

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 407 138**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.06.2004 E 04756455 (4)**

97 Fecha y número de publicación de la concesión europea: **20.03.2013 EP 1641943**

54 Título: **Métodos y sistemas relacionados con nucleótidos terminadores 2'**

30 Prioridad:

30.06.2003 US 483861 P

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

11.06.2013

73 Titular/es:

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

72 Inventor/es:

**GELFAND, DAVID, H.;
REICHERT, FRED, L.;
BODEPUDI, VEERAI AH;
GUPTA, AMAR;
WILL, STEPHEN, G. y
MYERS, THOMAS, B.**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 407 138 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

Métodos y sistemas relacionados con nucleótidos terminadores 2'.

5 CAMPO DE LA INVENCION

La invención se refiere de manera general a la química y biología molecular de los ácidos nucleicos. Más concretamente, la invención proporciona métodos de secuenciación y marcaje de ácidos nucleicos, además de otros aspectos relacionados en los que participan nucleótidos terminadores 2'.

10

ANTECEDENTES DE LA INVENCION

La secuenciación de ácidos nucleicos implica la determinación de la secuencia de nucleótidos de una molécula particular de ácidos nucleicos. El conocimiento de la secuencia de una molécula de ácidos nucleicos típicamente resulta fundamental para elucidar la función de la molécula y para facilitar la manipulación de la misma. Además, las variaciones en los genomas individuales con frecuencia explican las diferencias de susceptibilidad a enfermedades y las respuestas farmacológicas al tratamiento. A título ilustrativo, los cambios en una única base de una molécula de ácidos nucleicos, que comúnmente se denominan polimorfismos de nucleótido único (SNP), pueden afectar al riesgo individual de presentar una enfermedad dada. Mediante la comparación de estas variaciones, por ejemplo, los investigadores están mejorando la comprensión de la utilidad médica de los SNP, incrementando de esta manera nuestra capacidad de diagnosticar, pronosticar y tratar con efectividad las enfermedades.

15

20

25

30

La tecnología de secuenciación de ácidos nucleicos se inició a finales de los años 60 con intentos para secuenciar el ARN. En particular, la secuencia del ARN ribosómico 5S de *Escherichia coli* (Brownlee *et al.*, "Nucleotide sequence of 5S-ribosomal RNA from *Escherichia coli*," *Nature* 215(102):735, 1967) y el ARN del bacteriófago R17 codificante de proteína de cubierta (Adams *et al.*, "Nucleotide sequence from the coat protein cistron of R17 bacteriophage RNA", *Nature* 223(210):1009, 1969) son algunos de los primeros ejemplos de secuenciación del ARN. Posteriormente, Sanger describió la secuenciación del ADN del bacteriófago f1 mediante síntesis cebada con ADN polimerasa (Sanger *et al.*, "Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage f1 DNA", *Proc. Natl. Acad. Sci. USA* 70(4):1209, 1973), mientras que Gilbert y Maxam informaron de la secuencia de nucleótidos de ADN del operador lac (Gilbert y Maxam, "The nucleotide sequence of the lac operator", *Proc. Natl. Acad. Sci. USA* 70(12):3581, 1973).

35

40

En 1977, Sanger describió la utilización de nucleósidos trifosfato modificados (incluyendo dideoxirribosa) en combinación con desoxirribonucleótidos para terminar el alargamiento de cadena (Sanger *et al.*, "DNA sequencing with chain-terminating inhibitors", *Biotechnology* 24:104, 1977). En ese mismo año, Maxam y Gilbert informaron de un método para secuenciar ADN que utilizaba el corte químico del ADN preferentemente en guaninas, en adeninas, en citosinas y timinas igualmente, y únicamente en citosinas (Maxam y Gilbert, "A new method for sequencing DNA", *Proc. Natl. Acad. Sci. USA* 74:560, 1977). Estos dos métodos aceleraron la secuenciación manual basada en la separación electroforética de fragmentos de ADN marcados con marcadores radioactivos y la posterior detección mediante autorradiografía.

45

50

El uso del método dideoxi de Sanger para la secuenciación del ADN se ha extendido mucho más ampliamente que el método de corte químico de Maxam-Gilbert. El método de Sanger incluye la síntesis de una nueva cadena de ADN partiendo de un sitio de cebado específico y finalizando con la incorporación de un nucleótido de terminación de cadena o terminador. En particular, una ADN polimerasa extiende un ácido nucleico cebador hibridado con una localización específica sobre un molde de ADN mediante la incorporación de desoxinucleótidos (dNTP) complementarios al molde. La síntesis de la nueva cadena de ADN continúa hasta que se termina aleatoriamente la reacción mediante la inclusión de un dideoxinucleótido (ddNTP). Estos análogos de nucleótido son incapaces de proporcionar soporte a la extensión posterior de la cadena debido a que la fracción ribosa de ddNTP no presenta el 3'-hidroxilo necesario para formar un enlace fosfodiéster con el dNTP siguiente. Esto produce una población de fragmentos de secuenciación truncados, cada uno con un extremo 5' definido o fijado y un extremo 3' variable. Entre las desventajas del método dideoxi se encuentra el coste asociado a la preparación de los ddNTP.

55

60

65

Dos metodologías de secuenciación automática utilizados frecuentemente son la secuenciación de pigmento-ácido nucleico cebador y de pigmento-terminador. Estos métodos resultan adecuados para la utilización con fracciones de marcaje fluorescente. Aunque la secuenciación también puede llevarse a cabo utilizando fracciones de marcaje radioactivo, la secuenciación basada en la fluorescencia resulta crecientemente preferente. Brevemente, en la secuenciación de pigmento-cebador, se utiliza un cebador marcado fluorescentemente en combinación con ddNTP no marcados. El procedimiento típicamente utiliza cuatro reacciones de síntesis y hasta cuatro carriles en un gel para cada molde que debe secuenciarse (uno correspondiente a cada uno de los productos de terminación específicos de base). Tras la extensión del ácido nucleico cebador, las mezclas de reacción de secuenciación que contienen productos de terminación dideoxinucleótido incorporados se someten rutinariamente a electroforesis en un gel de secuenciación de ADN. Tras la separación mediante electroforesis, los productos marcados fluorescentemente se excitan en el gel con un láser y se detecta la fluorescencia con un detector apropiado. En sistemas automáticos, un detector escanea el fondo del gel durante la electroforesis con el fin de detectar la fracción

de marcaje que se ha utilizado, a medida que las reacciones recorren la matriz de gel (Smith *et al.*, "Fluorescence detection in automated DNA sequence analysis", Nature 321:674, 1986). En una modificación de este método, se marca cada uno de los cuatro cebadores con un marcador fluorescente diferente. Tras completarse las cuatro reacciones de secuenciación diferentes, las mezclas se combinan y la reacción se somete a análisis en gel en un solo carril, y se detectan individualmente las diferentes etiquetas fluorescentes (una para cada uno de los cuatro diferentes productos de terminación específicos de base).

Alternativamente se utilizan métodos de secuenciación de pigmento-terminador. En este método, se utiliza una ADN polimerasa para incorporar los dNTP y ddNTP marcados fluorescentemente en el extremo en crecimiento de un cebador de ADN (Lee *et al.*, "DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye terminators and probability analysis of termination fragments", Nucleic Acid Res. 20:2471, 1992). Este procedimiento ofrece la ventaja de que no resulta necesario sintetizar cebadores marcados con pigmento. Además, las reacciones de pigmento-terminador resultan más convenientes en el aspecto de que la totalidad de las cuatro reacciones puede llevarse a cabo en el mismo tubo.

Entre otros métodos de deconvolución de mezclas de reacción de secuenciación se incluyen la utilización de espectrometría de iones en fase gaseosa. Por ejemplo, la espectrometría de masas de tiempo de vuelo y desorción/ionización por láser asistida por matriz (EM MALDI-TOF) es un enfoque que se ha utilizado con éxito en la secuenciación de alto rendimiento y análisis de genotipado de SNP (ver, por ejemplo, Sauer *et al.*, "Facile method for automated genotyping of single nucleotide polymorphisms by mass spectrometry", Nucleic Acids Res. 30(5):e22, 2002).

