poblaciones microbianas (y sin cultivar).

Alternativamente, se puede compartimentalizar la genoteca de ácidos nucleicos junto con los componentes de un sistema de transcripción/traducción *in vitro* (tal como se describe con mayor detalle en el presente documento) y la variante de polimerasa expresada *in vitro* dentro del compartimento.

5 Cada compartimento comprende también ácido húmico. Es deseable añadir el ácido húmico a una concentración suficiente para proporcionar una presión de selección, de manera que las ADN polimerasas resistentes a ácido húmico se puedan seleccionar. Hay que subrayar que la concentración de ácido húmico no deberá ser tan alta como para que se produzca la total inhibición de la actividad de polimerasa y, en consecuencia, no se pueda seleccionar polimerasas resistentes a ácido húmico.

10 Cada compartimento comprende también componentes para la reacción PCR, por ejemplo, nucleótido trifosfatos (dNTP), tampón, magnesio y cebadores de oligonucleótido. Los cebadores de oligonucleótido pueden tener secuencias que corresponden a las secuencias que flanquean el gen polimerasa (es decir, dentro del ADN genómico o vector) o secuencias dentro del gen polimerasa. El ciclo térmico de PCR se inicia entonces para permitir que cualquiera de las variantes de polimerasa que tengan actividad polimerasa amplíe la secuencia de ácido nucleico.

15 Las polimerasas activas pueden amplificar sus secuencias de ácidos nucleicos correspondientes, al mismo tiempo que las secuencias de ácidos nucleicos que codifican polimerasas débilmente activas o inactivas se repliquen débilmente o no se repliquen en absoluto. En general, el número de copias final de cada miembro de la genoteca de ácidos nucleicos será según lo esperado proporcional al nivel de actividad de la variante de polimerasa que codifica. Los ácidos nucleicos que codifican polimerasas activas se sobre expresarán y los ácidos nucleicos que codifican polimerasas inactivas o débilmente activas se infra-representarán. A continuación, se pueden clonar y secuenciar, etc. las secuencias amplificadas resultantes, y determinar la capacidad de replicación de cada miembro.

#### 25 PCR de transcriptasa inversa

La RT-PCR se utiliza para amplificar dianas de ARN. En este proceso, se utiliza la enzima transcriptasa inversa para convertir ARN a ADN complementario (ADNc), que se puede amplificar después utilizando PCR. Se ha demostrado que este método es útil para la detección de virus de ARN.

30 Los métodos de la divulgación pueden emplear RT-PCR y la ADN polimerasa obtenida por ingeniería genética de la presente invención se puede utilizar en RT-PCR. La agrupación de ácidos nucleicos que codifica la ADN polimerasas y sus variantes se puede proporcionar en forma de genoteca de ARN. Esta genoteca podría generarse *in vivo* en bacterias, células de mamífero, levaduras, etc., que están compartimentalizadas, o por transcripción *in vitro* de ADN compartimentalizado. El ARN podría codificar la ADN polimerasa co-compartimentalizada que ha sido expresada *in vivo* (y liberada en emulsión junto con ARN a través de los medios que se describen más adelante) o *in vitro*. Otros componentes necesarios para la amplificación (polimerasas y/o transcriptasa inversas, dNTP, cebadores) están también compartimentalizados. A la presión seleccionada de ácido húmico, el producto de ADNc de la reacción de transcripción inversa sirve como matriz para amplificación por PCR.

#### 40 Otras técnicas de amplificación

45 En la presente divulgación se pueden explotar otras tecnologías de amplificación alternativas. Por ejemplo, la amplificación por círculo rodante (Lizardi et al., (1998) Nat Genet 19:225) es una tecnología de amplificación disponible en el mercado (RCAT™) que se activa mediante ADN polimerasa y que puede replicar sondas de oligonucleótido circulares con cinética lineal o geométrica en condiciones isotérmicas.

50 En presencia de dos cebadores adecuadamente diseñados, tiene lugar una amplificación geométrica a través del desplazamiento e hiper ramificación de la cadena de ADN para generar  $10^{12}$  o más copias de cada círculo en 1 hora.

Si se utiliza un único cebador, RCAT genera en pocos minutos una cadena lineal de miles de copias de ADN unidas en tándem de una diana unida covalentemente a esa diana.

55 Otra técnica más, la amplificación por desplazamiento de cadena (SDA; Walker et al., (1992) PNAS (USA) 80:392) comienza con una secuencia única definida específicamente para una diana específica. Pero a diferencia de otras técnicas que se basan en el ciclo térmico, SDA es un proceso isotérmico en el que se utiliza una serie de cebadores, ADN polimerasas y una enzima de restricción para amplificar exponencialmente la secuencia de ácido nucleico única.

60 SDA comprende tanto una fase de generación de diana como una fase de amplificación exponencial.

65 En la generación de diana, el ADN de doble cadena se desnaturaliza térmicamente creando dos copias mono-catenarias. Una serie de cebadores especialmente fabricados se combinan con ADN polimerasas (cebadores de amplificación para copiar la secuencia base y cebadores de interposición para desplazar las cadenas recién creadas) para formar dianas alteradas con capacidad de amplificación exponencial.

El proceso de amplificación exponencial empieza con dianas alteradas (cadenas de ADN parciales monocatenarias con sitios de reconocimiento de enzima restringidos) desde la fase de generación de diana.

Se une un cebador de amplificación a cada cadena en su secuencia de ADN complementaria. La ADN polimerasa utiliza entonces el cebador para identificar una localización para extender el cebador desde su extremo 3' utilizando la diana alterada como matriz para añadir nucleótidos individuales. El cebador extendido forma así un segmento de ADN de doble cadena que contiene un sitio de reconocimiento de enzima de restricción completo en cada extremo.

A continuación, se une una enzima de restricción al segmento de ADN de doble cadena en su sitio de reconocimiento. La enzima de restricción se disocia desde el sitio de reconocimiento después de haber segmentado solamente una cadena del segmento de doble cara, formando una mella. La ADN polimerasa reconoce la mella y extiende la cadena desde el sitio desplazando la cadena creada previamente. El sitio de reconocimiento se mella y se restaura repetidamente de esta forma con la enzima de restricción y la ADN polimerasa con desplazamiento continuo de las cadenas de ADN que contienen el segmento diana.

Cada cadena desplazada queda disponible entonces para hibridarse con cebadores de amplificación como se ha descrito. El proceso continúa con la repetición de mellas, extensión y desplazamiento de nuevas cadenas de ADN con el resultado de amplificación exponencial de la diana de ADN original.

#### Evolución dirigida

En una realización preferente, la presente divulgación proporciona un método para generar una ADN polimerasa por ingeniería genética que comprende las etapas de:

- (a) proporcionar una agrupación de ácidos nucleicos que comprende miembros que codifican cada uno de ellos individualmente una ADN polimerasa o variante de ADN polimerasa;
- (b) proporcionar ácido húmico;
- (c) subdividir la agrupación de ácidos nucleicos en compartimentos, de manera que cada compartimento comprende esencialmente un miembro ácido nucleico de la agrupación junto con la ADN polimerasas o variante codificada por el miembro de ácido nucleico, y ácido húmico;
- (d) dejar que tenga lugar el procesamiento del miembro de ácido nucleico, y
- (e) detectar el procesamiento del miembro de ácido nucleico mediante la ADN polimerasa; y opcionalmente, repetir la serie de las etapas (a) a (f) una o más veces.

Las técnicas de evolución dirigida y autorreplicación compartimentalizada se detallan en las patentes británicas GB 97143002 y GB 98063936 y GB 01275643, a nombre de los autores de la presente invención.

En su forma más simple CSR implica la segregación de genes que codifican y dirigen la producción de ADN polimerasas dentro de compartimentos acuosos diferenciados, espacialmente separados de una nueva emulsión agua en aceite térmicamente estable. Provista con nucleótido trifosfatos y cebadores flanqueantes apropiados, las polimerasas replican únicamente sus propios genes. En consecuencia, solamente se replican genes que codifican las polimerasas activas, mientras que las variantes inactivas que no pueden copiar sus genes desaparecen de la agrupación de genes. Por analogía con los sistemas biológicos, entre las variantes adaptadas de forma diferenciada, las más activas (las más idóneas) producen la mayoría de "vástagos", según lo cual se correlaciona directamente el número de copias post-selección con la productividad enzimática.

Por tanto, al exponer repertorios de genes de ADN polimerasa (diversificados a través de mutación dirigida y aleatoria) a autoamplificación y alterando las condiciones en las que puede tener lugar la autoamplificación, se puede utilizar el sistema para aislar y obtener por ingeniería genética polimerasas con una mejor resistencia al ácido húmico.

La encapsulación de PCR ha sido descrita anteriormente para vesículas lipídicas (Oberholzer, T., Albrizio, M. & Luisi, P. L. (1995) Chem. Biol. 2, 677-82) y células y tejido fijos (Haase, A. T., Retzel, E. F. & Staskus, K. A. (1990) Proc. Natl. Acad. Sci. USA 87, 4971-5; Embleton, M. J., Gorochoy, G., Jones, P. T. & Winter, G. (1992) Nucleic Acids) pero de baja eficiencia.

#### Principios que subyacen bajo la tecnología CST

##### Microcápsulas

Los compartimentos o "microcápsulas" utilizados de acuerdo con el método de la divulgación requieren propiedades físicas apropiadas para permitir que funcione dicha divulgación.

En primer lugar, para asegurar que los ácidos nucleicos y los productos génicos no se difunden entre las microcápsulas, se debe aislar el contenido de cada microcápsula del entorno de que rodea a las microcápsulas, de manera que no exista ningún cambio o muy pocos cambios en los ácidos nucleicos y los productos génicos entre las microcápsulas a lo largo de la escala de tiempo del experimento.

En segundo lugar, el método de la presente divulgación requiere que haya solo un número limitado de ácidos nucleicos por microcápsula. Esto asegura que el producto génico de un ácido nucleico individual esté aislado de otros ácidos nucleicos. Por tanto, el acoplamiento entre el ácido nucleico y el producto génico será muy específico. El factor de enriquecimiento es máximo con una media de uno o menos ácidos nucleicos por cada microcápsula, siendo la unión entre ácido nucleico y la actividad del producto génico codificado lo más hermética posible, ya que el producto génico de un ácido nucleico individual estará aislado de los productos de todos los demás ácidos nucleicos. Sin embargo, incluso aunque se aplique la situación teóricamente óptima de, por término medio, un solo ácido nucleico o menos por microcápsula, una proporción de 5, 10, 50, 100 o 1000 o más ácidos nucleicos por microcápsula será beneficiosa para clasificar una genoteca grande. Las rondas de clasificación posteriores, incluyendo encapsulación renovada con distribución de ácidos nucleicos que difieren, permitirá una clasificación más rigurosa de los ácidos nucleicos. Preferentemente, existe un solo ácido nucleico, o menos por microcápsula.

En tercer lugar, la formación y la composición de las microcápsulas no deberán suprimir la función de la maquinaria, la expresión de los ácidos nucleicos y la actividad de los productos génicos.

En consecuencia, sea cual sea el sistema de microencapsulación utilizado deberá cumplir estos tres requisitos. El (los) sistema(s) apropiados) pueden variar dependiendo de la naturaleza precisa de los requerimientos en cada aplicación de la divulgación, tal como será evidente para las personas especializadas.

Hay disponible una amplia variedad de procedimientos de microencapsulación (véase Benita, 1996) y se pueden utilizar para crear las microcápsulas de acuerdo con la presente divulgación. De hecho, en la bibliografía se han identificado más de 200 métodos de microencapsulación (Finch, 1993).

Entre ellos se incluyen vesículas acuosas envueltas con membrana como por ejemplo vesículas lipídicas (liposomas) (New, 1990) y vesículas de tensioactivo no iónico (van Hal *et al.*, 1996). Se trata de cápsulas membranosas cerradas de una o múltiples capas dobles de moléculas ensambladas no covalentemente, estando separada cada capa doble de su vecina por un compartimiento acuoso. En el caso de liposomas, la membrana está compuesta de moléculas de lípido; estas son normalmente fosfolípidos, si bien se pueden incorporar también esteroides como colesterol en las membranas (New, 1990). Se pueden realizar diversas reacciones bioquímicas catalizadas por enzima, incluyendo polimerización de ARN y ADN, dentro de los liposomas (Chakrabarti *et al.*, 1994; Oberholzer *et al.*, 1995a; Oberholzer *et al.*, 1995b; Walde *et al.*, 1994; Wick & Luisi, 1996).

