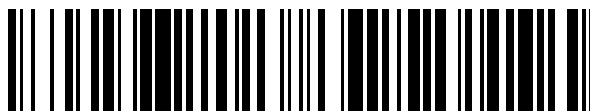


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 578 634**

51 Int. Cl.:

C12N 9/12

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.05.2001** **E 10167769 (8)**

97 Fecha y número de publicación de la concesión europea: **06.04.2016** **EP 2295551**

54 Título: **Transcriptasas inversas termoestables y usos de las mismas**

30 Prioridad:

26.05.2000 US 207196 P

15.03.2001 US 808124

01.05.2001 US 845157

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

28.07.2016

73 Titular/es:

LIFE TECHNOLOGIES CORPORATION (100.0%)

5791 Van Allen Way

Carlsbad, CA 92008, US

72 Inventor/es:

SMITH, MICHAEL D.;
POTTER, ROBERT JASON;
DHARIWAL, GULSHAN;
GERARD, GARY F. y
ROSENTHAL, KIM

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 578 634 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

Transcriptasas inversas termoestables y usos de las mismas

Antecedentes de la invención

Campo de la invención

- 5 La presente invención pertenece a los campos de la biología molecular y de la biología celular. La invención se refiere en general a enzimas transcriptasas inversas y a métodos para la transcripción inversa de moléculas de ácido nucleico, especialmente moléculas de ARN mensajero. Específicamente, la invención se refiere a enzimas transcriptasas inversas que han sido mutadas o modificadas para aumentar la termoestabilidad, reducir la actividad de desoxinucleotidil transferasa terminal, y/o aumentar la fidelidad, y a métodos para producir, amplificar o
- 10 secuenciar moléculas de ácido nucleico (particularmente moléculas de cADN) utilizando estas enzimas o composiciones de transcriptasas inversas. La invención se refiere también a moléculas de ácido nucleico producidas por estos métodos y al uso de tales moléculas de ácido nucleico para producir los polipéptidos deseados. La invención se refiere también a kits o composiciones que comprenden tales enzimas.

Técnica relacionada

15 cADN y genotecas de cADN

- Examinar la estructura y fisiología de un organismo, tejido o célula, es deseable a menudo para determinar su contenido genético. El marco genético de un organismo está codificado en la secuencia de doble hebra de las bases nucleotídicas del ácido desoxirribonucleico (ADN) que está contenido en las células somáticas y germinales del organismo. El contenido genético de un segmento particular de ADN, o gen, se manifiesta típicamente por la
- 20 producción de la proteína que codifica el gen. Para producir una proteína, se produce una copia complementaria de una hebra de la doble hélice de ADN mediante las enzimas ARN polimerasas, dando como resultado una secuencia específica de ácido ribonucleico (ARN). Este tipo particular de ARN, puesto que contiene el mensaje genético del ADN para la producción de una proteína, se llama ARN mensajero (mARN).

- Dentro de una célula, tejido u organismo dado, existen miríadas de especies de mARN, codificando cada una de ellas una proteína separada y específica. Este hecho proporciona una herramienta potente a los investigadores interesados en el estudio de la expresión genética en un tejido o célula. Las moléculas de mARN se pueden aislar y se pueden manipular después por diferentes técnicas de biología molecular, permitiendo de este modo la elucidación de todo el contenido funcional genético de una célula, tejido u organismo.

- Un método común para el estudio de la expresión génica es la producción de clones de ADN complementario (cADN). En esta técnica, las moléculas de mARN de un organismo se aíslan de un extracto de las células o tejidos del organismo. Este aislamiento emplea a menudo matrices de cromatografía sólidas, tales como celulosa o agarosa, con las cuales han sido complejados oligómeros de timidina (T). Puesto que el 3' termini en la mayor parte de las moléculas eucarióticas de mARN contiene una cadena de bases de adenosina (A), y puesto que la base A se empareja con T, las moléculas de mARN se pueden purificar rápidamente de otras moléculas y sustancias en el
- 30 extracto de tejido o células. A partir de estas moléculas de mARN purificadas, se pueden hacer copias de cADN utilizando la enzima transcriptasa inversa (RT), lo que da como resultado la producción de moléculas de cADN de una sola hebra. Esta reacción se denomina típicamente reacción de la primera hebra. Los cADNs de una sola hebra se pueden convertir entonces en una copia de ADN completo de doble hebra (es decir, un cADN de doble hebra) del mARN original (y por lo tanto de la secuencia original del ADN de doble hebra, que codifica este mARN, contenido en el genoma del organismo) por la acción de una ADN-polimerasa. Los cADNs de doble hebra específicos de la proteína se pueden insertar entonces en un plásmido o vector viral, que se introduce después en una célula hospedante, bacteriana, de levadura, animal o vegetal. Las células hospedantes se hacen crecer entonces en medio de cultivo, dando como resultado una población de células hospedantes que contienen (o en muchos casos, que expresan) el gen de interés.

- Este procedimiento completo, desde el aislamiento del mARN de un organismo o tejido fuente hasta la inserción del cADN en un plásmido o vector para el crecimiento de poblaciones celulares que contienen el gen aislado, se denomina "clonación de cADN." El conjunto de los cADNs preparados a partir de una fuente dada de mARNs se denomina una "genoteca de cADN." Los clones de cADN en una genoteca de cADN corresponden a los genes transcritos en el tejido fuente. El análisis de una genoteca de cADN puede aportar mucha información sobre el
- 45 modelo de expresión génica en el organismo o tejido del cual se derivó la genoteca.

Enzimas transcriptasas inversas retrovirales

- Se han estudiado a fondo tres formas prototípicas de transcriptasa inversa retroviral. La transcriptasa inversa del virus de la leucemia murina de Moloney (M-MLV) contiene una única subunidad de 78 kDa con actividad de ADN-polimerasa dependiente de ARN y actividad de RNasa H (ribonucleasa H). Esta enzima ha sido clonada y expresada
- 55 en una forma totalmente activa en *E. coli* (revisado en Prasad, V.R., Reverse Transcriptase, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, p. 135 (1993)). La transcriptasa inversa del virus de la

inmunodeficiencia humana (VIH) es un heterodímero de subunidades p66 y p51 en las que la subunidad más pequeña se deriva de la más grande por escisión proteolítica. La subunidad p66 tiene ambos, un dominio de ADN-polimerasa dependiente de ARN y un dominio de RNasa H, mientras que la subunidad p51 tiene solamente un dominio de ADN-polimerasa. La transcriptasa inversa de p66/p51 del VIH activa ha sido clonada y expresada satisfactoriamente en una serie de hospedantes de expresión, incluyendo *E. coli* (revisado en Le Grice, S.F.J., Reverse Transcriptase, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, p. 163 (1993)). Dentro del heterodímero p66/p51 de VIH, la subunidad 51-kD es catalíticamente inactiva, y la subunidad 66-kD tiene actividad tanto de ADN-polimerasa como de RNasa H (Le Grice, S.F.J., et al., EMBO Journal 10:3905 (1991); Hostomsky, Z., et al., J. Virol. 66:3179 (1992)). La transcriptasa inversa del virus de la leucosis-sarcoma aviar (ASLV), que incluye pero no se limita a la transcriptasa inversa del virus de sarcoma de Rous (RSV), transcriptasa inversa del virus de la mieloblastosis aviar (AMV), transcriptasa inversa del virus de la eritroblastosis aviar (AEV) y virus auxiliar MCAV, transcriptasa inversa del virus de la mielocitomatosis aviar MC29 y virus auxiliar MCAV, transcriptasa inversa del virus de la reticuloendoteliosis aviar (REV-T) y virus auxiliar REV-A, transcriptasa inversa del virus del sarcoma aviar UR2 y virus auxiliar UR2AV, transcriptasa inversa del virus del sarcoma aviar Y73 y virus auxiliar YAV, transcriptasa inversa del virus asociado a Rous (RAV), y transcriptasa inversa del virus asociado a la mieloblastosis (MAV), es también un heterodímero de dos subunidades, α (aproximadamente 62 kDa) y β (aproximadamente 94 kDa), en el cual α se deriva de β por escisión proteolítica (revisado en Prasad, V.R., Reverse Transcriptase, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1993), p. 135). La transcriptasa inversa de ASLV puede existir en dos formas estructurales adicionales catalíticamente activas, $\beta\beta$ y $\alpha\beta$ (Hizi, A. and Joklik, W.K., J. Biol. Chem. 252: 2281 (1977)). El análisis de sedimentación da a entender que $\alpha\beta$ y $\beta\beta$ son dímeros y que la forma α existe en equilibrio entre las formas monoméricas y diméricas (Grandgenett, D.P., et al., Proc. Nat. Acad. Sci. USA 70:230 (1973); Hizi, A. and Joklik, W.K., J. Biol. Chem. 252:2281 (1977); y Soltis, D.A. and Skalka, A.M., Proc. Nat. Acad. Sci. USA 85:3372 (1988)). Las transcriptasas inversas de ASLV $\alpha\beta$ y $\beta\beta$ son los únicos ejemplos conocidos de transcriptasa inversa retroviral que incluye tres actividades diferentes en el mismo complejo proteínico: actividades de ADN-polimerasa, RNasa H, y ADN-endonucleasa (integrasa) (revisado en Skalka, A.M., Reverse Transcriptase, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1993), p. 193). La forma α carece del dominio y actividad de integrasa.

Diversas formas de las subunidades individuales de la transcriptasa inversa de ASLV han sido clonadas y expresadas. Estas incluyen un polipéptido precursor de 98 kDa que normalmente es procesado proteolíticamente a β y un polipéptido de 4 kDa separado del extremo carboxi β (Alexander, F., et al., J. Virol. 61:534 (1987) y Anderson, D. et al., Focus 17:53 (1995)), y la subunidad β madura (Weis, J.H. and Salstrom, J.S., Patente de Estados Unidos N° 4.663.290 (1987); y Soltis, D.A. and Skalka, A.M., Proc. Nat. Acad. Sci. USA 85:3372 (1988)). (Véase también Werner S. and Wohrl B.M., Eur. J. Biochem. 267:4740-4744 (2000); Werner S. and Wohrl B.M., J. Virol. 74:3245-3252 (2000); Werner S. and Wohrl B.M., J. Biol. Chem. 274:26329-26336 (1999).) La transcriptasa inversa $\alpha\beta$ heterodimérica de RSV ha sido purificada también de células de *E. coli* que expresan un gen RSV P clonado (Chernov, A.P., et al., Biomed. Sci. 2:49 (1991)).

Eficiencia de la transcripción inversa

Como se ha indicado antes, la conversión de mRNA en cADN por transcripción inversa mediada por transcriptasa inversa, es una etapa esencial en el estudio de las proteínas expresadas de genes clonados. Sin embargo, el uso de transcriptasa inversa sin modificar para catalizar la transcripción inversa es ineficaz por una serie de razones. En primer lugar, la transcriptasa inversa algunas veces degrada un molde de ARN antes de que se inicie o se complete la reacción de primera hebra, principalmente debido a la actividad intrínseca de RNasa H presente en la transcriptasa inversa. En adición, el cebado defectuoso de la molécula del molde de mRNA puede llevar a la introducción de errores en la primera hebra de cADN mientras que la estructura secundaria de la propia molécula de mRNA puede hacer a algunos mARNs refractarios a la síntesis de la primera hebra.

La eliminación de la actividad de RNasa H de la transcriptasa inversa puede eliminar el primer problema y mejorar la eficiencia de la transcripción inversa (Gerard, G.F., et al., FOCUS 11(4):60 (1989); Gerard, G.F., et al., FOCUS 14(3):91 (1992)). Sin embargo, dichas transcriptasas inversas (formas "RNasa H") no resuelven los problemas adicionales de cebado defectuoso y estructura secundaria de mRNA.

Otro factor que influye en la eficiencia de la transcripción inversa es la capacidad del ARN para formar estructuras secundarias. Dichas estructuras secundarias se pueden formar, por ejemplo, cuando regiones de las moléculas de ARN tienen suficiente complementariedad para hibridarse y formar ARN de doble hebra. Generalmente, la formación de estructuras secundarias de ARN se puede reducir aumentando la temperatura de las soluciones que contienen las moléculas de ARN. Por lo tanto, en muchos casos, es deseable la transcripción inversa del ARN a temperaturas por encima de 37 °C. Sin embargo, las transcriptasas inversas conocidas en la técnica, generalmente pierden actividad cuando se incuban a temperaturas muy superiores a 37 °C (p. ej., 50 °C).

Los métodos para generar e identificar enzimas que tienen la fidelidad alterada son conocidos en la técnica (véase p. ej. Publicación PCT No. WO 98/23733 o Cadwell R. et al., PCR Methods and Applications 2:28-33). Por ejemplo, Kaushik N. et al., (Biochemistry 9:2617-2627) han identificado un residuo catalíticamente esencial de transcriptasa inversa de MLV por mutagénesis y Halvas E.K. et al. (J. Virol. 74:312-319 (2000)) han desarrollado un ensayo para

identificar los determinantes estructurales en la transcriptasa inversa de MLV importantes para la fidelidad de la enzima.

Han sido descritas ciertas formas RNasa H de la transcriptasa inversa de M-MLV que presentan un aumento o mejora de la termoestabilidad en comparación con la enzima natural (véase la publicación PCT No. WO 98/47912; Gerard G.F. et al., Molecular Biotechnology, 8:61-77 (1997) o el documento JP 2000/139457) y los métodos de producción de dichos mutantes de transcriptasa inversa que carecen de actividad RNasa H están descritos p. ej. en la patente de Estados Unidos N° 5.405.776.

Sumario de la invención

La presente invención proporciona enzimas transcriptasas inversas, composiciones que comprenden dichas enzimas y métodos útiles para resolver las limitaciones de la transcripción inversa expuestas antes. En general, la invención proporciona composiciones para uso en la transcripción inversa de una molécula de ácido nucleico que comprende uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, quince, etc.) polipéptidos que tienen la actividad de transcriptasa inversa de la invención. Dichas composiciones pueden comprender además uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, etc.) nucleótidos, un tampón adecuado, y/o uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, quince, etc.) ADN-polimerasas. Las composiciones de la invención pueden comprender también uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, quince, etc.) cebadores oligonucleótidos.

La transcriptasa inversa de la invención es una transcriptasa inversa de M-MLV en donde la histidina de la posición natural 204 está reemplazada con arginina para aumentar o mejorar la termoestabilidad. En realizaciones específicas, las transcriptasas inversas descritas pueden ser de cadena única (subunidad única) o de multi-cadenas (multi-subunidades) y pueden tener actividad de RNasa H reducida o sustancialmente reducida. Las enzimas descritas se seleccionan del grupo que consiste en RNasa H-transcriptasa inversa del virus de la leucemia murina de Moloney (M-MLV), RNasa H-transcriptasa inversa del virus de sarcoma de Rous (RSV), RNasa H-transcriptasa inversa del virus de mieloblastosis aviar (AMV), RNasa H-transcriptasa inversa del virus asociado a Rous (RAV), RNasa H-transcriptasa inversa del virus asociado a mieloblastosis (MAV) u otras RNasa H-transcriptasas inversas de ASLV y RNasa H-transcriptasa inversa del virus de la inmunodeficiencia humana (VIH) y mutantes de las mismas. En las composiciones preferidas, las transcriptasas inversas se presentan en concentraciones de trabajo.

En ciertos aspectos, las transcriptasas inversas descritas han sido modificadas o mutadas para aumentar o mejorar la termoestabilidad. Los ejemplos de dichas transcriptasas inversas incluyen enzimas que tienen una o más modificaciones o mutaciones en posiciones que corresponden a aminoácidos seleccionados del grupo que consiste en:

- (a) leucina 52 de la transcriptasa inversa de M-MLV;
- (b) tirosina 64 de la transcriptasa inversa de M-MLV;
- (c) lisina 152 de la transcriptasa inversa de M-MLV;
- (d) histidina 204 de la transcriptasa inversa de M-MLV;
- (e) metionina 289 de la transcriptasa inversa de M-MLV; y
- (f) treonina 306 de la transcriptasa inversa de M-MLV.

En realizaciones específicas, la descripción se dirige a las transcriptasas inversas de M-MLV en donde la leucina 52 está reemplazada con prolina, la tirosina 64 está reemplazada con arginina, la lisina 152 está reemplazada con metionina, la histidina 204 está reemplazada con arginina, la metionina 289 está reemplazada con leucina y/o la treonina 306 está reemplazada con lisina o con arginina. Se describen además transcriptasas inversas, aparte de la transcriptasa inversa de M-MLV, que contienen alteraciones que corresponden a las indicadas anteriormente.

En aspectos adicionales, la invención incluye también transcriptasas inversas termoestables que retienen al menos aproximadamente 50 %, al menos aproximadamente 60 %, al menos aproximadamente 70 %, al menos aproximadamente 85 %, al menos aproximadamente 95 %, al menos aproximadamente 97 %, al menos aproximadamente 98 %, al menos aproximadamente 99 %, o al menos aproximadamente 100 % de actividad de transcriptasa inversa después de calentamiento a 50 °C durante 5 minutos.

Como se ha indicado antes, las enzimas descritas incluyen transcriptasas inversas que presentan actividad de transcriptasa inversa o después de la formación de multímeros (*p. ej.*, dímeros) o como moléculas de proteínas individuales (es decir, en forma monomérica). Los ejemplos de transcriptasas inversas que presentan actividad de transcriptasa inversa después de la formación de multímeros incluyen las transcriptasas inversas de AMV, RSV y VIH. Un ejemplo de una transcriptasa inversa que presenta actividad de transcriptasa inversa como proteínas individuales, separadas (es decir, en forma monomérica) es la transcriptasa inversa de M-MLV.

Las transcriptasas inversas multiméricas pueden formar homo-multímeros o hetero-multímeros. En otras palabras, las subunidades del complejo proteínico multimérico pueden ser idénticas o diferentes. Un ejemplo de una

transcriptasa inversa hetero-dimérica es la transcriptasa inversa de AMV, que está compuesta de dos subunidades que difieren en la secuencia primaria de aminoácidos. Más específicamente, como ya se ha expuesto, la transcriptasa inversa de AMV puede estar compuesta de dos subunidades en donde una de estas subunidades se genera por un proceso proteolítico de la otra. Por lo tanto, la transcriptasa inversa de AMV dimérica puede estar compuesta de subunidades de diferente tamaño que comparten regiones de identidad de la secuencia de aminoácidos.

La presente descripción se refiere en particular a transcriptasas inversas mutantes o modificadas en donde se han realizado uno o más (p. ej., uno, dos, tres, cuatro, cinco, diez, doce, quince, veinte, etc.) cambios de aminoácidos que hacen a la enzima más termoestable en la síntesis de ácido nucleico, en comparación con las transcriptasas inversas no mutadas o no modificadas. Los sitios para mutación o modificación para producir las enzimas transcriptasas inversas termoeestables de la presente invención y/o transcriptasas inversas que presentan otras características (p. ej., aumento de la fidelidad, reducción de la actividad TdT, etc.) se listan para algunas transcriptasas inversas en la Tabla 1. Las modificaciones descritas en la Tabla 1 producen preferiblemente transcriptasas inversas termoeestables. Sitios similares o equivalentes o sitios correspondientes en otras transcriptasas inversas pueden ser mutados o modificados para producir transcriptasas inversas termoeestables adicionales, así como transcriptasas inversas que presentan otras características (p. ej., aumento de la fidelidad, reducción de la actividad TdT, etc.).

Tabla 1

Transcriptasa inversa (RT)	Aminoácidos
M-MLV	L52, Y64, L135, H143, K152, Q165, G181, H204, I218, N249, M289, T306, A517, D524, T544, V546, W548, E562, H577, D583, L604, S606, G608, F625, L626, H629, H631, H638, G641
AMV	V2, L4, W12, P14, H16, T17, W20, I21, Q23, W24, L26, P27, G29, V32, Q36, L42, Q43, L44, G45, H46, I47, P49, S50, L51, S52, C53, W54, F59, I61, A64, S65, G66, S67, Y68, L70, L71, A76, A79, P83, A86, V87, Q88, Q89, G90, A91, W101, P102, L108, Q120, S131, V132, N133, N134, Q135, P137, A138, Q142, Q148, T151, Y180, M181, S190, H191, G193, A196, I201, S202, P214, V217, Q218, P221, G222, Q224, L226, G227, Y228, G231, T233, Y234, A236, P237, G239, L240, P244, I246, T248, W250, Q252, G257, Q260, W261, P264, L266, G267, L272, Y277, Q279, L280, G282, S283, P285, N286, A288, N292, L293, M297, I302, V303, L305, S306, T308, L311, L320, I332, G333, V334, G336, Q337, G338, P345, W348, L349, F350, S351, P354, A357, F358, A360, W361, L362, V364, L365, T366, T370, A374, V377, G381, C392, P400, G402, L405, G412, I414, F423, I425, A426, P428, L433, H440, P441, V443, G444, P445, A451, S453, S454, T455, H456, G458, V459, V460, W462, W468, I470, I473, A474, L476, G477, A478, S479, V480, Q481, Q482, L483, A491, W495, P496, T497, T498, P499, T500, A507, F508, M512, L513, G520, V521, P522, S523, T524, A525, A527, F528, L534, S535, Q536, S538, V543, S548, H549, S550, V552, P553, F556, T557, N560, A562
RSV	V2, L4, W12, P14, H16, T17, W20, I21, Q23, W24, L26, P27, G29, V32, Q36, L42, Q43, L44, G45, H46, I47, P49, S50, L51, S52, C53, W54, F59, I61, A64, S65, G66, S67, Y68, L70, L71, A76, A79, P83, A86, V87, Q88, Q89, G90, A91, W101, P102, L108, Q120, S131, V132, N133, N134, Q135, P137, A138, Q142, Q148, T151, Y180, M181, S190, H191, G193, A196, I201, S202, P214, V217, Q218, P221, G222, Q224, L226, G227, Y228, G231, T233, Y234, A236, P237, G239, L240, P244, I246, T248, W250, Q252, G257, Q260, W261, P264, L266, G267, L272, Y277, Q279, L280, G282, S283, P285, N286, A288, N292, L293, M297, I302, V303, L305, S306, T308, L311, L320, I332, G333, V334, G336, Q337, G338, P345, W348, L349, F350, S351, P354, A357, F358, A360, W361, L362, V364, L365, T366, T370, A374, V377, G381, C392, P400, G402, L405, G412, I414, F423, I425, A426, P428, L433, H440, P441, V443, G444, P445, A451, S453, S454, T455, H456, G458, V459, V460, W462, W468, I470, I473, A474, L476, G477, A478, S479, V480, Q481, Q482, L483, A491, W495, P496, T497, T498, P499, T500, A507, F508, M512, L513, G520, V521, P522, S523, T524, A525, A527, F528, L534, S535, Q536, S538, V543, S548, H549, S550, V552, P553, F556, T557, N560, A562
VIH	I1, P3, L11, P13, G14, M15, Q22, W23, L25, T26, T38, G44, I46, S47, G50, P51, N53, P54, Y55, F60, I62, S67, T68, W70, L73, V89, Q90L91, G92, I93, S104, V110, G111, S133, I134, N135, N136, P139, G140, I141, Q144, N146, Q150" Y182, M183, I194, G195, Q106, T199, Q206, L209, P216, Q221, P224, P225, L227, M229, G230, Y231, H234, Q241, P242, V244, L245, S250, T252, N254, Q257, G261, N264, W265, Q268, P271, G272, Q277, C279, L281, L282, G284, T285, A287, L288, T289, V291, P293, L294, T295, L300, A303, I308, L309, P312, H314, Y317, L324, I328, Q329, G332, Q333, G334, Y341, P344, F345, Y353, M356, G358, A359, H360, T361, Q372, T376, V380, Q392, W405,

Transcriptasa inversa (RT)	Aminoácidos
	Q406, A407, F415, V416, N417, T418, P419, P420, L424, W425, P432, V434, G435, A436, A444, A445, N446, T449, L451, N459, G461, Q463, V465, V466, P467, L468, T469, N470, T471, T472, N473, Q474, Y482, Q486, S488, G489, L490, Q499, Y500, G503, I504, S512, S514, L516, N518, Q519, Q523, I525, W534, P536, A537, H538, G540, I541, G542, Q546, L550, S552, A553, V554, I555

Los expertos en la técnica apreciarán que un aislado de virus diferente puede codificar una enzima transcriptasa inversa que tiene un aminoácido diferente en las posiciones identificadas antes. Dichos aislados pueden ser modificados para producir las transcriptasas inversas termoestables de la presente invención.

- 5 Las transcriptasas inversas termoestables de la invención pueden tener también una o más propiedades: (a) actividad de RNasa H reducida o sustancialmente reducida, (b) actividad de desoxinucleotidil-transferasa terminal reducida o sustancialmente reducida, y/o (c) aumento de la fidelidad.

Las enzimas de la invención que tienen actividad de desoxinucleotidil-transferasa terminal reducida o sustancialmente reducida, pueden tener una o más modificaciones o mutaciones en las posiciones que corresponden a los aminoácidos seleccionados del grupo que consiste en:

(a) tirosina 133 de la transcriptasa inversa de M-MLV;

(b) treonina 197 de la transcriptasa inversa de M-MLV; y

(c) fenilalanina 309 de la transcriptasa inversa de M-MLV.

15 En realizaciones específicas, la descripción se dirige a las transcriptasas inversas de M-MLV en donde la tirosina 133 está reemplazada con alanina, la treonina 197 está reemplazada con ácido glutámico, y/o la fenilalanina 309 está reemplazada con asparagina. Se describen además transcriptasas inversas, aparte de la transcriptasa inversa de M-MLV, que contienen alteraciones que corresponden a las indicadas anteriormente.

20 En otras realizaciones determinadas, las transcriptasas inversas pueden no incluir combinaciones específicas de mutaciones, tales como transcriptasas inversas que combinan una mutación de la histidina 204 a arginina (H204R) y una mutación de la treonina 306 a lisina (T306K); transcriptasas inversas que combinan H204R y T306K con una mutación de la fenilalanina 309 a asparagina (F309N); y transcriptasas inversas que combinan H204R, T306K y F309N con una mutación de la valina 223 a histidina (V223H).

Adicionalmente, las enzimas que han presentado un aumento de la fidelidad pueden tener una o más modificaciones o mutaciones en posiciones que corresponden a los aminoácidos seleccionados del grupo que consiste en:

25 (a) tirosina 64 de la transcriptasa inversa de M-MLV;

(b) arginina 116 de la transcriptasa inversa de M-MLV;

(c) glutamina 190 de la transcriptasa inversa de M-MLV; y

(d) valina 223 de la transcriptasa inversa de M-MLV.