A partir de lo anteriormente expuesto, resulta evidente que resultan deseables métodos de secuenciación y genotipado adicionales de los ácidos nucleicos. La presente invención proporciona nuevos métodos de secuenciación de ácidos nucleicos que utilizan nucleótidos terminadores 2', así como una diversidad de características adicionales, incluyendo enfoques al marcaje de ácidos nucleicos que resultarán evidentes a partir de una revisión completa de la exposición siguiente.

DESCRIPCIÓN RESUMIDA DE LA INVENCIÓN

La presente invención proporciona métodos de secuenciación y marcaje de ácidos nucleicos que utilizan ácidos nucleicos terminadores 2', por ejemplo en lugar de ddNTP, acilonucleótidos trifosfato u otros tipos de terminadores de extensión de ácidos nucleicos. Entre los nucleótidos terminadores 2' de la invención, que presentan anillos de azúcar intactos (por ejemplo anillos de azúcar pentosa) o anillos análogos de azúcar (por ejemplo anillos carbocíclicos, etc.) se incluyen grupos bloqueantes (por ejemplo un grupo bloqueante cargado negativamente, un grupo bloqueante voluminoso y/o similar) en posiciones 2' de dichas fracciones de azúcar. Además, los biocatalizadores que incorporan nucleótidos comprenden la capacidad de extender cebadores u otros ácidos nucleicos con dichos nucleótidos terminadores 2' (por ejemplo un 2'-fosfato-3'-hidroxil-NTP o NDP, etc.) en el extremo 3' del ácido nucleico cebador, por ejemplo de una manera dirigida por un molde (es decir, la incorporación de los nucleótidos terminadores 2' en los ácidos nucleicos cebadores). Determinados biocatalizadores que incorporan nucleótidos a los que se hace referencia en la presente memoria, tales como desoxinucleotidil-transferasa terminal (TdT, EC 2.7.7.31), polinucleótido fosforilasa (PNPasa, EC 2.7.7.8), etc., son generalmente capaces de extender los ácidos nucleicos de una manera independiente de molde. Tras la incorporación de un nucleótido terminador 2' en el extremo 3'-terminal de un ácido nucleico cebador, el ácido nucleico típicamente deja de ser extensible por un biocatalizador que incorpora nucleótidos de la invención. Además, un ácido nucleico cebador extendido que comprende un nucleótido terminador 2' también es generalmente resistente a la actividad enzimática correctora de errores (por ejemplo una actividad de exonucleasa 3'-5', etc.). De esta manera, un biocatalizador que incorpora nucleótidos utilizado en un método de la invención opcionalmente incluye una actividad exonucleasa 3'-5', por ejemplo para mejorar la fidelidad de la secuencia respecto a enfoques que utilizan catalizadores que presentan actividades correctoras de errores reducidas o nulas. Además de métodos, la invención proporciona además mezclas de reacción, kits, sistemas, ordenadores y medios legibles por ordenador relacionados con los nucleótidos 2' descritos en la presente memoria. La presente invención proporciona una alternativa económica a los métodos de terminador preexistentes. Los nucleótidos terminadores 2' de la invención se sustituyen fácilmente en diversos protocolos de secuenciación, marcaje terminal u otros sin sacrificar la facilidad de uso.

Más concretamente, un aspecto de la presente invención se refiere a un método de extensión de un cebador de ácidos nucleicos. El método incluye incubar un ácido nucleico molde (por ejemplo ADN, ARN, etc.) con por lo menos un biocatalizador que incorpora nucleótidos, por lo menos un nucleótido terminador 2' (por ejemplo un nucleósido 2'-monofosfato-3'-hidroxil-5'-trifosfato, etc.) y por lo menos un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico molde. El cebador de ácidos nucleicos generalmente comprende ADN. El biocatalizador que incorpora nucleótidos extiende el cebador de ácidos nucleicos para producir por lo menos un cebador de ácidos nucleicos extendido, incorporando el nucleótido terminador 2' en un extremo terminal del cebador de ácidos nucleicos extendido. En algunas realizaciones, el ácido nucleico de molde se incuba con el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' y el cebador de ácidos nucleicos en solución, mientras que en otras, el cebador de ácidos nucleicos o el ácido nucleico molde se une covalentemente a un soporte sólido.

En determinadas realizaciones de la invención, el método incluye además detectar una masa molecular del cebador de ácidos nucleicos extendido o un fragmento del mismo. En estas realizaciones, un genotipo del ácido nucleico molde es determinable a partir de la masa molecular detectada del cebador de ácidos nucleicos extendido o fragmento del mismo. La masa molecular típicamente se detecta utilizando espectrometría de iones en fase gaseosa (por ejemplo espectrometría de masas MALDI-TOF u otra versión de la espectrometría de iones en fase gaseosa).

El nucleótido terminador 2', el cebador de ácidos nucleicos extendido y/o el cebador de ácidos nucleicos opcionalmente comprende por lo menos un marcaje (por ejemplo un pigmento fluorescente, un isótopo radioactivo, un grupo modificador de la masa, etc.). En estas realizaciones, el método incluye además generalmente detectar una señal detectable producida por el marcaje (por ejemplo espectrofotométricamente, etc.), de manera que es determinable un genotipo del ácido nucleico de molde a partir de la señal detectable. Por ejemplo, el marcaje se une opcionalmente, por ejemplo a una base heterocíclica del nucleótido terminador 2', una fracción azúcar del nucleótido terminador 2' y/o un grupo fosfato del nucleótido terminador 2'. Opcionalmente, un conector une el marcaje al nucleótido terminador 2'.

El método de extensión de un cebador de ácidos nucleicos opcionalmente incluye además incubar el ácido nucleico de molde con por lo menos un nucleótido extensible (por ejemplo un ribonucleótido, un desoxirribonucleótido y/o similar). En estas realizaciones, el biocatalizador que incorpora nucleótidos típicamente produce múltiples cebadores de ácidos nucleicos extendidos diferentes y el método incluye generalmente además la resolución de múltiples cebadores de ácidos nucleicos extendidos diferentes, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico de molde es determinable a partir de los cebadores de ácidos nucleicos extendidos que se han resuelto. Por ejemplo, los cebadores de ácidos nucleicos extendidos opcionalmente se resuelven mediante la determinación de las masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos. En determinadas realizaciones, los cebadores de ácidos nucleicos extendidos comprenden además marcajes y los cebadores de ácidos nucleicos extendidos se resuelven mediante separación de los cebadores de ácidos nucleicos extendidos marcados unos de otros y la detección de las señales detectables producidas por los marcajes. A título ilustrativo, los cebadores de ácidos nucleicos extendidos marcados se separan mediante por lo menos una técnica de separación, tal como electroforesis, cromatografía y espectrometría de iones en fase gaseosa (por ejemplo espectrometría de masas MALDI-TOF u otra versión de espectrometría de iones en fase gaseosa).

En otros aspectos, la invención proporciona un método para extender un ácido nucleico, por ejemplo para marcar el extremo del ácido nucleico y/o para otras aplicaciones. El método incluye incubar por lo menos un ácido nucleico con por lo menos un biocatalizador que incorpora nucleótidos (por ejemplo una transferasa terminal, una polinucleótido fosforilasa, etc.) y por lo menos un nucleótido terminador 2' marcado. El biocatalizador que incorpora nucleótidos extiende los ácidos nucleicos para producir por lo menos un ácido nucleico extendido mediante la incorporación del nucleótido terminador 2' en un extremo terminal (por ejemplo un extremo 3'-terminal) del ácido nucleico. En determinadas realizaciones, el método incluye además la hibridación del ácido nucleico extendido con otro ácido nucleico y la detección de una señal detectable producida por el marcaje.