Con un sistema de vesícula con envoltura de membrana, gran parte de la fase acuosa queda fuera de las vesículas y por lo tanto no está compartimentalizada. Esta fase acuosa continua deberá ser eliminada, o se deberá inhibir o destruir los sistemas biológicos en ella (por ejemplo, por digestión de ácidos con ADNasa o ARNasa) para que las reacciones se limiten a las microcápsulas (Luisi *et al.*, 1987).

Se han demostrado también reacciones bioquímicas catalizadas por enzima en microcápsulas generadas a través de diversos métodos. Muchas enzimas son activas en soluciones micelares inversas (Bru & Walde, 1991; Bru & Walde, 1993; Creagh *et al.*, 1993; Haber *et al.*, 1993; Kumar *et al.*, 1989; Luisi & B., 1987; Mao & Walde, 1991; Mao *et al.*, 1992; Pérez *et al.*, 1992; Walde *et al.*, 1994; Walde *et al.*, 1993; Walde *et al.*, 1988) como por ejemplo el sistema AOT-isocianato-agua (Menger & Yamada, 1979).

Las microcápsulas también se pueden generar por polimerización interfacial o complejación interfacial (Whateley, 1996). Las microcápsulas de esta clase tienen membranas rígidas, no permeables o membranas semi-permeables. Las microcápsulas semipermeables rebordadas con membranas de nitrato de celulosa, las membranas de poliamida y las membranas de lípido-poliámida pueden soportar reacciones bioquímicas, incluyendo sistemas multi-enzima (Chang, 1987; Chang, 1992; Lim, 1984). Se ha demostrado que las microcápsulas alginato / polilisina (Lim & Sun, 1980), que se pueden formar en condiciones muy suaves, y son perfectamente biocompatibles, de manera que proporcionan, por ejemplo, un método eficaz de encapsulación de tejidos y células vivas (Chang, 1992; Sun *et al.*, 1992).

Se pueden emplear sistemas de microencapsulación no membranosos basados en un reparto en fases de un entorno acuoso en un sistema coloidal como pueda ser una emulsión.

Preferentemente, las microcápsulas de la presente divulgación se forman a partir de emulsiones; sistemas heterogéneos de dos fases líquidas inmiscibles estando una de las fases dispersada en la otra como gotas de un tamaño microscópico o coloidal (Becher, 1957; Sherman, 1968; Lissant, 1974; Lissant, 1984).

## 60 Emulsiones

Las emulsiones se pueden producir a partir de cualquier combinación adecuada de líquidos inmiscibles. Preferentemente, la emulsión de la presente divulgación tiene agua (que contiene los componentes bioquímicos) como fase presente en forma de gotitas finalmente divididas (la fase discontinua o interna dispersa) y un líquido hidrófobo inmiscible (un aceite) como matriz en la que se suspenden estas gotas (la fase externa o continua no dispersada). Dichas emulsiones se denominan "agua en aceite" (W/O). Esto tiene la ventaja de que la fase acuosa

en su totalidad que contiene los componentes bioquímicos se compartimentaliza en gotas separadas (la fase interna). La fase externa, que es un aceite hidrófobo, no contiene por lo general ninguno de los componentes bioquímicos y, por tanto, es inerte.

5 La emulsión se puede establecer por adición de uno o más agentes superficialmente activos (tensoactivos). Dichos tensoactivos se denominan agentes emulsionantes y actúan en la interfaz agua/aceite para prevenir (o al menos retrasar) la separación de fases. Se pueden utilizar muchos aceites y muchos emulsionantes para generar emulsiones agua en aceite; una reciente compilación enumera más de 16.000 tensoactivos, muchos de los cuales se utilizan como agentes emulsionantes (Ash and Ash, 1993). Entre los aceites adecuados se incluyen aceite mineral blanco y tensoactivos no iónicos (Schick, 1966) como monooleato de sorbitano (Span™ 80; ICI) y monooleato de polioxietileno sorbitano (Tween™ 80; ICI) y Triton-X-100.

15 El uso de tensoactivos aniónicos puede ser también beneficioso. Entre los tensoactivos adecuados se incluyen colato sódico y taurocolato sódico. Son particularmente preferentes desoxicolato sódico, preferentemente a una concentración de 0,5 % p/v o inferior. La inclusión de dichos tensoactivos puede aumentar en algunos casos la expresión de ácido nucleicos y/o la actividad de productos génicos. La adición de algunos tensoactivos aniónicos a la mezcla de reacción no emulsionada anula completamente la traducción. Durante la emulsificación, sin embargo, se transfiere el tensoactivo desde la fase acuosa a la interfaz y se restaura la actividad. La adición de un agente tensoactivo a las mezclas que se van a emulsionar asegura que las reacciones tienen lugar solamente después de la compartimentalización.

25 La creación de una emulsión requiere por lo general la aplicación de energía mecánica para forzar las fases entre sí. Existen diversas formas de hacerlo mediante el empleo de diversos dispositivos mecánicos, incluyendo agitadores (como varillas magnéticas de agitación, propulsores y agitadores de turbina, dispositivos de pala y whisk), homogeneizadores (incluyendo homogeneizadores rotor/estator, homogeneizadores de válvula a alta presión y homogeneizadores de chorro), molinos coloidales, dispositivos de emulsificación de membrana y ultrasonido (Becher 1957, Dickingson, 1994)

30 Las microcápsulas acuosas formadas en las emulsiones agua en aceite son por lo general estables con un mínimo cambio, si lo hay, de los ácidos nucleicos o los productos génicos entre las microcápsulas. Por otra parte, los inventores han demostrado que tienen lugar varias reacciones bioquímicas en las microcápsulas de emulsión. Asimismo, también están activos en las microcápsulas de emulsión complicados procesos bioquímicos, sobre todo transcripción y traducción génica. Existe tecnología para crear emulsiones con volúmenes en todo el camino hasta llegar a la escala industrial de miles de litros (Becher, 1957; Sherman, 1968; Lissant, 1974; Lissant, 1984).

35 El tamaño de microcápsula preferente variará dependiendo de los requisitos precisos de cada proceso de selección individual que se vaya a realizar de acuerdo con la presente divulgación. En todos los casos, existirá un equilibrio óptimo entre el tamaño de la genoteca, el enriquecimiento requerido y la concentración de componentes requerida en cada microcápsula para conseguir una expresión eficiente y reactividad de los productos génicos.

#### 40 Expresión dentro de las microcápsulas

45 Los procesos de expresión deben tener lugar dentro de cada microcápsula individual proporcionada por la presente divulgación. Tanto la transcripción *in vitro*, como la traducción/transcripción acoplada son menos eficientes a concentraciones de ADN sub-nanomolares. Teniendo en cuenta el requerimiento de que esté presente únicamente un número limitado de moléculas de ADN en cada microcápsula, se establece por tanto un límite superior en la práctica en cuanto al posible tamaño de la microcápsula. Preferentemente, el volumen medio de las microcápsulas es inferior a  $5,2 \times 10^{-16} \text{ m}^3$ , (que corresponde a una microcápsula esférica de diámetro inferior a  $10 \mu\text{m}$ , Más preferentemente menos de  $6,5 \times 10^{-17} \text{ m}^3$  ( $5 \mu\text{m}$ ), Más preferentemente aproximadamente  $4,2 \times 10^{-18} \text{ m}^3$  ( $2 \mu\text{m}$ ) e idealmente aproximadamente  $9 \times 10^{-18} \text{ m}^3$  ( $2,6 \mu\text{m}$ ).

55 La concentración eficaz de ADN o ARN en las microcápsulas puede aumentarse artificialmente a través de diversos métodos conocidos entre las personas versadas en la técnica. Entre ellos se incluyen, por ejemplo, la adición de sustancias químicas de exclusión de volumen como polietileno glicoles (PEG) y diversas técnicas de amplificación génica, incluyendo transcripción utilizando ARN polimerasas, incluyendo las derivadas de bacterias como *E. coli* (Roberts, 1969; Blattner and Dahlberg, 1972; Roberts *et al.*, 1975; Rosenberg *et al.*, 1975), eucariotas p.ej. (Weil *et al.*, 1979; Manley *et al.*, 1983) y bacteriófagos como T7, T3 y SP6 (Melton *et al.*, 1984); la reacción en cadena de la polimerasa (PCR) (Saiki *et al.*, 1988); amplificación de Q $\beta$  replicasa (Miele *et al.*, 1983; Cahill *et al.*, 1991; Chetverin and Spirin, 1995; Katanaev *et al.*, 1995); la reacción en cadena de la ligasa LCR) (Landegren *et al.*, 1988; Barany, 1991); y sistema de replicación de secuencia auto-sostenida (Fahy *et al.*, 1991) y amplificación de desplazamiento de cadena (Walker *et al.*, 1992). Se puede recurrir incluso a técnicas de amplificación génica que requieren ciclos térmicos, tales como PCR y LCR si las emulsiones y los sistemas de transcripción *in vitro* o transcripción-traducción acoplada son termoestables (por ejemplo, los sistemas de transcripción – traducción acoplada podrían derivarse de un organismo termoestable como *Thermus aquaticus*).

65

El aumento de la concentración del ácido nucleico local eficaz da cabida al uso de microcápsulas más grandes de manera eficaz. Esto permite un límite superior en la práctica preferente para el volumen de la microcápsula de aproximadamente  $5,2 \times 10^{-16} \text{ m}^3$  (que corresponde a una esfera de diámetro 10  $\mu\text{m}$ ).

5 El tamaño de la microcápsula debe ser lo suficientemente grande como para acomodar todos los componentes requeridos de las reacciones bioquímicas que deben tener lugar dentro de la microcápsula. Por ejemplo, tanto reacciones de transcripción como las reacciones de transcripción-traducción *in vitro*, requieren una concentración de nucleósido trifosfato de aproximadamente 2 mM.

10 Por ejemplo, para transcribir un gen a una única molécula de ARN corta de 500 bases de longitud, sería necesario un mínimo de 500 moléculas de nucleósido trifosfato por microcápsula ( $8,33 \times 10^{-22}$  moles). Para constituir una solución de 2 mM, este número de moléculas ha de estar incluido en una microcápsula de un volumen de  $4,17 \times 10^{-19}$  litros ( $4,17 \times 10^{-22} \text{ m}^3$ ) que, si es esférico, deberá tener un diámetro de 93nm.

15 Por otra parte, en particular en el caso de reacciones que impliquen traducción, debe señalarse que los ribosomas necesarios para que tenga lugar la traducción son de aproximadamente 20 nm de diámetro. Por lo tanto, el límite inferior preferente para las microcápsulas es un diámetro de aproximadamente 100 nm.

20 Por lo tanto, el volumen de la microcápsula es preferentemente del orden comprendido entre  $5,2 \times 10^{-22} \text{ m}^3$  y  $5,2 \times 10^{-16} \text{ m}^3$  que corresponde a una esfera de un diámetro entre 0,1  $\mu\text{m}$  y 10  $\mu\text{m}$ , más preferentemente entre aproximadamente  $5,2 \times 10^{-19} \text{ m}^3$  y  $6,5 \times 10^{-17} \text{ m}^3$  (1  $\mu\text{m}$  y 5  $\mu\text{m}$ ). Lo más ventajoso son diámetros de esfera de aproximadamente 2,6  $\mu\text{m}$ .