30 En realizaciones específicas, las transcriptasas inversas pueden no incluir transcriptasas inversas de M-MLV, transcriptasas inversas de HTV, transcriptasas inversas de AMV, y/o transcriptasas inversas de RSV. Así, por ejemplo, en ciertas realizaciones, la descripción se dirige a transcriptasas inversas con aumento de la termoestabilidad que no son una transcriptasa inversa de VIH. En otras realizaciones, la descripción se dirige a transcriptasas inversas con aumento de la termoestabilidad que no son una transcriptasa inversa de M-MLV. En aún otras realizaciones, la descripción se dirige a transcriptasas inversas con aumento de la termoestabilidad que no son una transcriptasa inversa de AMV. En otras realizaciones más, la descripción se dirige a transcriptasas inversas con aumento de la termoestabilidad que no son una transcriptasa inversa de RSV.

40 La presente invención se dirige también a moléculas de ácido nucleico (*p. ej.*, vectores) que contienen un gen o ácido nucleico que codifica las transcriptasas inversas mutantes o modificadas de la presente invención y a las células hospedantes que contienen dicho ADN u otras moléculas de ácido nucleico. Se puede utilizar cualquier número de hospedantes para expresar el gen o la molécula de ácido nucleico de interés, incluyendo células procariotas y eucariotas. En realizaciones específicas, se utilizan las células procariotas para expresar las transcriptasas inversas de la invención. Un ejemplo de un hospedante procariota apropiado para uso con la presente invención es *Escherichia coli*. Los ejemplos de hospedantes eucariotas apropiados para uso con la presente invención incluyen células fúngicas (*p. ej.*, células de *Saccharomyces cerevisiae*, células de *Pichia pastoris*, etc.),

células de plantas, y células de animales (*p. ej.*, células de *Drosophila melanogaster*, células de *Spodoptera frugiperda* Sf9 y Sf21, células de *Trichoplusia* High-Five, células de *C. elegans*, células de *Xenopus laevis*, células CHO, células COS, células VERO, células BHK, etc.).

La invención se refiere también a un método para producir las transcriptasas inversas de la invención, comprendiendo dicho método:

(a) cultivar una célula hospedante que comprende un gen u otra molécula de ácido nucleico que codifica una transcriptasa inversa de la invención (preferiblemente dicho gen u otra molécula de ácido nucleico de la transcriptasa inversa, está contenido en un vector dentro de la célula hospedante);

(b) expresar el gen o molécula de ácido nucleico; y

(c) aislar dicha transcriptasa inversa de la célula hospedante.

La invención se dirige también a métodos para preparar una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, etc.) moléculas de ácido nucleico, que comprenden mezclar uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, etc.) moldes de ácido nucleico (preferiblemente uno o más moldes de ARN y lo más preferiblemente uno o más moldes de ARN mensajero) con una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, quince, etc.) transcriptasas inversas de la invención e incubar la mezcla en condiciones suficientes para preparar una primera molécula o moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes de ácido nucleico. En algunas realizaciones, la mezcla se incuba a una temperatura elevada. En realizaciones específicas, la temperatura elevada puede ser de aproximadamente 40 °C o superior, de aproximadamente 45 °C o superior, de aproximadamente 50 °C o superior, de aproximadamente 51 °C o superior, de aproximadamente 52 °C o superior, de aproximadamente 53 °C o superior, de aproximadamente 54 °C o superior, de aproximadamente 55 °C o superior, de aproximadamente 56 °C o superior, de aproximadamente 57 °C o superior, de aproximadamente 58 °C o superior, de aproximadamente 59 °C o superior, de aproximadamente 60 °C o superior, de aproximadamente 61 °C o superior, de aproximadamente 62 °C o superior, de aproximadamente 63 °C o superior, de aproximadamente 64 °C o superior, de aproximadamente 65 °C o superior, de aproximadamente 66 °C o superior, de aproximadamente 67 °C o superior, de aproximadamente 68 °C o superior, de aproximadamente 69 °C o superior, de aproximadamente 70 °C o superior, de aproximadamente 71 °C o superior, de aproximadamente 72 °C o superior, de aproximadamente 73 °C o superior, de aproximadamente 74 °C o superior, de aproximadamente 75 °C o superior, de aproximadamente 76 °C o superior, de aproximadamente 77 °C o superior, o de aproximadamente 78 °C o superior; o a un intervalo de temperatura de aproximadamente 37 °C a aproximadamente 75 °C, de aproximadamente 40 °C a aproximadamente 75 °C, de aproximadamente 45 °C a aproximadamente 75 °C, de aproximadamente 50 °C a aproximadamente 75 °C, de aproximadamente 51 °C a aproximadamente 75 °C, de aproximadamente 52 °C a aproximadamente 75 °C, de aproximadamente 53 °C a aproximadamente 75 °C, de aproximadamente 54 °C a aproximadamente 75 °C, de aproximadamente 55 °C a aproximadamente 75 °C. En otras realizaciones, la temperatura elevada está dentro del intervalo de aproximadamente 50 °C a aproximadamente 70 °C, de aproximadamente 51 °C a aproximadamente 70 °C, de aproximadamente 52 °C a aproximadamente 70 °C, de aproximadamente 53 °C a aproximadamente 70 °C, de aproximadamente 54 °C a aproximadamente 70 °C, de aproximadamente 55 °C a aproximadamente 70 °C, de aproximadamente 56 °C a aproximadamente 70 °C, de aproximadamente 57 °C a aproximadamente 70 °C, de aproximadamente 58 °C a aproximadamente 70 °C, de aproximadamente 59 °C a aproximadamente 70 °C, de aproximadamente 60 °C a aproximadamente 70 °C, de aproximadamente 61 °C a aproximadamente 70 °C, de aproximadamente 62 °C a aproximadamente 70 °C, de aproximadamente 63 °C a aproximadamente 70 °C, de aproximadamente 64 °C a aproximadamente 70 °C, de aproximadamente 65 °C a aproximadamente 70 °C, de aproximadamente 66 °C a aproximadamente 70 °C, de aproximadamente 67 °C a aproximadamente 70 °C, de aproximadamente 68 °C a aproximadamente 70 °C, de aproximadamente 69 °C a aproximadamente 70 °C, de aproximadamente 70 °C a aproximadamente 70 °C, de aproximadamente 71 °C a aproximadamente 70 °C, de aproximadamente 72 °C a aproximadamente 70 °C, de aproximadamente 73 °C a aproximadamente 70 °C, de aproximadamente 74 °C a aproximadamente 70 °C, de aproximadamente 75 °C a aproximadamente 70 °C. En otras realizaciones, la temperatura elevada puede estar dentro del intervalo de aproximadamente 45 °C a aproximadamente 60 °C, de aproximadamente 46 °C a aproximadamente 60 °C, de aproximadamente 47 °C a aproximadamente 60 °C, de aproximadamente 48 °C a aproximadamente 60 °C, de aproximadamente 49 °C a aproximadamente 60 °C, de aproximadamente 50 °C a aproximadamente 60 °C, de aproximadamente 51 °C a aproximadamente 60 °C, de aproximadamente 52 °C a aproximadamente 60 °C, de aproximadamente 53 °C a aproximadamente 60 °C, o de aproximadamente 54 °C a aproximadamente 60 °C. En realizaciones específicas adicionales, la primera molécula de ácido nucleico es cADN de una única hebra.

Los moldes de ácido nucleico adecuados para la transcripción inversa según este aspecto de la invención incluyen cualquier molécula de ácido nucleico o población de moléculas de ácido nucleico (preferiblemente ARN y lo más preferiblemente mARN), particularmente las derivadas de una célula o tejido. En un aspecto específico, se utiliza una población de moléculas de mARN (un número de diferentes moléculas de mARN, típicamente obtenidas de un tipo particular de célula o tejido) para preparar una genoteca de cADN. Los ejemplos de fuentes celulares de moldes de ácido nucleico incluyen células bacterianas, células fúngicas, células de plantas y células de animales.

La invención se refiere también a métodos para preparar una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, etc.) moléculas de ácido nucleico de doble hebra. Tales métodos comprenden (a) mezclar uno o más moldes de ácido nucleico (preferiblemente ARN o mARN, y más preferiblemente una población de moldes de mARN) con una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, quince, etc.) transcriptasas inversas de la invención; (b) incubar la mezcla en condiciones suficientes para preparar una primera molécula o moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes; y (c) incubar la primera molécula o moléculas de ácido nucleico en condiciones suficientes para preparar una segunda molécula o moléculas de ácido nucleico complementarias de todo o parte de la primera molécula o moléculas de ácido nucleico, formando de este modo una

- o más moléculas de ácido nucleico de doble hebra que comprenden la primera y la segunda moléculas de ácido nucleico. En algunas realizaciones, la incubación de la etapa (b) se realiza a una temperatura elevada. En realizaciones específicas, la temperatura elevada puede ser de aproximadamente 40 °C o superior, de aproximadamente 45 °C o superior, de aproximadamente 50 °C o superior, de aproximadamente 51 °C o superior, de aproximadamente 52 °C o superior, de aproximadamente 53 °C o superior, de aproximadamente 54 °C o superior, de aproximadamente 55 °C o superior, de aproximadamente 56 °C o superior, de aproximadamente 57 °C o superior, de aproximadamente 58 °C o superior, de aproximadamente 59 °C o superior, de aproximadamente 60 °C o superior, de aproximadamente 61 °C o superior, de aproximadamente 62 °C o superior, de aproximadamente 63 °C o superior, de aproximadamente 64 °C o superior, de aproximadamente 65 °C o superior, de aproximadamente 66 °C o superior, de aproximadamente 67 °C o superior, de aproximadamente 68 °C o superior, de aproximadamente 69 °C o superior, de aproximadamente 70 °C o superior, de aproximadamente 71 °C o superior, de aproximadamente 72 °C o superior, de aproximadamente 73 °C o superior, de aproximadamente 74 °C o superior, de aproximadamente 75 °C o superior, de aproximadamente 76 °C o superior, de aproximadamente 77 °C o superior, o de aproximadamente 78 °C o superior; o a un intervalo de temperatura de aproximadamente 37 °C a aproximadamente 75 °C, de aproximadamente 40 °C a aproximadamente 75 °C, de aproximadamente 45 °C a aproximadamente 75 °C, de aproximadamente 50 °C a aproximadamente 75 °C, de aproximadamente 51 °C a aproximadamente 75 °C, de aproximadamente 52 °C a aproximadamente 75 °C, de aproximadamente 53 °C a aproximadamente 75 °C, de aproximadamente 54 °C a aproximadamente 75 °C, o de aproximadamente 55 °C a aproximadamente 75 °C. En algunas realizaciones, la temperatura elevada está dentro del intervalo de aproximadamente 50 °C a aproximadamente 70 °C, de aproximadamente 51 °C a aproximadamente 70 °C, de aproximadamente 52 °C a aproximadamente 70 °C, de aproximadamente 53 °C a aproximadamente 70 °C, de aproximadamente 54 °C a aproximadamente 70 °C, de aproximadamente 55 °C a aproximadamente 70 °C, de aproximadamente 56 °C a aproximadamente 65 °C, de aproximadamente 57 °C a aproximadamente 65 °C, de aproximadamente 58 °C a aproximadamente 64 °C, o de aproximadamente 59 °C a aproximadamente 62 °C. En otras realizaciones, la temperatura elevada puede estar dentro del intervalo de aproximadamente 45 °C a aproximadamente 60 °C, de aproximadamente 46 °C a aproximadamente 60 °C, de aproximadamente 47 °C a aproximadamente 60 °C, de aproximadamente 48 °C a aproximadamente 60 °C, de aproximadamente 49 °C a aproximadamente 60 °C, de aproximadamente 50 °C a aproximadamente 60 °C, de aproximadamente 51 °C a aproximadamente 60 °C, de aproximadamente 52 °C a aproximadamente 60 °C, de aproximadamente 53 °C a aproximadamente 60 °C, o de aproximadamente 54 °C a aproximadamente 60 °C. Tales métodos pueden incluir el uso de una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, etc.) ADN-polimerasas como parte del procedimiento de preparación de una o más moléculas de ácido nucleico de doble hebra. Dichas ADN-polimerasas son preferiblemente ADN-polimerasas termoestables y lo más preferiblemente la síntesis del ácido nucleico conseguida con dichas ADN-polimerasas se realiza a temperaturas elevadas, esto es, superiores a 37 °C. La descripción se refiere también a composiciones útiles para preparar dichas moléculas de ácido nucleico de doble hebra. Tales composiciones comprenden una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, etc.) transcriptasas inversas y opcionalmente una o más ADN-polimerasas, un tampón adecuado, uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, etc.) cebadores, y/o uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, etc.) nucleótidos.
- La invención se refiere también a métodos para amplificar una molécula de ácido nucleico. Dichos métodos de amplificación comprenden mezclar la molécula o moléculas de ácido nucleico de doble hebra producidas como se ha descrito antes con una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, etc.) ADN-polimerasas (preferiblemente ADN-polimerasas termoestables) e incubar la mezcla en condiciones suficientes para amplificar la molécula de ácido nucleico de doble hebra. En una primera realización, la invención se refiere a un método para amplificar una molécula de ácido nucleico, comprendiendo el método (a) mezclar uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, veinte, etc.) moldes de ácido nucleico (preferiblemente uno o más moldes de ARN o mARN y más preferiblemente una población de moldes de mARN) con una o más transcriptasas inversas de la invención y con una o más ADN-polimerasas y (b) incubar la mezcla en condiciones suficientes para amplificar las moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes. En algunas realizaciones, la incubación de la etapa (b) se realiza a una temperatura elevada. En realizaciones específicas, la temperatura elevada puede ser de aproximadamente 40 °C o superior, 45 °C o superior, 50 °C o superior, 51 °C o superior, aproximadamente 52 °C o superior, aproximadamente 53 °C o superior, aproximadamente 54 °C o superior, aproximadamente 55 °C o superior, aproximadamente 56 °C o superior, aproximadamente 57 °C o superior, aproximadamente 58 °C o superior, aproximadamente 59 °C o superior, aproximadamente 60 °C o superior, aproximadamente 61 °C o superior, aproximadamente 62 °C o superior, aproximadamente 63 °C o superior, aproximadamente 64 °C o superior, aproximadamente 65 °C o superior, aproximadamente 66 °C o superior, aproximadamente 67 °C o superior, aproximadamente 68 °C o superior, aproximadamente 69 °C o superior, aproximadamente 70 °C o superior, aproximadamente 71 °C o superior, aproximadamente 72 °C o superior, aproximadamente 73 °C o superior, aproximadamente 74 °C o superior, aproximadamente 75 °C o superior, aproximadamente 76 °C o superior, aproximadamente 77 °C o superior, o aproximadamente 78 °C o superior; o a un intervalo de temperatura de aproximadamente 37 °C a aproximadamente 75 °C, de aproximadamente 40 °C a aproximadamente 75 °C, de aproximadamente 45 °C a aproximadamente 75 °C, de aproximadamente 50 °C a aproximadamente 75 °C, de aproximadamente 51 °C a aproximadamente 75 °C, de aproximadamente 52 °C a aproximadamente 75 °C, de aproximadamente 53 °C a aproximadamente 75 °C, de aproximadamente 54 °C a aproximadamente 75 °C, de aproximadamente 55 °C a aproximadamente 75 °C. En algunas realizaciones, la

temperatura elevada está dentro del intervalo de aproximadamente 50 °C a aproximadamente 70 °C, de aproximadamente 51 °C a aproximadamente 70 °C, de aproximadamente 52 °C a aproximadamente 70 °C, de aproximadamente 53 °C a aproximadamente 70 °C, de aproximadamente 54 °C a aproximadamente 70 °C, de aproximadamente 55 °C a aproximadamente 70 °C, de aproximadamente 55 °C a aproximadamente 65 °C, de aproximadamente 56 °C a aproximadamente 65 °C, de aproximadamente 56 °C a aproximadamente 64 °C o de aproximadamente 56 °C a aproximadamente 62 °C. En otras realizaciones, la temperatura elevada puede estar dentro del intervalo de aproximadamente 45 °C a aproximadamente 60 °C, de aproximadamente 46 °C a aproximadamente 60 °C, de aproximadamente 47 °C a aproximadamente 60 °C, de aproximadamente 48 °C a aproximadamente 60 °C, de aproximadamente 49 °C a aproximadamente 60 °C, de aproximadamente 50 °C a aproximadamente 60 °C, de aproximadamente 51 °C a aproximadamente 60 °C, de aproximadamente 52 °C a aproximadamente 60 °C, de aproximadamente 53 °C a aproximadamente 60 °C, o de aproximadamente 54 °C a aproximadamente 60 °C. Preferiblemente, las transcriptasas inversas (1) tienen actividad RNasa H reducida o sustancialmente reducida, (2) tienen actividad TdT reducida o sustancialmente reducida, y/o (3) presentan aumento de la fidelidad. Preferiblemente, las ADN-polimerasas comprenden una primera ADN-polimerasa que tiene actividad de 3' exonucleasa y una segunda ADN-polimerasa que tiene actividad de 3' exonucleasa sustancialmente reducida.

La descripción se refiere también a composiciones que comprenden una o más transcriptasas inversas y una o más ADN-polimerasas para uso en reacciones de amplificación. Tales composiciones pueden comprender además uno o más nucleótidos y/o un tampón adecuado para amplificación. Las composiciones pueden comprender también uno o más cebadores oligonucleótidos.

La descripción se dirige también a moléculas de ácido nucleico (particularmente moléculas de cADN de una hebra o de doble hebra) o moléculas amplificadas de ácido nucleico producidas según los métodos descritos anteriormente y a vectores (particularmente vectores de expresión) que comprenden estas moléculas de ácido nucleico o moléculas amplificadas de ácido nucleico.

La descripción se dirige además a células hospedantes recombinantes que comprenden las moléculas de ácido nucleico, moléculas amplificadas de ácido nucleico o vectores, descritos anteriormente. Los ejemplos de dichas células hospedantes incluyen células bacterianas, células de levaduras, células de plantas y células de animales (incluyendo células de insectos y células de mamíferos).

La invención se dirige adicionalmente a métodos para producir polipéptidos codificados por las moléculas de ácido nucleico producidas por los métodos de la invención. Dichos métodos comprenden cultivar las células hospedantes recombinantes descritas anteriormente y aislar los polipéptidos codificados, y a los polipéptidos producidos por dichos métodos.

La invención se refiere también a métodos para secuenciar una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, etc.) moléculas de ácido nucleico utilizando las composiciones o enzimas de la invención. Dichos métodos comprenden (a) mezclar una o más moléculas de ácido nucleico (*p. ej.*, una o más moléculas de ARN o ADN) a secuenciar con uno o más cebadores, una o más transcriptasas inversas de la invención, uno o más nucleótidos y uno o más agentes de terminación, tal como uno o más didesoxinucleósidos trifosfatos; (b) incubar la mezcla en condiciones suficientes para sintetizar una población de moléculas de ácido nucleico complementarias de todo o parte de la una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, cincuenta, cien, doscientas, etc.) moléculas de ácido nucleico a secuenciar; y (c) separar la población de moléculas de ácido nucleico para determinar la secuencia de nucleótidos de todo o parte de la una o más moléculas de ácido nucleico a secuenciar. Dichos métodos pueden comprender también (a) mezclar una molécula de ácido nucleico (*p. ej.*, una o más moléculas de ARN o ADN) a secuenciar con uno o más cebadores, una o más transcriptasas inversas de la invención, uno o más nucleótidos y uno o más agentes de terminación, tal como uno o más didesoxinucleósidos trifosfatos; (b) incubar la mezcla en condiciones suficientes para sintetizar una población de moléculas de ácido nucleico complementarias de todo o parte de la molécula de ácido nucleico a secuenciar; y (c) separar los miembros de la población de moléculas de ácido nucleico para determinar la secuencia de nucleótidos de todo o parte de la molécula de ácido nucleico a secuenciar. En algunas realizaciones, dicha incubación se puede llevar a cabo a temperaturas elevadas como se describe en este documento.

La invención se refiere también a kits para uso en métodos de la invención. Tales kits se pueden utilizar para preparar, secuenciar o amplificar moléculas de ácido nucleico (de única hebra o doble hebra), preferiblemente a las temperaturas elevadas descritas en la presente memoria. Los kits comprenden un soporte, tal como una caja o cartón, que tiene en estrecho confinamiento en la misma, uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, etc.) recipientes, tales como viales, tubos, frascos y similares. En los kits de la invención, un primer recipiente contiene una o más de las enzimas transcriptasas inversas de la presente invención. Los kits de la invención pueden comprender también, en el mismo o diferentes recipientes, una o más ADN-polimerasas (preferiblemente ADN-polimerasas termoestables), uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, etc.) tampones adecuados para la síntesis de ácido nucleico, uno o más nucleótidos y uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, etc.) cebadores oligonucleótidos. Alternativamente, los componentes del kit se pueden dividir en recipientes separados (*p. ej.*, un recipiente para cada enzima y/o componente). Los kits de la invención pueden comprender también instrucciones o protocolos para llevar a cabo los métodos de la invención. En los kits a modo de ejemplo, las transcriptasas inversas tienen actividad de RNasa H reducida o sustancialmente

reducida, y se seleccionan lo más preferiblemente del grupo que consiste en RNasa H-transcriptasa inversa de M-MLV, RNasa H- transcriptasa inversa de RSV, RNasa H- transcriptasa inversa de AMV, RNasa H- transcriptasa inversa de RAV, RNasa H- transcriptasa inversa de MAV y RNasa H- transcriptasa inversa de VIH. En otros kits preferidos, las transcriptasas inversas tienen actividad TdT reducida o sustancialmente reducida, y/o presentan aumento de la fidelidad, como se describe en otra parte en este documento.

En kits adicionales preferidos de la invención, las enzimas (transcriptasas inversas y/o ADN-polimerasas) de los recipientes están presentes en concentraciones de trabajo.

Por lo tanto, la invención se dirige además a kits para uso en la transcripción inversa, amplificación o secuenciación de una molécula de ácido nucleico, comprendiendo el kit una o más transcriptasas inversas termoestables.

En realizaciones específicas, las transcriptasas inversas de los kits descritos pueden tener una o más modificaciones o mutaciones en posiciones que corresponden a los aminoácidos seleccionados del grupo que consiste en:

(a) leucina 52 de la transcriptasa inversa de M-MLV;

(b) tirosina 64 de la transcriptasa inversa de M-MLV;

(c) lisina 152 de la transcriptasa inversa de M-MLV;

(d) arginina 204 de la transcriptasa inversa de M-MLV;

(e) metionina 289 de la transcriptasa inversa de M-MLV; y

(f) treonina 306 de la transcriptasa inversa de M-MLV.

Las transcriptasas inversas de la descripción incluyen cualquier transcriptasa inversa que tiene aumento de la termoestabilidad. Dichas transcriptasas inversas incluyen transcriptasas inversas retrovirales, transcriptasas inversas bacterianas, transcriptasas inversas de retrotransposones (p. ej., transcriptasas inversas de los retrotransposones Ty1 y/o Ty3), y ADN-polimerasas que tienen actividad de transcriptasa inversa. Las transcriptasas inversas descritas incluyen transcriptasas inversas de una sola subunidad y de multi-subunidades y preferiblemente transcriptasas inversas retrovirales. En particular, la descripción se refiere a transcriptasas inversas de M-MLV y transcriptasas inversas de ASLV (tales como AMV-RT y RSV-RT). Dichas transcriptasas inversas preferiblemente tienen actividad de RNasa H reducida o sustancialmente reducida.

En otros ejemplos determinados, las transcriptasas inversas pueden no incluir combinaciones específicas de mutaciones, tales como las transcriptasas inversas que combinan una mutación de la histidina 204 a arginina (H204R) y una mutación de la treonina 306 a lisina (T306K); las transcriptasas inversas que combinan H204R y T306K con una mutación de fenilalanina 309 a asparagina (F309N); y transcriptasas inversas que combinan H204R, T306K y F309N con una mutación de valina 223 a histidina (V223H).

Otras realizaciones de la presente invención serán evidentes para los expertos en la técnica a la vista de los siguientes dibujos y descripción de la invención, y de las reivindicaciones.

Breve descripción de los dibujos/figuras

La Figura 1 es un mapa del plásmido pBAD-6-His-M-MLV H- (F1).

La Figura 2 es una representación lineal de la secuencia codificadora de la transcriptasa inversa de M-MLV que muestra las localizaciones de los sitios de escisión de la enzima de restricción utilizados para generar los segmentos del gen utilizado para generar mutaciones.

La Figura 3 representa una fosfoimagen de barrido de un ensayo de extensión utilizando (1) transcriptasa inversa SuperScript™II, y (2) F309N. El cebador de 18-meros marcado con ³²P, anillado a un molde de ADN de 47-meros (5 nM) se extendió por unidades iguales de transcriptasa inversa a 37 °C durante 30 minutos como se ve en las reacciones de extensión con los 4 nucleótidos. Las reacciones de extensión se analizaron por electroforesis en gel desnaturalizante al 6 %. P, cebador no extendido.

La Figura 4 representa una fosfoimagen de barrido que muestra un ensayo de extensión de TdT de transcriptasa inversa SuperScript™II y los mutantes F309N, T197E, y Y133A. El cebador de 18-meros marcado con ³²P, anillado a un molde de ADN de 47-meros (5 nM) se extendió con unidades decrecientes de transcriptasa inversa (pista (1) 646 unidades, pista (2) 200 unidades, pista (3) 50 unidades, y pista (4) 20 unidades) a 37 °C durante 30 minutos con los cuatro nucleótidos (véase la sección de Métodos más adelante en el ejemplo 3). Las reacciones de extensión se analizaron por electroforesis en gel desnaturalizante al 6 %. En este ensayo, la extensión más allá de los moldes de 47 nucleótidos se considera adición no dirigida al molde o actividad TdT. P, cebador no extendido.

La Figura 5 representa una fosfoimagen de barrido que muestra ensayos de inserción defectuosa de transcriptasa inversa SuperScript™II (1) y transcriptasa inversa F309N de proteína mutante (2) con molde de ADN. El cebador de 18-meros marcado con [³²P], anillado a un molde de ADN de 47-meros (5 nM) se extendió con unidades iguales de proteína de transcriptasa inversa a 37 °C durante 30 min. como se ve en las reacciones de extensión con los cuatro nucleótidos. Las reacciones de extensión se realizaron también en presencia de solamente 3 dNTPs complementarios; minus dCTP, minus dATP, minus TTP, y minus dGTP. Las reacciones de extensión se analizaron por electroforesis en gel desnaturante al 6 %. En este ensayo, una eficiencia más alta de elongación del cebador terminado con sólo tres nucleótidos reflejará una fidelidad más baja de la transcriptasa inversa SuperScript™II ensayada. P, cebador no extendido.