En algunas realizaciones, el ácido nucleico comprende un cebador de ácidos nucleicos que es por lo menos parcialmente complementario respecto a por lo menos una subsecuencia de un ácido nucleico de molde, y el método comprende incubar el ácido nucleico de molde con el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' marcado y el cebador de ácidos nucleicos. En estas realizaciones, el biocatalizador que incorpora nucleótidos típicamente comprende un enzima seleccionado de entre, por ejemplo, una polimerasa, una transferasa terminal, una transcriptasa inversa, una polinucleótido fosforilasa, una telomerasa y similar. A título ilustrativo, el biocatalizador que incorpora nucleótidos opcionalmente comprende un enzima modificado (por ejemplo una ADN polimerasa CS5 G46E E678G, una ADN polimerasa CS6 G46E E678G, una ADN polimerasa Taq E615G, una polimerasa Δ ZO5R, una ADN polimerasa CS5 G46E L329A E678G, etc.). Típicamente, el método incluye además incubar el ácido nucleico de molde con por lo menos un nucleótido extensible. En estas realizaciones, el biocatalizador que incorpora nucleótidos generalmente produce múltiples cebadores de ácidos nucleicos extendidos diferentes y el método comprende resolver los múltiples cebadores de ácidos nucleicos extendidos. Por lo menos una parte de una secuencia de bases del ácido nucleico de molde típicamente es determinable a partir de los cebadores de ácidos nucleicos extendidos resueltos. Típicamente, los cebadores de ácidos nucleicos extendidos se resuelven mediante la determinación de las masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos. Por ejemplo, los cebadores de ácidos nucleicos extendidos opcionalmente se resuelven mediante la separación de los cebadores de ácidos nucleicos extendidos unos respecto a otros y la detección de las señales detectables producidas por estos marcajes.

En otro aspecto, la invención se refiere a un método de inhibición de la extensión adicional de un ácido nucleico extendido, por ejemplo para tratar un huésped infectado con un agente patogénico o similar. El método incluye poner en contacto por lo menos un ácido nucleico (por ejemplo ADN microbiano, ARN vírico, etc.) con por lo menos un biocatalizador que incorpora nucleótidos y por lo menos un nucleósido o nucleótido terminador 2', o una sal farmacéuticamente aceptable del mismo. El ácido nucleico generalmente comprende ADN o ARN. Además, el nucleósido o nucleótido terminador 2', o sal farmacéuticamente aceptable del mismo, no puede ser extendido por el biocatalizador que incorpora nucleótidos. El biocatalizador que incorpora nucleótidos extiende el ácido nucleico,

produciendo por lo menos un ácido nucleico extendido mediante la incorporación del nucleósido o nucleótido terminador 2', o la sal farmacéuticamente aceptable del mismo, en un extremo terminal del ácido nucleico, inhibiendo de esta manera la extensión adicional del ácido nucleico extendido. A título ilustrativo, en el caso de que el ácido nucleico comprenda ADN microbiano, el biocatalizador que incorpora nucleótidos y el nucleósido o nucleótido terminador 2', o la sal farmacéuticamente aceptable del mismo, se pone en contacto generalmente en un huésped infectado con un microbio que comprende el ADN microbiano.

En otro aspecto, la invención proporciona un método para secuenciar un ácido nucleico diana. El método incluye: (a) incubar el ácido nucleico diana con una o más polimerasas, uno o más 2'-monofosfato-3'-hidroxil-nucleósidos (por ejemplo 2'-monofosfato-3'-hidroxil-5'-trifosfato nucleósidos, 2'-monofosfato-3'-hidroxil-5'-difosfato nucleósidos, etc.), uno o más nucleótidos extensibles y uno o más cebadores que son complementarios respecto a por lo menos una subsecuencia del ácido nucleico diana. Las polimerasas extienden los cebadores, produciendo productos de extensión de los cebadores que incorporan los 2'-monofosfato-3'-hidroxil-nucleósidos en los extremos 3'-terminales de los productos de extensión de cebador. Algunas realizaciones, (a) comprenden incubar el ácido nucleico diana, las polimerasas, los nucleótidos extensibles y los cebadores de ácidos nucleicos con por lo menos dos 2'-monofosfato-3'-hidroxil-nucleósidos diferentes. Otras realizaciones, (a) comprenden múltiples reacciones separadas en las que por lo menos dos de las reacciones comprenden 2'-monofosfato-3'-hidroxil-nucleósidos diferentes. En estas realizaciones, los 2'-monofosfato-3'-hidroxil-nucleósidos diferentes comprenden opcionalmente diferentes marcajes. El método incluye además: (b) identificar los 2'-monofosfato-3'-hidroxil-nucleósidos en los productos de extensión de cebador, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico diana sea determinable a partir de los 2'-monofosfato-3'-hidroxil-nucleósidos identificados. Por ejemplo, (b) opcionalmente comprende determinar las masas moleculares de los productos de extensión de cebador o los fragmentos 3'-terminales de los mismos y la secuencia del ácido nucleico diana a partir de las masas moleculares. Las masas moleculares generalmente se determinan utilizando espectrometría de iones en fase gaseosa. En algunas realizaciones, los productos de extensión de cebador comprenden marcajes y (b) comprenden separar los productos de extensión de cebador unos de otros y detectar las señales detectables producidas por los marcajes. Los productos de extensión de cebador típicamente se separan mediante una o más técnicas de separación, incluyendo, por ejemplo, la electroforesis, la cromatografía, la espectrometría de iones en fase gaseosa, etc.

En todavía otros aspectos, la invención proporciona una mezcla de reacción que comprende por lo menos un nucleótido terminador 2' marcado tal como se indica en la presente memoria (por ejemplo un 2'-monofosfato-3'-hidroxil-nucleósido, tal como un 2'-monofosfato-3'-hidroxil-5'-trifosfato-nucleósido, un 2'-monofosfato-3'-hidroxil-5'-difosfato-nucleósido, etc.) y por lo menos un biocatalizador que incorpora nucleótidos, tal como se indica en la presente memoria. En algunas realizaciones, la mezcla de reacción incluye además por lo menos una pirofosfatasa (por ejemplo una pirofosfatasa termoestable, etc.). La mezcla de reacción opcionalmente incluye además uno o más nucleótidos extensibles (por ejemplo ribonucleótidos, desoxirribonucleótidos y/o similares). Opcionalmente se marca por lo menos uno de los nucleótidos extensibles. En determinadas realizaciones, la mezcla de reacción incluye además un ácido nucleico de molde y un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico de molde. Opcionalmente, el ácido nucleico de molde o el cebador de ácidos nucleicos se une (por ejemplo covalente o no covalentemente) a un soporte sólido. En algunas de dichas realizaciones, el cebador comprende un marcaje. Por ejemplo, un marcaje utilizado tal como se indica en la presente memoria opcionalmente comprende un pigmento fluorescente (por ejemplo seleccionado de entre los pigmentos de la familia de la fluoresceína, los pigmentos de la familia de la polihalofluoresceína, los pigmentos de la familia de la hexaclorofluoresceína, los pigmentos de la familia de la coumarina, los pigmentos de la familia de la rodamina, los pigmentos de la familia de la cianina, los pigmentos de la familia de la oxazina, los pigmentos de la familia de la tiazina, los pigmentos de la familia de la escuaraina, los pigmentos de la familia de los lantánidos quelados y los pigmentos de la familia de BODIPY®).