25 No es coincidencia que las dimensiones preferentes de los compartimentos (gotas o de 2,6  $\mu\text{m}$  de diámetro medio) se asemejen íntimamente con las de las bacterias, por ejemplo, en *Escherichia* son bastones de 1,1-1,5 x 2,0-6,0  $\mu\text{m}$  y *Azotobacter* son células ovoides de 1,5-2,0  $\mu\text{m}$  de diámetro. En su forma más simple, la evolución de Darwin se basa en un mecanismo "un genotipo un fenotipo". La concentración de un único gen compartimentalizado, o genoma, desciende de 0,4 nM en un compartimento de 2  $\mu\text{m}$  de diámetro, a 25 pM en un compartimento de 5  $\mu\text{m}$  de diámetro. La maquinaria de transcripción/traducción procariota ha evolucionado para funcionar en compartimentos de ~1-2  $\mu\text{m}$  de diámetro, mientras que los genes en solitario están a concentraciones aproximadamente nanomolares. Un gen en solitario, en un compartimento de 2,6  $\mu\text{m}$  de diámetro está a una concentración de 0,2 nM. Esta concentración génica es lo suficientemente alta para una traducción eficiente. La compartimentalización en un volumen asegura también que, aun en el caso de que se forme solamente una sola molécula del producto génico, esté presente a aproximadamente 0,2 nM, lo que es importante si el producto génico ha de tener una actividad modificadora del propio ácido nucleico. El volumen de la microcápsula deberá seleccionarse pues teniendo en cuenta no solamente los requerimientos para la transcripción y traducción del ácido nucleico/ácido nucleico, sino también la actividad modificadora requerida para el producto génico en el método de la divulgación.

40 El tamaño de las microcapsulas de emulsión debería variar simplemente adaptando las condiciones de emulsión utilizadas para formar la emulsión de acuerdo con los requerimientos del sistema de selección. Cuanto más grande sea el tamaño de la microcápsula, mayor será el volumen requerido para encapsular un ácido nucleico/genoteca de ácidos nucleicos dada, ya que el factor limitante en definitiva será el tamaño de la microcápsula y por lo tanto el número de microcápsulas posible por unidad de volumen.

45 El tamaño de las microcápsulas se selecciona no solo teniendo en cuenta los requerimientos del sistema de transcripción/traducción, sino también los del sistema de selección empleado para la construcción ácido nucleico/ácido nucleico. Por lo tanto, es posible que los componentes del sistema de selección, como por ejemplo un sistema de modificación química, requieran volúmenes de reacción y/o concentraciones de reactivo que no sean las óptimas para la transcripción/traducción. Tal como se ha expuesto en el presente documento, dichos requerimientos se pueden acomodar a través de una etapa de re-encapsulación secundaria; asimismo, se pueden acomodar seleccionando el tamaño de la microcápsula para aumentar al máximo la transcripción/traducción y la selección en conjunto. Es de preferencia la determinación empírica del volumen óptimo de microcápsula y de la concentración del reactivo, por ejemplo, tal como se exponen en el presente documento.

55 Un "ácido nucleico" de acuerdo con la presente divulgación es como se ha descrito anteriormente. Preferentemente, un ácido nucleico es una molécula o construcción seleccionada del grupo que consiste en una molécula de ADN, una molécula de ARN, una molécula de ácido nucleico parcial o totalmente artificial que consiste exclusivamente en bases sintéticas o una mezcla de bases sintéticas y naturales, cualquiera de los mencionados unido a un polipéptido, y cualquiera de los mencionados unido a cualquier otro grupo molecular o construcción. Ventajosamente, el otro grupo molecular o construcción se puede seleccionar del grupo que consiste en ácidos nucleicos, sustancias poliméricas, perlas en partículas, por ejemplo, perlas de poliestireno, sustancias magnéticas, como perlas magnéticas, marcadores como fluoroforos o marcadores isotópicos, reactivos químicos, agentes de unión como macrocitos y similares.

60

La porción del ácido nucleico del ácido nucleico puede comprender secuencias reguladoras adecuadas, como por ejemplo las requeridas para la expresión eficiente del producto génico, por ejemplo promotores, potenciadores, secuencias de inicio de la traducción, secuencias de poliadenilación, sitios de empalme, y similares.

## 5 Selección de producto

Se puede conectar un ligando o sustrato al ácido nucleico a través de diversos medios, que serán evidentes para las personas especializadas en la técnica (véase, por ejemplo, Hermanson, 1996). Cualquier marca será suficiente para permitir la posterior selección del ácido nucleico. La clasificación se puede realizar a través de cualquier método que permita una separación preferencial, amplificación o supervivencia de los ácidos nucleicos marcados. Entre los ejemplos se incluyen selección por unión (incluyendo técnicas basadas en separación magnética, como por ejemplo, utilizando Dynabeads™), y por resistencia a la degradación (por ejemplo por nucleasas, incluyendo endonucleasas de restricción).

15 Una forma en la que se puede unir la molécula de ácido nucleico a un ligando o sustrato es a través de biotilación. Esto se puede realizar por amplificación por PCR con un cebador 5'-biotilación, de manera que se unan covalentemente biotina y ácido nucleico.

20 El ligando o sustrato que se seleccione se puede unir al ácido nucleico modificado a través de diversos medios que serán evidentes para las personas especializadas en la técnica. Un ácido nucleico biotilado puede acoplarse a una microperla de poliestireno (0,035 a 0,2 um de diámetro) que está revestida con avidina o estreptavidina, que se unirá por tanto al ácido nucleico con una afinidad muy alta. Esta perla se puede derivar con sustrato o ligando a través de un método adecuado, como adición de sustrato biotilado o acoplamiento covalente.

25 Alternativamente, el ácido nucleico biotilado se puede acoplar a avidina o estreptavidina formando complejo con una molécula de proteína grande como tiroglobulina (669 Kd) o ferritina (440 Kd). Este complejo se puede derivar con sustrato o ligando, por ejemplo, por acoplamiento covalente con el grupo alfa-amino de lisinas o a través de interacción no covalente, como biotina-avidina. El sustrato puede estar presente en forma no unida al ácido nucleico, pero conteniendo una "marca" inactiva que requiere una posterior etapa para su activación, como por ejemplo fotoactivación (p.ej. de un análogo de biotina "enjalado" (Sundberg *et al.*, 1995; Pirrung and Huang, 1996)). El catalizador que se selecciona entonces convierte el sustrato en el producto. La "marca" podría activarse entonces y el sustrato "marcado" y/o producto unido a la molécula de unión a marca (p.ej., avidina o estreptavidina) formar complejo con el ácido nucleico. La relación entre sustrato y producto unido al ácido nucleico a través de la "marca" por tanto reflejará la proporción del sustrato y el producto en solución.

35 Cuando se detienen todas las reacciones y se combinan las microcápsulas, los ácidos nucleicos que codifican enzimas activas pueden enriquecerse utilizando un anticuerpo u otra molécula que se une, o que reacciona específicamente con la "marca" Aunque tanto los sustratos como el producto tienen la marca molecular, únicamente los ácidos nucleicos que codifican el producto génico activo se co-purificarán.

40 Los términos "aislamiento", "clasificación" y "selección" así como las variaciones de los mismos, se utilizan el presente documento. Aislamiento, de acuerdo con la presente divulgación se refiere al proceso de separar una entidad de una población heterogénea, por ejemplo una muestra, de manera que quede libre de al menos una sustancia con la que estaba asociada antes del proceso de aislamiento. En una realización preferente, el aislamiento se refiere a la purificación de una entidad esencialmente para homogeneidad. La clasificación de una entidad se refiere a un proceso para aislar preferentemente las entidades deseadas en relación con las entidades no deseadas. En la medida en que esto se refiera al aislamiento de las entidades deseadas, los términos "aislamiento" y "clasificación" son equivalentes. El método de la presente divulgación permite la clasificación de los ácidos nucleicos deseados a partir de agrupaciones (genotecas o repertorios) de ácidos nucleicos que contienen el ácido nucleico deseado. Selección se utiliza para referirse al proceso (incluyendo el proceso de clasificación) de aislar una entidad de acuerdo con una propiedad del mismo en particular.

## Microcápsulas /clasificación

55 Además de los ácidos nucleicos que se han descrito, las microcápsulas de acuerdo con la divulgación comprenderán además los componentes requeridos para que tenga lugar el proceso de clasificación. Otros componentes del sistema comprenderán por ejemplo los necesarios para la transcripción y/o traducción del ácido nucleico. Estos se seleccionan para los requisitos para un sistema específico entre los siguientes: un tampón adecuado, un sistema de transcripción/replicación *in vitro* y/o un sistema de traducción *in vitro* que contiene todos los ingredientes necesarios, enzimas y co-factores, ARN polimerasas, nucleótidos, ácidos nucleicos (naturales o sintéticos), ARN de transferencia, ribosomas y aminoácidos y los sustratos de la reacción de interés para permitir la selección del producto génico modificado.

65 Un tampón adecuado será aquel en el que todos los componentes deseados del sistema biológico estén activos y por tanto dependerá de los requisitos de cada sistema de reacción concreto. Los tampones adecuados para

reacciones biológicas y/o químicas son conocidos en la técnica, así como las recetas que se proporcionan en diversos manuales de laboratorio, como Sambrook *et al.*, 1989.

5 El sistema de traducción *in vitro* comprenderá normalmente un extracto de célula, normalmente de bacteria (Zubay, 1973; Zubay, 1980; Lesley *et al.*, 1991; Lesley, 1995), reticulocitos de conejo (Pelham and Jackson, 1976), o germen de trigo (Anderson *et al.*, 1983). Muchos sistemas adecuados están disponibles en el mercado (por ejemplo de Promega) incluyendo algunos que permiten transcripción/traducción acoplada (todos los sistemas bacterianos y el sistema de reticulocito y el extracto de germen de trigo TNT™ de Promega). La mezcla de aminoácidos utilizada puede incluir aminoácidos sintéticos, si se desea, para aumentar el posible número o variedad de proteínas producidas en la genoteca. Esto se puede llevar a cabo cargando ARNt con aminoácidos artificiales y utilizando estos ARNt para la traducción *in vitro* de las proteínas que se van a seleccionar (Ellman *et al.*, 1991; Benner, 1994; Mendel *et al.*, 1995).

15 Después de cada ronda de selección, se puede evaluar el enriquecimiento de la agrupación de ácidos nucleicos para los que codifican las moléculas de interés según las transcripción/replicación *in vitro* o transcripción /traducción acoplada no compartimentalizadas. La agrupación seleccionada se clona en un vector de plásmido adecuado y se produce ARN o proteína recombinante a partir de los clones adecuados para posterior purificación y ensayo.

#### 20 Identificación de microcápsula

Se puede identificar las microcápsulas en virtud de un cambio inducido por el producto génico deseado que tiene lugar o se manifiesta en sí mismo en la superficie de la microcápsula o se puede detectar desde el exterior, tal como se describe en la sección "Clasificación de microcápsula". Dicho cambio, cuando se identifica, sirve para desencadenar la modificación del gen dentro del compartimento. En un aspecto preferente de la divulgación, la identificación de la microcápsula depende de un cambio en las propiedades ópticas de la microcápsula que resultan de una reacción que lleva a luminiscencia, fosforescencia o fluorescencia dentro de la microcápsula. La modificación del gen dentro de las microcápsulas se desencadenaría gracias a la identificación de la luminiscencia, fosforescencia o fluorescencia. Por ejemplo, la identificación de luminiscencia, fosforescencia o fluorescencia puede desencadenar el bombardeo del compartimento con fotones (u otras partículas u ondas) que conduzca a la modificación del ácido nucleico. Se ha descrito un procedimiento similar anteriormente para la clasificación rápida de células (Keij *et al.*, 1994). La modificación del ácido nucleico puede ser el resultado por ejemplo del acoplamiento de una "marca" molecular, enjaulada por un grupo protector fotolábil con ácidos nucleicos: el bombardeo con fotones de una longitud de onda apropiada lleva a la eliminación de la jaula. A continuación, se combinan las microcápsulas y se agrupan los ácidos nucleicos en combinación en un entorno. Los ácidos nucleicos que codifican productos génicos que presentan la actividad deseada se pueden seleccionar por purificación de afinidad utilizando una molécula que se une específicamente o que reacciona específicamente con la "marca".