La Figura 6 representa una fosfoimagen de barrido que muestra un ensayo de inserción defectuosa de transcriptasa inversa SuperScript™II (1) y transcriptasa inversa T197A/F309N de proteína mutante (2) y V223H/F309N (3) con molde de ADN. El cebador de 18-meros marcado con [³²P], anillado a un molde de ADN de 47-meros (5 nM) se extendió con unidades iguales de proteína de transcriptasa inversa a 37 °C durante 30 min. como se ve en las reacciones de extensión con los cuatro nucleótidos. Las reacciones de extensión se realizaron también en presencia de solamente 3 dNTPs complementarios; minus dATP, y minus dCTP. Las reacciones de extensión se analizaron por electroforesis en gel desnaturante al 6 %. En este ensayo, una eficiencia más alta de elongación del cebador terminado con sólo tres nucleótidos reflejará una fidelidad más baja de la transcriptasa inversa SuperScript™II ensayada. P, cebador no extendido.

Descripción detallada de la invención

En la descripción que sigue, se utilizan una serie de términos utilizados en ADN recombinante, virología e inmunología. Para proporcionar una comprensión más clara y consistente de la memoria descriptiva y de las reivindicaciones, incluyendo el alcance a dar a dichos términos, se proporcionan las siguientes definiciones.

Vector de clonación. Como se usa en la presente memoria, "vector de clonación" significa una molécula de ácido nucleico tal como un plásmido, cósmido, fago, fagémido u otra molécula de ácido nucleico que es capaz de replicarse de modo autónomo en una célula hospedante, y que se caracteriza por uno o un pequeño número de sitios de reconocimiento de la endonucleasa de restricción en los cuales dichas secuencias de ácido nucleico pueden ser cortadas de una manera determinable, y en los cuales se puede insertar ADN para llevar a cabo su replicación y clonación. El vector de clonación puede contener además un marcador adecuado para uso en la identificación de células transformadas con el vector de clonación. Los marcadores, por ejemplo, son genes que confieren un fenotipo reconocible a las células hospedantes en las que se expresan dichos marcadores. Los marcadores utilizados comúnmente incluyen, pero no se limitan a, genes de resistencia a antibióticos tales como resistencia a tetraciclina o resistencia a ampicilina.

Vector de expresión. Como se usa en la presente memoria, "vector de expresión" significa una molécula de ácido nucleico similar a un vector de clonación pero que puede comprender adicionalmente secuencias de ácido nucleico capaces de aumentar y/o controlar la expresión de un gen u otra molécula de ácido nucleico que ha sido clonada en él, después de la transformación en un hospedante. Las secuencias de ácido nucleico adicionales pueden comprender secuencias promotoras, secuencias de unión represoras y similares. El gen clonado o la molécula de ácido nucleico, están usualmente ligados operativamente a una o más (p. ej., una, dos, tres, cuatro, etc.) de dichas secuencias de control tal como las secuencias promotoras.

Hospedante recombinante. Como se usa en la presente memoria, "recombinante" significa cualquier organismo procariota o eucariota o microorganismo que contiene los genes clonados o moléculas de ácido nucleico deseados, por ejemplo, en un vector de expresión, vector de clonación o cualquier molécula de ácido nucleico. El término "hospedante recombinante" quiere incluir también aquellas células hospedantes que han sido manipuladas genéticamente para contener el gen deseado u otra molécula de ácido nucleico en el cromosoma o genoma del hospedante.

Hospedante. Como se usa en la presente memoria, "hospedante" significa cualquier organismo procariota o eucariota que es el receptor de un vector de expresión, vector de clonación o cualquier molécula de ácido nucleico, replicable. La molécula de ácido nucleico puede contener, pero no se limita a, un gen estructural, un promotor y/o un origen de replicación.

Promotor. Como se usa en la presente memoria, "promotor" significa una secuencia de ácido nucleico generalmente descrita como la región 5' de un gen, en localización proximal al codón de partida que es capaz de dirigir la transcripción de un gen u otra molécula de ácido nucleico. En la región promotora, se inicia la transcripción de un gen o genes o ácido nucleico o ácidos nucleicos, adyacentes.

Gen. Como se usa en la presente memoria, "gen" significa una secuencia de ácido nucleico que contiene la información necesaria para la expresión de un polipéptido o proteína. Incluye el gen promotor y el gen estructural, así como otras secuencias implicadas en la expresión de las proteínas.

Gen estructural. Como se usa en la presente memoria, "gen estructural" significa un ADN u otra secuencia de ácido nucleico que es transcrita en el ARN mensajero que se traduce después en una secuencia de aminoácidos característica de un polipéptido específico.

- 5 **Ligado operativamente.** Como se usa en la presente memoria, "ligado operativamente" significa que un elemento de ácido nucleico está posicionado de tal manera que influye en la iniciación de la expresión del polipéptido codificado por el gen estructural u otra molécula de ácido nucleico.

Expresión. Como se usa en la presente memoria, "expresión" se refiere al proceso por el cual un gen u otra molécula de ácido nucleico produce un polipéptido. Incluye la transcripción del gen o molécula de ácido nucleico en el ARN mensajero (mARN) y la traducción de dicho mARN en polipéptido o polipéptidos.

- 10 **Sustancialmente puro.** Como se usa en la presente memoria, "sustancialmente puro" significa que el material deseado está esencialmente libre de componentes celulares contaminantes que se asocian con el material deseado, en la naturaleza. Los componentes celulares contaminantes pueden incluir, pero no se limitan a, actividades enzimáticas tales como fosfatasa, exonucleasas, endonucleasas o enzimas ADN-polimerasas indeseables. Preferiblemente, las transcriptasas inversas de la invención son sustancialmente puras.

- 15 **Cebador.** Como se usa en la presente memoria, "cebador" se refiere a un oligonucleótido de una sola hebra que se extiende mediante unión covalente de los monómeros de nucleótidos durante la amplificación o polimerización de una molécula de ADN.

- 20 **Molde.** El término "molde" como se usa en la presente memoria, se refiere a una molécula de ácido nucleico de doble hebra o de simple hebra que se va a amplificar, copiar o secuenciar. En el caso de una molécula de ADN de doble hebra, la desnaturalización de sus hebras para formar una primera y una segunda hebras de hebra única se puede llevar a cabo antes de que estas moléculas sean amplificadas, copiadas o secuenciadas. Un cebador complementario a una parte de un molde de ácido nucleico se hibrida en condiciones apropiadas y entonces una ácido nucleico-polimerasa, tal como las enzimas transcriptasas inversas de la invención, puede añadir monómeros de nucleótidos al cebador sintetizando de este modo una molécula de ácido nucleico complementario a dicho molde o a una parte del mismo. La molécula de ácido nucleico nuevamente sintetizada, según la invención, puede ser igual o más corta en longitud que el molde original. El desajuste en la incorporación durante la síntesis o extensión de la molécula de ácido nucleico nuevamente sintetizada puede dar como resultado uno o una serie de pares de bases desajustados. Por lo tanto, la molécula de ácido nucleico sintetizada no es necesario que sea exactamente complementaria al molde.

- 30 **Incorporación.** El término incorporación como se usa en la presente memoria, significa llegar a ser parte de una molécula de ácido nucleico o cebador.

Oligonucleótido. "Oligonucleótido" se refiere a una molécula sintética o natural que comprende una secuencia, unida covalentemente, de nucleótidos que están unidos por un enlace de fosfodiéster entre la posición 3' de la pentosa de un nucleótido y la posición 5' de la pentosa del nucleótido adyacente.

- 35 **Nucleótido.** Como se usa en la presente memoria, "nucleótido" se refiere a una combinación de base-azúcar-fosfato. Los nucleótidos son unidades monoméricas de una secuencia de ácido nucleico (ADN y ARN) y los desoxirribonucleótidos son incorporados al ADN por las ADN-polimerasas. El término nucleótido incluye desoxirribonucleósidos trifosfatos tales como dATP, dCTP, dITP, dUTP, dGTP, dTTP, o derivados de los mismos. Tales derivados incluyen, por ejemplo, [(αS)dATP, 7-desaza-dGTP y 7-desaza-dATP. El término nucleótido como se usa en la presente memoria, se refiere también a didesoxirribonucleósidos trifosfatos (ddNTPs) y sus derivados. Ejemplos ilustrativos de didesoxirribonucleósidos trifosfatos incluyen, pero no se limitan a, ddATP, ddCTP, ddGTP, ddITP, y ddTTP. Según la presente invención, un "nucleótido" puede estar sin marcar o marcado detectablemente por técnicas bien conocidas. Las marcas detectables incluyen, por ejemplo, isótopos radioactivos, marcas fluorescentes, marcas quimioluminiscentes, marcas bioluminiscentes y marcas enzimáticas.

- 45 **Hibridación.** Como se usa en la presente memoria, hibridación (hibridar) se refiere al emparejamiento de dos moléculas de ácido nucleico complementarias de una sola hebra (ARN y/o ADN) para dar una molécula de doble hebra. Como podrán apreciar los expertos en la técnica, dos moléculas de ácido nucleico se pueden hibridar, aunque el emparejamiento de las bases no sea totalmente complementario. En consecuencia, las bases desajustadas no evitan la hibridación de dos moléculas de ácido nucleico siempre que se usen las condiciones apropiadas, bien conocidas en la técnica.

- 50 **Transcriptasa inversa termoestable.** Para los fines de esta descripción, una transcriptasa inversa termoestable se define como una transcriptasa inversa que retiene un mayor porcentaje de su actividad después de un tratamiento térmico que el porcentaje retenido por una transcriptasa inversa que tiene termoestabilidad natural después de un tratamiento idéntico. Por lo tanto, una transcriptasa inversa que tiene una termoestabilidad aumentada/mejorada se define como una polimerasa que tiene un aumento de la termoestabilidad, preferiblemente de aproximadamente 1,2 a aproximadamente 10.000 veces, de aproximadamente 1,5 a aproximadamente 10.000 veces, de aproximadamente 2 a aproximadamente 5000 veces, o de aproximadamente 2 a aproximadamente 2000 veces (preferiblemente mayor que aproximadamente 5 veces, más preferiblemente mayor que aproximadamente 10 veces,

5 todavía más preferiblemente mayor que aproximadamente 50 veces, aún más preferiblemente mayor que
aproximadamente 100 veces, aún más preferiblemente mayor que aproximadamente 500 veces, y lo más
preferiblemente mayor que aproximadamente 1000 veces) de retención de la actividad después de un tratamiento
térmico suficiente para causar una reducción en la actividad de una transcriptasa inversa que es natural para la
10 termoestabilidad. Preferiblemente, la transcriptasa inversa mutante o modificada de la invención se compara con la
correspondiente transcriptasa inversa no modificada o natural para determinar la mejora o aumento relativo de la
termoestabilidad. Por ejemplo, después de un tratamiento térmico a 52 °C durante 5 minutos, una transcriptasa
inversa termoestable puede retener aproximadamente el 90 % de la actividad presente antes del tratamiento
térmico, mientras que una transcriptasa inversa que es natural para la termoestabilidad, puede retener el 10 % de su
15 actividad original. Asimismo, después de un tratamiento térmico a 53 °C durante cinco minutos, una transcriptasa
inversa termoestable puede retener aproximadamente un 80 % de su actividad original, mientras que una
transcriptasa inversa que es natural para la termoestabilidad puede no tener ninguna actividad medible.
Similarmente, después de un tratamiento térmico a 50 °C durante cinco minutos, una transcriptasa inversa
termoestable puede retener aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, o aproximadamente 95 % de su actividad original, mientras que una transcriptasa inversa que es natural para la termoestabilidad puede no tener ninguna actividad medible o puede retener 10 %, 15 % o 20 % de su actividad original. En el primer caso (esto es, después de tratamiento térmico a 52 °C durante 5 minutos), la transcriptasa inversa termoestable se podría decir que es 9 veces más termoestable que la transcriptasa inversa natural. Ejemplos de condiciones que se pueden utilizar para medir la termoestabilidad de las transcriptasas inversas se indican más adelante en el ejemplo 2.

La termoestabilidad de una transcriptasa inversa se puede determinar comparando la actividad residual de una muestra de la transcriptasa inversa que ha sido sometida a un tratamiento térmico, esto es, incubada a 52 °C durante un período de tiempo dado, por ejemplo, cinco minutos, con una muestra control de la misma transcriptasa inversa que ha sido incubada a temperatura ambiente durante el mismo período de tiempo que el tratamiento térmico. Típicamente la actividad residual se puede medir siguiendo la incorporación de un desoxirribonucleótido radiomarcado a un cebador oligodesoxirribonucleótido utilizando un molde de oligorribonucleótido complementario. Por ejemplo, se puede ensayar la capacidad de la transcriptasa inversa para incorporar $[\alpha\text{-}^{32}\text{P}]\text{-dGTP}$ a un cebador de oligo-dG utilizando un molde de poli(riboC) para determinar la actividad residual de la transcriptasa inversa.

En otro aspecto, las transcriptasas inversas termoestables se definen como cualquier transcriptasa inversa que se inactiva a una temperatura más alta en comparación con la correspondiente transcriptasa inversa natural, no mutada, o no modificada. Preferiblemente, la temperatura de inactivación para las transcriptasas inversas termoestables de la invención es de aproximadamente 2 °C a aproximadamente 50 °C (p. ej., aproximadamente 2 °C, aproximadamente 4 °C, aproximadamente 5 °C, aproximadamente 8 °C, aproximadamente 10 °C, aproximadamente 12 °C, aproximadamente 14 °C, aproximadamente 16 °C, aproximadamente 18 °C, aproximadamente 20 °C, aproximadamente 24 °C, aproximadamente 26 °C, aproximadamente 28 °C, aproximadamente 30 °C, aproximadamente 33 °C, aproximadamente 35 °C, aproximadamente 38 °C, aproximadamente 40 °C, aproximadamente 42 °C, aproximadamente 44 °C, aproximadamente 46 °C, aproximadamente 48 °C, o aproximadamente 50 °C) más alta que la temperatura de inactivación para la correspondiente transcriptasa inversa natural, no mutada, o no modificada. Más preferiblemente, la temperatura de inactivación para las transcriptasas inversas de la invención es de aproximadamente 5 °C a aproximadamente 50 °C, de aproximadamente 5 °C a aproximadamente 40 °C, de aproximadamente 5 °C a aproximadamente 30 °C, o de aproximadamente 5 °C a aproximadamente 25 °C mayor que la temperatura de inactivación para la correspondiente transcriptasa inversa natural, no mutada, o no modificada; cuando se comparan en las mismas condiciones.

La diferencia en la temperatura de inactivación para la transcriptasa inversa de la invención comparada con su correspondiente transcriptasa inversa natural, no mutada, o no modificada, se puede determinar tratando muestras de dichas transcriptasas inversas a diferentes temperaturas durante un período de tiempo definido y midiendo después la actividad residual de transcriptasa inversa, si la hubiera, una vez que las muestras han sido tratadas por calor. La determinación de la diferencia o delta en la temperatura de inactivación entre la transcriptasa inversa de ensayo comparada con el control natural, no mutado, o no modificado, se realiza comparando la diferencia de temperatura a la que cada transcriptasa inversa es inactivada (es decir, ninguna actividad residual de transcriptasa inversa es medible en el ensayo particular utilizado). Como se podrá apreciar, se pueden utilizar cualquier número de ensayos de transcriptasa inversa para determinar la diferencia o delta de las temperaturas de inactivación para todas las transcriptasas inversas ensayadas.

Actividad de extensión terminal. Como se usa en la presente memoria, actividad de extensión terminal se refiere a la capacidad de una transcriptasa inversa (RT) para añadir bases adicionales al extremo 3' de una hebra de cADN nuevamente sintetizada, más allá del extremo 5' del molde de ADN o mARN. La actividad de extensión terminal puede añadir bases específicamente (con preferencia un nucleótido) o aleatoriamente.

La actividad de extensión terminal es conocida también como actividad de desoxinucleotidil-transferasa terminal (TdT). Una transcriptasa inversa que tiene actividad TdT reducida, sustancialmente reducida, o eliminada, se define como cualquier transcriptasa inversa que tiene una actividad TdT más baja que la actividad específica de la correspondiente enzima natural, no mutada o no modificada, particularmente, inferior a aproximadamente 90 % de la

actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 85 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 80 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 75 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 50 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 25 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 15 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a 10 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, inferior a aproximadamente 5 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada, o inferior a aproximadamente 1 % de la actividad específica de la correspondiente enzima natural, no mutada o no modificada. La actividad TdT eliminada, se define como un nivel de actividad que es indetectable por los métodos de ensayo indicados aquí en el ejemplo 3.

Como se indica más adelante en el ejemplo 3, se conocen en la técnica transcriptasas inversas que extienden las moléculas de ácido nucleico 2-3 nucleótidos más allá del extremo de los moldes (*p. ej.*, moldes de ARN o ADN). Además, en cualquier mezcla de reacción en la que tiene lugar la transcripción inversa, pueden estar presentes mezclas de moléculas que contienen diferentes números de nucleótidos que se extienden más allá del extremo del molde. La actividad TdT se determina en esta memoria en referencia al número o porcentaje de moléculas que contienen uno o más nucleótidos que se extienden más allá del extremo del molde. Por ejemplo, si una transcriptasa inversa natural añade 1 o más nucleótidos más allá del extremo de un molde al 90 % de las moléculas generadas durante la transcripción inversa y una transcriptasa inversa modificada añade 1 o más nucleótidos más allá del extremo de un molde al 45 % de las moléculas en las mismas o similares condiciones, entonces se podría decir que la transcriptasa inversa modificada presenta una reducción del 50 % en la actividad TdT en comparación con la enzima natural. Además, se ha generado un mutante F309N, T306K, H204R de SuperScript™II de M-MLV que presenta aproximadamente 0 % de la actividad TdT presentada por SuperScript™II cuando se utiliza ADN como un molde y aproximadamente 10-20 % de la actividad TdT presentada por SuperScript™ cuando se utiliza ARN como un molde. Sin embargo, otras realizaciones determinadas de la descripción excluyen específicamente al mutante F309N, T306K, H204R.

Fidelidad. La fidelidad se refiere a la certeza de polimerización, o a la capacidad de la transcriptasa inversa para discriminar los sustratos correctos de los incorrectos, (*p. ej.*, nucleótidos) cuando se sintetizan moléculas de ácido nucleico que son complementarias a un molde. Cuanto más alta es la fidelidad de una transcriptasa inversa, menos nucleótidos incorpora erróneamente la transcriptasa inversa en la hebra creciente durante la síntesis de ácido nucleico; esto es, un aumento o mejora en la fidelidad tiene como resultado una transcriptasa inversa más exacta con reducción de la tasa de error o reducción de la tasa de incorporación errónea.

Una transcriptasa inversa que tiene una fidelidad aumentada/mejorada/más alta, se define como una polimerasa que tiene algún aumento en la fidelidad, preferiblemente una reducción en el número de nucleótidos incorporados erróneamente de aproximadamente 1,2 a aproximadamente 10.000 veces, aproximadamente 1,5 a aproximadamente 10.000 veces, aproximadamente 2 a aproximadamente 5000 veces, o aproximadamente 2 a aproximadamente 2000 veces (preferiblemente mayor que aproximadamente 5 veces, más preferiblemente mayor que aproximadamente 10 veces, aún más preferiblemente mayor que aproximadamente 50 veces, aún más preferiblemente mayor que aproximadamente 100 veces, aún más preferiblemente mayor que aproximadamente 500 veces y lo más preferiblemente mayor que aproximadamente 1000 veces), durante la síntesis de cualquier molécula dada de ácido nucleico de una longitud dada en comparación con la transcriptasa inversa control. Preferiblemente, la transcriptasa inversa mutada o modificada de la invención, se compara con la correspondiente transcriptasa inversa no modificada o natural para determinar la mejora o aumento relativo en la fidelidad. Por ejemplo, una transcriptasa inversa mutada puede incorporar erróneamente un nucleótido en la síntesis de un segmento de una molécula de ácido nucleico de 1000 bases en comparación con una transcriptasa inversa no mutada que incorpora erróneamente 10 nucleótidos en el segmento del mismo tamaño. Se podría decir que dicha transcriptasa inversa mutante tiene un aumento de la fidelidad de 10 veces.

La fidelidad se puede medir también por la reducción de la incidencia de desplazamiento del marco de lectura, como se describe más adelante en el ejemplo 5. Una transcriptasa inversa que tiene aumento de la fidelidad se define como una polimerasa o transcriptasa inversa que tiene algún aumento en la fidelidad con respecto al desplazamiento del marco de lectura, en comparación con una transcriptasa inversa control (*p. ej.*, una transcriptasa inversa natural), por ejemplo, una transcriptasa inversa que tiene un aumento de la fidelidad mayor que aproximadamente 1,2 veces con respecto al desplazamiento del marco de lectura, que tiene un aumento de la fidelidad mayor que aproximadamente 1,5 veces con respecto al desplazamiento del marco de lectura, que tiene un aumento de la fidelidad mayor que aproximadamente 5 veces con respecto al desplazamiento del marco de lectura, que tiene un aumento de la fidelidad mayor que aproximadamente 10 veces con respecto al desplazamiento del marco de lectura, que tiene un aumento de la fidelidad mayor que aproximadamente 20 veces con respecto al desplazamiento del marco de lectura, que tiene un aumento de la fidelidad mayor que aproximadamente 30 veces con respecto al desplazamiento del marco de lectura, o que tiene un aumento de la fidelidad mayor que aproximadamente 40 veces con respecto al desplazamiento del marco de lectura.

Una transcriptasa inversa que tiene una fidelidad aumentada/mejorada/más alta, con respecto al desplazamiento del marco de lectura, se puede definir también como una transcriptasa inversa o polimerasa que tiene algún aumento en la fidelidad, tal como de aproximadamente 1,5 a aproximadamente 10.000 veces, de aproximadamente 2 a aproximadamente 5.000 veces, de aproximadamente 2 a aproximadamente 2000 veces, de aproximadamente 1,5 a aproximadamente 40 veces, de aproximadamente 5 a aproximadamente 40 veces, de aproximadamente 10 a aproximadamente 40 veces, de aproximadamente 20 a aproximadamente 40 veces, de aproximadamente 30 a aproximadamente 40 veces, de aproximadamente 5 a aproximadamente 30 veces, de aproximadamente 10 a aproximadamente 30 veces, de aproximadamente 15 a aproximadamente 30 veces, de aproximadamente 20 a aproximadamente 30 veces, de aproximadamente 5 a aproximadamente 20 veces, de aproximadamente 10 a aproximadamente 20 veces, de aproximadamente 15 a aproximadamente 20 veces, de aproximadamente 10 a aproximadamente 100 veces, de aproximadamente 15 a aproximadamente 100 veces, de aproximadamente 20 a aproximadamente 100 veces, de aproximadamente 30 a aproximadamente 100 veces, o de aproximadamente 50 a aproximadamente 100 veces de aumento de la fidelidad con respecto al desplazamiento del marco de lectura.

Una transcriptasa inversa que tiene una reducción de la incorporación errónea, se define en la presente memoria como una transcriptasa inversa mutada o modificada que tiene aproximadamente o menos del 90 %, que tiene aproximadamente o menos del 85 %, que tiene aproximadamente o menos del 75 %, que tiene aproximadamente o menos del 70 %, que tiene aproximadamente o menos del 60 %, o preferiblemente que tiene aproximadamente o menos del 50 %, preferiblemente que tiene aproximadamente o menos del 25 %, más preferiblemente que tiene aproximadamente o menos del 10 %, y lo más preferiblemente que tiene aproximadamente o menos del 1 % de incorporación errónea relativa en comparación con la correspondiente enzima natural, no mutada o no modificada.

La fidelidad o la tasa de incorporación errónea de una transcriptasa inversa se puede determinar por secuenciación o por otros métodos conocidos en la técnica (Eckert & Kunkel, 1990, Nucl. Acids Res. 18:3739-3744). En un ejemplo, la secuencia de una molécula de ADN sintetizada por las transcriptasas inversas no mutadas y mutadas, se puede comparar con la secuencia esperada (conocida). De este modo, se puede determinar el número de errores (incorporación errónea o desplazamientos del marco de lectura) para cada enzima y se puede comparar. En otro ejemplo, las transcriptasas inversas no mutadas y mutadas, se pueden utilizar para secuenciar una molécula de ADN que tiene una secuencia conocida. Se puede comparar el número de errores de secuenciación (incorporación errónea o desplazamientos del marco de lectura) para determinar la fidelidad o la tasa de incorporación errónea de las enzimas. Otros medios para determinar la fidelidad o la tasa de incorporación errónea incluyen un ensayo de complementación directa utilizando un molde de ARN como se describe más adelante y se ha descrito previamente en Boyer J.C. et al. Methods Enzymol. 275:523 (1996), y se indica en los ejemplos. Otros métodos para determinar la fidelidad o la tasa de incorporación errónea serán reconocidos por los expertos en la técnica.

Salto de hebra. El salto de hebra, como se usa en la presente memoria, se refiere a un tipo de mutación aleatoria causada por una transcriptasa inversa "que salta" más de uno (p. ej., dos, cinco, diez, cincuenta, cien, etc.) nucleótidos en el molde de mARN, dando como resultado la delección de los correspondientes nucleótidos en el cADN resultante.

Dominio de la mano. El dominio de la mano, como se usa en la presente memoria, se refiere a aquellos aminoácidos que están en el área o áreas que controlan la interacción del molde, cebador, o nucleótido de la transcriptasa inversa. Este dominio se caracteriza además por un grupo de tres regiones de estructura secundaria en una enzima transcriptasa inversa, las regiones del pulgar, dedos y palma. La región del pulgar se define como situada entre los aminoácidos 240-315 de la transcriptasa inversa de VIH, o entre los aminoácidos 280-355 de la transcriptasa inversa de M-MLV. La región de los dedos se define como situada entre los aminoácidos 1-85 y 120-154 de la transcriptasa inversa de VIH, o entre los aminoácidos 1-124 y 161-193 de la transcriptasa inversa de M-MLV. La región de la palma se define como situada entre los aminoácidos 86-199 y 155-239 de la transcriptasa inversa de VIH, o entre los aminoácidos 126-160 y 193-279 de la transcriptasa inversa de M-MLV. Estas áreas se definen de modo general, y los aminoácidos que definen los N-termini y C-termini son aproximados. También pueden ser definidas las regiones correspondientes para otras transcriptasas inversas. Las transcriptasas inversas preferidas tienen una o más modificaciones o mutaciones dentro del dominio de la mano. Más particularmente, las transcriptasas inversas comprenden una o más mutaciones o modificaciones dentro de una o más regiones, incluyendo las regiones del pulgar, dedos y palma.

En general, la descripción proporciona composiciones para uso en la transcripción inversa de una molécula de ácido nucleico que comprenden una transcriptasa inversa con una o más (p. ej., una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones o modificaciones que hacen que la transcriptasa inversa sea más termoestable. La descripción proporciona también composiciones para uso en la transcripción inversa de una molécula de ácido nucleico que comprenden una transcriptasa inversa con una o más mutaciones o modificaciones que hacen a la transcriptasa inversa más eficiente, que tenga una fidelidad más alta, y/o que tenga actividad TdT más baja. La descripción proporciona además composiciones que comprenden una transcriptasa inversa con una o más mutaciones o modificaciones que hacen a la transcriptasa inversa más termoestable y más eficiente.