En otros aspecto, la invención proporciona un kit para extender un ácido nucleico (por ejemplo para marcar el ácido nucleico, para secuenciar ácidos nucleicos diana, etc.). El kit incluye: (a) por lo menos un biocatalizador que incorpora nucleótidos tal como se indica en la presente memoria, y (b) por lo menos un nucleótido terminador 2' marcado tal como se indica en la presente memoria. Por ejemplo, el nucleótido terminador 2' comprende por lo menos un marcaje (por ejemplo enzimas (por ejemplo fosfatasa alcalina y peroxidasa de rábano picante) y sustratos enzimáticos, fracciones radioactivas, fracciones fluorescentes, cromóforos, marcajes quimioluminiscentes, marcajes electroquimioluminiscentes, tales como Origin™ (Igen), grupos modificadores de la masa, ligandos con parejas de unión específicas, etc.). En algunas realizaciones, el kit incluye además uno o más nucleótidos extensibles y, opcionalmente, por lo menos uno los nucleótidos extensibles comprende un marcaje. Opcionalmente, el kit incluye además por lo menos una pirofosfatasa, tal como una pirofosfatasa termoestable. Típicamente, el kit incluye además: (c) un juego de instrucciones para extender el cebador de ácidos nucleicos con el biocatalizador que incorpora nucleótidos y el nucleótido terminador 2'. Además, el kit opcionalmente incluye además: (d) por lo menos un recipiente para empaquetar el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' y el juego de instrucciones. En determinadas realizaciones, el kit incluye además un ácido nucleico de molde y un cebador de ácidos nucleicos, en el que el cebador de ácidos nucleicos es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico de molde. Opcionalmente, el ácido nucleico de molde o el cebador de ácidos nucleicos se une a un soporte sólido. En algunas de dichas realizaciones, el cebador comprende un marcaje, tal como un isótopo radioactivo, un pigmento fluorescente, un grupo modificador de la masa, o similar.

En otros aspectos, la invención se refiere a un sistema para extender un cebador de ácidos nucleicos. El sistema incluye: (a) por lo menos un recipiente que comprende un nucleótido terminador 2' marcado. Típicamente, el sistema comprende una pluralidad de recipientes. El sistema incluye además: (b) por lo menos un modulador térmico operablemente conectado al recipiente para modular la temperatura en el mismo, y/o (c) por lo menos un componente de transferencia de líquidos que transfiere líquido hacia y/o desde el recipiente. El sistema opcionalmente incluye además por lo menos un detector operablemente conectado al recipiente para detectar señales detectables producidas en el recipiente. El sistema típicamente incluye además por lo menos un controlador operablemente conectado al modulador térmico para llevar a cabo la modulación de la temperatura en el recipiente y/o al componente de transferencia de líquidos para llevar a cabo la transferencia de líquido hacia y/o desde el recipiente.

En otros aspectos, la invención proporciona un ordenador o medio legible por ordenador que comprende un conjunto de datos que comprende por lo menos un carácter correspondiente a por lo menos un nucleótido terminador 2' marcado tal como se indica en la presente memoria. Típicamente, el conjunto de datos comprende una pluralidad de cadenas de caracteres correspondiente a una pluralidad de secuencias de ácidos nucleicos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las figuras 1A-D ilustran esquemáticamente nucleótidos terminadores 2' según determinadas realizaciones de la invención.

Las figuras 2A y B ilustran esquemáticamente nucleótidos terminadores 2' según algunas realizaciones de la invención.

Las figuras 3A-C ilustran esquemáticamente tetrafosfatos marcados con pigmento según diversas realizaciones de la invención.

Las figuras 4A y B muestran esquemáticamente nucleótidos tetrafosfato marcados según determinadas realizaciones de la invención.

La figura 5 ilustra esquemáticamente un marcaje unido a un nucleótido tetrafosfato mediante un conector según una realización de la invención.

Las figuras 6A-D muestran esquemáticamente diversos nucleótidos terminadores 2' que presentan unidos pigmentos fluorescentes según determinadas realizaciones de la invención.

La figura 7 es un perfil espectral que muestra los datos de un análisis de secuencias de un molde de ADN M13mp18 utilizando nucleótidos terminadores 2' no marcados y un cebador marcado con pigmento fluorescente.

Las figuras 8A y B son perfiles espectrales que muestran los datos de un análisis de secuencias de un molde de ADN M13mp18 utilizando un cebador no marcado y un nucleótido terminador 2' marcado con pigmento fluorescente.

DESCRIPCIÓN DETALLADA DE LA INVENCION

I. DEFINICIONES

Antes de describir la presente invención en detalle, debe entenderse que la presente invención no se encuentra limitada a métodos, mezclas de reacción, sistemas, ordenadores o medios legibles por ordenador particulares, los cuales, evidentemente, pueden variar. También debe entenderse que la terminología utilizada en la presente memoria presenta el propósito de describir realizaciones particulares únicamente y no pretende ser limitativa. Además, a menos que se indique lo contrario, todos los términos técnicos y científicos utilizados en la presente memoria presentan los mismos significados entendidos comúnmente por el experto ordinario en la materia a la que se refiere la presente invención. En la descripción y reivindicación de la presente invención, se utilizan las variantes terminológicas y gramaticales siguientes según las definiciones proporcionadas a continuación.

La expresión "ácido nucleico" se refiere a nucleótidos (por ejemplo ribonucleótidos, desoxirribonucleótidos, nucleótidos terminadores 2', dideoxinucleótidos, etc.) y polímeros (por ejemplo que comprenden ácidos desoxirribonucleicos (ADN), ácidos ribonucleicos (ARN), híbridos de ADN-ARN, oligonucleótidos, polinucleótidos, genes, ADNc, aptámeros, ácidos nucleicos antisentido, ARN interfirientes (ARNi), balizas moleculares, sondas de ácidos nucleicos, ácidos péptido-nucleicos (APN), conjugados de APN-ADN, conjugados de APN-ARN, etc.) que comprenden dichos nucleótidos unidos covalentemente entre sí, de una manera lineal o ramificada.

Un ácido nucleico típicamente es de cadena sencilla o de doble cadena y generalmente contiene enlaces fosfodiéster, aunque en algunos casos, tal como se indica de manera general en la presente memoria, se incluyen análogos de ácidos nucleicos que pueden presentar esqueletos alternativos, incluyendo, por ejemplo y sin limitación,

fosforamida (Beaucage *et al.*, Tetrahedron 49(10):1925, 1993) y referencias en la misma; Letsinger, J. Org. Chem. 35:3800, 1970; Sprinzl *et al.*, Eur. J. Biochem. 81:579, 1977; Letsinger *et al.*, Nucl. Acids Res. 14:3487, 1986; Sawai *et al.*, Chem. Lett. 805, 1984; Letsinger *et al.*, J. Am. Chem. Soc. 110:4470, 1988; and Pauwels *et al.*, Chemica Scripta 26: 1419, 1986; fosforotioato (Mag *et al.*, Nucleic Acids Res. 19:1437, 1991; y la patente US nº 5.644.048); fosforoditioato (Briu *et al.*, J. Am. Chem. Soc. 111:2321, 1989), enlaces O-metilfosforoamidita (ver Eckstein, Oligonucleotides and Analogies: A Practical Approach, Oxford University Press, 1992), y esqueletos y enlaces de ácidos péptido-nucleicos (ver Egholm, J. Am. Chem. Soc. 114:1895, 1992; Meier *et al.*, Chem. Int. Ed. Engl. 31:1008, 1992; Nielsen, Nature 365:566, 1993; Carlsson *et al.*, Nature 380: 207, 1996). Entre otros ácidos nucleicos análogos se incluyen aquellos con esqueletos cargados positivamente (Denpcy *et al.*, Proc. Natl. Acad. Sci. USA 92: 6097, 1995); esqueletos no iónicos (patentes US nº 5.386.023, nº 5.637.684, nº 5.602.240, nº 5.216.141 y nº 4.469.863; Angew, Chem. Intl. Ed. English 30: 423, 1991; Letsinger *et al.*, J. Am. Chem. Soc. 110:4470, 1988; Letsinger *et al.*, Nucleoside & Nucleotide 13:1597, 1994; capítulos 2 y 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", editores: Y. S. Sanghvi y P. Dan Cook; Mesmaeker *et al.*, Bioorganic & Medicinal Chem. Lett. 4: 395, 1994; Jeffs *et al.*, J. Biomolecular NMR 34:17, 1994; Tetrahedron Lett. 37:743, 1996) y esqueletos no de ribosa, incluyendo los indicados en las patentes US nº 5.235.033 y nº 5.034.506, y los capítulos 6 y 7, ASC Symposium Series 580, Carbohydrate Modifications in Antisense Research, editores: Y.S. Sanghvi y P. Dan Cook. Los ácidos nucleicos que contienen uno o más azúcares carbocíclicos también se encuentran incluidos dentro de la definición de ácidos nucleicos (ver Jenkins *et al.*, Chem. Soc. Rev. páginas 169-176, 1995). También se describen varios análogos de ácidos nucleicos en, por ejemplo, Rawls C. y E. News, 2 de junio de 1997, página 35. Estas modificaciones del esqueleto de ribosa-fosfato pueden llevarse a cabo para facilitar la adición de fracciones adicionales, tales como marcajes, o para alterar la estabilidad y vida media de dichas moléculas en medios fisiológicos.