#### 35 Procedimiento en varias etapas

40 Se podrá apreciar asimismo que de acuerdo con la presente invención, no es necesario que todos los procesos de transcripción/replicación y/o traducción y selección tengan lugar en una única etapa, teniendo lugar todas las reacciones en una microcápsula. El procedimiento de selección puede comprender dos o más etapas. En primer lugar, la transcripción/replicación y/o traducción de cada ácido nucleico de la genoteca de ácidos nucleicos puede tener lugar en una primera microcápsula. Cada producto génico se une entonces al ácido nucleico que lo codifica (que reside en la misma microcápsula). A continuación, se rompen las microcápsulas y opcionalmente, se purifican los ácidos nucleicos unidos a sus productos génicos correspondientes. Alternativamente, se pueden unir los ácidos nucleicos a sus correspondientes productos génicos empleando métodos que no se basan en la encapsulación. Por ejemplo, despliegue en fagos (Smith, G.P., 1985), despliegue de liposoma (Mattheakkis *et al.*, 1994), fusión RNA-péptido (Roberts and Szostak, 1997) o fusión de péptido represor lac (Cull, *et al.*, 1992).

50 En la segunda etapa del procedimiento, se coloca cada ácido nucleico purificado unido a su producto génico en una segunda microcápsula que contiene compontes de la reacción que se va a seleccionar. Se inicia entonces esta reacción. Una vez completadas las reacciones, se vuelven a romper las microcápsulas y se seleccionan los ácidos nucleicos modificados. En el caso de reacciones complicadas en múltiples etapas en las que participan muchos componentes por separado y etapas de reacción, es posible realizar una o más de las etapas que intervienen entre la etapa inicial de creación y la unión del producto génico con el ácido nucleico y la etapa final de generar el cambio seleccionable en el ácido nucleico.

#### 60 Genotecas de secuencias de ácidos nucleicos

En este punto, los términos "genoteca", "repertorio" y "agrupación" se utilizan de acuerdo con su significado habitual en la técnica, de modo que una genoteca de ácidos nucleicos codifica un repertorio de productos génicos. La selección de un ácido nucleico/ácido nucleico de una genoteca de ácidos nucleicos (por ejemplo una genoteca de mutante taq) de acuerdo con la presente invención requerirá en la mayoría de los casos el rastreo de un gran número de ácidos nucleicos variantes. Las genotecas de ácidos nucleicos se pueden crear de diversas formas, incluyendo las siguientes.

Las agrupaciones de ácidos nucleicos de origen natural se pueden clonar desde ADN genómico o ADNc (Sambrook *et al.*, 1989); por ejemplo, se ha demostrado que genotecas de mutante Taq u otras genotecas de ADN polimerasa obtenidas por amplificación por PCR de repertorios de Taq o ADN polimerasas son fuentes muy eficaces de fragmentos de ADN polimerasa.

5 Las genotecas de genes también se pueden obtener por codificación de todos (véase por ejemplo Smith, 1985; Parmley and Smith, 1988) o parte de los genes (véase por ejemplo Lowman *et al.*, 1991) o agrupaciones de genes (véase por ejemplo Nissim *et al.*, 1994) mediante un oligonucleótido sintético adulterado o aleatorizado. Las genotecas se pueden obtener también introduciendo mutaciones en un ácido nucleico o agrupación de ácidos nucleicos “aleatorizados” a través de diversas técnicas *in vivo*, incluyendo, el uso de “cepas de mutación” de bacterias como *E. coli mutD5* (Liao *et al.*, 1986; Yamagishi *et al.*, 1990; Low *et al.*, 1996). Se pueden introducir también mutaciones aleatorias tanto *in vivo* como *in vitro* por mutágenos químicos e ionización o irradiación de UV (véase Friedberg *et al.*, 1995), o incorporación de análogos de base mutagénica (Freese, 1959; Zaccolo *et al.*, 1996). También se pueden introducir mutaciones “al azar” en genes *in vitro* durante la polimerización, por ejemplo, utilizando polimerasas propensas a error. La PCR propensa a error introduce errores de copia aleatorios imponiendo condiciones de reacción imperfectas y por tanto mutagénicas o “descuidadas (por ejemplo, añadiendo Mn<sup>2+</sup> o Mg<sup>2+</sup> a la mezcla de reacción (Cadwell and Joyce, 1991, PCR Meth. Appl. 2:28-33; Leung *et al.*, 1989, Technique 1:11-13). Se ha demostrado que este método es útil para generar genotecas aleatorizadas de secuencias de nucleótidos. De acuerdo con el método de la divulgación, el término “aleatorio” puede referirse a posiciones aleatorias con un repertorio aleatorio de aminoácidos en esas posiciones o puede ser posiciones seleccionadas (predeterminadas) con un repertorio aleatorio de aminoácidos en las posiciones seleccionadas.

Se puede introducir una mayor diversificación utilizando recombinación de homólogo ya sea *in vivo* (véase Kowalczykowski *et al.*, 1994) o *in vitro* (Stemmer, 1994a; Stemmer, 1994b)). Un ejemplo de una técnica de recombinación homóloga *in vitro* para generar diversidad génica es barajado génico.

El barajado génico implica la fragmentación aleatoria de varios ADN mutantes seguido de su re-ensamblaje por PCR en moléculas de longitud completa (Smith, Nature, 370: 324 [1994]). Entre los ejemplos de los diversos procedimientos de barajado génico se incluyen sin limitarse a ellos, ensamblaje seguido de tratamiento con ADNasa, el proceso de extensión escalonado (STEP) y recombinación *in vitro* con cebado aleatorio. En el método mediado por ADNasa, se segmentan segmentos de ADN aislada de una agrupación de mutantes positivos en fragmentos aleatorios con ADNasa I y se somete a múltiples rondas de PCR sin añadir cebador. Las longitudes de los fragmentos aleatorios se aproximan a las del segmento sin segmentar a medida que prosiguen los ciclos de PCR con el resultado de que las mutaciones presentes en diferentes clones se mezclan y se acumulan en algunas de las secuencias resultantes. Los múltiples ciclos de selección y el barajado han llevado a un potenciamiento funcional de varias enzimas (Stemmer, Nature, 370: 398 [1994]; Stemmer, Proc. Natl. Acad. Sci. USA, 91: 10747 [1994]; Cramer *et al.*, Nat. Biotech., 14: 315 [1996]; Zhang *et al.*, Proc. Natl. Acad. Sci. USA, 94: 4504 [1997]; and Cramer *et al.*, Nat. Biotech., 15: 436 [1997]).

Se ha descrito una modificación del barajado génico, el protocolo de extensión escalonada (StEP) (WO 98/42832; Shao *et al.*, 1998; Zhao *et al.*, 1997; Zhao *et al.*, 1998). StEP implica el cebado de polinucleótidos matriz con cebadores aleatorios o flanqueantes. Se re-ensamblan los cebadores extendidos en ciclos extremadamente rápidos de PCR, que generan productos de extensión cada vez más largos. En cada ciclo, los productos de cebador/extensión se pueden hibridar con diferentes matrices sobre la base de la complementariedad de secuencia. La matriz que conecta diferentes secuencias crea “casetes de recombinación”. Se continúa el proceso hasta que se crean genes de longitud completa.

Se ha descrito asimismo una modificación de la tecnología StEP (patente estadounidense No. 5.965.408). Al igual que StEP, los cebadores aleatorios se hibridan con una(s) diana(s) para barajado. Se extienden los cebadores aleatorios hasta que se detienen por “impedimentos” como dímeros de purina. La terminación prematura se facilita por bloqueo de la polimerasa con aductos asociados con la matriz. Se aíslan los fragmentos y se utilizan en una reacción de PCR por separado para crear fragmentos que se solapan más largos.

Se conoce una amplia gama de técnicas en la especialidad para rastrear productos génicos de genotecas combinatorias obtenidas por mutaciones puntuales y para rastrear genotecas de ADNc para productos génicos que tienen una propiedad determinada. Dichas técnicas serán por lo general adaptables para un rastreo rápido de las genotecas generadas por la mutagénesis combinatoria o recombinación de homólogos o variantes de ADN polimerasa. Las técnicas más usadas para rastrear genotecas grandes comprenden normalmente la clonación de la genoteca en vectores de expresión replicables, la transformación de células apropiadas con la librería de vectores resultantes y la expresión de los genes combinatorios en condiciones en las que la detección de una actividad deseada facilita un aislamiento relativamente fácil del vector que codifica el gen cuyo producto ha sido detectado. Las técnicas de evolución dirigida para detección y selección de la actividad de ADN polimerasa deseada han sido descritas ya.

Vectores y células huésped

Se pueden utilizar vectores y células huésped adecuados para hospedar el ácido nucleico que codifica las ADN polimerasa candidato o sus genotecas, o la ADN polimerasa obtenida por ingeniería genéticas de la presente divulgación. Las células huésped y los vectores se pueden utilizar también para expresar y aislar polipéptidos de ADN polimerasas candidatos o ADN polimerasa obtenida por ingeniería genéticas de la presente divulgación. La célula huésped adecuada puede emplearse también para aislar genes de ADN polimerasa de tipo silvestre y, alternativamente, o adicionalmente, para expresar polipéptidos de ADN polimerasa de tipo silvestre para su uso en los métodos de la presente divulgación.

Vectores

Los vectores de expresión se pueden construir a partir de un vector de partida como por ejemplo un vector disponible en el mercado. Los vectores preferentes son aquellos que son compatibles con la célula huésped, bacteriana, de insecto o de mamífero. Entre dichos vectores se incluyen, entre otros, pCR11, pCR3, y pcDNA3.1 (Invitrogen, San Diego, Calif.), pBSII (Stratagene, La Jolla, Calif.), pET15 (Novagen, Madison, Wis.), pGEX (Pharmacia Biotech, Piscataway, N.J.), pEGFP-N2 (Clontech, Palo Alto, Calif.), pETL (BlueBacII, Invitrogen), pDSR-alfa (PCT Pub. No. WO 90/14363) y pFastBacDual (Gibco-BRL, Grand Island, N.Y.).

Otros vectores adecuados incluyen, pero sin limitarse a ellos, cósmidos, plásmidos, o virus modificados, pero se podrá apreciar que el sistema de vectores debe ser compatible con la célula huésped seleccionada. Dichos virus incluyen, sin limitarse a ellos plásmidos como derivados de plásmido Bluescript® (un fagémido a base de ColE1 de un alto número de copias, Stratagene Cloning Systems, La Jolla Calif.), plásmidos de clonación por PCR diseñados para clonación de productos de PCR Taq-amplificados (p.ej. TOPO™ TA Cloning® Kit, derivados de plásmido PCR2.1®, Invitrogen, Carlsbad, Calif.), pASK75, y vectores de mamífero, levaduras o virus como sistema de expresión de baculovirus (derivados de plásmido pBacPAK, Clontech, Palo Alto, Calif.).

Los vectores pueden incluir también un elemento regulador de transcripción (un promotor) unido operativamente a la secuencia de ADN polimerasa. El promotor puede contener opcionalmente porciones de operador y/o sitios de unión a ribosoma. Los ejemplos no exhaustivos de promotores bacterianos compatibles con *E. coli* incluyen: promotor trc, promotor alfa-lactamasa (penicilinas); promotor lactosa; promotor triptifano (trp), promotor arabinosa BAD operon; promotor PI lambda derivado y sitio de unión a ribosoma gen N; y el promotor tac híbrido derivado de secuencias de los promotores trp y lac UV5.

Una vez construido el vector e insertado la molécula de ácido nucleico que codifica un polipéptido de ADN polimerasa en el sitio apropiado del vector, el vector completado puede insertarse en una célula huésped adecuada para amplificación y/o expresión de polipéptido. La transformación de un vector de expresión para polipéptido de ADN polimerasa en una célula huésped seleccionada, se puede llevar a cabo a través de métodos muy conocidos, incluyendo métodos como transfección, infección, cloruro de calcio, electroporación, micro-inyección, lipofección, método DEAE-dextrano u otras técnicas conocidas. El método seleccionado dependerá en parte del tipo de células huésped que se utilice. Estos métodos y otros métodos adecuados son conocidos entre las personas especializadas en la técnica y se exponen por ejemplo en Sambrook et al., supra.