Las enzimas en estas composiciones están presentes preferiblemente en concentraciones de trabajo y tienen también preferiblemente una actividad RNasa H reducida o sustancialmente reducida. Alternativamente, las transcriptasas inversas utilizadas en las composiciones descritas pueden tener actividad RNasa H. Las

transcriptasas inversas preferidas incluyen transcriptasas inversas retrovirales tales como la transcriptasa inversa de M-MLV, transcriptasa inversa de VIH, transcriptasa inversa de RSV, transcriptasa inversa de AMV, transcriptasa inversa de RAV, y la transcriptasa inversa de MAV u otras transcriptasas inversas de ASLV o sus correspondientes derivados de RNasa H. Las transcriptasas inversas adicionales que se pueden utilizar para preparar las composiciones descritas incluyen transcriptasas inversas bacterianas (*p. ej.*, transcriptasa inversa de *Escherichia coli*) (véase, *p. ej.*, Mao et al., Biochem. Biophys. Res. Commun. 227:489-93 (1996)) y transcriptasas inversas de *Saccharomyces cerevisiae* (*p. ej.*, transcriptasas inversas de los retrotransposones Ty1 o Ty3) (véase, *p. ej.*, Cristofari et al., Jour. Biol. Chem. 274:36643-36648 (1999); Mules et al., Jour. Virol. 72:6490-6503 (1998)).

Se puede hacer una serie de mutaciones a las transcriptasas inversas y, en un aspecto preferido, se pueden hacer múltiples mutaciones para producir un aumento de la termoestabilidad. Dichas mutaciones incluyen mutaciones puntuales, mutaciones de desplazamiento del marco de lectura, delecciones e inserciones, siendo preferidas una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones puntuales. Las mutaciones se pueden introducir en las transcriptasas inversas de la presente invención utilizando cualquier metodología conocida por los expertos en la técnica. Las mutaciones pueden ser introducidas aleatoriamente, por ejemplo, llevando a cabo una reacción de la PCR en presencia de manganeso como un cofactor iónico de metal divalente. Alternativamente, se puede utilizar la mutagénesis dirigida por oligonucleótidos para crear las polimerasas mutantes lo que permite todas las clases posibles de cambios de pares de bases en cualquier sitio determinado a lo largo de la molécula de ADN codificadora. En general, esta técnica implica el anillamiento de un oligonucleótido complementario (excepto para uno o más desajustes) con una secuencia de nucleótidos de una sola hebra que codifica la transcriptasa inversa de interés. El oligonucleótido desajustado es extendido entonces por la ADN-polimerasa, generando una molécula de ADN de doble hebra que contiene el cambio deseado en la secuencia de una hebra. Los cambios en la secuencia pueden dar como resultado, por ejemplo, la delección, sustitución, o inserción de un aminoácido. El polinucleótido de doble hebra puede ser insertado entonces en un vector de expresión apropiado, y de este modo se puede producir un polipéptido mutante o modificado. La mutagénesis dirigida por oligonucleótidos descrita antes, puede ser llevada a cabo, por ejemplo, por la PCR.

La invención se dirige también a métodos para transcripción inversa de una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, etc.) moléculas de ácido nucleico, que comprenden mezclar uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, doce, quince, veinte, etc.) moldes de ácido nucleico, que son preferiblemente ARN o ARN mensajero (mARN) y más preferiblemente una población de moléculas de mARN, con una o más transcriptasas inversas de la presente invención e incubar la mezcla en condiciones suficientes para preparar una molécula o moléculas de ácido nucleico complementarias de todo o parte del uno o más (*p. ej.*, uno, dos, tres, cuatro, cinco, diez, quince, veinte, treinta, etc.) moldes. Para preparar la molécula o moléculas de ácido nucleico complementarias al uno o más moldes, se utilizan preferiblemente un cebador (*p. ej.*, un cebador oligo(dT)) y uno o más nucleótidos, para la síntesis de ácido nucleico en la dirección 5' a 3'. Las moléculas de ácido nucleico adecuadas para la transcripción inversa según este aspecto de la invención incluyen cualquier molécula de ácido nucleico, particularmente las derivadas de una célula procariota o eucariota. Dichas células pueden incluir células normales, células enfermas, células transformadas, células establecidas, células progenitoras, células precursoras, células fetales, células embrionarias no humanas, células bacterianas, células de levaduras, células animales (incluyendo células humanas), células aviares, células de plantas y similares, o tejido aislado de una planta o un animal (*p. ej.*, ser humano, vaca, cerdo, ratón, oveja, caballo, mono, canino, felino, rata, conejo, ave, pez, insecto, etc.). Dichas moléculas de ácido nucleico pueden ser aisladas también de virus.

La invención proporciona además métodos para amplificar o secuenciar una molécula de ácido nucleico que comprenden poner en contacto la molécula de ácido nucleico con una transcriptasa inversa de la presente invención. Dichos métodos preferidos comprenden una o más reacciones en cadena de la polimerasa (PCRs).

Fuentes de transcriptasas inversas

Las enzimas para uso en las composiciones, métodos y kits de la descripción, incluyen cualquier enzima que tiene actividad de transcriptasa inversa. Dichas enzimas incluyen, pero no se limitan a, transcriptasa inversa retroviral, transcriptasa inversa retrotransposón, transcriptasa inversa de la hepatitis B, transcriptasa inversa del virus del mosaico de la coliflor, transcriptasa inversa bacteriana, ADN-polimerasa *Tth*, ADN-polimerasa *Taq* (Saiki, R.K., et al., Science 239:487-491 (1988); Patentes de Estados Unidos Números 4.889.818 y 4.965.188), ADN-polimerasa *Tne* (Publicación PCT No. WO 96/10640), ADN-polimerasa *Tma* (Patente de Estados Unidos Nº 5.374.553) y mutantes, fragmentos, variantes o derivados de las mismas (véase, *p. ej.*, las patentes de Estados Unidos de propiedad común 5.948.614 y 6.015.668). Preferiblemente, las transcriptasas inversas descritas incluyen transcriptasas inversas retrovirales tales como la transcriptasa inversa de M-MLV, transcriptasa inversa de AMV, transcriptasa inversa de RSV, transcriptasa inversa de RAV, transcriptasa inversa de MAV, y en general las transcriptasas inversas de ASLV. Como entenderán los expertos en la técnica, se pueden obtener transcriptasas inversas modificadas por métodos de ingeniería recombinante o genética que son rutinarios y bien conocidos en la técnica. Las transcriptasas inversas mutantes se pueden obtener, por ejemplo, mediante la mutación del gen o genes que codifican la transcriptasa inversa de interés, por mutagénesis dirigida al sitio o mutagénesis aleatoria. Tales mutaciones pueden incluir mutaciones puntuales, mutaciones por delección y mutaciones por inserción. Por ejemplo, se pueden utilizar una o más mutaciones puntuales (*p. ej.*, sustitución de uno o más aminoácidos con uno o más aminoácidos diferentes) para construir las transcriptasas inversas mutantes de la invención.

La descripción incluye además transcriptasas inversas que son idénticas en un 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, o 99 % a nivel de aminoácidos con una transcriptasa inversa natural (*p. ej.*, transcriptasa inversa de M-MLV, transcriptasa inversa de AMV, transcriptasa inversa de RSV, transcriptasa inversa de VIH, etc.) y presentan aumento de termoestabilidad. También se describen transcriptasas inversas que son idénticas en un 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, o 99 % a nivel de aminoácidos con una transcriptasa inversa que comprende la secuencia de aminoácidos indicada más adelante en la Tabla 3 (SEQ ID NO:2) y presentan aumento de termoestabilidad y/o más eficiencia (*p. ej.*, que tienen fidelidad más alta y/o que tienen actividad TdT más baja). La descripción incluye también moléculas de ácido nucleico que codifican las transcriptasas inversas descritas antes.

La descripción incluye también fragmentos de transcriptasas inversas que comprenden al menos 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, o 700 residuos de aminoácidos y retienen una o más actividades asociadas con las transcriptasas inversas. Dichos fragmentos se pueden obtener por mutación por delección, mediante métodos recombinantes que son rutinarios y bien conocidos en la técnica, o por digestión enzimática de la transcriptasa inversa o transcriptasas inversas de interés utilizando cualquiera de una serie de enzimas proteolíticas bien conocidas. La descripción incluye además moléculas de ácido nucleico que codifican las transcriptasas inversas mutantes y los fragmentos de transcriptasa inversa, descritos antes.

Los fragmentos de transcriptasa inversa de la descripción comprenden también los aminoácidos 1-355, 1-498, 1-500, y 1-550 de la transcriptasa inversa de M-MLV, así como los correspondientes fragmentos de otras transcriptasas inversas. Los fragmentos de transcriptasa inversa de la descripción comprenden además polipéptidos que son idénticos en un 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, o 99 % a uno o más de los fragmentos indicados antes.

Por una proteína o fragmento de proteína que tiene una secuencia de aminoácidos “idéntica” al menos, por ejemplo, en un 70 % a una secuencia de aminoácidos de referencia, se quiere decir que la secuencia de aminoácidos de la proteína es idéntica a la secuencia de referencia excepto que la secuencia de la proteína puede incluir hasta 30 alteraciones de aminoácidos por cada 100 aminoácidos de la secuencia de aminoácidos de la proteína de referencia. En otras palabras, para obtener una proteína que tiene una secuencia de aminoácidos al menos idéntica en un 70 % a una secuencia de aminoácidos de referencia, hasta un 30 % de los residuos de aminoácidos de la secuencia de referencia pueden estar delecionados o sustituidos con otro aminoácido, o una serie de aminoácidos hasta un 30 % de los residuos de aminoácidos totales en la secuencia de referencia pueden estar insertados en la secuencia de referencia. Estas alteraciones de la secuencia de referencia pueden tener lugar en las posiciones amino (N-)terminal y/o carboxi (C-) terminal de la secuencia de aminoácidos de referencia y/o en cualquier lugar entre estas posiciones terminales, entremezcladas o individualmente entre los residuos en la secuencia de referencia y/o en uno o más grupos contiguos dentro de la secuencia de referencia. Desde un punto de vista práctico, si una secuencia de aminoácidos dada es, por ejemplo, idéntica al menos en un 70 % a la secuencia de aminoácidos de una proteína de referencia, se puede determinar convencionalmente utilizando programas de ordenador conocidos tales como los descritos antes para las determinaciones de identidad de secuencia de ácido nucleico, o utilizando el programa CLUSTAL W (Thompson, J.D., et al., Nucleic Acids Res. 22:4673-4680 (1994)).

Las enzimas preferidas para uso en la invención incluyen aquellas que tienen actividad de RNasa H reducida o sustancialmente reducida. Tales enzimas que tienen actividad de RNasa H reducida o sustancialmente reducida se pueden obtener, por ejemplo, por mutación del dominio de RNasa H dentro de la transcriptasa inversa de interés, por ejemplo, introduciendo una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones puntuales, una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones por delección, y/o una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones por inserción como se ha descrito antes.

Por una enzima con “actividad RNasa H sustancialmente reducida” se entiende que la enzima tiene una actividad RNasa H inferior a aproximadamente 30 %, inferior a aproximadamente 25 %, inferior a aproximadamente 20 %, más preferiblemente inferior a aproximadamente 15 %, inferior a aproximadamente 10 %, inferior a aproximadamente 7,5 %, o inferior a aproximadamente 5 %, y lo más preferiblemente inferior a aproximadamente 5 % o inferior a aproximadamente 2 %, de la actividad RNasa H de la correspondiente enzima natural o enzima RNasa H⁺, tales como las transcriptasas inversas naturales de virus de la leucemia murina de Moloney (M-MLV), de virus de la mieloblastosis aviar (AMV) o de virus del sarcoma de Rous (RSV).

Las transcriptasas inversas que tienen actividad RNasa H reducida o sustancialmente reducida, han sido descritas previamente (véase la patente de Estados Unidos 5.668.005, patente de Estados Unidos 6.063.608, y Publicación PCT No. WO 98/47912). La actividad de RNasa H de cualquier enzima se puede determinar por una variedad de ensayos, tales como los descritos, por ejemplo, en la patente de Estados Unidos N° 5.244.797, en Kotewicz, M.L., et al., Nucl. Acids Res. 16:265 (1988), en Gerard, G.F., et al., FOCUS 14(5):91 (1992), y en la patente de Estados Unidos N° 5.668.005.

Las enzimas descritas incluyen, pero no se limitan a, RNasa H- transcriptasa inversa de M-MLV, RNasa H- transcriptasa inversa de RSV, RNasa H- transcriptasa inversa de AMV, RNasa H- transcriptasa inversa de RAV, RNasa H- transcriptasa inversa de MAV y RNasa H- transcriptasa inversa de VIH. Los expertos en la técnica

entenderán, sin embargo, que cualquier enzima capaz de producir una molécula de ADN a partir de una molécula de ácido ribonucleico (esto es, que tiene actividad de transcriptasa inversa) que tiene actividad RNasa H reducida o sustancialmente reducida, puede ser utilizada de modo equivalente en las composiciones, métodos y kits descritos.

- Las enzimas para uso en la invención incluyen también aquellas en las que la actividad de desoxinucleotidil transferasa terminal (TdT) ha sido reducida, sustancialmente reducida, o eliminada. Tales enzimas que tienen actividad desoxinucleotidil transferasa terminal reducida o sustancialmente reducida, o en las que la actividad TdT ha sido eliminada, se pueden obtener, por ejemplo, por mutación de los residuos de aminoácidos dentro de la transcriptasa inversa de interés que están en estrecha proximidad o en contacto con el molde-cebador, por ejemplo, introduciendo una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, doce, quince, veinte, treinta, etc.) mutaciones puntuales, una o más mutaciones por delección, y/o una o más mutaciones por inserción. Las transcriptasas inversas que presentan reducción de la actividad TdT están descritas en la Publicación de Estados Unidos No. 2003/0003452 presentada el 15 de marzo de 2001, e incluyen transcriptasas inversas con una o más alteraciones en posiciones de aminoácidos equivalentes o correspondientes a Y64, M289, F309, T197 y/o Y133 de la transcriptasa inversa de M-MLV.
- En un aspecto, las sustituciones de aminoácidos se realizan en una o más de las posiciones identificadas antes (esto es, posiciones de aminoácidos equivalentes o correspondientes a Y64, M289, F309, T197 o Y133 de la transcriptasa inversa de M-MLV). Por lo tanto, los aminoácidos en estas posiciones pueden ser reemplazados con otros aminoácidos incluyendo Ala, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, y Val. Ejemplos específicos de transcriptasas inversas que presentan una actividad TdT reducida o sustancialmente reducida, incluyen las transcriptasas inversas de M-MLV (*p. ej.*, SuperScript™II) en las que (1) el residuo de fenilalanina en la posición 309 ha sido reemplazado con asparagina, (2) el residuo de treonina en la posición 197 ha sido reemplazado con alanina o con ácido glutámico, y/o (3) el residuo de tirosina en la posición 133 ha sido reemplazado con alanina.
- Las enzimas para uso en la invención incluyen también aquellas que presentan aumento de la fidelidad. Las transcriptasas inversas que presentan aumento de la fidelidad están descritas en la solicitud de Estados Unidos N° 60/189.454, presentada el 15 de marzo de 2000, y en la publicación de Estados Unidos No. 2003/0003452 presentada el 15 de marzo de 2001, e incluyen transcriptasas inversas con una o más alteraciones en las posiciones equivalentes o correspondientes a las indicadas a continuación en la Tabla 2.

Tabla 2:

Transcriptasa inversa (RT)	Aminoácido
M-MLV	Y64 (<i>p. ej.</i> , Y64W, Y64R), R116 (<i>p. ej.</i> , R116M), K152 (<i>p. ej.</i> , K152R), Q190 (<i>p. ej.</i> , Q190F), T197 (<i>p. ej.</i> , T197A, T197E), V223 (<i>p. ej.</i> , V223H, V223I, V223F), D124, H126, Y133 (<i>p. ej.</i> , Y133A, Y133H), F309 (<i>p. ej.</i> , F309N, F309R)
AMV	W25, R76, K110, Q149, T156, M182
RSV	W25, R76, K110, Q149, T156, M182
VIH	W24, R78, G112, Q151, A158, M184

- En algunas realizaciones de la invención, las sustituciones de aminoácidos se hacen en una o más de las posiciones identificadas antes. Por lo tanto, los aminoácidos en estas posiciones pueden ser reemplazados con otros aminoácidos incluyendo Ala, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, y Val. Ejemplos específicos de transcriptasas inversas que presentan aumento de la fidelidad incluyen la transcriptasa inversa de M-MLV en la cual (1) el residuo de valina en la posición 223 ha sido reemplazado con histidina, fenilalanina o isoleucina, (2) el residuo de arginina en la posición 116 ha sido reemplazado con metionina, (3) el residuo de lisina en la posición 152 ha sido reemplazado con arginina, (4) el residuo de ácido glutámico en la posición 190 ha sido reemplazado con fenilalanina, (5) el residuo de treonina en la posición 197 ha sido reemplazado con alanina o ácido glutámico, (6) el residuo de fenilalanina en la posición 309 ha sido reemplazado con asparagina o arginina, (7) el residuo de tirosina en la posición 133 ha sido reemplazado con histidina o alanina, y/o (8) el residuo de tirosina en la posición 64 ha sido reemplazado con triptófano o arginina.

Por lo tanto, en realizaciones específicas, la invención incluye transcriptasas inversas que presentan un aumento en la termoestabilidad y, opcionalmente, presentan también una o más de las siguientes características: (1) actividad de RNasa H reducida o sustancialmente reducida, (2) actividad TdT reducida o sustancialmente reducida, y/o (3) aumento de la fidelidad.

- La invención se refiere también a los mutantes de la transcriptasa inversa, en donde las mutaciones o sustituciones han sido hechas en una región reconocida de la enzima transcriptasa inversa. Tales regiones incluyen, pero no se

limitan a, las regiones de los dedos, palma y pulgar. Por lo tanto, la invención incluye transcriptasas inversas que presentan aumento de la termoestabilidad (así como otras propiedades), como se ha descrito en otro lugar en esta memoria, y tienen una o más (*p. ej.*, una, dos, tres, cuatro, cinco, diez, quince, etc.) mutaciones o modificaciones en el dominio de la mano y, más específicamente, en una o más regiones incluyendo las regiones de los dedos, palma y/o pulgar.

Los polipéptidos que tienen actividad de transcriptasa inversa para uso en la invención se pueden aislar de sus fuentes naturales virales o bacterianas según procedimientos estándar de aislamiento y purificación de proteínas naturales que son bien conocidos por los expertos en la técnica (véase, *p. ej.*, Houts, G.E., et al., J. Virol. 29:517 (1979)). En adición, los polipéptidos que tienen actividad de transcriptasa inversa se pueden preparar por técnicas de ADN recombinante que son familiares para los expertos en la técnica (véase, *p. ej.*, Kotewicz, M.L., et al., Nucl. Acids Res. 16:265 (1988); Soltis, D.A., and Skalka, A.M., Proc. Natl. Acad. Sci. USA 85:3372-3376 (1988)).

En un aspecto de la invención, las transcriptasas inversas mutantes o modificadas se preparan por técnicas recombinantes. Una serie de genes de transcriptasa inversa clonados están disponibles o se pueden obtener utilizando técnicas recombinantes estándar (véase la patente de Estados Unidos Nº 5,668,005 y la publicación PCT Nº WO 98/47912).

Para clonar un gen u otra molécula de ácido nucleico que codifica una transcriptasa inversa que será modificada de acuerdo con la invención, se puede utilizar el ADN aislado que contiene el gen o el marco de lectura abierto de la transcriptasa inversa para construir una genoteca de ADN recombinante. Cualquier vector, bien conocido en la técnica, se puede utilizar para clonar la transcriptasa inversa de interés. Sin embargo, el vector utilizado debe ser compatible con el hospedante en el cual el vector recombinante será transformado.

Los vectores procariotas para construir la genoteca de plásmidos incluyen plásmidos tales como aquellos capaces de replicación en *E. coli* tales como, por ejemplo, los vectores pBR322, ColE1, pSC101, pUC (pUC18, pUC19, etc.: en: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1982); y Sambrook et al., en: Molecular Cloning A Laboratory Manual (2d ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989)). Los plásmidos de *Bacillus* incluyen pC194, pUB110, pE194, pC221, pC217, etc. Tales plásmidos están descritos por Glyczan, T. en: The Molecular Biology Bacilli, Academic Press, York (1982), 307-329. Los plásmidos de *Streptomyces* adecuados incluyen pU101 (Kendall et al., J. Bacteriol. 169:4177-4183 (1987)). Los plásmidos de *Pseudomonas* son revisados por John et al., (Rad. Insec. Dis. 8:693-704 (1986)), e Igaki, (Jpn. J. Bacteriol. 33:729-742 (1978)). Se pueden utilizar también en un amplio intervalo de hospedantes, plásmidos o cósmidos, tal como pCP13 (Darzins and Chakrabarty, J. Bacteriol. 159:9-18 (1984)) para la presente invención. Los vectores preferidos para clonar los genes y las moléculas de ácido nucleico de la presente invención son vectores procariotas. Preferiblemente, los vectores pCP13 y pUC se utilizan para clonar los genes de la presente invención.

Hospedantes adecuados para clonar los genes y moléculas de ácido nucleico de la transcriptasa inversa de interés son los hospedantes procariotas. Un ejemplo de un hospedante procariota es *E. coli*. Sin embargo, los genes y moléculas de ácido nucleico de la transcriptasa inversa deseada de la presente invención pueden ser clonados en otros hospedantes procariotas incluyendo, pero sin limitarse a, hospedantes de los géneros *Escherichia*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia*, y *Proteus*. Los hospedantes bacterianos de particular interés incluyen *E. coli* DH10B, que se puede obtener de Invitrogen Corp. (Carlsbad, CA).

Los hospedantes eucariotas para clonación y expresión de la transcriptasa inversa de interés incluyen células de levaduras, hongos, y de mamífero. La expresión de la transcriptasa inversa deseada en tales células eucariotas puede requerir el uso de regiones reguladoras eucariotas que incluyen promotores eucariotas. La clonación y expresión del gen o molécula de ácido nucleico de la transcriptasa inversa en células eucariotas se pueden llevar a cabo por técnicas bien conocidas utilizando sistemas de vectores eucarióticos bien conocidos.

Una vez que se ha construido la genoteca de ADN en un vector particular, un hospedante apropiado se transforma por técnicas bien conocidas. Las células transformadas se distribuyen en placas a una densidad para producir aproximadamente 200-300 colonias transformadas por placa de Petri. Para selección de la transcriptasa inversa, se criban entonces las colonias en cuanto a la expresión de una transcriptasa inversa termoestable como se describe en los ejemplos más adelante. Brevemente, los cultivos de toda una noche de colonias transformantes individuales se ensayan directamente en cuanto a actividad de transcriptasa inversa termoestable utilizando un desoxinucleótido marcado y se analizan en cuanto a la presencia de producto marcado. Si se detecta actividad de transcriptasa inversa termoestable, el mutante es secuenciado para determinar qué aminoácidos mantienen la actividad de transcriptasa inversa. El gen o molécula de ácido nucleico que codifica una transcriptasa inversa de la presente invención puede ser clonado utilizando métodos bien conocidos por los expertos en la técnica.

Modificaciones o mutaciones de polimerasas

Se pueden hacer una o más mutaciones en cualquier transcriptasa inversa con el fin de aumentar la termoestabilidad de la enzima, o conferir a la enzima otras propiedades descritas en la presente memoria. Dichas mutaciones incluyen mutaciones puntuales, mutaciones de desplazamiento del marco de lectura, delecciones e

inserciones. Preferiblemente, se utilizan una o más mutaciones puntuales, que dan como resultado una o más sustituciones de aminoácidos, para producir transcriptasas inversas que tienen termoestabilidad aumentada o mejorada. En un aspecto preferido de la invención, se pueden hacer una o más mutaciones en posiciones equivalentes o correspondientes a la posición H204 (*p. ej.*, H204R) y/o T306 (*p. ej.*, T306K o T306R) de la transcriptasa inversa de M-MLV para producir el resultado deseado en otras transcriptasas inversas de interés. En otros ejemplos determinados, las transcriptasas inversas pueden no incluir combinaciones específicas de mutaciones, tales como transcriptasas inversas que combinan una mutación de la histidina 204 a arginina (H204R) y una mutación de la treonina 306 a lisina (T306K); transcriptasas inversas que combinan H204R y T306K con una mutación de la fenilalanina 309 a asparagina (F309N); y transcriptasas inversas que combinan H204R, T306K y F309N con una mutación de la valina 223 a histidina (V223H).

En realizaciones específicas, se pueden hacer una o más mutaciones en posiciones equivalentes o correspondientes a la posición L52, Y64, R116, Y133, K152 Q190, T197, H204, V223, M289, T306 y/o F309 de la transcriptasa inversa de M-MLV para producir un resultado deseado (*p. ej.*, aumento de la termoestabilidad, aumento de la fidelidad, reducción de la actividad TdT, etc.). Por lo tanto, en realizaciones específicas, utilizando las posiciones de aminoácidos de la transcriptasa inversa de M-MLV como un marco de referencia, las proteínas descritas incluyen transcriptasas inversas (*p. ej.*, transcriptasa inversa de M-MLV, transcriptasa inversa de AMV, transcriptasa inversa de VIH, transcriptasa inversa de RSV, etc.) que tienen una o más de las siguientes alteraciones: L52P, Y64S, Y64W, Y64R, R116M, Y133A, Y133H, K152R, K152M, Q190F, T197R, T197E, T197A, T197K, H204R, V223H, V223F, V223I, M289L, T306K, T306R, F309R, y/o F309N, con la excepción de aquellas mutaciones o combinaciones de mutaciones específicamente excluidas en ciertas realizaciones como se describe en otro lugar en esta memoria, así como composiciones que contienen estas proteínas, moléculas de ácido nucleico que codifican estas proteínas, y células hospedantes que contienen estas moléculas de ácido nucleico.

Las mutaciones en las transcriptasas inversas que alteran las propiedades de termoestabilidad de estas proteínas pueden estar presentes conjuntamente con alteraciones que tienen poco o ningún efecto sobre las actividades normalmente asociadas con las transcriptasas inversas (*p. ej.*, actividad de RNasa H, actividad de transcriptasa inversa, actividad de desoxinucleotidil transferasa terminal (TdTasa), etc.) o alteran sustancialmente una o más actividades normalmente asociadas con las transcriptasas inversas. Un ejemplo de una transcriptasa inversa que tiene una combinación de mutaciones de este tipo es una transcriptasa inversa de M-MLV que tiene las siguientes alteraciones: K152M, V223H.