Además de dichas bases heterocíclicas naturales que se encuentran presentes típicamente en los ácidos nucleicos (por ejemplo adenina, guanina, timina, citosina y uracilo), los análogos de ácidos nucleicos también incluyen los que presentan bases heterocíclicas no naturales, muchas de las cuales se describen, o se hace referencia a ellas, en la presente memoria. En particular, se describen adicionalmente muchas bases no naturales en, por ejemplo, Seela *et al.*, Helv. Chim. Acta 74:1790, 1991; Grein *et al.*, Bioorg. Med. Chem. Lett. 4:971-976, 1994, y Seela *et al.*, Helv. Chim. Acta 82:1640, 1999. A título ilustrativo adicional, opcionalmente se incluyen determinadas bases utilizadas en los nucleótidos que actúan como modificadores de la temperatura de fusión (T_f). Por ejemplo, entre algunos de ellos se incluyen las 7-deazapurinas (por ejemplo la 7-deazaguanina, la 7-deazaadenina, etc.), las pirazolo[3,4-d]pirimidinas, propinil-dN (por ejemplo propinil-dU, propinil-dC, etc.) y similares. Ver, por ejemplo, la patente US nº 5.990.303, titulada "Synthesis of 7-deaza-2'-deoxyguanosine nucleotides", concedida el 23 de noviembre de 1996 a Seela. Entre otras bases heterocíclicas representativas se incluyen, por ejemplo, hipoxantina, inosina, xantina, derivados 8-aza de la 2-aminopurina, 2,6-diaminopurina, 2-amino-6-cloropurina, hipoxantina, inosina y xantina; derivados 7-deaza-8-aza de adenina, guanina, 2-aminopurina, 2,6-diaminopurina, 2-amino-6-cloropurina, hipoxantina, inosina y xantina; 6-azacitosina; 5-fluorocitosina; 5-clorocitosina; 5-yodocitosina; 5-bromocitosina; 5-metilcitosina; 5-propinilcitosina; 5-bromoviniluracilo; 5-fluorouracilo; 5-clorouracilo; 5-yodouracilo; 5-bromouracilo; 5-trifluorometiluracilo; 5-metoximetiluracilo; 5-etiniluracilo; 5-propiniluracilo y similares.

Un "nucleósido" se refiere a un componente ácido nucleico que comprende una base o grupo básico (por ejemplo que comprende por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo y/o similar) unido covalentemente a una fracción de azúcar (por ejemplo un azúcar ribosa, etc.) o un derivado de una fracción de azúcar, o un equivalente funcional de una fracción de azúcar (por ejemplo un análogo, tal como un anillo carbocíclico). Por ejemplo, en el caso de que un nucleósido incluya una fracción de azúcar, la base típicamente se une a una posición 1' de dicha fracción de azúcar. Tal como se ha indicado anteriormente, una base puede ser natural (por ejemplo una base purina, tal como adenina (A) o guanina (G); una base pirimidina, tal como timina (T), citosina (C) o uracilo (U)), o no natural (por ejemplo una base 7-deazapurina, una base pirazolo[3,4-d]pirimidina, una base propinil-dN, etc.). Entre los nucleósidos ejemplares se incluyen ribonucleósidos, desoxirribonucleósidos, dideoxirribonucleósidos, nucleósidos carbocíclicos, etc.

Un "nucleótido" se refiere a un éster de un nucleósido, por ejemplo un éster fosfato de un nucleósido. Por ejemplo, un nucleótido puede incluir 1, 2, 3 ó más grupos fosfato unidos covalentemente a una posición 5' de una fracción de azúcar del nucleósido.

Un "oligonucleótido" se refiere a un ácido nucleico que incluye por lo menos dos nucleótidos, típicamente más de tres nucleótidos, y más típicamente más de diez nucleótidos. El tamaño exacto de un oligonucleótido generalmente depende de diversos factores, incluyendo la función o uso último del oligonucleótido. Los oligonucleótidos se preparan opcionalmente mediante cualquier método, incluyendo, por ejemplo, la clonación y digestión de restricción de secuencias apropiadas, o la síntesis química directa mediante un método tal como el método del fosfotriéster de Narang *et al.* (Meth. Enzymol. 68:90-99, 1979); el método del fosfodiéster de Brown *et al.*, Meth. Enzymol. 68:109-151, 1979; el método de la dietilfosforamidita de Beaucage *et al.*, Tetrahedron Lett. 22:1859-1862, 1981); el método del triéster de Matteucci *et al.*, J. Am. Chem. Soc. 103:3185-3191, 1981; los métodos de síntesis automatizada, o el método en soporte sólido de la patente US nº 4.458.066, entre otros métodos conocidos de la técnica.

Un "cebador de ácidos nucleicos" típicamente es un ácido nucleico que puede hibridarse con un ácido nucleico

molde y que permite la extensión o alargamiento de la cadena utilizando, por ejemplo, un biocatalizador que incorpora nucleótidos, tal como una polimerasa termoestable, bajo condiciones de reacción apropiadas. Un cebador de ácidos nucleicos típicamente es un oligonucleótido natural o sintético (por ejemplo un oligodesoxirribonucleótido de cadena sencilla, etc.). Aunque opcionalmente se utilizan otras longitudes de cebador de ácidos nucleicos, típicamente se encuentran comprendidas entre 15 y 35 nucleótidos. Las moléculas de cebador de ácidos nucleicos cortas generalmente requieren temperaturas más bajas para formar complejos híbridos suficientemente estables con los ácidos nucleicos de molde. Un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a una subsecuencia de un ácido nucleico de molde típicamente es suficiente para hibridarse con el ácido nucleico de molde para que se produzca la extensión. Un cebador de ácidos nucleicos puede marcarse, si se desea, mediante la incorporación de un marcaje detectable, por ejemplo mediante técnicas espectroscópicas, fotoquímicas, bioquímicas, inmunoquímicas o químicas. A título ilustrativo, entre los marcajes útiles se incluyen isótopos radioactivos, pigmentos fluorescentes, reactivos electrodenso, enzimas (utilizados comúnmente en los ensayos de ELISA), biotina o haptenos y proteínas para las que se dispone de antisueros o anticuerpos monoclonales. Muchos de dichos marcajes y otros marcajes se describen en mayor detalle en la presente memoria y/o son conocidos de la técnica. Además, un cebador de ácidos nucleicos puede proporcionar simplemente un sustrato para un biocatalizador que incorpora nucleótidos de una manera independiente del molde.

Un "cebador de ácidos nucleicos extendido" se refiere a un cebador de ácidos nucleicos al que se han añadido o incorporado de otra manera uno o más nucleótidos adicionales (por ejemplo se han unido covalentemente).