Células huésped

Se entiende por "célula huésped" o "célula obtenida por ingeniería genética recombinante" una célula que contiene un vector y que soporta la replicación y/o expresión del vector de expresión. Las células huésped pueden ser células huésped procariontas (como *E. coli*) o células huésped eucariotas (como célula de levadura, insecto o vertebrado). La célula huésped, cuando se cultiva en condiciones apropiadas sintetiza un polipéptido de ADN polimerasa que se puede recoger posteriormente desde el medio de cultivo (si la célula huésped lo secreta al medio) o directamente desde la célula huésped que lo produce (si no lo secreta). La selección de la célula huésped apropiada dependerá de diversos factores, como los niveles de expresión deseada, las modificaciones de polipéptido que son deseables o necesarias para la actividad (como glucosilación o fosforilación) y la facilidad de plegamiento en una molécula biológicamente activa.

Son de particular interés como células huésped las células bacterianas. Por ejemplo, las diversas cepas de *E. coli* (p.ej. HB101, DH5α, DH10, y MC1061) son células huésped conocidas en el campo de la biotecnología. Se pueden emplear asimismo varias cepas de *B. subtilis*, *Pseudomonas* spp., otras *Bacillus* spp., *Streptomyces* spp., y similares, en los métodos empleados en la presente divulgación.

Las células huésped se pueden utilizar para expresar el candidato heterólogo o la ADN polimerasa obtenida por ingeniería genéticas de la presente divulgación. Tal como se utiliza en el presente documento "heterólogo" haciendo referencia a un ácido nucleico es un ácido nucleico que se origina a partir de una especie extraña, o si lo hace a partir de la misma especie, está sustancialmente modificado desde su forma nativa en cuanto a la composición y/o el locus genómico por intervención humana deliberada. Por ejemplo, un promotor unido operativamente a un gen estructural heterólogo es de una especie diferente a aquella de la que se deriva el gen estructural, o, si es de la

misma especie, uno o ambos están sustancialmente modificados desde su forma original. Una proteína heterogénea se puede originar a partir de especies extrañas, o, si lo hace desde la misma especie, está sustancialmente modificada desde su forma original por intervención humana deliberada.

- 5 Los sistemas de replicación preferibles incluyen M13, ColEI, SV40, baculovirus, lambda, adenovirus, y similares. Se ha aislado un gran número de regiones reguladoras de iniciación y terminación de la transcripción y se ha demostrado que son eficaces en la transcripción y traducción de proteínas heterólogas de varios huéspedes. En la técnica se conocen ejemplos de dichas regiones, métodos de aislamiento, formas de manipulación, etc. En las condiciones de expresión apropiadas, se pueden utilizar las células huésped como fuente de ADN polimerasas producidas por recombinación o péptidos y polipéptidos derivados.

### Ejemplos

15 Ejemplo 1: Genotecas de quimeras de polimerasa

Se construyeron genotecas de variantes génicas de polimerasa quimérica utilizando la técnica de barajado Step de PCR (Zhao et al., (1998) Nature Biotechnol. 16, 258-261).

- 20 Para una primera genoteca 3T: Se amplificaron genes de polimerasa de tipo silvestre de *Thermus aquaticus* (Taq) y T8 (una variante Taq 11 veces más termoestable seleccionada previamente (Ghadessy et al. Proc Natl Acad Sci U S A. 2001 Abr 10; 98(8):4552-7), *Thermus thermophilus* (Tth) y *Thermus flavus* (Tfl) desde ADN genómico y se clonaron en pASK75 (Skerra 1994) y se sometieron a ensayo para determinar su actividad. Se barajaron estos genes utilizando Step, después se volvieron a clonar en pASK75 y se transformaron en *E. coli* TG1 para dar la genoteca 3T.

- 25 Para una segunda genoteca más diversa 8T, se amplificaron genes Pol I de ADN genómico de *Thermus brockianus*, *Thermus filiformis*, *Thermus scotoductus* y *Thermus oshimai* por PCR y se clonaron en el vector pAsk75 vector.

- 30 Se generó T8 por Step como en el caso anterior incluyendo los genes Pol I de *Thermus thermophilus*, *Thermus aquaticus*, *Thermus flavus*, *Thermus brockianus*, *Thermus filiformis*, *Thermus scotoductus* y *Thermus oshimai* así como *Deinococcus radiodurans* (una bacteria resistente a radiación) que había sido clonada previamente en pAsk75 en el laboratorio de los autores de la invención.

- 35 Se puntuó el tamaño de la genoteca por ensayos de dilución y determinando la proporción de clones que contenían inserto por rastreo de PCR y fue aproximadamente  $1 \times 10^8$  en ambos casos. Un digesto de restricción de diagnóstico de 20 clones produjo 20 patrones de restricción únicos, lo que indica que la genoteca era diversa.

El posterior secuenciado de quimeras seleccionadas demostró una media de 4 a 6 cruzados por gen.

- 40 Ejemplo 2: Producción de ácido húmico

Se rompió una muestra de suelo de turba en pequeños trozos y se añadió agua. A continuación, se calentó la muestra a 50 °C durante 1 hora para ayudar a la solubilización.

- 45 Se centrifugaron las muestras resultantes a 13.000 rpm durante 30 minutos y se recuperó la fase acuosa. A continuación, se redujo el volumen 10 veces utilizando un concentrador.

- 50 Se sometió a ensayo la actividad inhibidora del ácido húmico resultante realizando 30 ciclos de PCR (94 °C 10 min), después 30 ciclos de 94 °C 30 s, 50 °C 30 s, 72 °C 1 min después 65 °C 10 min en presencia de una serie de doble dilución de ácido húmico en 60 % ácido húmico a 0,03 % (12 puntos). Se llevó a cabo la PCR (tampón 1xSuperTaq, 0,2 mM dNTP, 1 µM cebadores (AAA AAT CTA GAT AAC GAG GGC AA y ACC ACC GAA CTG CGG GTG ACG CCA AGC G), 1 µl SuperTaq, 2,5 µl de un crecimiento durante toda la noche de células *E. coli* y 0,01 µl de pAsk75 como matriz (100 µm de reserva), agua y ácido húmico según se necesitó) en presencia de residuo celular de *E. coli* ya que se sabe que ADN y proteína contrarrestan hasta cierto punto el efecto inhibitor que tiene el ácido húmico sobre las polimerasas.

Se observó que el ácido húmico inhibió totalmente la PCR a una concentración de 5 % y superior.

- 60 Ejemplo 3: Selección de clones resistentes a ácido húmico

Se llevó a cabo la emulsificación CSR y la selección en la genoteca StEP Taq, Tth y Tfl descrita esencialmente (Ghadessy et al. 2001), pero con la adición de ácido húmico a la fase acuosa de la emulsión como fuente de presión selectiva. La cantidad máxima de ácido húmico que produce una selección positiva fue 20 %.

- 65 Los cebadores utilizados fueron (5'- GTA AAA CGA CGG CCA GTA CCA CCG AAC TGC GGG TGA CGC CAA GCG-3', and 5'- CAG GAA ACA GCT ATG ACA AAA ATC TAG ATA ACG AGG GCA A-3').

Se extrajo con éter la fase acuosa, se purificó por PCR (Qiagen, Chatsworth, CA) con 35 % GnHCl adicional, se digirió con DpnI para eliminar el ADN de plásmido metilado, se trató con ExoSAP (USB) para eliminar los cebadores residuales, se reamplificó con cebadores fuera del nido (CAG GAA ACA GCT ATG AC y GTA AAA CGA CGG CCA GT), se volvieron a clonar y se transformaron en *E. coli* tal como se ha indicado.

Se rastrearon los clones resultantes y se clasificaron por rangos utilizando un ensayo PCR. Brevemente, se añadieron 2,5 µl de células inducidas a 20 µl de mezcla de PCR (tampón 1xSuperTaq, 0,2 mM dNTP, 1 µM, 0,01 µl de pAsk75 (100 µM de reserva), agua y ácido húmico según fue necesario) con los cebadores correspondientes (AAA AAT CTA GAT AAC GAG GGC AA y ACC ACC GAA CTG CGG GTG ACG CCA AGC G). Se rastrearon 6 placas a varias concentraciones de ácido húmico (10 %, 5 %) y se aisló un total de 14 polimerasas que funcionaban en PCR en condiciones en las que el tipo silvestre no (es decir 5 o 10 % de ácido húmico): P1H2, P2E2, P3D5, P4D10, P4F12, P5E1, P5H2, P6A9, P6A10, P6C10, P6D1, P6F3, P6F4.

Ejemplo 4: Clasificación de clones seleccionados:

Se sembraron en estrías clones de polimerasa: P1H2, P2E2, P3D5, P4D10, P4F8, P4F12, P5E1, P5H2, P6A9, P6A10, P6C10, P6D1, P6F3 sobre placas de agar selectivas y se cultivaron durante toda la noche a 37 °C, se diluyeron 1/100 en 2xTY/ Amp incubado a 37°C hasta una D.O.<sub>595</sub>= 0,5 (aprox. 2 horas). Se añadió tetraciclina anhidra a una concentración final de 0,04 µg/ml y se indujeron los cultivos durante 4 horas a 37°C, agitando. Se centrifugaron las células, se descartó el sobrenadante y se volvió a suspender el sedimento de células en ¼ del volumen de tampón 1x Taq seguido de incubación a 85 °C durante 10 minutos. Se aclaró el lisado por centrifugación. Se normalizaron las polimerasas y se clasificaron por rangos de actividad en PCR esencialmente tal como explica Ghadessy et al 2001 utilizando un programa de PCR (94 °C 1 min, 30 x (94 °C 30 seg, 50 °C 30 seg, 72 °C 1 min), 65°C durante 2 min) utilizando cebadores 1: 5'-ACC ACC GAA CTG CGG GTG ACG CCA AG-3' y 2: 5'-GGG TAC GTG GAG ACC CTC TTC GGC C-3' y 10 ng de vector pASK-Taq como matriz. Se determinó la resistencia a la inhibición de ácido húmico utilizando una dilución en serie de ácido húmico de extracto de turba (HuAc P) (véase lo anterior) y ácido húmico disponible en el mercado (Fluka, código de producto: 53680; Lote: 1102067 34505220) (HuAc F (disuelto en tampón 1xTaq para saturación (es decir, límite de solubilidad) (Tabla 1).

Tabla a: Actividad en ácido húmico desde dos fuentes diferentes

Polimerasa	actividad	f <sup>a</sup>	Actividad en HuAc P	f <sub>HuAcP</sub> <sup>b</sup>	Actividad en HuAc F 1/50	f <sub>HuAcF</sub> <sup>c</sup>	f <sub>HuAcP</sub> / f <sub>HuAcF</sub> <sup>d</sup>
			1/10				
P2E2	1/128	4	1/16	0,25	1/64	0,03125	8
P1H2	1/64	2	1/8	1	1/32	0,25	4
P6A10	1/64	2	1/16	0,5	1/32	0,25	2
P3D5	1/8	0,25	1/32	2	1/128	0,5	4
P4D10	1/128	4	1/2	2	1/16	0,25	8
P4F8	1/32	1	1/32	0,5	1/64	0,25	2
P5H2	1/16	0,5	1/16	2	1/64	0,5	4
P6A9	1/16	0,5	1/16	2	1/64	0,5	4
<b>P4F12 (Hu1)</b>	1/32	1	1/4	<b>4</b>	1/8	<b>2</b>	<b>2</b>
P5E1	1/64	2	1/8	1	1/16	0,5	2
P6C10	1/32	1	1/8	2	1/32	0,5	4
P6D1	1/32	1	1/16	1	1/64	0,25	4
<b>P6F3</b>	1/8	0,25	1/8	<b>8</b>	1/32	<b>2</b>	<b>4</b>
Taqwt	1/32	1	1/16	1	1/16	1	1

<sup>a</sup> f: actividad relativa con respecto a Taqwt

<sup>b</sup> f<sub>HuAcP</sub>: actividad relativa con respecto a Taqwt en ácido húmico extraído de turba (HuAc P)

<sup>c</sup> f<sub>HuAcF</sub>: actividad relativa con respecto a Taqwt en ácido húmico de Fluka (HuAc F)

<sup>d</sup> f<sub>HuAcP</sub> / f<sub>HuAcF</sub>: actividad relativa en HuAc P con respecto a HuAc F

Los clones presentan universalmente resistencias mayores a los efectos de inhibición de HuAc P para lo que han sido seleccionados. P4F12 (Hu1) y P6F3 presentan el nivel de resistencia máximo (f<sub>HuAcP</sub>) reteniendo la actividad a 4- resp. 8-veces la concentración de HuAc P a la que se inhibe completamente Taqwt en PCR. La resistencia a HuAc F es baja o ausente. Solamente P4F12 (Hu1) y P6F3 presentan un aumento de la resistencia (2x) en comparación con wtTaq con ácido húmico disponible en el comercio (HuAc F).