Una mutación que se ha demostrado que aumenta la fidelidad de SuperScript™II (Invitrogen Corp. (Carlsbad, CA) Catalog No. 18064-022) es V223H (véase la solicitud de Estados Unidos N° 60/189.454, presentada el 15 de marzo de 2000, y la publicación de Estados Unidos N° 2003/0003452 presentada el 15 de marzo de 2001). Sin embargo, la alteración V223H reduce la termoestabilidad de esta enzima. Se identificó un mutante, K152M, que suprime el efecto desestabilizante de enzimas que tienen la mutación V223H. Por lo tanto, la invención incluye la transcriptasa inversa de M-MLV que contiene alteraciones en las posiciones K152 y V223 y presenta tanto aumento de la fidelidad como aumento de la termoestabilidad. Los ejemplos específicos de tales transcriptasas inversas son aquellos en los que K152 está reemplazado con metionina y V223 está reemplazado con histidina. Se describen también otras transcriptasas inversas (*p. ej.*, transcriptasa inversa de AMV, transcriptasa inversa de VIH, transcriptasa inversa de RSV, etc.) con las correspondientes alteraciones.

SuperScript™II es una RNasa H- transcriptasa inversa de M-MLV que tiene las siguientes sustituciones: D524G, E562Q, y D583N (véase las patentes de Estados Unidos Números 5.017.492, 5.244.797, 5.405.776, 5.668.005, y 6.063.608).

Se hacen una o más sustituciones de aminoácidos en una o más posiciones seleccionadas para cualquier transcriptasa inversa de interés. Por lo tanto, los aminoácidos en las posiciones seleccionadas pueden ser reemplazados con otros aminoácidos incluyendo Ala, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, y Val. En algunas realizaciones preferidas, el aminoácido seleccionado será un residuo de superficie sin carga y será reemplazado por un residuo cargado. En algunas realizaciones preferidas, el residuo de superficie sin carga puede ser reemplazado por un aminoácido cargado positivamente (*p. ej.* lisina o arginina).

Las posiciones correspondientes de la transcriptasa inversa de M-MLV identificadas antes, pueden ser fácilmente identificadas para otras transcriptasas inversas por los expertos en la técnica. Por lo tanto, dada la región definida y los ensayos descritos en la presente solicitud, un experto en la técnica puede hacer una o una serie de modificaciones que darían como resultado un aumento de la termoestabilidad de cualquier transcriptasa inversa de interés. Los residuos a modificar de acuerdo con la presente invención pueden incluir los listados en la Tabla 1 anterior.

Son conocidas las secuencias de nucleótidos para la transcriptasa inversa de M-MLV (Shinnick et al., 1981, Nature 293:543-548; Georgiadis et al., 1995, Structure 3:879-892), transcriptasa inversa de AMV (Joliot et al., 1993, Virology 195:812-819), transcriptasa inversa de RSV (Schwartz et al., 1983, célula 32:853-859), y transcriptasa inversa de VIH (Wong-Staal et al., 1985, Nature 313:277-284).

Preferiblemente, se utiliza la mutagénesis dirigida por oligonucleótidos para crear las transcriptasas inversas mutantes lo que permite todas las clases posibles de cambios de pares de bases en cualquier sitio determinado a lo largo de la molécula de ADN codificadora.

Aumento de la expresión de transcriptasas inversas

Para optimizar la expresión de las transcriptasas inversas de la presente invención, son bien conocidos promotores inducibles o constitutivos y se pueden utilizar para expresar niveles altos de un gen estructural de transcriptasa inversa en un hospedante recombinante. Similarmente, se pueden utilizar vectores con alto número de copias, bien conocidos en la técnica, para alcanzar altos niveles de expresión. Los vectores que tienen un alto número de copias inducibles pueden ser útiles también para aumentar la expresión de las transcriptasas inversas de la invención en un hospedante recombinante.

Para expresar el gen estructural deseado en una célula procariota (tal como, *E. coli*, *B. subtilis*, *Pseudomonas*, etc.), es necesario ligar operativamente el gen estructural deseado a un promotor procariota funcional. Sin embargo, el promotor natural del gen de la transcriptasa inversa puede funcionar en hospedantes procariotas permitiendo la expresión del gen de la transcriptasa inversa. Por lo tanto, el promotor natural u otros promotores se pueden utilizar para expresar el gen de la transcriptasa inversa. Dichos otros promotores que se pueden utilizar para aumentar la expresión incluyen promotores constitutivos o regulables (es decir, inducibles o desrepresibles). Los ejemplos de promotores constitutivos incluyen el promotor *int* del bacteriófago λ , y el promotor *bla* del gen de la β -lactamasa de pBR322. Ejemplos de promotores procariotas inducibles incluyen los promotores principales derecho e izquierdo del bacteriófago λ (P_R y P_L), *trp*, *recA*, *lacZ*, *lacI*, *tet*, *gal*, *trc*, *ara* BAD (Guzman, et al., 1995, J. Bacteriol. 177(14):4121-4130) y los promotores *tac* de *E. coli*. Los promotores de *B. subtilis* incluyen α -amilasa (Ulmanen et al., J. Bacteriol 162:176-182 (1985)) y los promotores del bacteriófago de *Bacillus* (Gryczan, T., en: The Molecular Biology of Bacilli, Academic Press, New York (1982)). Los promotores de *Streptomyces* están descritos por Ward et al., Mol. Gen. Genet. 203:468478 (1986)). Los promotores procariotas están también revisados por Glick, J. Ind. Microbiol. 1:277-282 (1987); Cenatiempo, Y., Biochimie 68:505-516 (1986); y Gottesman, Ann. Rev. gent. 18:415-442 (1984). La expresión en una célula procariota requiere también la presencia de un sitio de unión ribosómico aguas arriba de la secuencia codificadora del gen. Dichos sitios de unión ribosómicos están descritos, por ejemplo, por Gold et al., Ann. Rev. Microbiol. 35:365404 (1981).

Para aumentar la expresión de las polimerasas de la invención en una célula eucariota, se pueden utilizar promotores y hospedantes eucariotas bien conocidos. El aumento de la expresión de las polimerasas se puede conseguir en un hospedante procariota. Un ejemplo de un hospedante procariota apropiado para uso con la presente invención es *Escherichia coli*.

Aislamiento y purificación de transcriptasas inversas

La enzima o enzimas de la presente invención se producen preferiblemente por crecimiento en cultivo del hospedante recombinante que contiene y expresa el gen de la transcriptasa inversa deseada. Sin embargo, la transcriptasa inversa de la presente invención se puede aislar de cualquier cepa que produzca la transcriptasa inversa de la presente invención. Los fragmentos de la transcriptasa inversa están incluidos también en la presente invención. Dichos fragmentos incluyen fragmentos proteolíticos y fragmentos que tienen actividad de transcriptasa inversa.

Se puede añadir al medio de cultivo, cualquier nutriente que pueda ser asimilado por un hospedante que contiene el gen clonado de la transcriptasa inversa. Las condiciones óptimas de cultivo se pueden seleccionar caso a caso de acuerdo con la cepa utilizada y la composición del medio de cultivo. Se pueden añadir también antibióticos al medio de cultivo para asegurar el mantenimiento del ADN vector que contiene el gen deseado a expresar. Las formulaciones de los medios han sido descritas en catálogos DSM o ATCC y Sambrook et al., en: Molecular Cloning, a Laboratory Manual (2nd ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

Las células hospedantes recombinantes que producen las transcriptasas inversas de esta invención se pueden separar del cultivo líquido, por ejemplo, por centrifugación. En general, las células microbianas recogidas se dispersan en un tampón adecuado, y después se rompen forzando por tratamiento ultrasónico o por otros procedimientos bien conocidos para permitir la extracción de las enzimas por la solución tampón. Después de la separación de los restos celulares por ultracentrifugación o centrifugación, se pueden purificar las transcriptasas inversas por técnicas estándar de purificación de proteínas tales como extracción, precipitación, cromatografía, cromatografía de afinidad, electroforesis o similares. Los ensayos para detectar la presencia de la transcriptasa inversa durante la purificación son bien conocidos en la técnica y se pueden utilizar durante los métodos bioquímicos de purificación convencionales, para determinar la presencia de estas enzimas.

En algunas realizaciones, las transcriptasas inversas de la presente invención se pueden modificar para contener una etiqueta de afinidad para facilitar la purificación de la transcriptasa inversa. Las etiquetas de afinidad adecuadas son bien conocidas por los expertos en la técnica e incluyen, pero no se limitan a, secuencias repetidas de aminoácidos tales como seis histidinas, epítomos tales como el epítomo de hemaglutinina y el epítomo myc, y otras secuencias de aminoácidos que permitan la purificación simplificada de la transcriptasa inversa.

Las transcriptasas inversas de la invención tienen preferiblemente actividades específicas mayores que aproximadamente 5 unidades/mg, más preferiblemente mayores que aproximadamente 50 unidades/mg, todavía más preferiblemente mayores que aproximadamente 100 unidades/mg, 250 unidades/mg, 500 unidades/mg, 1000 unidades/mg, 5000 unidades/mg o 10.000 unidades/mg, y lo más preferiblemente mayores que aproximadamente 15.000 unidades/mg, mayores que aproximadamente 16.000 unidades/mg, mayores que aproximadamente 17.000 unidades/mg, mayores que aproximadamente 18.000 unidades/mg, mayores que aproximadamente 19.000 unidades/mg y mayores que aproximadamente 20.000 unidades/mg. En algunas realizaciones, las transcriptasas inversas de la presente invención pueden tener actividades específicas mayores que aproximadamente 50.000 unidades/mg, mayores que aproximadamente 100.000 unidades/mg, mayores que aproximadamente 150.000 unidades/mg, mayores que aproximadamente 200.000 unidades/mg, mayores que aproximadamente 250.000 unidades/mg y mayores que aproximadamente 300.000 unidades/mg. Los intervalos preferidos de actividades específicas para las transcriptasas inversas de la invención incluyen una actividad específica de aproximadamente 5 unidades/mg a aproximadamente 350.000 unidades/mg, una actividad específica de aproximadamente 5 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 50 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 100 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 250 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 500 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 1000 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 5000 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 10.000 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 25.000 unidades/mg a aproximadamente 300.000 unidades/mg, una actividad específica de aproximadamente 100 unidades/mg a aproximadamente 500 unidades/mg, una actividad específica de aproximadamente 100 unidades/mg a aproximadamente 400 unidades/mg, y una actividad específica de aproximadamente 200 unidades/mg a aproximadamente 500 unidades/mg. Otros intervalos preferidos de actividades específicas incluyen una actividad específica de aproximadamente 200.000 unidades/mg a aproximadamente 350.000 unidades/mg, una actividad específica de aproximadamente 225.000 unidades/mg a aproximadamente 300.000 unidades/mg, y una actividad específica de aproximadamente 250.000 unidades/mg a aproximadamente 300.000 unidades/mg. Preferiblemente, el extremo inferior del intervalo de actividad específica puede variar de 50, 100, 200, 300, 400, 500, 700, 900, 1000, 5000, 10.000, 20.000, 30.000, 35.000, 40.000, 45.000, 50.000, 55.000, 60.000, 65.000, 70.000, 75.000, y 80.000 unidades/mg, mientras que el extremo superior del intervalo puede variar de 350.000, 300.000, 250.000, 200.000, 150.000, 140.000, 130.000, 120.000, 110.000, 100.000, y 90.000 unidades/mg. En algunas realizaciones de la presente invención, la actividad específica de la transcriptasa inversa termoestable preparada de acuerdo con la presente invención puede ser más alta que la actividad específica de una transcriptasa inversa no termoestable. En algunas realizaciones, la actividad específica de la transcriptasa inversa termoestable puede ser 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 % o más, más alta que la actividad específica de una transcriptasa inversa no termoestable correspondiente. En algunas realizaciones preferidas, la actividad específica de la transcriptasa inversa termoestable según la presente invención puede ser 30 % o más, más alta que la actividad específica de una transcriptasa inversa no termoestable correspondiente. De acuerdo con la invención, la actividad específica es una medida de la actividad enzimática (en unidades) de la proteína o enzima con respecto a la cantidad total de proteína o enzima utilizada en una reacción. La medida de una actividad específica puede ser determinada por métodos estándar bien conocidos por los expertos en la técnica.

Las transcriptasas inversas de la invención se pueden utilizar para hacer moléculas de ácido nucleico a partir de uno o más moldes. Dichos métodos comprenden mezclar uno o más moldes de ácido nucleico (*p. ej.*, mARN, y más preferiblemente una población de moléculas de mARN) con una o más de las transcriptasas inversas de la invención e incubar la mezcla en condiciones suficientes para preparar una o más moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes de ácido nucleico.

La invención se refiere también a métodos para la amplificación de una o más moléculas de ácido nucleico que comprenden mezclar uno o más moldes de ácido nucleico con una de las transcriptasas inversas de la invención, e incubar la mezcla en condiciones suficientes para amplificar una o más moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes de ácido nucleico. Dichos métodos de amplificación pueden comprender además el uso de una o más ADN-polimerasas y se pueden emplear como en las reacciones de RT-PCR estándar.

La invención se refiere también a métodos para la secuenciación de una o más moléculas de ácido nucleico que comprenden (a) mezclar una o más moléculas de ácido nucleico a secuenciar con una o más moléculas de ácido nucleico cebador, una o más transcriptasas inversas de la invención, uno o más nucleótidos y uno o más agentes de terminación; (b) incubar la mezcla en condiciones suficientes para sintetizar una población de moléculas de ácido nucleico complementarias de todo o parte de la una o más moléculas de ácido nucleico a secuenciar; y (c) separar la población de moléculas de ácido nucleico para determinar la secuencia de nucleótidos de todo o parte de la una o más moléculas de ácido nucleico a secuenciar.

La descripción se refiere también a moléculas de ácido nucleico producidas por tales métodos (que pueden ser moléculas de cADN de longitud completa), vectores (particularmente vectores de expresión) que comprenden estas moléculas de ácido nucleico y células hospedantes que comprenden estos vectores y moléculas de ácido nucleico.

Fuentes de ADN-polimerasa

Una variedad de ADN-polimerasas son útiles de acuerdo con la presente invención. Tales polimerasas incluyen, pero no se limitan a, ADN-polimerasa de *Thermus thermophilus* (*Tth*), ADN-polimerasa de *Thermus aquaticus* (*Taq*), ADN-polimerasa de *Thermotoga neapolitana* (*Tne*), ADN-polimerasa de *Thermotoga maritima* (*Tma*), ADN-polimerasa de *Thermococcus litoralis* (*Tli* o VENT™), ADN-polimerasa de *Pyrococcus furiosus* (*Pfu*), ADN-polimerasa de *Pyrococcus species* GB-D (DEEPVENT™), ADN-polimerasa de *Pyrococcus woosii* (*Pwo*), ADN-polimerasa de *Bacillus sterothophilus* (*Bst*), ADN-polimerasa de *Bacillus caldophilus* (*Bca*), ADN-polimerasa de *Sulfolobus acidocaldarius* (*Sac*), ADN-polimerasa de *Thermoplasma acidophilum* (*Tac*), ADN-polimerasa de *Thermus flavus* (*Tfl/Tub*), ADN-polimerasa de *Thermus ruber* (*Tru*), ADN-polimerasa de *Thermus brockianus* (DYNAZYME™), ADN-polimerasa de *Methanobacterium thermoautotrophicum* (*Mth*), ADN-polimerasa de *Mycobacterium* spp. (*Mtb*, *Mlep*), y mutantes, variantes y derivados de las mismas.

Las ADN-polimerasas utilizadas según la invención pueden ser cualquier enzima que pueda sintetizar una molécula de ADN a partir de un molde de ácido nucleico, típicamente en la dirección 5' a 3'. Dichas polimerasas pueden ser mesófilas o termófilas, pero son preferiblemente termófilas. Las polimerasas mesófilas incluyen ADN-polimerasa T5, ADN-polimerasa T7, ADN-polimerasa fragmento de Klenow, ADN-polimerasa III, y similares. Las ADN-polimerasas preferidas son ADN-polimerasas termoestables tales como *Taq*, *Tne*, *Tma*, *Pfu*, VENT™, DEEPVENT™, *Tth* y mutantes, variantes y derivados de las mismas (Patente de Estados Unidos Nº 5.436.149; Patente de Estados Unidos Nº 5.512.462; Publicación PCT No. WO 92/06188; Publicación PCT No. WO 92/06200; Publicación PCT No. WO 96/10640; Barnes, W.M., Gene 112:29-35 (1992); Lawyer, F.C., et al., PCR Meth. Appl. 2:275-287 (1993); Flaman, J.-M., et al., Nucl. Acids Res. 22(15):3259-3260 (1994)). Para la amplificación de moléculas de ácido nucleico largas (p. ej., moléculas de ácido nucleico más largas que aproximadamente 3-5 Kb de longitud), se utilizan típicamente al menos dos ADN-polimerasas (una que carece sustancialmente de actividad de 3' exonucleasa y la otra que tiene actividad de 3' exonucleasa). Véase la patente de Estados Unidos Nº 5.436.149; patente de Estados Unidos Nº 5.512.462; Barnes, W.M., Gene 112:29-35 (1992); Publicación PCT No. WO 98/06736. Los ejemplos de ADN-polimerasas que carecen sustancialmente de actividad de 3' exonucleasa incluyen, pero no se limitan a, ADN-polimerasas *Taq*, *Tne(exo)*, *Tma*, *Pfu(exo)*, *Pwo* y *Tth*, y mutantes, variantes y derivados de las mismas. Ejemplos no limitantes de ADN-polimerasas que tienen actividad de 3' exonucleasa incluyen *Pfu*, DEEPVENT™ y *Tli/VENT™* y mutantes, variantes y derivados de las mismas.

Formulación de las composiciones de enzima

Para formar las composiciones de la presente invención, se mezclan preferiblemente una o más transcriptasas inversas en una solución salina tamponada. Se pueden añadir opcionalmente una o más ADN-polimerasas y/o uno o más nucleótidos, y/o uno o más cebadores, para preparar las composiciones de la invención. Más preferiblemente, las enzimas se proporcionan en concentraciones de trabajo, en soluciones salinas tamponadas estables. Los términos "estable" y "estabilidad" como se usan en la presente memoria, generalmente significan la retención por una composición, tal como una composición de enzima, de al menos 70 %, preferiblemente al menos 80 %, y lo más preferiblemente al menos 90 %, de la actividad enzimática original (en unidades) después de que la enzima o composición que contiene la enzima haya sido almacenada durante aproximadamente una semana a una temperatura de aproximadamente 4 °C, aproximadamente dos a seis meses a una temperatura de aproximadamente -20 °C, y aproximadamente seis meses o más a una temperatura de aproximadamente -80 °C. Como se usa en la presente memoria, el término "concentración de trabajo" significa la concentración de una enzima que está cerca de la concentración óptima o en la concentración óptima, utilizada en una solución para realizar una función particular (tal como transcripción inversa de ácidos nucleicos).

El agua utilizada para formar las composiciones de la presente invención es preferiblemente destilada, desionizada y esterilizada por filtración (a través de un filtro de 0,1-0,2 micras), y está libre de contaminación por las enzimas DNasa y RNasa. Dicha agua está disponible comercialmente, por ejemplo, de Sigma Chemical Company (Saint Louis, Missouri), o se puede preparar si es necesario según métodos bien conocidos por los expertos en la técnica.

En adición a los componentes enzimáticos, las presentes composiciones comprenden preferiblemente uno o más tampones y cofactores necesarios para la síntesis de una molécula de ácido nucleico tal como una molécula de cADN. Los tampones particularmente preferidos para uso en la formación de las presentes composiciones son el acetato, sulfato, hidrócloruro, fosfato o formas de ácido libre de tris-(hidroximetil)aminometano (TRIS®), aunque se pueden utilizar tampones alternativos de la misma fuerza iónica y pKa aproximadas que el TRIS® con resultados equivalentes. Además de las sales tampones, se incluyen en las composiciones sales cofactores tales como las de potasio (preferiblemente cloruro de potasio o acetato de potasio) y magnesio (preferiblemente cloruro de magnesio o acetato de magnesio). La adición de uno o más carbohidratos y/o azúcares a las composiciones y/o a las mezclas de la reacción de síntesis puede ser también ventajosa para reforzar el aumento de la estabilidad de las composiciones y/o mezclas de reacción durante el almacenaje. Dichos carbohidratos o azúcares preferidos para inclusión en las composiciones y/o mezclas de la reacción de síntesis de la invención incluyen, pero no se limitan a, sacarosa, trehalosa, glicerol, y similares. Además, dichos carbohidratos y/o azúcares se pueden añadir a los tampones de almacenaje para las enzimas utilizadas en la producción de las composiciones y kits enzimáticos de la invención. Dichos carbohidratos y/o azúcares están comercialmente disponibles de una serie de fuentes, incluyendo Sigma (St. Louis, MO).

A menudo es preferible disolver primero las sales tampón, las sales cofactores y los carbohidratos o azúcares a las concentraciones de trabajo en agua y ajustar el pH de la solución antes de la adición de las enzimas. De este modo, las enzimas sensibles al pH estarán menos expuestas a la inactivación mediada por ácidos o compuestos alcalinos durante la formulación de las presentes composiciones.

Para formular la solución de sales tampones, se combina una sal tampón que es preferiblemente una sal de Tris(hidroximetil)aminometano (TRIS®), y lo más preferiblemente su sal hidrocloreto, con una cantidad suficiente de agua para dar una solución que tiene una concentración de TRIS® 5-150 milimolar, preferiblemente 10-60 milimolar, y lo más preferiblemente aproximadamente 20-60 milimolar. A esta solución, se puede añadir una sal de magnesio (preferiblemente el cloruro o la sal acetato del mismo) para proporcionar una concentración de trabajo de la misma 1-10 milimolar, preferiblemente 1,5-8,0 milimolar, y lo más preferiblemente aproximadamente 3-7,5 milimolar. Se puede añadir también a la solución una sal de potasio (preferiblemente una sal cloruro o acetato de potasio), a una concentración de trabajo 10-100 milimolar y lo más preferiblemente aproximadamente 75 milimolar. Se puede añadir a la solución un agente reductor tal como ditiotreitól, preferiblemente en una concentración final de aproximadamente 1-100 mM, más preferiblemente una concentración de aproximadamente 5-50 mM o aproximadamente 7,5-20 mM, y lo más preferiblemente a una concentración de aproximadamente 10 mM. Las concentraciones preferidas de carbohidratos y/o azúcares para inclusión en las composiciones de la invención varían de aproximadamente 5 % (p/v) a aproximadamente 30 % (p/v), aproximadamente 7,5 % (p/v) a aproximadamente 25 % (p/v), aproximadamente 10 % (p/v) a aproximadamente 25 % (p/v), aproximadamente 10 % (p/v) a aproximadamente 20 % (p/v), y preferiblemente aproximadamente 10 % (p/v) a aproximadamente 15 % (p/v). Se puede añadir también una pequeña cantidad de una sal de etilendiaminotetraacetato (EDTA), tal como EDTA disódico (preferiblemente aproximadamente 0,1 milimolar), aunque la inclusión de EDTA no parece ser esencial para la función o estabilidad de las composiciones de la presente invención. Después de la adición de todos los tampones y sales, esta solución salina tamponada se mezcla bien hasta que se disuelven todas las sales, y se ajusta el pH utilizando métodos conocidos en la técnica hasta un valor de pH de 7,4 a 9,2, preferiblemente 8,0 a 9,0, y lo más preferiblemente aproximadamente 8,4.

A estas soluciones salinas tamponadas, se añaden las enzimas (transcriptasas inversas y/o ADN-polimerasas) para producir las composiciones de la presente invención. Las transcriptasas inversas de M-MLV se añaden preferiblemente a una concentración de trabajo en la solución de aproximadamente 1000 a aproximadamente 50.000 unidades por mililitro, de aproximadamente 2000 a aproximadamente 30.000 unidades por mililitro, de aproximadamente 2500 a aproximadamente 25.000 unidades por mililitro, de aproximadamente 3000 a aproximadamente 22.500 unidades por mililitro, de aproximadamente 4000 a aproximadamente 20.000 unidades por mililitro, y lo más preferiblemente a una concentración de trabajo de aproximadamente 5000 a aproximadamente 20.000 unidades por mililitro. Las transcriptasas inversas de AMV, transcriptasas inversas de RSV y transcriptasas inversas de VIH, se añaden preferiblemente a una concentración de trabajo en la solución de aproximadamente 100 a aproximadamente 5000 unidades por mililitro, de aproximadamente 125 a aproximadamente 4000 unidades por mililitro, de aproximadamente 150 a aproximadamente 3000 unidades por mililitro, de aproximadamente 200 a aproximadamente 2500 unidades por mililitro, de aproximadamente 225 a aproximadamente 2000 unidades por mililitro, y lo más preferiblemente a una concentración de trabajo de aproximadamente 250 a aproximadamente 1000 unidades por mililitro. Las enzimas del grupo termófilo de ADN-polimerasa (*Taq*, *Tne*, *Tma*, *Pfu*, VENT, DEEPVENT, *Tth* y mutantes, variantes y derivados de las mismas) se añaden preferiblemente a una concentración de trabajo en la solución de aproximadamente 100 a aproximadamente 1000 unidades por mililitro, de aproximadamente 125 a aproximadamente 750 unidades por mililitro, de aproximadamente 150 a aproximadamente 700 unidades por mililitro, de aproximadamente 200 a aproximadamente 650 unidades por mililitro, de aproximadamente 225 a aproximadamente 550 unidades por mililitro, y lo más preferiblemente a una concentración de trabajo de aproximadamente 250 a aproximadamente 500 unidades por mililitro. Las enzimas se pueden añadir a la solución en cualquier orden, o se pueden añadir simultáneamente.

Las composiciones de la invención pueden comprender además uno o más nucleótidos, que son preferiblemente desoxinucleósidos trifosfatos (dNTPs) o didesoxinucleósidos trifosfatos (ddNTPs). Los componentes dNTP de las presentes composiciones sirven como "bloques de construcción" para los ácidos nucleicos nuevamente sintetizados, que se incorporan allí por la acción de las polimerasas, y los ddNTPs pueden ser utilizados en métodos de secuenciación según la invención. Los ejemplos de nucleótidos adecuados para uso en las presentes composiciones incluyen, pero no se limitan a, dUTP, dATP, dTTP, dCTP, dGTP, dITP, 7-deaza-dGTP, α -tio-dATP, α -tio-dTTP, α -tio-dGTP, α -tio-dCTP, ddUTP, ddATP, ddTTP, ddCTP, ddGTP, ddITP, 7-deaza-ddGTP, α -tio-ddATP, α -tio-ddTTP, α -tio-ddGTP, α -tio-ddCTP o derivados de los mismos, todos los cuales están comercialmente disponibles de fuentes que incluyen Invitrogen Corp. (Carlsbad, CA), New England BioLabs (Beverly, Massachusetts) y Sigma Chemical Company (Saint Louis, Missouri). Los nucleótidos pueden estar sin marcar, o pueden estar marcados detectablemente por acoplamiento de los mismos por métodos conocidos en la técnica con radioisótopos (*p. ej.*, ^3H , ^{14}C , ^{32}P o ^{35}S), vitaminas (*p. ej.*, biotina), restos fluorescentes (*p. ej.*, fluoresceína, rodamina, rojo Texas, o ficoeritrina), marcas quimioluminiscentes (*p. ej.*, utilizando los sistemas de quimioluminiscencia PHOTO-GENE™ o ACES™, disponibles comercialmente de Invitrogen Corp. (Carlsbad, CA)), dioxigenina y similares. Los nucleótidos marcados también se pueden obtener comercialmente, por ejemplo, de Invitrogen Corp. (Carlsbad, CA) o Sigma Chemical Company (Saint Louis, Missouri). En las presentes composiciones, se añaden los nucleótidos para dar una concentración de trabajo de cada nucleótido aproximadamente 10-4000 micromolar, aproximadamente 50-2000

micromolar, aproximadamente 100-1500 micromolar, o aproximadamente 200-1200 micromolar, y lo más preferiblemente una concentración aproximadamente 1000 micromolar.