Un "ácido nucleico molde" se refiere a un ácido nucleico al que puede hibridarse y ser extendido un cebador de ácidos nucleicos. De acuerdo con lo anteriormente expuesto, entre los ácidos nucleicos de molde se incluyen subsecuencias que son por lo menos parcialmente complementarias a los ácidos nucleicos molde. Los ácidos nucleicos molde pueden derivarse esencialmente de cualquier fuente. A título ilustrativo, los ácidos nucleicos de molde opcionalmente se derivan o se aíslan a partir de, por ejemplo, microorganismos en cultivo, microorganismos no en cultivo, mezclas biológicas complejas, tejidos, sueros, sueros o tejidos agrupados, asociaciones multiespecíficas, restos biológicos fosilizados o de otro tipo no vivo, aislados del medio ambiente, tierras, aguas subterráneas, instalaciones de gestión de residuos, ambientes marinos abisales o similares. Además, los ácidos nucleicos de molde opcionalmente incluyen o se derivan a partir de, por ejemplo, moléculas individuales de ADNc, juegos clonados de ADNc, bibliotecas de ADNc, ARN extraídos, ARN naturales, ARN transcritos *in vitro*, ADN genómicos caracterizados o no caracterizados, ADN genómicos clonados, bibliotecas de ADN genómico, ADN o ARN fragmento enzimáticamente, ADN o ARN fragmentado químicamente, ADN o ARN físicamente fragmentado, o similares. Los ácidos nucleicos de molde también pueden sintetizarse químicamente utilizando técnicas conocidas de la técnica. Además, los ácidos nucleicos de molde opcionalmente corresponden a por lo menos una parte de un gen o son complementarios de la misma. Tal como se utiliza en la presente memoria, un "gen" se refiere a cualquier segmento de ADN asociado a una función biológica. De esta manera, entre los genes se incluyen secuencias codificantes y opcionalmente las secuencias reguladoras requeridas para su expresión. Entre los genes opcionalmente también se incluyen segmentos de ADN no expresados que forman, por ejemplo, secuencias de reconocimiento para otras proteínas.

Los ácidos nucleicos se han "extendido" o "alargado" cuando se han incorporado nucleótidos adicionales (u otras moléculas análogas) a los ácidos nucleicos. Por ejemplo, un ácido nucleico se extiende opcionalmente con un biocatalizador que incorpora nucleótidos, tal como una polimerasa, que típicamente añade nucleótidos al extremo 3'-terminal de un ácido nucleico.

Un "nucleótido extensible" se refiere a un nucleótido al que puede añadirse o unirse covalentemente por lo menos otro nucleótido, por ejemplo en una reacción catalizada por un biocatalizador que incorpora nucleótidos tras incorporar el nucleótido extensible en un polímero de nucleótidos. Entre los ejemplos de nucleótidos extensibles se incluyen los desoxirribonucleótidos y los ribonucleótidos. Un nucleótido extensible típicamente se extiende mediante la adición de otro nucleótido a una posición 3' de la fracción azúcar del nucleótido extensible.

Un nucleótido "no extensible" se refiere a un nucleótido, que tras la incorporación en un ácido nucleico evita la extensión adicional del ácido nucleico, por ejemplo por como mínimo un biocatalizador que incorpora nucleótidos.

Un "nucleótido terminador 2'" se refiere a un análogo de nucleótido que comprende un grupo bloqueante (GB) en la posición 2' de la fracción sacárida del nucleótido. Un "grupo bloqueante" se refiere a un grupo o fracción química que típicamente evita la extensión de un ácido nucleico (es decir, un nucleótido terminador 2' típicamente no es extensible por uno o más biocatalizadores que incorporan nucleótidos). Es decir, tras incorporar un nucleótido terminador 2' en un ácido nucleico (por ejemplo en un extremo 3'-terminal del ácido nucleico), el grupo bloqueante evita la extensión adicional de un ácido nucleico por parte de por lo menos un biocatalizador que incorpora nucleótidos, seleccionado de entre, por ejemplo, ADN polimerasa CS5 G46E E678G, ADN polimerasa CS5 G46E L329A E678G, ADN polimerasa CS6 G46E E678G, ADN polimerasa Δ ZO5R, polimerasa ZO5, ADN polimerasa Taq E615G, una polimerasa de *Thermus flavus* (Tfl) (por ejemplo una polimerasa Tfl modificada que incorpora los nucleótidos terminadores 2' indicados en la presente memoria), polimerasa de *Thermatoga maritima* o Tma-25, la polimerasa Tma-30, ADN polimerasas de *Thermus thermophilus* (Tth), polimerasa SPS-17 de una especie de *Thermus*, polimerasa Taq E615G, polimerasa ZO5R de *Thermus*, ADN polimerasa de T7, ADN polimerasa I de

Kornberg o ADN polimerasa I de *E. coli*, ADN polimerasa Klenow, ADN polimerasa Taq, ADN polimerasa microcócica, ADN polimerasa α , transcriptasa inversa, transcriptasa inversa del AMV, transcriptasa inversa M-MuLV, ADN polimerasa, ARN polimerasa, ARN polimerasa de *E. coli*, ARN polimerasa SP6, ARN polimerasa de T3, ADN polimerasa de T4, ARN polimerasa de T7, ARN polimerasa II, transferasa terminal, polinucleótido fosforilasa (PNP),
 5 ADN polimerasa que incorpora ribonucleótidos y/o similares. Un grupo bloqueante ejemplar es un grupo fosfato. En la presente memoria también se describen otros grupos bloqueantes representativos. Entre los nucleótidos terminadores 2' ejemplares se incluyen los nucleósidos 2'-monofosfato-3'-hidroxilo-5'-trifosfato y los nucleósidos 2'-monofosfato-3'-hidroxilo-5'-difosfato. También se describen adicionalmente otros nucleótidos terminadores 2' en la
 10 presente memoria y en, por ejemplo, la solicitud provisional de patente US nº 60/519.661, titulada "Synthesis and compositions of 2'-terminator nucleotides", presentada el 12 de noviembre de 2003, de Gelfand *et al.*

Un "nucleótido tetrafosfato" se refiere a un nucleótido que incluye cuatro grupos fosfato. Entre los nucleótidos tetrafosfato se incluyen los nucleósidos 2'-monofosfato-5'-trifosfato y los nucleósidos 3'-monofosfato-5'-trifosfato.

15 Un "grupo bloqueante cargado negativamente" se refiere a un grupo bloqueante que comprende por lo menos una carga negativa, la cual por lo menos contribuye a la propiedad de no extensibilidad del nucleótido al que se encuentra unido, por ejemplo mediante repulsión electrostática de los nucleótidos entrantes. A título ilustrativo, entre los grupos bloqueantes cargados negativamente en las posiciones 2' de los nucleótidos de la invención se incluyen
 20 opcionalmente los grupos fosfato, carboxi u otros grupos a los que se hace referencia en la presente memoria, los cuales típicamente comprenden por lo menos una carga negativa con la ionización. En determinadas realizaciones, múltiples factores pueden contribuir a la propiedad de no extensibilidad de un nucleótido de la invención, incluyendo, por ejemplo, la carga y tamaño del grupo bloqueante.

Un "grupo bloqueante voluminoso" se refiere a un grupo bloqueante que comprende un tamaño suficiente para impedir estéricamente la incorporación de un nucleótido entrante, contribuyendo por lo menos de esta manera a la
 25 propiedad de no extensibilidad del nucleótido al que se encuentra unido el grupo bloqueante. Tal como se ha indicado anteriormente, en algunas realizaciones de la invención, múltiples factores pueden contribuir a la propiedad de no extensibilidad de un nucleótido terminador 2', incluyendo, por ejemplo, la carga y tamaño del grupo
 30 bloqueante.

Una "fracción" o "grupo" se refiere a una de las partes en las que algo, tal como una molécula, se divide (por ejemplo un grupo funcional, un grupo sustituyente, o similar). Por ejemplo, un nucleótido típicamente comprende un grupo
 35 básico (por ejemplo adenina, timina, citosina, guanina, uracilo o un grupo básico análogo), una fracción sacárida (por ejemplo una fracción que comprende un anillo de azúcar o un análogo del mismo) y uno o más grupos fosfato.

Un grupo "modificador de la masa" modifica la masa, típicamente medida en términos de peso molecular en daltons, de una molécula que comprende el grupo. Por ejemplo, los grupos modificadores de la masa que incrementan el
 40 poder de discriminación de tamaño o secuencia entre por lo menos dos ácidos nucleicos con diferencias de una sola base, pueden utilizarse para facilitar la secuenciación utilizando, por ejemplo, determinaciones del peso molecular.