Esto refleja la selección de polimerasas en cuenta a su resistencia a HuAc P y no HuAc F. La actividad relativa en HuAc P con respecto a HuAc F (f<sub>HuAcP</sub> / f<sub>HuAcF</sub>) para Taqwt es 1, mientras que la mayoría de los clones seleccionados presentan una f<sub>HuAcP</sub> / f<sub>HuAcF</sub> 2-8. HuAc P y HuAc F están claramente diferenciados y reflejan la naturaleza heterogénea de las sustancias húmicas. Las selecciones futuras pueden alternar diferentes preparaciones de ácido húmico para asegurar una resistencia general a ácidos húmicos, si bien P4F12 (Hu1) y P6F3 despliegan ya un bajo nivel de resistencia general.

Ejemplo 5: Selección de polimerasa resistente a la inhibición con suelo:

5 Utilizando una selección de CSR normal, tal como describe (Ghadessy et al., 2001), se seleccionaron tres polimerasas que presentan una mayor resistencia a la inhibición del suelo en comparación con polimerasa de *Thermus aquaticus* de tipo silvestre. Se seleccionaron los clones después de dos rondas de CSR.

La muestra de suelo utilizada en los experimentos fue recogida en Cambridge y presentó un pH ligeramente alcalino.

10 Para la primera ronda de CSR se utilizó una parte alícuota de la muestra de suelo para establecer una suspensión espesa de suelo en tampón 1x Supertaq, que se añadió después a la reacción como reactivo inhibidor. En una reacción de control con polimerasa Supertaq, se pudo observar el primer producto a una concentración de 0,3 % de la suspensión espesa de barro. La primera ronda de CSR se llevó a cabo después en presencia de 2,5 % de suspensión espesa de suelo.

15 Para la segunda ronda, se transfirió una parte alícuota de muestra de suelo en tampón 1 x Supertaq y a continuación se incubó esta suspensión espesa de suelo durante 2 horas a 50 °C, seguido de 20 minutos a 90 °C. A continuación, se centrifugó el extracto a 8.000 rpm durante 10 minutos y se mantuvo el sobrenadante como solución inhibidora (-20 °C).

20 A continuación, se determinó la concentración de inhibición utilizando polimerasa Supertaq. La polimerasa comienza a ser inhibida en torno a un 3%, con una inhibición casi completa a un 6 % de concentración. Para la segunda ronda de CSR, la concentración inhibidora se estableció en 5 %.

25 En presencia del inhibidor, los clones resultantes soil3, soil4 y soil5 parecen ser el doble (soil3, soil4), respectivamente o el triple (soil5) de activos que polimerasa de tipo silvestre de *Thermus aquaticus*.

Todas las polimerasas se conservaron en glicerol en nitrógeno líquido.

#### LISTADO DE SECUENCIAS

30 <110> Medical Research Council  
 <120> Polimerasa  
 35 <130> P024088WO  
 <160> 18  
 <170> PatentIn versión 3.4  
 40 <210> 1  
 <211> 2502  
 <212> ADN  
 <213> Artificial  
 45 <220>  
 <223> Polimerasa  
 50 <400> 1

# ES 2 627 274 T3

atggcgatgc ttccctcttt tgagcccaaa ggcoggggcc tcctggtgga oggcccaccac 60  
ctggcctacc gcaccttctt cgcctgaag ggctcaoca cgagccgggg cgaaccggtg 120  
caggcgtct acggtttcgc caagagctc ctcaaggccc tgaaggagga cgggtacaag 180  
gcgctcttcg tggctcttga cgccaaggcc cctccttcc gccacgaggc ctaogaggcc 240  
tacaaggcgg ggagggcccc gacccccgag gacctcccc gccagctcgc cctcatcaag 300  
gagctggtgg acctcctggg gtttaccgc ctogaggtcc aaggctacga ggcggacgac 360  
gtcctcgcca cctggccaa gaaggcggaa aaagaagggt acgaggtgog catcctcacc 420  
gocgaccggg acctctacca gctcgtctcc gacgcgtog cogtcctcca ccccgagggc 480  
cacctcatca cccggagtg gctttgggag aagtacggcc tcaggccgga gcagtgggtg 540  
gacttcogcg cctcgtggg ggaccctcc gacaacctcc ccggggtaaa gggcatoggg 600

# ES 2 627 274 T3

gagaagaccy cctcaagct cctcaaggag tggggaagcc tggaaaatct cctcaagaac 660  
 ctggatcggg taaagccgga aaacgtccgg gagaagatca aggcccacct ggaagacctc 720  
 aggtctctct tggagctctc ccgggtgctt accgacctcc ccctggagggt ggacctcgcc 780  
 caggggcggg agcccgaccg ggaagggtt agggccttcc tggagaggct ggagttcggc 840  
 agcctcctcc atgagttcgg ccttctggaa agccccaagg ccctggagga ggcccctgg 900  
 cccccaccg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960  
 gatcttctgg ccctggcgcg cgcacggggt ggtcgggtcc accgggcccc cgagccttat 1020  
 aaagccctca gggacttgaa ggaggcgcg ggccttctcg ccaaagacct gagcgttctg 1080  
 gcctaaggg aaggccttg cctcccgcc ggcgacgacc ccctgctctt cgcctacctc 1140  
 ctggaccctt ccaacaccac ccccgagggg gtggcccggc gctaccgagg ggagtgagc 1200  
 gaggagggcg gggagcggc cgcctttcc gagaggctct tcgccaacct gtgggggagg 1260  
 cttgaggggg aggagaggct ctttgctt taccgggagg tggataggcc ctttccgct 1320  
 gtctggccc acatggaggc cacaggggtg cgcctggacg tggcctatct cagggccttg 1380  
 tccttgagg tggcogagga gatgcccgc ctcgaggcog aggtcttccg cctggcggc 1440  
 cacccttca acctcaactc ccgggaccag ctggaaggg tcctctttga cgagctaggg 1500  
 cttcccgcca tggcaagac ggagaggacc ggcaagcgt ccaccagcgc cgcctcctg 1560  
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtacog ggagctcacc 1620  
 aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcgc 1680  
 ctccaccccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740  
 aacctccaga acatcccgt ccgcaccccg cttgggcaga ggatccgcog ggccttcac 1800  
 gcogaggagg ggtggtatt ggtggcctg gactatagcc agatagagct cagggtgctg 1860

ES 2 627 274 T3

gcccacctct ccggcgacga gaacctgata cgggtcttcc aggaggggog ggacatccac 1920  
 acggagacog ccagctggat gttoggtgtc cccccggagg cagtggacc cctgatgogc 1980  
 cggggggcca agaoggtgaa cttoggggtc ctctaogga tgtcogcca taggctctcc 2040  
 caggagcttt ccatoccta cgaggaggog gtggccttta tagagogcta cttocaaagc 2100  
 ttocccaagg tgogggcctg gatagaaaag aocctggagg aggggaggaa gcggggctac 2160  
 gtggaaaacc tcttoggag aaggogctac gtgcoogacc tcaacgcoog ggtgaagagc 2220  
 gtcaggagg cgcggagcg catggccttc aacatgcoog tccagggcac cgcgcogac 2280  
 ctcatgaagc togcatggt gaagctcttc cccgcctcc gggagatggg ggoccgcatg 2340  
 ctctccagg tcaogacga gctoctcctg gaggoccccc aagcoggggc ogaggaggtg 2400  
 gggctttgg ccaaggaggc catggagaag gcctatcccc togcogtaoc octggaggtg 2460  
 gaggtgggga toggggagga ctggctttcc gccaaagggtt ag 2502

5 <210> 2  
 <211> 833  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Polimerasa

<400> 2

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

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu  
 20 25 30

ES 2 627 274 T3

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys  
 35 40 45

Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe Val  
 50 55 60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala  
 65 70 75 80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Leu Pro Arg Gln Leu  
 85 90 95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu Glu  
 100 105 110

Val Gln Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys  
 115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp  
 130 135 140

Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu Gly  
 145 150 155 160

His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
 165 170 175

Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp Asn  
 180 185 190

ES 2 627 274 T3

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu  
 195 200 205

Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Val  
 210 215 220

Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp Leu  
 225 230 235 240

Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu Glu  
 245 250 255

Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg Ala  
 260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu  
 275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu  
 290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala  
 305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala  
 325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu  
 340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu

ES 2 627 274 T3

355	360	365
Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser		
370	375	380
Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr		
385	390	395
Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn		
	405	410
Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg		
	420	425
Glu Val Asp Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr		
	435	440
Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val		
	450	455
Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly		
465	470	475
His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe		
	485	490
Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Arg Thr Gly Lys		
	500	505
Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro		
	515	520
		525

ES 2 627 274 T3

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser  
 530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg  
 545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser  
 565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly  
 580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val  
 595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser  
 610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His  
 625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp  
 645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr  
 660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu  
 675 680 685

ES 2 627 274 T3

Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val  
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr  
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala  
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met  
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys  
755 760 765

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val  
770 775 780

His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val  
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val  
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys  
820 825 830

Gly

<210> 3  
<211> 2502  
<212> ADN  
<213> Artificial

5

<220>  
<223> Polimerasa

10

ES 2 627 274 T3

<400> 3

atggcgatgc ttccctctt tgagccaaa ggccgggtcc tcttgggtga cggccaccac 60  
ctggcctacc gcaccttctt cgcctgaag ggctcacca cgagcgggg cgaaccggtg 120  
caggcgtct acggtctgc caagagctc ctcaaggccc tgaaggagga cgggtacaag 180  
gocgtcttgc tggctttga cgccaaggcc cctccttcc gccacgaggc ctacgaggcc 240  
tacaaggcgg ggaggcccc gaccccgag gacttcccc gccagctgc cctcatcaag 300  
gagctggtgg acctcctggg gtttaccgt ctogaggtcc cggctacga ggcggaagac 360  
gttctcgcca ccctggcaa gaaggcgaa aaggaggggt acgaggtgag catcctcacc 420  
gccgacggc acctctacca actcgtctcc gaccgcgtc cgtctctca ccccgaggc 480  
cacctcatca ccccgagtg gctttggag aagtacggc taaggcogga gcagtgggtg 540  
gacttcggc ccctcgtgg ggacctcc gacaacctc cgggggtcaa gggcatoggg 600  
gagaagacc ccctcaagct cctcaaggag tggggaagcc tggaaaacct cctcaagaac 660  
ctggacggg taaagccaga aaacgtcgg gagaagatca aggccacct ggaagacctc 720  
aggctctct tggagctct cgggtgctc accgacctc ccctggaggt ggaacctgcc 780  
cagggcggg agcccgacc ggaaggctt agggccttc tggagaggct tgagtttggc 840  
agcctctcc acgagttcg cttctggaa agccccagg ccctggagga ggccccctgg 900  
ccccgcggg aaggggctt cgtgggctt gtgcttccc gcaaggcgc catgtgggcc 960