Para reducir el deterioro de los componentes, el almacenaje de las composiciones reactivas es preferiblemente a aproximadamente 4 °C para un máximo de un día, o lo más preferiblemente a -20 °C para un máximo de un año.

- 5 En otro aspecto, las composiciones y transcriptasas inversas de la invención se pueden preparar y conservar en forma seca en presencia de uno o más carbohidratos, azúcares, o polímeros sintéticos. Los carbohidratos, azúcares o polímeros preferidos para la preparación de composiciones o transcriptasas inversas secas incluyen, pero no se limitan a, sacarosa, trehalosa, y polivinilpirrolidona (PVP) o combinaciones de las mismas. Véase, *p. ej.*, las patentes de Estados Unidos números. 5.098.893, 4.891.319, y 5.556.771. Dichas composiciones y enzimas secas se pueden
- 10 conservar a diferentes temperaturas durante períodos extensos de tiempo sin deterioro significativo de las enzimas o componentes de las composiciones de la invención. Preferiblemente, las transcriptasas inversas o composiciones secas se conservan a 4 °C o a -20 °C.

Producción/fuentes de moléculas de cADN

- 15 De acuerdo con la invención, las moléculas de cADN (de una sola hebra o de doble hebra) se pueden preparar a partir de una variedad de moléculas molde de ácido nucleico. Las moléculas de ácido nucleico para uso en la presente invención incluyen moléculas de ADN y ARN de una sola hebra o de doble hebra, así como híbridos ADN:ARN de doble hebra. Las moléculas más preferidas de ácido nucleico incluyen moléculas de ARN mensajero (mARN), ARN de transferencia (tARN) y ARN ribosómico (rARN), aunque las moléculas de mARN son el molde preferido según la invención.

- 20 Las moléculas de ácido nucleico que se utilizan para preparar moléculas de cADN según los métodos de la presente invención se pueden preparar sintéticamente según los métodos estándar de síntesis de química orgánica que serán familiares para los expertos en la técnica. Más preferiblemente, las moléculas de ácido nucleico se pueden obtener de fuentes naturales, tales como una variedad de células, tejidos, órganos u organismos. Las células que se pueden utilizar como fuentes de moléculas de ácido nucleico pueden ser procariotas (células bacterianas, incluyendo pero
- 25 sin limitarse a aquellas de especies de los géneros *Escherichia*, *Bacillus*, *Serratia*, *Salmonella*, *Staphylococcus*, *Streptococcus*, *Clostridium*, *Chlamydia*, *Neisseria*, *Treponema*, *Mycoplasma*, *Borrelia*, *Legionella*, *Pseudomonas*, *Mycobacterium*, *Helicobacter*, *Erwinia*, *Agrobacterium*, *Rhizobium*, *Xanthomonas* y *Streptomyces*) o eucariotas (incluyendo hongos (especialmente levaduras), plantas, protozoos y otros parásitos, y animales incluyendo insectos (particularmente células de *Drosophila* spp), nematodos (particularmente células de *Caenorhabditis elegans*), y
- 30 mamíferos (particularmente células humanas)).

- Las células somáticas de mamíferos que se pueden utilizar como fuentes de ácidos nucleicos incluyen células sanguíneas (reticulocitos y leucocitos), células endoteliales, células epiteliales, células neuronales (de los sistemas nerviosos central o periférico), células musculares (incluyendo miocitos y mioblastos del músculo esquelético, liso o cardíaco), células del tejido conjuntivo (incluyendo fibroblastos, adipocitos, condrocitos, condroblastos, osteocitos y
- 35 osteoblastos) y otras células estromales (*p. ej.*, macrófagos, células dendríticas, células de Schwann). Las células germinales de los mamíferos (espermatoцитos y oocitos) se pueden utilizar también como fuentes de ácido nucleicos para uso en la invención, como se pueden utilizar también las células madre progenitoras, precursoras y no humanas que dan lugar a las anteriores células somáticas y germinales. También son adecuadas para uso como fuentes de ácido nucleico los tejidos u órganos de mamíferos tales como los derivados del cerebro, riñones, hígado,
- 40 páncreas, sangre, médula ósea, músculo, nervios, piel, fuentes de tejido genitourinario, circulatorio, linfoide, gastrointestinal y conjuntivo, así como las derivadas de un embrión o feto de mamífero (no humano).

- Cualquiera de las células, tejidos y órganos procariotas o eucariotas anteriores, pueden ser normales, enfermas, transformadas, establecidas, progenitoras, precursoras, fetales o embrionarias no humanas. Las células enfermas pueden incluir, por ejemplo, aquellas implicadas en enfermedades infecciosas (causadas por bacterias, hongos o
- 45 levaduras, virus (incluyendo SIDA, VIH, HTLV, herpes, hepatitis y similares) o parásitos), en patologías genéticas o bioquímicas (*p. ej.*, fibrosis quística, hemofilia, enfermedad de Alzheimer, distrofia muscular o esclerosis múltiple) o en procesos cancerosos. Las líneas de células animales transformadas o establecidas pueden incluir, por ejemplo, células COS, células CHO, células VERO, células BHK, células HeLa, células HepG2, células K562, células 293, células L929, células F9, y similares. Otras células, líneas celulares, tejidos, órganos y organismos adecuados como
- 50 fuentes de ácido nucleicos para uso en la presente invención serán evidentes para los expertos en la técnica.

- Una vez que se obtienen las células, tejidos, órganos u otras muestras de partida, se pueden aislar de ellas las moléculas de ácido nucleico (tal como mARN) por métodos que son bien conocidos en la técnica (véase, *p. ej.*, Maniatis, T., et al., *Cell* 15:687-701 (1978); Okayama, H., and Berg, P., *Mol. Cell. Biol.* 2:161-170 (1982); Gubler, U., and Hoffman, B.J., *Gene* 25:263-269 (1983)). Las moléculas de ácido nucleico aisladas de este modo, se pueden
- 55 utilizar entonces para preparar moléculas de cADN y genotecas de cADN.

Las moléculas de cADN o las genotecas de cADN se producen mezclando una o más moléculas de ácido nucleico obtenidas como se ha descrito antes, que es preferiblemente una o más moléculas de mARN tal como una población de moléculas de mARN, con un polipéptido que tiene actividad de transcriptasa inversa de la presente

invención, o con una o más de las composiciones de la invención, en condiciones que favorecen la transcripción inversa de la molécula de ácido nucleico por la acción de las enzimas o las composiciones para formar una o más moléculas de cADN (de una sola hebra o de doble hebra). Por lo tanto, el método de la invención comprende (a) mezclar uno o más moldes de ácido nucleico (preferiblemente uno o más moldes de ARN o mARN, tal como una población de moléculas de mARN) con una o más transcriptasas inversas de la invención y (b) incubar la mezcla en condiciones suficientes para formar una o más moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes. Dichos métodos pueden incluir el uso de una o más ADN-polimerasas, uno o más nucleótidos, uno o más cebadores, uno o más tampones, y similares. La invención se puede utilizar conjuntamente con métodos de síntesis de cADN tales como los descritos en los ejemplos más adelante, u otros que son bien conocidos en la técnica (véase, *p. ej.*, Gubler, U., and Hoffman, B.J., *Gene* 25:263-269 (1983); Krug, M.S., and Berger, S.L., *Meth. Enzymol.* 152:316-325 (1987); Sambrook, J., et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 8.60-8.63 (1989); Publicación PCT No. WO 99/15702; Publicación PCT No. WO 98/47912; y Publicación PCT No. WO 98/51699), para producir moléculas o genotecas de cADN.

Otros métodos de síntesis de cADN que se pueden utilizar ventajosamente en la presente invención serán fácilmente evidentes para los expertos en la técnica

Habiendo obtenido las moléculas o genotecas de cADN según los presentes métodos, se pueden aislar estos cADNs para posterior análisis o manipulación. Metodologías detalladas para la purificación de los cADN se enseñan en el manual GENETRAPPER™ (Invitrogen Corp. (Carlsbad, CA)), aunque se pueden utilizar también métodos estándar alternativos de aislamiento de cADN que son conocidos en la técnica (véase, *p. ej.*, Sambrook, J., et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 8.60-8.63 (1989)).

En otros aspectos de la invención, la invención puede ser utilizada en métodos para amplificar y secuenciar moléculas de ácido nucleico. Los métodos de amplificación de ácido nucleico según este aspecto de la invención pueden ser reacciones de una etapa (*p. ej.*, RT-PCR de una etapa) o de dos etapas (*p. ej.*, RT-PCR de dos etapas). Según la invención, las reacciones tipo RT-PCR de una etapa se pueden realizar en un tubo con lo que se reduce la posibilidad de contaminación. Dichas reacciones de una etapa comprenden (a) mezclar un molde de ácido nucleico (*p. ej.*, mARN) con una o más transcriptasas inversas de la presente invención y con una o más ADN-polimerasas y (b) incubar la mezcla en condiciones suficientes para amplificar una molécula de ácido nucleico complementaria de todo o parte del molde. Dicha amplificación se puede llevar a cabo por la actividad de la transcriptasa inversa sola o en combinación con la actividad de la ADN-polimerasa. Las reacciones de RT-PCR de dos etapas se pueden llevar a cabo en dos etapas separadas. Dicho método comprende (a) mezclar un molde de ácido nucleico (*p. ej.*, mARN) con una transcriptasa inversa de la presente invención, (b) incubar la mezcla en condiciones suficientes para formar una molécula de ácido nucleico (*p. ej.*, una molécula de ADN) complementaria de todo o parte del molde, (c) mezclar la molécula de ácido nucleico con una o más ADN-polimerasas y (d) incubar la mezcla de la etapa (c) en condiciones suficientes para amplificar la molécula de ácido nucleico. Para amplificación de moléculas de ácido nucleico largas (es decir, mayores que aproximadamente 3-5 Kb de longitud), se puede utilizar una combinación de ADN-polimerasas, tales como una ADN-polimerasa que tiene actividad de 3' exonucleasa y otra ADN-polimerasa que tiene actividad de 3' exonucleasa sustancialmente reducida.

Los métodos de secuenciación de ácido nucleico según este aspecto de la invención pueden comprender tanto la secuenciación en ciclo (secuenciación en combinación con amplificación) como las reacciones de secuenciación estándar. El método de secuenciación de la invención comprende por tanto (a) mezclar una molécula de ácido nucleico a secuenciar con uno o más cebadores, una o más transcriptasas inversas de la invención, uno o más nucleótidos y uno o más agentes de terminación, (b) incubar la mezcla en condiciones suficientes para sintetizar una población de moléculas de ácido nucleico complementarias de todo o parte de la molécula a secuenciar, y (c) separar la población para determinar la secuencia de nucleótidos de toda o parte de la molécula a secuenciar. Según la invención, se pueden utilizar una o más ADN-polimerasas (preferiblemente ADN-polimerasas termoestables) en combinación con las transcriptasas inversas de la invención o separadas de ellas.

Los métodos de amplificación que se pueden utilizar de acuerdo con la presente invención incluyen PCR (Patentes de Estados Unidos Números 4.683.195 y 4.683.202), amplificación por desplazamiento de hebra (SDA; Patente de Estados Unidos N° 5.455.166; documento EP 0 684 315), y amplificación basada en la secuencia de ácido nucleico (NASBA; Patente de Estados Unidos N° 5.409.818; documento EP 0 329 822), así como técnicas más complejas de huellas de ácido nucleico basadas en la PCR tales como análisis de ADN polimórfico amplificado aleatoriamente (RAPD) (Williams, J.G.K., et al., *Nucl. Acids Res.* 18(22):6531-6535, 1990), PCR con cebadores arbitrarios (AP-PCR; Welsh, J., and McClelland, M., *Nucl. Acids Res.* 18(24):7213-7218, 1990), huellas de amplificación de ADN (DAF; Caetano-Anollés et al., *Bio/Technology* 9:553-557, 1991), PCR de microsatélites o amplificación dirigida de la región minisatélite de ADN (DAMD; Heath, D.D., et al., *Nucl. Acids Res.* 21(24): 5782-5785, 1993), y análisis de polimorfismos en la longitud de fragmentos amplificados (AFLP) (documento EP 0 534 858; Vos, P., et al., *Nucl. Acids Res.* 23(21):4407-4414, 1995; Lin, J.J., and Kuo, J., *FOCUS* 17(2):66-70, 1995). Las técnicas de secuenciación de ácido nucleico que pueden emplear las presentes composiciones incluyen métodos de secuenciación de dideoxi tales como los descritos en las patentes de Estados Unidos Números 4.962.022 y 5.498.523. En un aspecto particularmente preferido, la invención se puede utilizar en métodos de amplificación o

secuenciación de una molécula de ácido nucleico que comprenden una o más reacciones en cadena de la polimerasa (PCRs), tales como cualquiera de los métodos basados en la PCR descritos antes.

Kits

En otra realización, la presente invención se puede montar en kits para uso en la transcripción inversa o amplificación de una molécula de ácido nucleico, o en kits para uso en la secuenciación de una molécula de ácido nucleico. Los kits según este aspecto de la invención comprenden un soporte, tal como una caja, cartón, tubo o similar, que tiene en estrecho confinamiento en el mismo uno o más recipientes, tales como viales, tubos, ampollas, frascos y similares, en donde un primer recipiente contiene uno o más polipéptidos de la presente invención que tienen actividad de transcriptasa inversa. Cuando se utiliza más de un polipéptido que tiene actividad de transcriptasa inversa, estos pueden estar en un recipiente único como mezclas de dos o más polipéptidos, o en recipientes separados. Los kits de la invención pueden comprender también (en el mismo recipiente o en recipientes separados) una o más ADN-polimerasas, un tampón adecuado, uno o más nucleótidos y/o uno o más cebadores.

En un aspecto específico de la invención, los kits de transcripción inversa y amplificación pueden comprender uno o más componentes (en mezclas o por separado) incluyendo uno o más polipéptidos que tienen actividad de transcriptasa inversa de la invención, uno o más nucleótidos necesarios para la síntesis de una molécula de ácido nucleico, y/o uno o más cebadores (p. ej., oligo(dT) para transcripción inversa). Dichos kits de transcripción inversa y amplificación pueden comprender además una o más ADN-polimerasas. Los kits de secuenciación de la invención pueden comprender uno o más polipéptidos que tienen actividad de transcriptasa inversa de la invención, y opcionalmente una o más ADN-polimerasas, uno o más agentes de terminación (p. ej., moléculas de didesoxinucleósido trifosfato) necesarios para la secuenciación de una molécula de ácido nucleico, uno o más nucleótidos y/o uno o más cebadores. Los polipéptidos preferidos que tienen actividad de transcriptasa inversa, ADN-polimerasas, nucleótidos, cebadores y otros componentes adecuados para uso en los kits de transcripción inversa, amplificación y secuenciación de la invención incluyen los descritos anteriormente. Los kits englobados por este aspecto de la presente invención pueden comprender además reactivos y compuestos adicionales necesarios para llevar a cabo los protocolos estándar de transcripción inversa, amplificación o secuenciación de ácido nucleico. Tales polipéptidos que tienen actividad de transcriptasa inversa de la invención, ADN-polimerasas, nucleótidos, cebadores, y reactivos, componentes o compuestos adicionales, pueden estar contenidos en uno o más recipientes, y pueden estar contenidos en dichos recipientes en una mezcla de dos o más de los componentes indicados antes o pueden estar contenidos en los kits de la invención en recipientes separados.

Uso de moléculas de ácido nucleico

Las moléculas de ácido nucleico o las genotecas de cADN preparadas por los métodos de la presente invención se pueden caracterizar además, por ejemplo por clonación y secuenciación (es decir, determinación de la secuencia de nucleótidos de la molécula de ácido nucleico), por los métodos de secuenciación de la invención o por otros que son estándar en la técnica (véase, p. ej., las patentes de Estados Unidos números 4.962.022 y 5.498.523, que se dirigen a los métodos de secuenciación de ADN). Alternativamente, estas moléculas de ácido nucleico se pueden utilizar para la fabricación de diferentes materiales en procesos industriales, tales como sondas de hibridación por métodos que son bien conocidos en la técnica. La producción de sondas de hibridación a partir de cADNs proporcionará, por ejemplo, a los que trabajan en el campo médico la capacidad para examinar las células o tejidos de un paciente en cuanto a la presencia de un marcador genético particular tal como un marcador de cáncer, de una enfermedad infecciosa o genética, o un marcador de desarrollo embrionario. Además, tales sondas de hibridación se pueden utilizar para aislar fragmentos de ADN de las genotecas de ADN o cADN preparadas a partir de una célula, tejido o organismo diferente, para posterior caracterización.

Las moléculas de ácido nucleico de la presente descripción se pueden utilizar también para preparar composiciones para uso en metodologías de ADN recombinante. Por consiguiente, la presente descripción se refiere a vectores recombinantes que comprenden las moléculas de cADN o moléculas de ácido nucleico amplificadas, a células hospedantes que se tratan genéticamente con los vectores recombinantes, a métodos para la producción de un polipéptido recombinante utilizando estos vectores y células hospedantes, y a polipéptidos recombinantes producidos utilizando estos métodos.

Los vectores recombinantes se pueden producir según este aspecto mediante la inserción, utilizando métodos que son bien conocidos en la técnica, de una o más de las moléculas de cADN o moléculas amplificadas de ácido nucleico preparadas según los presentes métodos en un vector. El vector utilizado en este aspecto puede ser, por ejemplo, un fago o un plásmido, y es preferiblemente un plásmido. Son preferidos los vectores que comprenden regiones control que actúan en cis con el ácido nucleico que codifica el polipéptido de interés. Los factores apropiados que actúan en trans pueden ser suministrados por el hospedante, suministrados por un vector complementario o suministrados por el propio vector después de introducción en el hospedante.

En ciertas realizaciones preferidas a este respecto, los vectores proporcionan la expresión específica (y se denominan por tanto "vectores de expresión"), que puede ser inducible y/o específica del tipo de célula. Son particularmente preferidos entre tales vectores aquellos inducibles por factores ambientales que son fáciles de manipular, tales como la temperatura y los aditivos nutrientes.

Los vectores de expresión útiles en la presente invención incluyen vectores derivados de cromosomas, episomas y virus, *p.* ej., vectores derivados de plásmidos o bacteriófagos bacterianos, y vectores derivados de combinaciones de los mismos, tales como cósmidos y fagémidos, e incluirán preferiblemente al menos un marcador seleccionable tal como un gen de resistencia a la tetraciclina o ampicilina para cultivar en una célula hospedante bacteriana. Antes de la inserción en un vector de expresión de este tipo, las moléculas de cADN o moléculas amplificadas de ácido nucleico de la invención deberán ser ligadas operativamente con un promotor apropiado, tal como el promotor fago lambda P_L, los promotores *lac*, *trp* y *tac* de *E. coli*. Otros promotores apropiados serán conocidos por los expertos.

Entre los vectores preferidos para uso en la presente invención se incluyen pQE70, pQE60 y pQE-9, disponibles de Qiagen; vectores pBS, vectores Phagescript, vectores Bluescript, pNH8A, pNH16a, pNH18A, pNH46A, disponibles de Stratagene; pcADN3 disponible de Invitrogen; pGEX, pTrxfus, pTrc99a, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 disponible de Pharmacia; y pSPORT1, pSPORT2 y pSV-SPORT1, disponible de Invitrogen Corp. (Carlsbad, CA). Otros vectores adecuados serán fácilmente evidentes para los expertos.

La invención proporciona también métodos para producir una célula hospedante recombinante que comprenden las moléculas de cADN, moléculas amplificadas de ácido nucleico o vectores recombinantes de la invención, así como las células hospedantes producidas por tales métodos. Las células hospedantes representativas (procariotas o eucariotas) que se pueden producir según la invención incluyen, pero no se limitan a, células bacterianas, células de levaduras, células de plantas y células de animales. Las células hospedantes bacterianas preferidas incluyen células de *Escherichia coli* (muy particularmente cepas DH10B y Stbl2 de *E. coli*, que están disponibles comercialmente (Invitrogen Corp. (Carlsbad, CA)), células de *Bacillus subtilis*, células de *Bacillus megaterium*, células de *Streptomyces* spp., células de *Erwinia* spp., células de *Klebsiella* spp. y células de *Salmonella typhimurium*. Las células hospedantes animales preferidas incluyen células de insectos (muy particularmente células de *Spodoptera frugiperda* Sf9 y Sf21 y células de *Trichoplusia* High-Five) y células de mamífero (muy particularmente células CHO, COS, VERO, BHK y humanas). Dichas células hospedantes se pueden preparar por técnicas bien conocidas de transformación, electroporación o transfección que serán familiares para los expertos en la técnica.

En adición, la invención proporciona métodos para producir un polipéptido recombinante, y los polipéptidos producidos por estos métodos. Según este aspecto de la invención, se puede producir un polipéptido recombinante cultivando cualquiera de las células hospedantes recombinantes anteriores en condiciones que favorecen la producción de un polipéptido desde ellas, y el aislamiento del polipéptido. Los métodos para cultivar las células hospedantes recombinantes, y para la producción y aislamiento de polipéptidos desde ellas, son bien conocidos por los expertos en la técnica.

Será evidente para un experto en las técnicas relevantes, que son obvias otras modificaciones y adaptaciones adecuadas a los métodos y aplicaciones descritos en la presente memoria y se pueden hacer sin separarse del alcance de la invención o de cualquier realización de la misma. Habiendo descrito ahora la presente invención en detalle, se entenderá más claramente la misma en referencia a los siguientes ejemplos, que se incluyen aquí con fines de ilustración solamente y no se pretende que sean limitantes de la invención.

Ejemplos

Ejemplo 1: Preparación de transcriptasas inversas mutantes

Se obtuvo el plásmido pBAD de Invitrogen y se insertó la secuencia codificadora de la transcriptasa inversa de M-MLV para producir el plásmido pBAD-6-His-M-MLV H-(F1). El plásmido pBAD-6-His-M-MLV H-(F1) se utilizó de ambos modos, como un vector de clonación y como un objetivo para la mutagénesis por PCR (Figura 1). El pBAD-6-His-M-MLV H-(F1) se replica en *E. coli* y confiere resistencia a la ampicilina a las células transformadas. El gen de la transcriptasa inversa de M-MLV se expresa a partir del promotor *ara* BAD que es inducido por la presencia de arabinosa. El promotor es reprimido por el producto del gen *araC*, que está presente en el plásmido. El hospedante utilizado, *E. coli* cepa DH10B, es un *araD* mutante y no puede metabolizar la arabinosa, haciendo que la arabinosa sea un inductor gratuito en las células DH10B transformadas con pBAD-6-His-M-MLV H-(F1). El plásmido contiene una 6 histidina que contiene la secuencia líder en el marco con la secuencia codificadora del gen de la transcriptasa inversa de M-MLV. Con referencia a la secuencia de este plásmido, proporcionada en la Tabla 3 (SEQ ID NOs:1 y 2), los nucleótidos 1-96 codifican la secuencia líder y los nucleótidos 97-99 codifican una metionina. Los expertos en la técnica apreciarán que la transcriptasa inversa de M-MLV natural se deriva por proteólisis a partir de una poliproteína precursora y por lo tanto la transcriptasa inversa de M-MLV natural no empieza con una metionina. Por lo tanto, el aminoácido número 1 de la transcriptasa inversa de M-MLV es la treonina después de la metionina codificada por los nucleótidos 97-99.