Un "anillo heterocíclico" se refiere a un anillo monocíclico o bicíclico que se encuentra saturado, insaturado o es aromático, y que comprende uno o más heteroátomos seleccionados independientemente de entre nitrógeno,
 45 oxígeno y azufre. Puede unirse un anillo heterocíclico a la fracción sacárida, o análogo de la misma, de un nucleótido de la invención mediante cualquier heteroátomo o átomo de carbono. Entre los anillos heterocíclicos ejemplares se incluyen morfolinilo, pirrolidinonilo, pirrolidinilo, piperidinilo, hidantoinilo, valerolactamilo, oxiranilo, oxetanilo, tetrahidrofuranilo, tetrahidropiranilo, tetrahidropiridinilo, tetrahidropirimidinilo, tetrahidrotiofenilo, tetrahidrotiopiranilo, tetrahidropirimidinilo, tetrahidrotiofenilo, tetrahidrotiopiranilo, furilo, benzofuranilo, tiofenilo, benzotiofenilo, pirrolilo, indolilo, isoindolilo, azaindolilo, piridilo, quinolinilo, isoquinolinilo, oxazolilo, isooxazolilo, benzoxazolilo, pirazolilo, imidazolilo, bencimidazolilo, tiazolilo, benzotiazolilo, isotiazolilo, piridazinilo, pirimidinilo,
 50 pirazinilo, triazinilo, cinnolinilo, ftalazinilo, quinazolinilo, y similares.

Un "anillo homocíclico" se refiere a un anillo carbocíclico saturado o insaturado (pero no aromático), tal como ciclopropano, ciclobutano, ciclopentano, ciclohexano, cicloheptano, ciclohexeno y similares.

Un "grupo alquilo" se refiere a una fracción de hidrocarburo saturado lineal, ramificada o cíclica e incluye todos los isómeros posicionales, por ejemplo metilo, etilo, propilo, 1-metiletilo, butilo, 1-metilpropilo, 2-metilpropilo, 1,1-
 55 dimetiletilo, pentilo, 1-metilbutilo, 2-metilbutilo, 3-metilbutilo, 2,2-dimetilpropilo, 1-etilpropilo, hexilo, 1,1-dimetilpropilo, 1,2-dimetilpropilo, 1-metilpentilo, 2-metilpentilo, 3-metilpentilo, 4-metilpentilo, 1,1-dimetilbutilo, 1,2-dimetilbutilo, 1,3-dimetilbutilo, 2,2-dimetilbutilo, 2,3-dimetilbutilo, 3,3-dimetilbutilo, 1-etilbutilo, 2-etilbutilo, 1,1,2-trimetilpropilo, 1,2,2-
 60 trimetilpropilo, 1-etil-1-metilpropilo y 1-etil-2-metilpropilo, n-hexilo, ciclohexilo, n-heptilo, n-octilo, 2-etilhexilo, n-nonilo, n-decilo y similares. Un grupo alquilo típicamente comprende aproximadamente 1 a 20 átomos de carbono y más típicamente comprende aproximadamente 2 a 15 átomos de carbono. Los grupos alquilo pueden encontrarse sustituidos o no sustituidos.

Un "grupo alquínico" se refiere a una fracción de hidrocarburo insaturado lineal, ramificado o cíclico que comprende uno o más dobles enlaces carbono-carbono. Entre los grupos alquénico ejemplares se incluyen etenilo, 2-propenilo,

2-butenilo, 3-butenilo, 1-metil-2-propenilo, 2-metil-2-propenilo, 2-pentenilo, 3-pentenilo, 4-pentenilo, 1-metil-2-butenilo, 2-metil-2-butenilo, 3-metil-2-butenilo, 1-metil-3-butenilo, 2-metil-3-butenilo, 3-metil-3-butenilo, 1,1-dimetil-2-propenilo, 1,2-dimetil-2-propenilo, 1-etil-2-propenilo, 2-hexenilo, 3-hexenilo, 4-hexenilo, 5-hexenilo, 1-metil-2-pentenilo, 2-metil-2-pentenilo, 3-metil-2-pentenilo, 4-metil-2-pentenilo, 1-metil-3-pentenilo, 2-metil-3-pentenilo, 3-metil-3-pentenilo, 4-metil-3-pentenilo, 1-metil-4-pentenilo, 2-metil-4-pentenilo, 3-metil-4-pentenilo, 4-metil-4-pentenilo, 1,1-dimetil-2-butenilo, 1,1-dimetil-3-butenilo, 1,2-dimetil-2-butenilo, 1,2-dimetil-3-butenilo, 1,3-dimetil-2-butenilo, 1,3-dimetil-3-butenilo, 2,2-dimetil-3-butenilo, 2,3-dimetil-2-butenilo, 2,3-dimetil-3-butenilo, 3,3-dimetil-2-butenilo, 1-etil-2-butenilo, 1-etil-3-butenilo, 2-etil-2-butenilo, 2-etil-3-butenilo, 1,1,2-trimetil-2-propenilo, 1-etil-1-metil-2-propenilo, 1-etil-2-metil-2-propenilo, y similares. Un grupo alqueno típicamente comprende aproximadamente 1 a 20 átomos de carbono y más típicamente comprende aproximadamente 2 a 15 átomos de carbono. Los grupos alqueno pueden encontrarse sustituidos o no sustituidos.

Un "grupo alquino" se refiere a una fracción de hidrocarburo insaturado lineal, ramificado o cíclico que comprende uno o más triples enlaces carbono-carbono. Entre los grupos alquino representativos se incluyen, por ejemplo, 2-propinilo, 2-butinilo, 3-butinilo, 1-metil-2-propinilo, 2-pentinilo, 3-pentinilo, 4-pentinilo, 1-metil-2-butinilo, 1-metil-3-butinilo, 2-metil-3-butinilo, 1,1-dimetil-2-propinilo, 1-etil-2-propinilo, 2-hexinilo, 3-hexinilo, 4-hexinilo, 5-hexinilo, 1-metil-2-pentinilo, 1-metil-3-pentinilo, 1-metil-4-pentinilo, 2-metil-3-pentinilo, 2-metil-4-pentinilo, 3-metil-4-pentinilo, 4-metil-2-pentinilo, 1,1-dimetil-2-butinilo, 1,1-dimetil-3-butinilo, 1,2-dimetil-3-butinilo, 2,2-dimetil-3-butinilo, 3,3-dimetil-1-butinilo, 1-etil-2-butinilo, 1-etil-3-butinilo, 2-etil-3-butinilo, 1-etil-1-metil-2-propinilo, y similares. Un grupo alquino típicamente comprende aproximadamente 1 a 20 átomos de carbono y más típicamente comprende aproximadamente 2 a 15 átomos de carbono. Los grupos alquino pueden encontrarse sustituidos o no sustituidos.

Un "grupo alcoxi" se refiere a un grupo alquilo que comprende un átomo de oxígeno e incluye, por ejemplo, metoxi, etoxi, propoxi, butoxi, pentoxi, heptiloxi, octiloxi y similares.

Un "grupo halo" se refiere a un grupo que comprende un átomo de halógeno, tal como F, Cl, Br o I.

Un "grupo arilo" se refiere a un grupo sustituyente de átomos o fracción que se deriva de un compuesto aromático. Entre los grupos arilo ejemplares se incluyen, por ejemplo, grupos fenilo, grupos bencilo, grupos toliilo, grupos xililo o similares. Entre los grupos arilo opcionalmente se incluyen múltiples anillos aromáticos (por ejemplo grupos difenilo, etc.). Además, un grupo arilo puede encontrarse sustituido o no sustituido.

Un "grupo ariloxi" se refiere a un grupo arilo que comprende un átomo de oxígeno e incluye, por ejemplo, fenoxi, clorofenoxi, metilfenoxi, metoxifenoxi, butilfenoxi, pentilfenoxi, benciloxi y similares.

Un "grupo alquilarilo" se refiere a un grupo que comprende fracciones alquilo y arilo.

Un "grupo éter" se refiere a una fracción lineal, ramificada o cíclica que comprende dos átomos de carbono unidos a un único átomo de oxígeno. Entre los grupos éter ejemplares se incluyen, por ejemplo, metoximetilo, metoxietilo, metoxipropilo, etoxietilo y similares.