ES 2 627 274 T3

gatcttctgg ccoctggccgc cgcaggggt ggtcgggtct accgggcccc cgagccttat 1020

aaagccctca gggacttgaa ggagggcgcg gggcttctcg ocaaagacct gagcgttctg 1080

gccctaaggg aaggccttgg cctcccgccc ggcgacgacc ccattgctcct cgcctacctc 1140

ctggaccctt ccaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

gaggaggcgg gggagcgggc cgccttttc gagaggctct togccaacct gtgggggag 1260

cttgaggggg aggagaggct cctttggctt taccgggagg tggataggcc cctttccgct 1320

gtcctggccc acatggaggc cacaggggtg cgcctggacg tggcctatct cagggccttg 1380

tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttcog cctggccggc 1440

cacccttca aactcaactc cgggaccag ctggaaaggg tcctctttga cgagctaggg 1500

cttcccgcca tggcaagac ggagaagacc ggcaagcgt ccaaccaggc cgcgtcctg 1560

gaggccctcc gcgagggcca cccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620

aagctgaaga gcacctacat tgacccttg ccggaacctc tccaccccag gacgggccc 1680

ctccacacc gcttcaacca gacggccaag gccacgggca ggctaagtag ctccgatccc 1740

aaactcaga acatccccgt ccgcaccccg ctcgggcaga ggatccgcog ggcttcatc 1800

gctgaggagg ggtggctatt ggtggctctg gactatagcc agatagagct cagggtgctg 1860

gccacctct cgggogacga gaacctgat cgggtcttcc aggaggggog ggacatccac 1920

acggaaaacc ccagctggat gttggcgtc ccccgggagg ccgtggacc cctgatgoc 1980

cgggoggcca agaccatcaa cttogggtt ctctaaggca tgtoggcca cgcctctcc 2040

caggagctag ccattcotta cgaggaggcc cgggccttca ttgagccta ctttcagagc 2100

ttccccagg tgccggcctg gattgagaag accctggagg agggcaggag gccggggtag 2160

gtggagacc tcttggccg ccgcogctac gtgccagacc tagaggccc ggtgaagagc 2220

ES 2 627 274 T3

gtgcgggagg cggccgagcg catggccttc aacatgctg tccagggcac cgccgcogac 2280  
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaacggg gcccaggatg 2340  
ctccttcagg tccaagacga gctggctctc gagaccccaa aagagagggc ggaggccgtg 2400  
gcccggtgg ccaaggaggt catggagggg gtgtatccc tggccgtgcc octggaggtg 2460  
gaggtgggga taggggagga ctggctctcc gccaaaggagt ga 2502

5 <210> 4  
<211> 833  
<212> PRT  
<213> Artificial  
10 <220>  
<223> Polimerasa  
<400> 4

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

ES 2 627 274 T3

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu  
 85 90 95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu Glu  
 100 105 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys  
 115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp  
 130 135 140

Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu Gly  
 145 150 155 160

His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
 165 170 175

Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp Asn  
 180 185 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu  
 195 200 205

Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Val  
 210 215 220

Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp Leu  
 225 230 235 240

ES 2 627 274 T3

Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu Glu  
 245 250 255

Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala  
 260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu  
 275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Glu  
 290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Ala Pro Met Trp Ala  
 305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val Tyr Arg Ala  
 325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu  
 340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu  
 355 360 365

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser  
 370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr  
 385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn

ES 2 627 274 T3

	405		410		415
Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg	420		425		430
Glu Val Asp Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr	435		440		445
Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val	450		455		460
Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly	465		470		475
His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe	485		490		495
Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys	500		505		510
Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro	515		520		525
Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser	530		535		540
Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg	545		550		555
Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser	565		570		575

ES 2 627 274 T3

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly  
 580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val  
 595 600 605

Val Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser  
 610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His  
 625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp  
 645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr  
 660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu  
 675 680 685

Glu Ala Arg Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val  
 690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr  
 705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala  
 725 730 735

ES 2 627 274 T3

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met  
 740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys  
 755 760 765

Leu Phe Pro Arg Leu Glu Glu Thr Gly Ala Arg Met Leu Leu Gln Val  
 770 775 780

His Asp Glu Leu Val Leu Glu Thr Pro Lys Glu Arg Ala Glu Ala Val  
 785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val  
 805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys  
 820 825 830

Glu

5 <210> 5  
 <211> 2499  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Polimerasa

<400> 5

atgogtggtatgcttctctcttttgagccc aagggcgcg toctctggt ggaaggccac 60

cacctggcct acgcacctt ocaagccctg aagggcctca ccaccagccg gggggagccg 120

ES 2 627 274 T3

gtgcaggcgg tctacggctt cgccaagagc ctctcaagg toctcaagga ggaoggggac 180

gogtgatog tggctttga cgccaaggcc coctcctec gccacgaggc ctacggggg 240

tacaaggcgg gcogggccc caogcoggag gactttccc ggcaactgc octcatcaag 300

gagctggtgg acctcctggg gctggcgc ctcgaggtcc cgggctacga ggoggaogac 360

gtcctggcca gcctggccaa gaaggoggaa aaggagggt acgaggtcc catcctcacc 420

gcogacaaag accttaccg gctccttcc gacogcatcc acgtcctcca cccogagggg 480

taoctcatca ccccgccctg gctttgggaa aagtacggcc tgaggocga ccagtgggc 540

gactacoggg coctgacogg ggacgagtc gacaaccttc cgggggtcaa gggcatoggg 600

gagaagacgg cgaggaagct tctggaggag tgggggagcc tggaagccct cctcaagaac 660

ctggaccggc tgaagcccgc catcogggag aagatcctgg ccacatgga ogatctgaag 720

ctctcctggg acctggccaa ggtggcacc gacctgccc tggaggtgga cttogccaaa 780

aggcgggagc cogaccgga gagccttagg gcctttctgg agaggcttga gtttggcagc 840

ctctccacg agttoggcct tctgaaaagc cccaaggccc tggaggaggc cccctggccc 900

cogccggaag gggccttctg gggctttgtg cttcccgca aggagccat gtggccgat 960

cttctggctc tggcgcgc cagggggggc cgggtccacc gggcccccga gcctataaa 1020

gcctcaggg acctgaagga ggcgogggg cttctcgca aagacctgag ogttctggcc 1080

ctgaggaag gccttggcct cccgcccgc gacgaccca tgctcctgc ctactcctg 1140

gaccttcca acaccaccoc ogagggggtg gcccgogct acggogggga gtggaogag 1200

gaggcgggg agcggcgc octtccgag aggtctctg ccaacctgtg ggggagcct 1260

gagggggagg agaggctcct ttggcttac cgggaggtgg agaggccct tccgctgtc 1320

ES 2 627 274 T3

ctggcccaca tggaggccac gggggtgccc ctggacgtgg cctatctcag ggccttgccc 1380

ctggaggtgg ccgaggagat cgcocgctc gaggcgagg tcttcgctt ggcggccac 1440

cccttcaacc tcaactcccg ggaccagctg gaaaggtcc tctttgacga gctagggctt 1500

ccggccatcg gcaagacgga gaagaccggc aagcgcctca ccagcggcg cgtcctggag 1560

gccttcgctg aggcaccacc catcgtggag aagatcctgc agtaccggga gctcaccacg 1620

ctgaagagca cctacattga ccccttgccc gacctcatcc acccaggac gggcggctc 1680

cacaccgct tcaaccagac ggccaaggcc acgggcaggc taagtagctc cgtaccacac 1740

ctccagaaca tcccctccg caccocgctt gggcagagga tcccgggc cttcatcgcc 1800

gaggaggggt ggctattggt ggcctggac tatagccaga tagagctcag ggtgctggcc 1860

caoctctccg gcgacgagaa cctgatcccg gtcttcagg agggcgggga catccacaag 1920

gagaccgcca gctggatggt cggcgtccc cgggagggcg tggacccct gatgcggcg 1980

ggggccaaga ccatcaactt cggggtccc taoggcctgt cggccaccg cctctcccag 2040

gagctagcca tcccttaaga ggaggcccag gccttcattg agcgtactt tcagagctc 2100

ccaaggtgc gggcctggat tgagaagacc ctggagagg gcaggaggcg ggggtactg 2160

gagaccctct tcggcggcg ccgctactg ccagacctag aggcgggt gaagagcgtg 2220

cgggagggcg ccgagcgcat ggccttcaac atgcccctc agggcaccg cgcggacctc 2280

atgaagctgg ctatggtgaa gctcttccc aggctggagg aatggggc caggatgctc 2340

cttcaggtcc acgacgact ggtcctcag gcccacaaag agagggcgga ggcctggcc 2400

cggtggcca aggaggtcat ggaggggtg tatcccctg ccgtgccctt ggaggtggag 2460

gtggggatag gggaggactg gcttccgcc aagggttag 2499

<210> 6  
 <211> 832  
 <212> PRT  
 <213> Artificial

ES 2 627 274 T3

<220>  
<223> Polimerasa

<400> 6

5

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

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly  
                  20                    25                    30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala  
                  35                    40                    45

Lys Ser Leu Leu Lys Val Leu Lys Glu Asp Gly Asp Ala Val Ile Val  
50                    55                    60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly  
65                    70                    75                    80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu  
                  85                    90                    95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu  
                  100                    105                    110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys  
                  115                    120                    125

ES 2 627 274 T3

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp  
 130 135 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly  
 145 150 155 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
 165 170 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn  
 180 185 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu  
 195 200 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu  
 210 215 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys  
 225 230 235 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val  
 245 250 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe  
 260 265 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu  
 275 280 285

ES 2 627 274 T3

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly  
 290 295 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp  
 305 310 315 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro  
 325 330 335

Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu  
 340 345 350

Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro  
 355 360 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn  
 370 375 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu  
 385 390 395 400

Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu  
 405 410 415

Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu  
 420 425 430

Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly  
 435 440 445

Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala





ES 2 627 274 T3

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala  
 785 790 795 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro  
 805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Gly  
 820 825 830

5 <210> 7  
 <211> 2499  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Polimerasa

<400> 7

atggcgatgc ttccctctt tgagccaaa gcccggtcc tctggtgga oggcactac 60  
 ctggcctacc gcaccttctt ggcctgaag gccctacca cgagcgggg ogaaccggtg 120  
 caggcgtct acgcttcgc caagagctc ctcaaggccc tgaaggagga oggtacaag 180  
 gccgtcttcg tggctttga gccaaggcc cctccttcc gccacgaggc ctaogaggcc 240  
 tacaaggcgg ggaggcccc gaccccgag gacttcccc gccagctgc cctcatcaag 300  
 gagctggtgg acctcctggg gttaccgc ctgaggtcc cggctaoga ggaggacgac 360  
 gttctgcca cctggccaa gaaggcga aaggagggg acgaggtgog catcctcacc 420  
 gccgacgtg acctctacca actgtctcc gacoggtcg ccgtcctcca cccogagggc 480  
 cacctcaica cccggagtg gctttgggag aagtaaggcc tcaggcogga gcagtgggtg 540  
 gacttccgg cctcgtggg ggacccctcc gacaacctcc ccgggtcaa gggcatcggg 600

ES 2 627 274 T3

gagaagacog ccoctcaagct octcaaggag tggggaagcc tggaaaacct cctcaagaac 660  
 ctggaccggc tgaagccgc catocgggag aagatcctgg occacatgga cgatctgaag 720  
 ctctoctggg acoctggcaa ggtgogcaoc gaocctgccc tggaggtgga ctocgcaaaa 780  
 aggcgggagc cogacogga gaggcttagg gcctttctgg agaggctga gtttggcagc 840  
 ctctccaag agttoggcct tetggaaagc cccaaggccc tggaggaggc ccoctggccc 900  
 cogcoggaag gggccttctg gggctttgtg ctttcccgca aggagcccat gtgggcgat 960  
 cttctggccc tggcgcgcg cagggggggc cgggtocacc gggcccccga gccttcaaaa 1020  
 gcctcaggg acctgaagga ggcgcggggg cttctcgcca aagacctgag cgttctggcc 1080  
 ctgagggag gcoctggcct ccgcccggc gacgaccca tgctoctgc ctacctctg 1140  
 gacccttoca acaccacccc cgagggggtg gcccgogct acggcggga gtggaoggag 1200  
 gaggcgggg agcgggcgc ccttccgag aggtctctcg ccaacctgtg ggggaggtt 1260  
 gagggggagg agaggctcct ttggctttac cgggaggtg agaggccct tccgctgtc 1320  
 ctggcccaca tggaggccac gggggtgoc ctggacgtg octatctag ggcoctgtcc 1380  
 ctggaggtg cogaggagat cgcocgctc gagcogagg tcttcogct gccogccac 1440  
 coctcaacc tcaactccg ggaccagctg gaaaggtcc tcttgacga gctagggctt 1500  
 ccggccatcg gcaagacgga gaagaccgc aagcgtcca ccagcgcgc cgtcctggag 1560  
 gcctcogcg aggcocacc catcgtggag aagatcctgc agtacggga gctaccaag 1620  
 ctgaagagca octacattga cccttgccg gacctcatcc acccaggac gggcgcctc 1680  
 cacaccgct tcaaccagac ggccaogcc acgggcaggc taagttagtc cgtcccaac 1740  
 ctocagaaca tcccgctcg cacocgctt gggcagagga tccgcgggc cttcatogcc 1800