La secuencia del gen de la transcriptasa inversa de M-MLV en pBAD-6-His-M-MLV H-(F1) que se utilizó en estos experimentos fue derivada de la secuencia de plásmido pRT601. El pRT601 se describe en las patentes de Estados Unidos números 5.668.005 y 5.017.492;

Tabla 3.

```

1   atgggggggtt ctcacatca tcacatcat ggtatggcta gcatgactgg tggacagcaa
    m g g s h h h h h h g m a s m t g g q q

61  atgggtcggg atctgtacga cgatgacgat aagcatatga ccctaaatat agaagatgag
    m g r d l y d d d d k h m t l n i e d e

121 tatcggctac atgagacctc aaaagagcca gatgtttctc taggggtccac atgggtgtct
    y r l h e t s k e p d v s l g s t w l s

181 gattttcttc aggcctgggc ggaaaccggg ggcatgggac tggcagttcg ccaagctcct
    d f p q a w a e t g g m g l a v r q a p

241 ctgatcatatc ttctgaaagc aacctctacc cccgtgtcca taaaacaata ccccatgtca
    l i i l l k a t s t p v s i k q y p m s

301 caagaagcca gactggggat caagcccccac atacagagac tgttggacca ggaataactg
    q e a r l g i k p h i q r l l d q g i l

361 gtaccttgcc agtccccctg gaacacgccc ctgctaccgg tcaagaaacc cgggactaat
    v p c q s p w n t p l l p v k k p g t n

421 gattacaggc ctgtccaaga tctgagagag gtcaacaaac gcgtagaaga catccacccc
    d y r p v q d l r e v n k r v e d i h p

481 accgtaccca acccctacaa cctcttgagt gggctcccac cgtcccacca gtggtacact
    t v p n p y n l l s g l p p s h q w y t

541 gttctagact taaaagatgc ctttttctgc ctgagactcc acccgacgtc tcagcctctc
    v l d l k d a f f c l r l h p t s q p l

601 ttgccttttg aatggagaga cccagagatg ggaatctctg gccaaactaac ctggaccaga
    f a f e w r d p e m g i s g q l t w t r

661 ctcccacagg gattcaaaaa cagtccacc ctgtttgatg aggcactgag cagagaccta
    l p q g f k n s p t l f d e a l r r d l

721 gcgacttcc ggatccagca cccagacttg atcctgctac agtacgtaga tgacttactg
    a d f r i q h p d l i l l q y v d d l l

781 ctggccgcca cttctgagct cgactgccaa caaggctactc gggccctgtt acaaacctta
    l a a t s e l d c q q g t r a l l q t l

841 ggagacctcg ggtatcgggc ctgcggccaag aaagcccaa tttgccagaa acaggtcaag
    g d l g y r a s a k k a q i c q k q v k

901 tatctggggt atcttctaaa agagggtcag agatggctga ctgaggccag aaaagagact
    y l g y l l k e g q r w l t e a r k e t

961 gtgatggggc agcctactcc gaagaccccg cggcaactaa gggagttcct agggacggca
    v m g q p t p k t p r q l r e f l g t a

1021 ggcttctgtc gcctctggat ccctgggttt gcagaaatgg cagccccctt gtaccctctc
    g f c r l w i p g f a e m a a p l y p l

1081 accaaaacgg ggactctgtt taattggggc ccagaccaac aaaaggccta tcaagaaatc
    t k t g t l f n w g p d q q k a y q e i

1141 aagcaagctc ttctaactgc cccagccctg gggttgccag atttgactaa gccctttgaa
    k q a l l t a p a l g l p d l t k p f e

1201 ctctttgtcg acgagaagca gggctacgcc aaagggtgtc taacgcaaaa actgggacct
    l f v d e k q g y a k g v l t q k l g p

```

```

1261 tggcgtcggc cggtggccta cctgtccaaa aagctagacc cagtagcagc tgggtggccc
    w r r p v a y l s k k l d p v a a g w p
1321 ccttgccctac ggatggtagc agccattgcc gtactgacaa aggatgcagg caagctaacc
    p c l r m v a a i a v l t k d a g k l t
1381 atgggacagc cactagtcac tctggccccc catgcagtag aggcactagt caaacaaccc
    m g q p l v i l a p h a v e a l v k q p
1441 cccgatcgat ggcttttcaa cgcccgcatg actcactatc aggccttgct tttggacacg
    p d r w l s n a r m t h y q a l l l d t
1501 gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg ctacactgct cccactgcct
    d r v q f g p v v a l n p a t l l p l p
1561 gaggaagggc tgcagcacia ctgccttgat atcctggccg aagcccacgg aaccgaccc
    e e g l q h n c l d i l a e a h g t r p
1621 gacctaacgg accagccgct cccagacgcc gaccacacct ggtacacggg tggatccagt
    d l t d q p l p d a d h t w y t g g s s
1681 ctcttgcaag agggacagcg taaggcggga gctgcggtga ccaccgagac cgaggtaatc
    l l q e g q r k a g a a v t t e t e v i
1741 tgggctaaag ccctgccagc cgggacatcc gctcagcggg ctcagctgat agcactcacc
    w a k a l p a g t s a q r a q l i a l t
1801 caggccctaa ggatggcaga aggtaagaag ctaaatgttt atacgaattc ccgttatgct
    q a l r m a e g k k l n v y t n s r y a
1861 tttgctactg cccatatcca tggagaaata tacagaaggc gtggggttgct cacatcagaa
    f a t a h i h g e i y r r r g l l t s e
1921 ggcaaagaga tcaaaaataa ggacgagata ttggccctac taaaagccct ctttctgccc
    g k e i k n k d e i l a l l k a l f l p
1981 aaaagactta gcataatcca ttgtccagga catcaaaagg gacacagcgc cgaggctaga
    k r l s i i h c p g h q k g h s a e a r
2041 ggcaaccgga tggctgacca agcggcccga aaggcagcca tcacagagaa tccagacacc
    g n r m a d q a a r k a a i t e n p d t
2101 tctaccctcc tcatagaaaa ttcacacccc aattcccgct taattaatta a
    s t l l i e n s s p n s r l i n -

```

La Tabla 4 proporciona una lista de las mutaciones puntuales introducidas en la secuencia de pRT601 codificadora de la transcriptasa inversa de M-MLV utilizada para producir el plásmido. La numeración de las mutaciones puntuales corresponde a la secuencia de nucleótidos presentada en la Tabla 3.

Tabla 4

Nucleótido # en la Tabla 3	cambio	Nucleótido # en la Tabla 3	cambio
411	a→c	993	a→g
459	g→a	1446	c→t
462	g→c	1449	c→a
543	g→t	1670	a→g
546	t→a	1675	a→t
585	c→g	1676	g→c
588	c→g	1783	g→c
589	a→t	1785	a→g
590	g→c	1845	t→g
639	a→t	1846	g→a

Nucleótido # en la Tabla 3	cambio	Nucleótido # en la Tabla 3	cambio
642	a→c	1849	a→t
710	a→g	1850	g→c
801	a→c	1950	c→a
990	t→g		

Las mutaciones que se introdujeron para preparar los mutantes RNasa H de la transcriptasa inversa de M-MLV son D524G, D583N, y E562Q. Las restantes mutaciones se introdujeron para insertar o eliminar los sitios de la enzima de restricción para facilitar la producción de segmentos de tamaño apropiado para la mutagénesis por PCR aleatoria. Este mutante de RNasa H se denomina en la presente memoria SuperScript™II o gen SuperScript™II.

La secuencia de la transcriptasa inversa de M-MLV se modificó por ingeniería genética para introducir sitios de escisión de la enzima de restricción como se muestra esquemáticamente en la Figura 2 sin cambiar los aminoácidos codificados por la secuencia. La secuencia se dividió en 5 segmentos y los oligonucleótidos fueron diseñados de modo que cada segmento pudiera ser amplificado.

Se prepararon segmentos a partir de pBAD-6-His-M-MLV H- (F1) mediante digestiones por la enzima de restricción y los segmentos se purificaron en gel fuera de la cadena principal del vector. Cada segmento se sometió a mutagénesis aleatoria por PCR en presencia de manganeso. Las condiciones de la PCR fueron estándar excepto que estaba presente MnCl₂ 0,25 mM, y la concentración de nucleótido trifosfato se limitó a 20 µM de cada dNTP (Tris-HCl 50 mM pH 8,3, KCl 50 mM, MgCl₂ 3 mM, dGTP 20 µM, dCTP 20 µM, dATP 20 µM, dTTP 20 µM, 1 unidad de ADN-polimerasa Taq por 100 µl de reacción). El producto de la PCR se extrajo con fenol-cloroformo, se precipitó con etanol y los segmentos mutados se clonaron en un vector del que se había separado el segmento dado. Las genotecas de transformantes para cada segmento mutado se cribaron en cuanto a variantes termoestables.

Ejemplo 2: Cribado de transcriptasas inversas termoestables

En este ejemplo se utilizaron las siguientes soluciones:

EG-por litro: 20 g de bacto-triptona, 10 g de bacto extracto de levadura, 2 ml de glicerol, 0,54 g de NaCl, 0,194 g de KCl

EG-arabinosa -150 ml de EG más 1,5 ml de ampicilina a 10 mg/ml y 1,5 ml de arabinosa al 20 % (p/v) (si las placas van a tener arabinosa)

Tampón 20X PEB-I - glucosa al 18 % (p/v), Tris-HCl 500 mM (pH 8.0), EDTA 200 mM.

Tampón de almacenamiento de quinasa - glicerol al 50 % (v/v), Tris-HCl 20 mM (pH 8.0), KCl 100 mM, βME 5 mM

Lisozima a 100 mg/ml - preparada en tampón de almacenamiento de quinasa y conservada a -20 °C

2X PLD - 5 ml de 20X PEB-I, 1 ml de DTT 1 M, 5 ml de Triton™ X-100 al 10 % v/v, 1 ml de lisozima a 100 mg/ml y 38 ml de agua

2X PZD - 0,5 ml de 20X PEB-I, 100 µl de DTT 1 M, 0,5 ml de Triton™ X-100 al 10 % (v/v), 10 µl de zimolasa y 3,9 ml de agua

Tampón de reacción 10X poli(C) -Tris-HCl 500 mM (pH 8.4), KCl 500 mM, MgCl₂ 100 mM

Mezcla de reacción 1,25X -1 ml de tampón de reacción 10X poli(C), 100 µl de DTT 1 M, 1 ml de poli(C)/oligo(dG) (30 mM/12 mM en nucleótido), 10 µl de dGTP 100 mM, 5,87 ml de agua y 20 µl de [α -³²P] dGTP a 10 µCi/µl

Se inocularon colonias transformantes individuales en los pocillos individuales de una placa de cultivo de 96 pocillos. Cada pocillo contenía 120 µl de medio EG-Ap (medio EG con 100 µg/ml de ampicilina) que contenía 0,2 % de arabinosa. Es preferible inocular en primer lugar una placa de 96 pocillos con medio selectivo sin el inductor, para dejar crecer dicha placa maestra durante la noche, y después hacer una réplica de la placa maestra en una placa de 96 pocillos con el inductor, y dejar crecer dicha placa durante la noche. Se dejaron crecer los cultivos durante la noche (*p. ej.*, 15-20 horas) a 37 °C sin agitación. Los cultivos de la noche se mezclaron con un volumen igual de 2X PLD a temperatura ambiente. Estos extractos se ensayaron algunas veces directamente para determinar la transcriptasa inversa antes de la etapa de calentamiento. Se calentaron los extractos poniéndolos en un baño de agua durante 5 o 10 minutos a temperaturas que variaron de 50 °C a 60 °C. Preferiblemente, se calentaron los cultivos durante 5 minutos a 52 °C. Después de la etapa de calentamiento, se mezclaron 10 µl del extracto con 40 µl de mezcla de reacción 1,25X RT. Se puso esta reacción en un baño de agua a 37 °C durante 10 minutos. Se colocó

una pequeña alícuota de la mezcla de reacción (5 µl) sobre una membrana de nilón cargada (Genescreen+, NEN). Se lavó la membrana dos veces con TCA al 10 % + 1 % de pirofosfato de sodio, se lavó con etanol, se secó, y se puso cerca de una pantalla de fósforo. Como alternativa, se puede lavar la membrana dos veces con pirofosfato de sodio al 4 % (pH 8.0), enjuagar con etanol, secar, y después poner cerca de una pantalla de fósforo. El producto radiactivo que ha sido atrapado sobre el filtro se detectó analizando la pantalla en un Posphorimager, utilizando el software ImageQuant (Molecular Devices).

Se seleccionaron los candidatos si presentaban más actividad de transcriptasa inversa (radiactividad) después de la etapa de inactivación por calor. Se cribaron estos candidatos una segunda vez para confirmar el fenotipo. Los candidatos que parecieron ser termoestables después del segundo cribado se hicieron crecer en pequeños cultivos y se ensayaron una tercera vez para determinar la actividad de transcriptasa inversa termoestable. Se secuenciaron los candidatos que fueron resistentes al calor de manera reproducible y se determinó la mutación en cada clon. Se diseñó un oligonucleótido correspondiente al sitio mutagenizado en el cual el codón para el aminoácido mutagenizado era aleatorio (NNK o NNN). Estos oligonucleótidos se utilizaron en la mutagénesis dirigida a un sitio para generar una genoteca en la cual se hicieron todas las posibles sustituciones en el sitio mutagenizado. Se hizo un cribado de esta genoteca para buscar la actividad de transcriptasa inversa termoestable, y se secuenciaron los clones más prometedores.

El cribado de mutantes en el segmento 2 (véase la Figura 2) dio como resultado la identificación de un mutante, H204R. El cribado de una genoteca mutagenizada en el sitio H204 dio como resultado varios mutantes, pero el único que era más termoestable que la transcriptasa inversa de M-MLV fue otro mutante H204R. Los mutantes H204R de la transcriptasa inversa de M-MLV tienen mayor termoestabilidad. El cribado de mutantes en el segmento 3 (véase la Figura 2) dio como resultado un mutante, T306K. La aleatorización de la posición T306 produjo mutantes termoestables que, cuando se secuenciaron, fueron T306R. Los dos mutantes T306K y T306R de la transcriptasa inversa de M-MLV tienen aproximadamente 1,5 veces mayor termoestabilidad.

Ejemplo 3: Mutantes de la transcriptasa inversa de TdT

Al comprobar la fidelidad de los mutantes de transcriptasa inversa (RT) para la extensión defectuosa en un ensayo de 3 dNTP, se observó que la transcriptasa inversa SuperScript™II se extendía 2-3 bases más allá del extremo del molde en presencia de 3 y 4 dNTPs. Esta extensión no dirigida al molde o actividad TdT se reduce en muchos mutantes, pero en unos pocos tales como F309N y T197E parece que esta actividad se reduce severamente o se elimina. Estos mutantes están probablemente en estrecha proximidad o en contacto con el molde-cebador como se determina por homología con la transcriptasa inversa de VIH y su estructura cristalina con el molde-cebador unido.

Métodos

Mutagénesis

Para F309N:

Se diseñaron cebadores correspondientes a la posición mutante F309 con la inserción silente de un sitio de restricción *Ngo*MIV en las posiciones de los aminoácidos 310-311. Los cebadores codificaron una secuencia NNK aleatoria para esta posición generando una genoteca aleatoria de F309 mutantes, donde N es cualquiera de las cuatro bases y K es T o G. Se utilizaron los cebadores junto con los cebadores internos de la transcriptasa inversa SuperScript™II en un sitio de restricción *Sst*I aguas arriba y en un sitio de restricción *Sal*I aguas abajo, en una reacción de la PCR estándar (10 ng de molde de transcriptasa inversa SuperScript™II, 2 µM de cada cebador, 48 µl de SuperMix (Invitrogen Corp. (Carlsbad, CA)) para 20 ciclos de 94 °C 15 segundos, 55 °C 15 segundos, 72 °C 30 segundos) para generar dos fragmentos de PCR. Estos fueron un fragmento *Sst*I-*Ngo*MIV de 240 pares de bases y un fragmento *Ngo*MIV-*Sal*I de 200 pares de bases. Se aislaron los fragmentos y se digirieron y ligaron juntos y después se insertaron en el clon original de la transcriptasa inversa SuperScript™II cortado con *Sst*I y *Sal*I. El producto de ligación resultante se transformó en células competentes DH10B de máxima eficiencia (Invitrogen Corp. (Carlsbad, CA)) para crear la genoteca de mutantes en el sitio F309. Esta genoteca fue puesta entonces en placas durante la noche para selección.

Para T197E y Y133A:

Los mutantes T197E y Y133A se prepararon por mutagénesis oligo-dirigida como se describe en Kunkel, T.A. et al. Methods Enzymol. 204:125 (1991). Brevemente, el gen de la transcriptasa inversa SuperScript™II se insertó en el vector pBADhisA (Invitrogen Corporation) y se denominó pBAD-SSII. Este plásmido se transformó en las células DH11S y se infectaron las células con el fago ayudante M13K07 a partir del cual se aisló el ADN de una sola hebra. Se diseñaron oligos correspondientes a cada mutación: T197E y Y133A. Cada oligo (100 µM) fue tratado con polinucleótido quinasa T4 (Invitrogen Corp. (Carlsbad, CA)) utilizando el tampón de reacción directo (Invitrogen Corp. (Carlsbad, CA)). El oligo fue anillado con un ADN pBAD-SSII de una sola hebra. Se añadieron a la reacción sobre hielo, ADN-polimerasa T7 nativa (USB) y ADN ligasa T4 (Invitrogen Corp. (Carlsbad, CA)) con tampón de síntesis (dNTPs 0,4 mM, Tris-HCl 17,5 mM, pH 7,5, MgCl₂ 5 mM, DTT 2,5 mM, y ATP 1 mM). Se incubaron las reacciones a 37 °C durante 30 minutos y se terminó añadiendo 1 µl de EDTA 0,5 M. Se transformaron las reacciones y se pusieron en placas con células DH10B. Se recogieron las colonias y se determinaron los mutantes por análisis con

enzima de restricción y se secuenciaron para confirmación utilizando un instrumento ABI 377 y un kit ABI Big Dye Terminator Cycle Sequencing Ready Reaction.

Selección de colonias que contienen transcriptasa inversa activa.

Se inocularon colonias transformantes individuales en los pocillos individuales de una placa de cultivo de 96 pocillos. Cada pocillo contenía 120 µl de medio (EG-Ap) conteniendo 0,2 % de arabinosa. Es preferible inocular en primer lugar una placa de 96 pocillos con medio selectivo sin el inductor, para cultivar dicha placa maestra durante la noche, y después hacer una réplica de la placa maestra en una placa de 96 pocillos con el inductor y cultivar dicha placa durante la noche. Se dejaron crecer los cultivos durante la noche a 37 °C sin agitación. Los cultivos después de una noche se mezclaron con un volumen igual de 2X PLD (glucosa al 1,8 %, Tris-HCl 50 mM, pH 8,0, EDTA 20 mM, DTT 20 mM, Triton™ X-100 al 1 %, lisozima a 2 mg/mL) a temperatura ambiente. Se ensayaron estos extractos directamente en cuanto a actividad de transcriptasa inversa mezclando 10 µl del extracto con 40 µl de mezcla de reacción 1,25X RT (Tris-HCl 62,5 mM, pH 8,4, KCl 62,5 mM, MgCl₂ 12,5 mM, DTT 12,5 mM, dGTP 1,25 mM, poliC/oligo dG (3,75 mM/1,5 mM en nucleótidos), [³²P] dGTP). Se puso esta reacción en un baño de agua a 37 °C durante 10 minutos. Se colocó una pequeña alícuota de la mezcla de reacción (5 µl) sobre una membrana de nilón cargada (Genescreen+, NEN). Se lavó la membrana dos veces con TCA al 10 % + 1 % de pirofosfato de sodio, se enjuagó con etanol, se secó, y se puso cerca de una pantalla de fósforo. El producto radiactivo que ha sido atrapado sobre el filtro se detectó analizando la pantalla en un Posphorimager, utilizando el software ImageQuant (Molecular Devices). Se seleccionaron los candidatos si presentaban actividad de transcriptasa inversa (radiactividad). Estos candidatos se cribaron una segunda vez para confirmar el fenotipo. Los candidatos confirmados se secuenciaron entonces para determinar qué aminoácidos mantenían actividad detectable de transcriptasa inversa.

Purificación de mutantes de transcriptasa inversa.

Se suspendió el sedimento celular que contenía transcriptasa inversa inducida en una relación de 2 mL de tampón de lisis (Tris-HCl 40 mM, pH 8,0, KCl 0,1 M, PMSF 1 mM)/1 gramo de sedimento celular. Se sonicó la suspensión sobre hielo y después se centrifugó a 27.000 g durante 30 minutos. Se filtró el extracto libre de células a través de una jeringa filtrante de 0,45 µ. El extracto libre de células se aplicó a una columna de Ni²⁺ HI-TRAP de 5 mL (Pharmacia) pre-equilibrada con 5 volúmenes de imidazol 5 mM en tampón A (Tris HCl 40 mM, pH 8,0, glicerol al 10 %, Triton X-100 al 0,01 %, KCl 0,1 M) a 1 mL/min. Se lavó la columna con 10 volúmenes de imidazol 5 mM en tampón A. Se eluyó la transcriptasa inversa lavando con 20 volúmenes de un gradiente de imidazol 5 mM a 1 M en tampón A. El eluato que contiene proteína transcriptasa inversa se aplicó a una columna Mono-S de 1 mL (Pharmacia) pre-equilibrada con 10 volúmenes de columna de KCl 50 mM en tampón B (Tris-HCl 40 mM, pH 8,0, glicerol al 10 %, Triton X-100 al 0,01 %, EDTA 0,1 mM, DTT 1 mM) a un caudal de 1,0 mL/min. Se lavó la columna con 10 volúmenes de KCl 50 mM en tampón B. Se eluyó la transcriptasa inversa con 20 volúmenes de un gradiente de KCl desde 50 mM a 1 M en tampón B. Las fracciones individuales se analizaron en cuanto a actividad de RT. La fracción que contiene el pico de actividad de RT se dializó frente al tampón de almacenamiento (Tris-HCl 40 mM, pH 8,0, glicerol al 50 %, Triton X-100 al 0,01 %, EDTA 0,1 mM, DTT 1 mM, KCl 0,1 M). Las transcriptasas inversas purificadas tenían más del 95 % de pureza, determinada por SDS-PAGE. Las concentraciones de proteína se determinaron utilizando el kit colorimétrico Biorad.

Método de ensayo de 3 dNTP.

Los procedimientos fueron modificados de los de Preston, B.D., et al. Science 242:1168 (1988). El molde-cebador de ADN se preparó por anillamiento de un molde de 47-meros (5'-GAGTTACAGTGT TTTTGTCCAGTCTGTAGCAGTGTGTGAATGGAAG-3') (SEQ ID NO:3) con un cebador de 18-meros (5'-CTTCCATTACACACTGC-3') (SEQ ID NO:4) marcado con [³²P] en el extremo 5' con polinucleótido-quinasa T4 (molde:cebador, 3:1). La mezcla de ensayo (10 µl) contenía molde-cebador 5 mM, transcriptasa inversa 50-200 nM como se especifica en las leyendas de la figura, 3 o 4 dNTPs (cada uno 250 µM), Tris-HCl 50 mM (pH 8,3), KCl 75 mM, MgCl₂ 3 mM, DTT 10 mM. Se incubaron las reacciones a 37 °C durante 30 minutos y terminaron por la adición de 5 µl de EDTA 40 mM, formamida al 99 %. Los productos de reacción se desnaturalizaron incubando a 95 °C durante 5 minutos y se analizaron por electroforesis sobre geles de poli(acrilamida) al 6 % con urea.

Para determinar si aparecía alguna actividad de TdT en la reacción de control del ensayo de 3 dNTP, que utiliza los 4 dNTPs, se repitió la reacción de control con cantidades variables de enzima, >600 unidades a 20 unidades, a 37 °C durante 30 minutos. Para SuperScript™II, T197E, y Y133A, se utilizaron 200, 100, 50, y 20 unidades. Para F309N, se utilizaron 646, 200, 50, y 20 unidades.

Resultados

Se llevó a cabo un ensayo de inserción defectuosa de transcriptasa inversa SuperScript™II F309N (H204R, T306K), de aquí en adelante denominada F309N, con molde de ADN. Este ensayo se empleó para comparar la capacidad de incorporación errónea del mutante con la de SuperScript™II. El ensayo es un ensayo de extensión del cebador utilizando molde-cebador de ADN sintético y conjuntos dNTP predispuestos que contenían solamente tres de cuatro dNTPs. Las reacciones se presentan sobre un gel en la Figura 3. Mientras se llevaba a cabo este procedimiento

para cribar a los mutantes con tasas más bajas de inserción defectuosa/extensión defectuosa se observó que la transcriptasa inversa SuperScript™II se extendía 2-3 nucleótidos más allá del extremo del molde y que algunas mutaciones reducían o parecían eliminar esta extensión no dirigida al molde o la actividad TdT. Como se muestra en la Figura 4, en presencia de los 4 dNTPs, la transcriptasa inversa SuperScript™II y el mutante F309N fueron capaces de extender el cebador aproximadamente igual, con la transcriptasa inversa SuperScript™II añadiendo 2 nucleótidos más allá del molde, y el F309N sin añadir ninguno más allá del extremo del molde. Para evaluar adicionalmente esta extensión no dirigida al molde se llevó a cabo la reacción de control para el ensayo de extensión defectuosa de 3 dNTPs conteniendo los 4 dNTPs con transcriptasa inversa SuperScript™II, F309N, T197E, y Y133A durante 30 minutos con cantidades variables de enzima. Los tres mutantes habían demostrado niveles muy reducidos de actividad TdT en cribados anteriores. Puesto que se había observado que 5 minutos con 20 unidades de enzima era tiempo más que suficiente para que se completara la extensión del cebador, una incubación de 30 minutos y 200 a 646 unidades de transcriptasa inversa fueron ambos un gran exceso sobre lo que era necesario para que se completara la reacción. Como se ve en la Figura 4, todas las reacciones de transcriptasa inversa en la cantidad más baja ensayada tuvieron productos de extensión similares a los de las reacciones en las concentraciones unitarias más altas demostrando que la reacción había llegado a completarse. Los 2 nucleótidos añadidos de la transcriptasa inversa SuperScript™II más allá del extremo del molde, F309N y T197E no se extendieron más allá del extremo del molde, y Y133A parece que tiene una pequeña cantidad de producto que es 1 nucleótido más allá del extremo del molde.

Ejemplo 4: Mutantes duales termoestables y TdT

La posición del aminoácido F309 en la transcriptasa inversa de M-MLV (RT) está alineada con la posición de W266 en la transcriptasa inversa de VIH. Esta posición está en la base del dominio del pulgar y se considera parte del tracto de unión con la ranura menor que interactúa con la ranura menor del molde-cebador. Se ha demostrado que las mutaciones H204R y T306K aumentan la termoestabilidad de la enzima. La mutación F309N en un clon H204R/T306K presenta una frecuencia de mutación 2,3x más baja en un ensayo directo de *lacZ* (Tabla 5) en molde de ARN y productos de extensión más cortos en un ensayo de extensión de 3 dNTP que la transcriptasa inversa SuperScript™II o H204R/T306K en la transcriptasa inversa SuperScript™II. Ambos hallazgos apoyan la reivindicación de una enzima con más alta fidelidad (Tabla 6).

Tabla 5

Frecuencia de mutación de los mutantes con alta fidelidad de la transcriptasa inversa de M-MLV			
Construido	placas totales	placas mutantes	MF(x 10 ⁻⁴)
SSII	15689	87	39
SSII (H204R, T306K)	14410	83	41
SSII (H204R, T306K, F309N)	11623	39	17
SSII (H204R, T306K, F309N, V223H)	11415	39	14

Tabla 5. La frecuencia de mutación de la transcriptasa inversa SuperScript™II y mutantes puntuales. Se determinó la frecuencia de mutación (MF) dividiendo el número de placas mutantes (azul claro o blanco) por el número total de placas. La frecuencia mutante previa del ADN de partida fue 17 x 10⁻⁴ para los 3 primeros constructos y 20 x 10⁻⁴ para el último constructo.

Tabla 6

Tasas de error (ER) de los mutantes con alta fidelidad de la transcriptasa inversa de M-MLV				
	M-MLV	SuperScript™II	F309N	V223H/F309N
ER global (oER)	1/17.000	1/15.000	1/34.000	1/41.000
Desajuste				
% de total	46	35	68 72	
ER (mER)	1/37.000	1/42.000	1/50.000	1/58.000
Desplazamiento del marco de lectura				

Tasas de error (ER) de los mutantes con alta fidelidad de la transcriptasa inversa de M-MLV				
	M-MLV	SuperScript™II	F309N	V223H/F309N
% de total	46	60	2122	
ER (rER)	1/37.000	1/25.000	1/162.000	1/188.000
Salto de hebra				
% de total	8	5	11 6	
ER (jER)	1/213.000	1/297.000	1/324.000	1/690.000

Métodos

Mutagénesis. Utilizando un protocolo de mutagénesis estándar dirigida a un sitio como se describe en el ejemplo 3, se anilló un cebador que contiene la mutación V223H con ADN de una sola hebra de SuperScript™II con las siguientes mutaciones: H204R, T306K, F309N. Se secuenciaron las colonias para confirmar la nueva combinación de V223H, H204R, T306K, y F309N.

Selección de colonias que contienen transcriptasa inversa activa. La selección de colonias se realizó como en el ejemplo 3.