ES 2 627 274 T3

gaggaggggt ggctattggt ggccttggac tatagccaga tagagctcag ggtgctggcc 1860  
cacctctctg gcgagagaa cctgatccgg gtcttccagg aggggcggga catccacacg 1920  
gagaccgcca gctgggtggt cggcgtcccc cgggaggccg tggacccoct gatgcgcccg 1980  
gcggccaaga ccatcaactt cggggtcttc tacggcatgt cggcccaccg cctctcccag 2040  
gagctagcca tcccttaoga ggaggcccag gccttcattg agcgtactt ccagagcttc 2100  
ccaaggtgc gggcctggat tgagaagacc ctggaggagg gcaggaggcg ggggtacgtg 2160  
gagaccctct toggccgccc ccgctacgtg ccagaccatg aggccgggtg gaagagcgtg 2220  
cgggaggcgg cogagcgcac ggccttcaac atgcccgctc agggcaccgc cggccacctc 2280  
atgaagctgg ctatggtgaa gctcttcccc aggctggagg aatgggggc caggatgctc 2340  
cttcaggctc acgacgagct ggtcctogag gccccaaaag agagggggga ggcggtggcc 2400  
cggctggcca aggaggtcct ggagggggtg tatcccctgg ccgtgcccct ggaggtggag 2460  
gtggggatag gggaggactg gctctccgcc aaggagtga 2499

5 <210> 8  
<211> 832  
<212> PRT  
<213> Artificial  
10 <220>  
<223> Polimerasa  
<400> 8

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

Asp Gly His Tyr Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu  
20 25 30

ES 2 627 274 T3

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys  
 35 40 45

Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe Val  
 50 55 60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala  
 65 70 75 80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu  
 85 90 95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu Glu  
 100 105 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys  
 115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp  
 130 135 140

Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu Gly  
 145 150 155 160

His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
 165 170 175

Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp Asn  
 180 185 190

ES 2 627 274 T3

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu  
 195 200 205

Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Leu  
 210 215 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys  
 225 230 235 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val  
 245 250 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe  
 260 265 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu  
 275 280 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly  
 290 295 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp  
 305 310 315 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro  
 325 330 335

Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu  
 340 345 350

ES 2 627 274 T3

Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro  
 355 360 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn  
 370 375 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu  
 385 390 395 400

Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu  
 405 410 415

Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu  
 420 425 430

Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly  
 435 440 445

Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala  
 450 455 460

Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His  
 465 470 475 480

Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp  
 485 490 495

Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg  
 500 505 510

Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile

ES 2 627 274 T3

515		520		525
Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr				
530		535		540
Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu				
545		550		555
His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser				
	565		570	575
Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln				
	580		585	590
Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala				
	595		600	605
Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly				
	610		615	620
Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr				
	625		630	635
Glu Thr Ala Ser Trp Val Phe Gly Val Pro Arg Glu Ala Val Asp Pro				
	645		650	655
Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly				
	660		665	670
Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu				
	675		680	685

ES 2 627 274 T3

Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg  
 690 695 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val  
 705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg  
 725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro  
 740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu  
 755 760 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His  
 770 775 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala  
 785 790 795 800

Arg Leu Ala Lys Glu Val Leu Glu Gly Val Tyr Pro Leu Ala Val Pro  
 805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu  
 820 825 830

- <210> 9
- <211> 2502
- <212> ADN
- <213> Artificial
- <220>
- <223> Polimerasa
- <400> 9

# ES 2 627 274 T3

atggcgatgc ttccctctt tgagcccaaa ggccgggtcc tctggtgga cggccaccac	60
ctggcctacc gcacctctt cgcctgaag ggctcacca cgagccggg cgaaccggtg	120
caggcggctc acggcttcgc caagagctc ctcaaggccc tgaaggagga cgggtacaag	180
gocgtcttcg tggctcttga cggcaaggcc cctccttcc gccacgaggc ctacgaggcc	240
tacaaggcgg ggaggcccc gaccocgag gacttcccc gccagctgc cctcatcaag	300
gagctggtgg acctcctggg gtttaccgc ctcgaggccc aaggetacga ggggacgac	360
gtcctcgcca ccttgccaa gaaggcggaa aaagaagggt acgaggtgcg cgtcctcacc	420
gccgaccggg acctctacca gctcgtctcc gaccgctcg ccgtcctcca ccccgagggc	480
cacctcatca ccccgagtg gctttgggag aagtaccgcc tcaggccgga gcagtgggtg	540
gacttcgctg cctcgtggg ggaccctcc gacaacctcc cgggggtcaa gggcatcggg	600
gagaagaccg cctcaagct cctcaaggag tggggaagcc tggaaaatct cctcaagaac	660
ctggatcggg taaagccgga aaacgtccgg gagaagatca agccccacct ggaagacctc	720
aggctctcct tggagctctc ccgggtcgc accgacctcc cctggagggt ggacctcggc	780
cagggcgggg agcccgaccg ggaaggcctt agggccttcc tggagaggct ggagttcggc	840
agcctcctcc atgagttcgg ccttctggaa agccccaaag cctggagga gggccctggg	900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc	960
gatcttctgg ccttgccgc cgcaggggg ggccgggtcc accgggcccc cgagccttat	1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg	1080

ES 2 627 274 T3

gocctgaggg aaggccttgg cctcccgccc gcogacgacc ccatgctoct cgcctacctc 1140  
 ctggaccott ccaacaccac cccogagggg gtggcccggc gctacggcgg ggagtggacg 1200  
 gaggagcggg gggagcgggc cgcctttcc gagaggctct tcgccaacct gtgggggagg 1260  
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttcogct 1320  
 gccttgccc acatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380  
 tcctggagg tggcogagga gatcggccc ctoagggccg aggtcttcog cctggccggc 1440  
 caccocctca acctcaactc cggggaccag ctggaaggg tcctctttga cgagctaggg 1500  
 cttcccgcca tcggcaagac ggagaagacc ggcaagcct ccaaccagcg cgcogtctcg 1560  
 gaggccctcc gogaggccca ccccatcgtg gagaagatcc tgcagtacog ggagctcacc 1620  
 aagctgaaga gcacctacat tgaccocctg ccggaocctc tcaccccag gaocggccgc 1680  
 ctocacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740  
 aacctcaga acatcccgt ccgcacccc cttgggcaga ggatccggcg ggccttcate 1800  
 gccgagggg ggtggctatt ggtggccctg gactatagcc agatagagct cagggctctg 1860  
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggcg ggacatcac 1920  
 acggagacc ccagctggat gttcggcgtc ccccgggagg ccgtggacc cctgatgcgc 1980  
 cggcgggcca agaccatcaa cttcggggtc ctctacggca tgcggccca cgcctctcc 2040  
 caggagctag ccatoctta cgaggaggcc caggccctca ttgagccta cttccagagc 2100  
 ttocccaagg tcggggcctg gattgagaag aacctggagg agggcaggag ggggggtac 2160  
 gtggagacc toctcggccg ccgcogctac gtgccagacc tagaggccc ggtgaagagc 2220  
 gtgcgggagg cggcogagcg catggccttc aacatgcocg tccagggcac cgcocccgac 2280

ES 2 627 274 T3

ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340  
 ctccctcagg tccacgacga gctggctctc gaggcccaa aagagagggc ggaggcogtg 2400  
 gcccgctgg ccaaggaggt catggagggg gtgtatccc tggcogtgc cctggaggtg 2460  
 gaggtgggga taggggagga ctggctctcc gccaggagt ga 2502

5 <210> 10  
 <211> 833  
 <212> PRT  
 <213> Artificial  
 10 <220>  
 <223> Polimerasa  
 <400> 10

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

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu  
 20 25 30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys  
 35 40 45

Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe Val  
 50 55 60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala  
 65 70 75 80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu  
 85 90 95

ES 2 627 274 T3

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu Glu  
100 105 110

Val Gln Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys  
115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Val Leu Thr Ala Asp Arg Asp  
130 135 140

Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu Gly  
145 150 155 160

His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
165 170 175

Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp Asn  
180 185 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu  
195 200 205

Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Val  
210 215 220

Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp Leu  
225 230 235 240

Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu Glu  
245 250 255

ES 2 627 274 T3

Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg Ala  
 260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu  
 275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu  
 290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala  
 305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala  
 325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu  
 340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu  
 355 360 365

Pro Pro Ala Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser  
 370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr  
 385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn  
 405 410 415

ES 2 627 274 T3

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg  
 420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Ala Leu Ala His Met Glu Ala Thr  
 435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val  
 450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly  
 465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe  
 485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys  
 500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro  
 515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser  
 530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg  
 545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser  
 565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly

ES 2 627 274 T3

	580		585		590
Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val					
	595		600		605
Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser					
	610		615		620
Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His					
	625		630		640
Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp					
	645		650		655
Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr					
	660		665		670
Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu					
	675		680		685
Glu Ala Gln Ala Leu Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val					
	690		695		700
Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr					
	705		710		720
Val Glu Thr Leu Leu Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala					
	725		730		735
Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met					
	740		745		750

ES 2 627 274 T3

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys  
 755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val  
 770 775 780

His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val  
 785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val  
 805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys  
 820 825 830

Glu

5 <210> 11  
 <211> 23  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Cebador oligonucleotídico

<400> 11  
 aaaaatctag ataacgaggg caa 23

15 <210> 12  
 <211> 28  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Cebador oligonucleotídico

<400> 12  
 accaccgaac tgccgggtgac gccaaagcg 28

25 <210> 13  
 <211> 45  
 <212> ADN  
 <213> Artificial

30 <220>  
 <223> Cebador oligonucleotídico

<400> 13  
 gtaaaacgac ggccagtacc accgaactgc ggggtgacgcc aagcg 45

ES 2 627 274 T3

<210> 14  
 <211> 40  
 <212> ADN  
 <213> Artificial  
 5  
 <220>  
 <223> Cebador oligonucleotídico  
 <400> 14  
 10 caggaaacag ctatgacaaa aatctagata acgagggcaa 40  
 <210> 15  
 <211> 17  
 <212> ADN  
 15 <213> Artificial  
 <220>  
 <223> Cebador oligonucleotídico  
 20 <400> 15  
 caggaaacag ctatgac 17  
 <210> 16  
 <211> 17  
 25 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Cebador oligonucleotídico  
 30 <400> 16  
 gtaaacgac ggccagt 17  
 <210> 17  
 <211> 26  
 35 <212> ADN  
 <213> Artificial  
 <220>  
 40 <223> Cebador oligonucleotídico  
 <400> 17  
 accaccgaac tgcgggtgac gccaag 26  
 45 <210> 18  
 <211> 25  
 <212> ADN  
 <213> Artificial  
 50 <220>  
 <223> Cebador oligonucleotídico  
 <400> 18  
 55 gggtagctgg agaccctctt cgcc 25

**REIVINDICACIONES**

- 5 1. Una polimerasa obtenida por ingeniería genética que tiene una secuencia de aminoácidos que comprende las SEQ ID NO 2, 4, 6, 8 o 10.
2. Una molécula de ácido nucleico aislada que codifica una ADN polimerasa obtenida por ingeniería genética de acuerdo con la reivindicación 1.
- 10 3. Una molécula de ácido nucleico aislada de acuerdo con la reivindicación 2, que comprende una secuencia de nucleótidos tal como se expone en cualquiera de las SEQ ID NO 1, 3, 5, 7 o 9.
4. Uso de una polimerasa obtenida por ingeniería genética de acuerdo con la reivindicación 1 para producir productos de extensión de cebador.
- 15 5. Un kit para amplificar ácidos nucleicos que comprende una polimerasa obtenida por ingeniería genética aislada de acuerdo con la reivindicación 1.