Purificación de RT mutantes. Se realizó la purificación como en el ejemplo 3.

Secuenciación de placas. Las placas del ensayo directo de *lacZ* se transfirieron desde la placa de agar blando a papel Whatmann 3MM y se dejaron secar durante al menos 1 hora. Se perforó entonces la placa y se añadió el disco de placa/papel directamente a una mezcla de reacción de secuenciación que contenía 4-8 µl de ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction (Perkin Elmer), 1 µl de cebador (GAAGATCGCACTCCAGCCAGC) (SEQ ID NO:5), y agua destilada hasta un volumen total de 20 µl. Se utilizó el protocolo de secuenciación de ciclo ABI a 96 °C 10 segundos, 50 °C 5 segundos, 60 °C 4 minutos durante 25 ciclos. Se separaron los discos de papel y se precipitaron las reacciones, después se volvieron a suspender en colorante de carga y se secuenciaron en una máquina de secuenciación ABI 377.

Se compararon las secuencias con la secuencia *alfa lacZ* natural y después se clasificaron como desplazamiento del marco de lectura (1 nucleótido de inserción o de delección), desajuste, o salto de hebra (una inserción o delección entre secuencias repetidas). Se determinó la tasa de error global para cada clase dividiendo la frecuencia de mutación por el número de sitios detectables (es decir, sitios cuya alteración produce un cambio fenotípico) (116) multiplicado por 0,5 (para excluir la contribución de la hebra única original) y después multiplicado por el porcentaje de mutantes observados en cada clase. $ER = MF/(\text{sitios detectables} \times 0,5) \times (\% \text{ en cada clase})$.

Método de ensayo de 3dNTP. Los ensayos de 3dNTP se realizaron como en el ejemplo 3.

Resultados

Se llevó a cabo un ensayo de inserción defectuosa de transcriptasa inversa SuperScript™II F309N (H204R T306K), de aquí en adelante denominada F309N, y de V223H F309N (H204R T306K), de aquí en adelante denominada V223H/F309N con molde de ADN. Este ensayo se empleó para comparar la capacidad de incorporación errónea del mutante con la de SuperScript™II. El ensayo es un ensayo de extensión del cebador utilizando molde-cebador de ADN sintético y conjuntos dNTP predispuestos que contenían solamente tres de los cuatro dNTPs. Las reacciones se presentan sobre un gel en la Figura 5 y en la Figura 6. En este ensayo, la eficiencia más alta de la extensión del cebador indica fidelidad más baja. Como se muestra en las figuras 5 y 6, en presencia de los 4 dNTPs, la transcriptasa inversa SuperScript™II y los mutantes F309N y V223H/F309N fueron capaces de extender el cebador aproximadamente igual, con alguna variación en la adición de nucleótidos no dirigidos al molde en el extremo del cebador. Sin embargo, cuando se incubaron con conjuntos de nucleótidos predispuestos, la transcriptasa inversa SuperScript™II fue capaz de catalizar una extensión sustancial más allá de los nucleótidos del molde para lo cual faltaba un dNTP complementario, indicando el uso de nucleótidos incorrectos y una fidelidad más baja. En la Figura 5, el mutante F309N (2) presentó productos de extensión más cortos que la transcriptasa inversa SuperScript™II en cada uno de los conjuntos predispuestos de tres dNTPs, indicando menos capacidad para incorporar nucleótidos incorrectos y por lo tanto una fidelidad más alta. En la Figura 6, el mutante V223H/F309N se extendió solo con los conjuntos dATP y dCTP. En cada caso V223H/F309N tenía también menos productos de extensión que SuperScript™II. Esto se corresponde con los resultados del ensayo de *lacZα* donde los mutantes F309N y V223H/F309N tenían una frecuencia de mutación más baja que la transcriptasa inversa SuperScript™II (17×10^{-4} y 14×10^{-4} a 39×10^{-4}). La transcriptasa inversa solo con las mutaciones H204R T306K sin F309N tiene una frecuencia de mutación similar a la transcriptasa inversa SuperScript™II (41×10^{-4} a 39×10^{-4}), lo que da a entender que estas

mutaciones no influyen en la fidelidad. Estos datos muestran una correlación entre el ensayo de inserción defectuosa sobre ADN y el ensayo de *lacZα* sobre ARN en donde los mutantes de fidelidad más alta tenían tanto productos de extensión más corta con conjuntos de dNTPs predispuestos como frecuencias de mutación más bajas en el ensayo de *lacZα*.

5 Ejemplo 5: Determinación de la tasa de error

Para determinar las tasas de error, se secuenciaron las placas mutantes del ensayo de *lacZ* utilizando métodos conocidos. Se clasificaron entonces las mutaciones en una de las siguientes categorías: desajustes para sucesos de inserción defectuosa, desplazamientos del marco de lectura para sucesos de inserción o delección única, o saltos para grandes inserciones o delecciones causadas por saltos entre secuencias similares. Se determinó entonces una tasa de error global para el ácido nucleico que codifica el péptido *lacZ-alfa* utilizando la siguiente ecuación:

ER (tasa de error) = MF (frecuencia de mutación) / (número de sitios detectables x 0,5), donde el número de sitios detectables es 116

No todas las bases mutadas en los ensayos dirigidos a *lacZ* dieron como resultado un cambio fenotípico detectable. Para determinar las tasas específicas de error para desajustes, desplazamiento del marco de lectura y saltos, se modificó la frecuencia de mutación multiplicando por el tanto por ciento del total de cada categoría de mutante, y después se utilizó para determinar la tasa específica de error. Lo que sigue es un mapa de secuencias del péptido *lacZα* en M13mp19 a partir de los ensayos de transcriptasa inversa SuperScript™II y la transcriptasa inversa SuperScript™II H203R T306K F309N de alta fidelidad. El subrayado indica deleciones; “^” indica inserciones de la base A, T, C, o G indicadas antes; A, T, C, o G mostradas antes de la secuencia completa indican desajustes.

Mapa de mutaciones introducidas por SuperScript™ II

```

      T      T      T C
AGCGCAACGC AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA
      1      1      4

      CG      C      CC
TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAAATTCAC ACAGGAAACA
      1

      C      CG      C
GCTATG ACC ATG ATT ACG^CCA AGC TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC CGG
      1

      T A      AAAA
      T A      AAA
      T A      A
      T A      A      C
GTA CCG AGC TCG AAT TCA CTG GCC GTC GTT^TTA CAA CGT CGT GAC TGG GAA AAC CCT GGC
      7      1      1      1

      TTTT
      TTTT
      C      TTTT
      C      TTT
      A T T
TC      C      C      T TC      T C T T      C      G      T
GTT ACC CAA CTT AAT CGC CTT GCA GCA CAT CCC^CCT^TTC^GCC AGC TGG CGT
      1      4

```

AAT AGC G (SEQ ID NO:6)

Tabla 7

Inserciones	40	38 %	60 % de desplazamiento del marco de lectura (inserción o delección)
Delecciones	23	22 %	
Desajustes	36	35 %	35 % de desajustes
Salto	5	5 %	5 % de saltos

Tabla 8

Tasa de error global (oER)	1/15.000	$(39 \times 10^{-4}) / (116 \times 0,5)$
Tasa de error de desajustes (mER)	1/42.500	$(0,35 \times 39 \times 10^{-4}) / (116 \times 0,5)$

Tasa de error de desplazamiento del marco de lectura (fER)	1/25.000	$(0,60 \times 39 \times 10^{-4}) / (116 \times 0,5)$
Tasa de error de saltos (jER)	1/297.000	$(0,05 \times 39 \times 10^{-4}) / (116 \times 0,5)$

Todas las publicaciones, patentes y solicitudes de patentes mencionadas en esta memoria descriptiva son indicativas del nivel de los expertos en la técnica a la que pertenece esta invención.

REIVINDICACIONES

1. Una transcriptasa inversa de M-MLV en donde la histidina en la posición 204 natural está reemplazada con arginina para aumentar o mejorar la termoestabilidad.
- 5 2. La transcriptasa inversa de la reivindicación 1, que comprende además una mutación o modificación en la posición correspondiente a la treonina 306.
3. La transcriptasa inversa de la reivindicación 2, en donde la treonina 306 está reemplazada o con lisina o con arginina.
4. La transcriptasa inversa de la reivindicación 1, que retiene al menos 50 % de su actividad de transcriptasa inversa después de calentamiento a 50 °C durante 5 minutos.
- 10 5. La transcriptasa inversa de la reivindicación 1, en donde la transcriptasa inversa tiene una o más propiedades seleccionadas del grupo que consiste en:
(a) actividad de RNasa H reducida o sustancialmente reducida;
(b) actividad de desoxinucleotidil transferasa terminal reducida o sustancialmente reducida; y
(c) aumento de la fidelidad.
- 15 6. La transcriptasa inversa de la reivindicación 5, en donde la transcriptasa inversa tiene una actividad de RNasa H reducida o sustancialmente reducida.
7. La transcriptasa inversa de la reivindicación 5, en donde la transcriptasa inversa tiene una actividad de desoxinucleotidil transferasa terminal reducida o sustancialmente reducida.
8. La transcriptasa inversa de la reivindicación 5, en donde la transcriptasa inversa tiene una fidelidad aumentada.
- 20 9. La transcriptasa inversa de la reivindicación 1, en donde la transcriptasa inversa es RNasa H-transcriptasa inversa de M-MLV
10. Un vector que comprende ácido nucleico que codifica la transcriptasa inversa de la reivindicación 1.
11. El vector de la reivindicación 10, en donde el vector es un vector de expresión y el ácido nucleico está ligado operativamente a un promotor.
- 25 12. El vector de la reivindicación 11, en donde el promotor se selecciona del grupo que consiste en un promotor λ -P_L, un promotor *tac*, un promotor *trp*, un promotor ara BAD y un promotor *trc*.
13. Una célula hospedante que comprende el vector de una cualquiera de las reivindicaciones 10 a 12.
14. Un método para producir una transcriptasa inversa, comprendiendo el método:
(a) cultivar la célula hospedante de la reivindicación 13;
- 30 (b) expresar el ácido nucleico; y
(c) aislar la transcriptasa inversa de la célula hospedante.
15. El método de la reivindicación 14, en donde la célula hospedante es una célula de *Escherichia coli*.
16. Un método para transcripción inversa de una o más moléculas de ácido nucleico que comprende:
(a) mezclar uno o más moldes de ácido nucleico con una o más transcriptasas inversas de la reivindicación 1; y
- 35 (b) incubar la mezcla de (a) en condiciones suficientes para preparar una o más primeras moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes.
17. El método de la reivindicación 16, en donde el molde de ácido nucleico es una molécula de ARN mensajero o una población de moléculas de mRNA.
- 40 18. El método de la reivindicación 17, comprendiendo además el método la etapa de incubar las una o más primeras moléculas de ácido nucleico en condiciones suficientes para preparar una o más segundas moléculas de ácido nucleico complementarias de todo o parte de la una o más primeras moléculas de ácido nucleico.
19. Un método de amplificación de una o más moléculas de ácido nucleico, comprendiendo el método:

(a) mezclar uno o más moldes de ácido nucleico con una o más transcriptasas inversas de la reivindicación 1 y una o más ADN-polimerasas; y

(b) incubar la mezcla de (a) en condiciones suficientes para amplificar una o más moléculas de ácido nucleico complementarias de todo o parte del uno o más moldes.

5 20. Un método para secuenciar una o más moléculas de ácido nucleico, comprendiendo el método:

(a) mezclar una o más-moléculas de ácido nucleico a secuenciar con uno o más cebadores, una o más transcriptasas inversas de la reivindicación 1, uno o más nucleótidos y uno o más agentes de terminación;

(b) incubar la mezcla de (a) en condiciones suficientes para sintetizar una población de moléculas complementarias de todo o parte de las una o más moléculas a secuenciar; y

10 (c) separar la población para determinar la secuencia de nucleótidos de todo o parte de las una o más moléculas a secuenciar.

21. Un método para secuenciar una molécula de ácido nucleico, comprendiendo el método:

(a) mezclar una molécula de ácido nucleico a secuenciar con uno o más cebadores, una o más transcriptasas inversas de la reivindicación 1, uno o más nucleótidos y uno o más agentes de terminación;

15 (b) incubar la mezcla de (a) en condiciones suficientes para sintetizar una población de moléculas complementarias de todo o parte de la molécula a secuenciar; y

(c) separar los miembros de la población para determinar la secuencia de nucleótidos de toda o parte de la molécula a secuenciar.

20 22. Un kit para utilización en transcripción inversa, amplificación o secuenciación de una molécula de ácido nucleico, comprendiendo el kit un primer recipiente que comprende una o más transcriptasas inversas de la reivindicación 1.

23. El kit de la reivindicación 22, que comprende además uno o más componentes seleccionados del grupo que consiste en uno o más nucleótidos, una o más ADN-polimerasas, un tampón adecuado, uno o más cebadores y uno o más agentes de terminación.

24. El kit de la reivindicación 23 en donde el agente de terminación es un didesoxinucleótido.

25

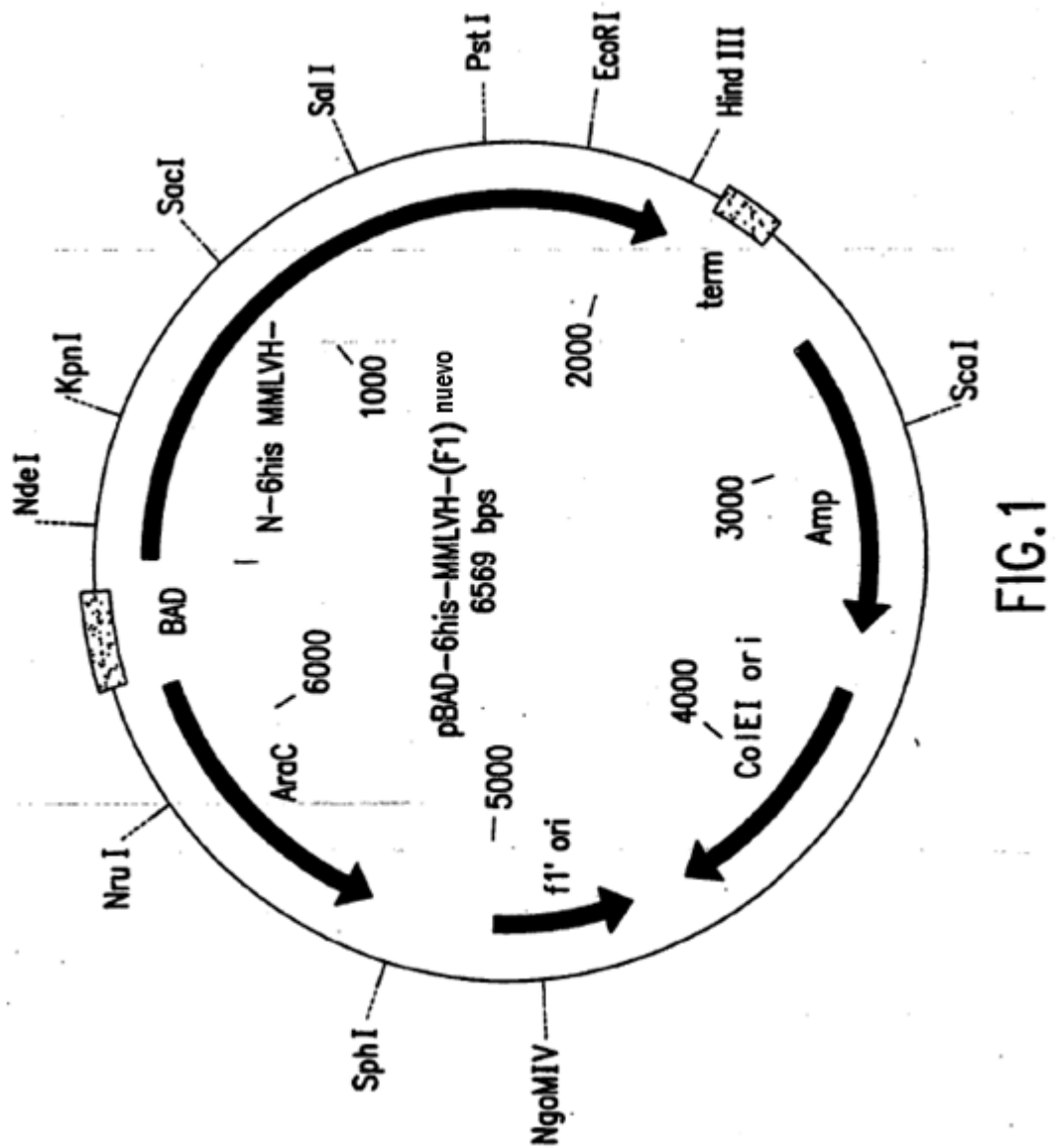


FIG.1

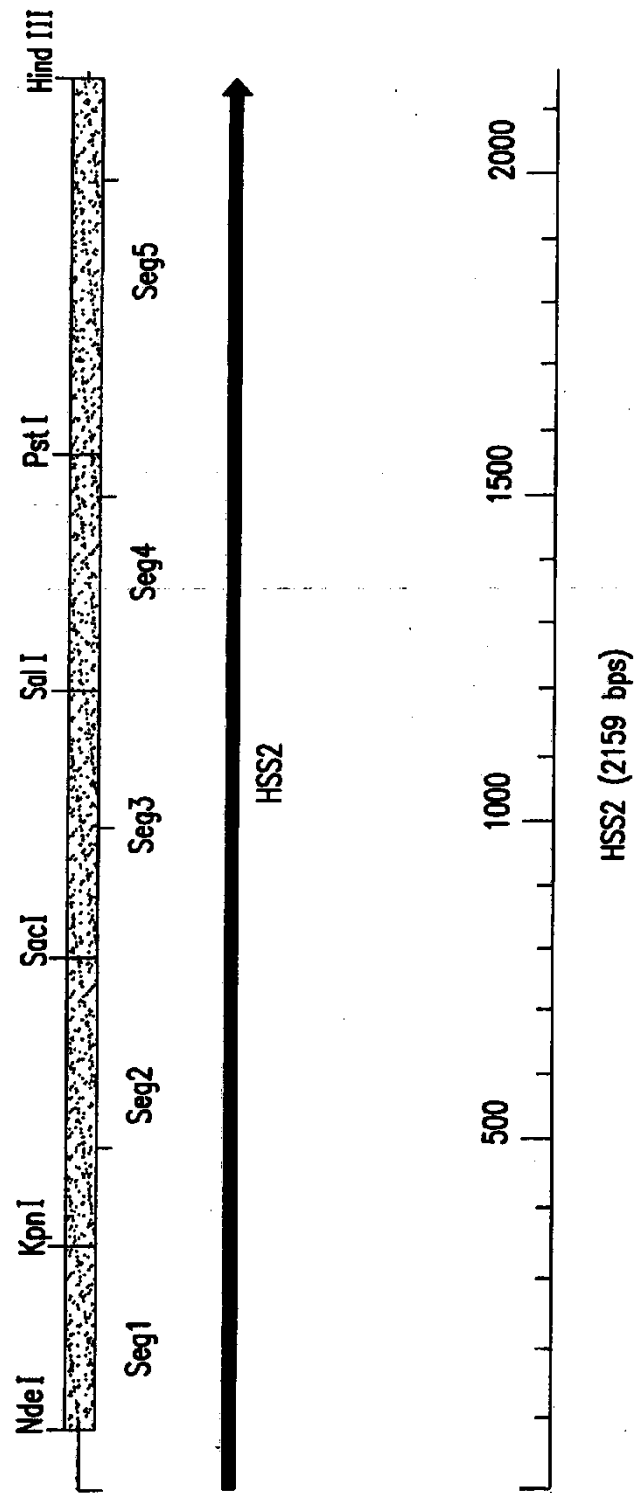


FIG.2



FIG.3

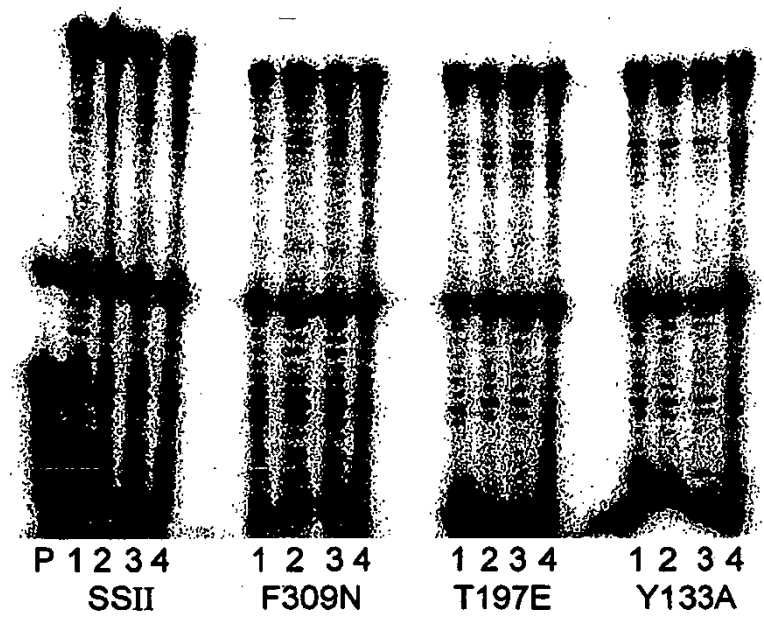


FIG.4

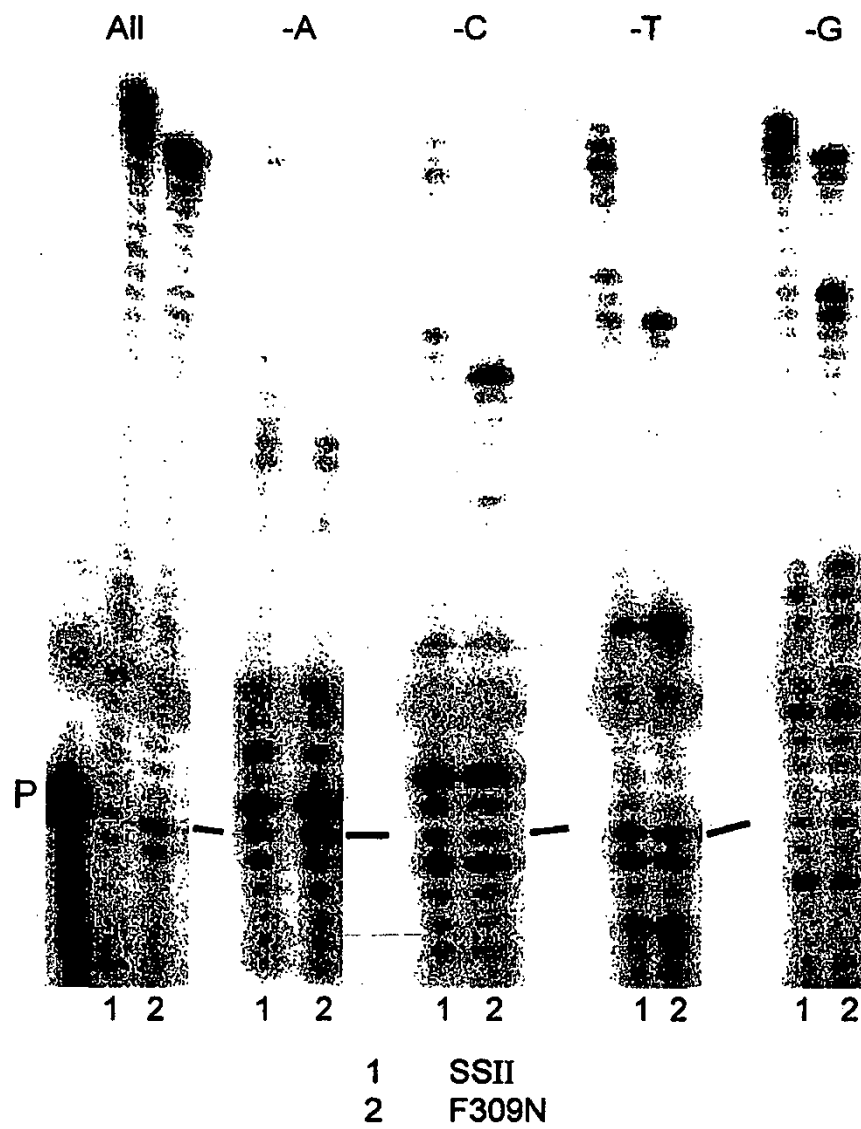


FIG.5

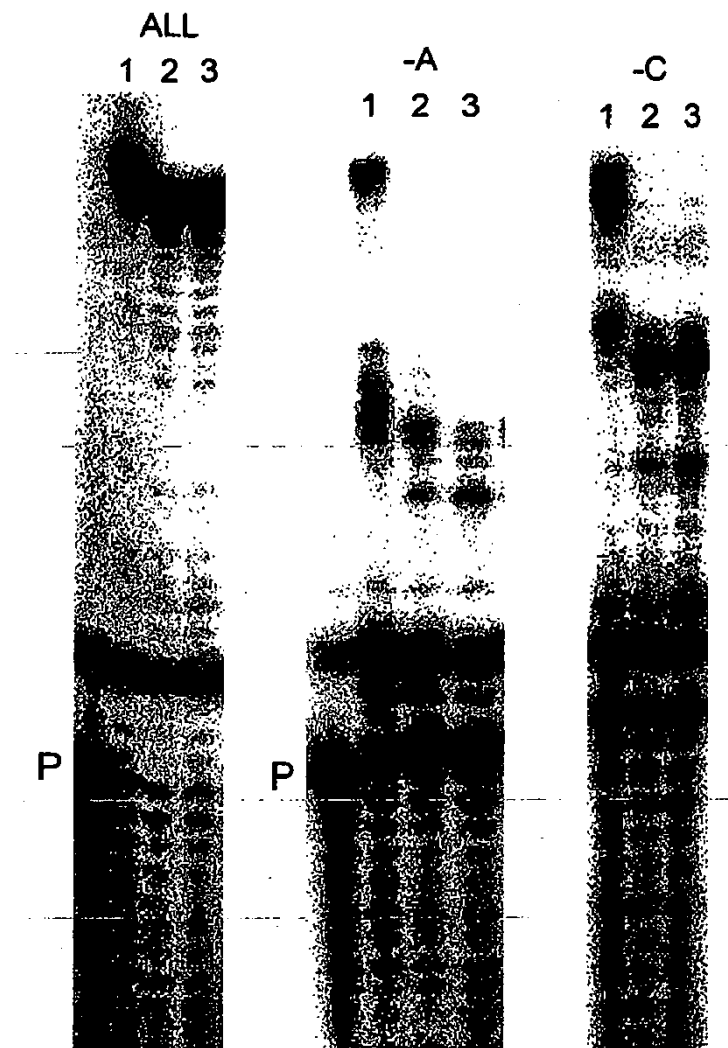


FIG.6