Un "grupo tioéter" se refiere a una fracción lineal, ramificada o cíclica que comprende dos átomos de carbono unidos a un único átomo de azufre e incluye, por ejemplo, metiletioetilo, metiltioetilo, metiltiopropilo y similares.

Un "grupo alquilamina" se refiere a un grupo amino que comprende por lo menos un grupo alquilo.

Un "grupo alquencilamina" se refiere a un grupo amino que comprende por lo menos un grupo alqueno.

Un "grupo alquencilamina" se refiere a un grupo amino que comprende por lo menos un grupo alquino.

Un "grupo éster" se refiere a una clase de compuestos orgánicos que incluye la fórmula general RCOOR', en la que R y R' se seleccionan independientemente de entre un grupo alquilo, un grupo alqueno, un grupo alquino, un grupo arilo o combinaciones de los mismos.

Un "poliaminoácido" se refiere a un compuesto o grupo que comprende dos o más residuos aminoácidos. Entre los poliaminoácidos ejemplares se incluyen péptidos, polipéptidos, proteínas y similares.

Un "heterooligo" se refiere a un oligonucleótido que comprende dos o más residuos nucleótidos diferentes.

Un "grupo heterooligo/poliaminoácido" se refiere a un grupo híbrido que comprende por lo menos una fracción heterooligo y por lo menos una fracción poliaminoácido.

Un "grupo aldehído" se refiere a un grupo orgánico que incluye la fórmula CHO.

Un "grupo alcohol" se refiere a un grupo orgánico que incluye por lo menos un grupo hidroxilo.

Un "grupo sililo" se refiere a una clase de compuestos orgánicos que incluye la fórmula general SiRR'R", en la que R, R' y R" son independientemente un H, un grupo alquilo, un grupo alquenoilo, un grupo alquinoilo, un grupo arilo o una combinación de dichos grupos.

5 Una "secuencia" de un ácido nucleico se refiere al orden e identidad de los nucleótidos en el ácido nucleico. Una secuencia típicamente se lee en la dirección 5' a 3'.

10 Una "secuencia de longitud completa" se refiere a una secuencia de ácidos nucleicos que comprende por lo menos sustancialmente el mismo número de nucleótidos que una secuencia de referencia o una secuencia de ácidos nucleicos que es por lo menos parcialmente complementaria a la secuencia de referencia. En determinadas realizaciones de la invención, por ejemplo un cebador de ácidos nucleicos extendido es complementario a una secuencia de longitud completa de un ácido nucleico molde u otra secuencia de referencia.

15 Una "subsecuencia" o "fragmento" se refiere a cualquier porción de una secuencia de ácidos nucleicos completa.

20 Un "genotipo" se refiere a la totalidad o a parte de la constitución genética de una célula o sujeto, o de un grupo de células o sujetos. Por ejemplo, un genotipo incluye las mutaciones y/o alelos particulares (por ejemplo polimorfismos, tales como polimorfismos de nucleótidos únicos (SNP) o similares) presentes en un locus dado o distribuidos en un genoma.

El término "unido" se refiere a interacciones entre las que se incluyen, aunque sin limitarse a ellas, el enlace covalente, el enlace iónico, la quimisorción, la fisorción y las combinaciones de los mismos.

25 Un "conector" o "espaciador" se refiere a una fracción química que une covalente o no covalentemente (por ejemplo iónicamente, etc.) un compuesto o grupo sustituyente a, por ejemplo, un soporte sólido, otro compuesto o grupo, o similar. Por ejemplo, un conector opcionalmente une un marcaje (por ejemplo un pigmento fluorescente, un isótopo radioactivo, etc.) a un nucleótido terminador 2' o similar. Los conectores típicamente son fracciones químicas bifuncionales, y en determinadas realizaciones comprenden uniones cortables que pueden cortarse mediante, por ejemplo, calor, un enzima, un agente químico, radiación electromagnética, etc., liberando materiales o compuestos a partir de, por ejemplo, un soporte sólido, otro compuesto, etc. Una elección cuidadosa de conector permite llevar a cabo el corte bajo condiciones apropiadas que son compatibles con la estabilidad del compuesto y del método de ensayo. Generalmente un conector no presenta actividad biológica específica aparte de, por ejemplo, la unión de especies químicas entre sí o la conservación de cierta distancia mínima u otra relación espacial entre dichas especies. Sin embargo, los constituyentes de un conector pueden seleccionarse para influir sobre cierta propiedad de las especies químicas unidas, tales como la conformación tridimensional, la carga neta, la hidrofobicidad, etc. Se proporciona una descripción adicional de las moléculas de conector en, por ejemplo, Lyttle *et al.*, *Nucleic Acids Res.* 24(14):2793, 1996; Shchepino *et al.*, *Nucleosides, Nucleotides & Nucleic Acids* 20:369, 2001; Doronina *et al.*, *Nucleosides, Nucleotides, & Nucleic Acids* 20:1007, 2001; Trawick *et al.*, *Bioconjugate Chem.* 12:900, 2001; Olejnik *et al.*, *Methods in Enzymology* 291:135, 1998; Pljevaljic *et al.*, *J. Am. Chem. Soc.* 125(12):3486, 2003; Ward, *et al.*, patente US n° 4.711.955; Stavrianopoulos, patente US n° 4.707.352, y Stavrianopoulos, patente US n° 4.707.440.

45 Un "biocatalizador que incorpora nucleótidos" se refiere a un catalizador que cataliza la incorporación de nucleótidos en un ácido nucleico. Los biocatalizadores que incorporan nucleótidos típicamente son enzimas. Un "enzima" es un catalizador basado en una proteína, que actúa reduciendo la energía de activación de una reacción química en la que participan otros compuestos o "sustratos". Un "enzima que incorpora nucleótidos" se refiere a un enzima que cataliza la incorporación de nucleótidos en un ácido nucleico. Entre los enzimas ejemplares que incorporan nucleótidos se incluyen, por ejemplo, ADN polimerasas, ARN polimerasas, transferasas terminales, transcriptasas inversas, telomerasas, polinucleótido fosforilasas, y similares. Otros biocatalizadores pueden estar basados en ADN ("ADNzimas") o en ARN ("ribozimas").

50 Un "enzima termoestable" se refiere a un enzima que es estable frente al calor (es decir, resiste a la degradación o a la desnaturalización) y conserva suficiente actividad catalítica al someterlo a temperaturas elevadas durante periodos seleccionados de tiempo. Por ejemplo, una polimerasa termoestable conserva suficiente actividad para llevar a cabo reacciones de extensión de cebador posteriores al someterla a temperaturas elevadas durante el tiempo necesario para llevar a cabo la desnaturalización de los ácidos nucleicos de doble cadena. Las condiciones de calentamiento necesarias para la desnaturalización son bien conocidas de la técnica y se encuentran ejemplificadas en las patentes US n° 4.683.202 y n° 4.683.195, las cuales se incorporan como referencia. Tal como se utiliza en la presente memoria, una polimerasa termoestable típicamente resulta adecuada para la utilización en una reacción de ciclado de la temperatura, tal como la reacción en cadena de la polimerasa ("PCR"). Para un enzima termoestable que incorpora nucleótidos, la actividad enzimática se refiere a la catálisis de la combinación de los nucleótidos de la manera adecuada para formar productos de extensión de cebador que sean complementarios a una cadena de ácido nucleico molde. Entre otros enzimas termoestables indicados en la presente memoria se incluyen pirofosfatasa termoestables, que de manera similar conservan suficiente actividad al someterlas a temperaturas elevadas, por ejemplo para minimizar la pirofosforólisis. De manera similar a los enzimas, los ADNzimas y los ribozimas también pueden ser termoestables.

Un enzima "modificado" se refiere a un enzima que comprende una secuencia de monómeros en la que por lo menos un monómero de la secuencia difiere de un monómero en una secuencia de referencia, tal como na forma nativa o de tipo salvaje del enzima u otra forma modificada del enzima, por ejemplo en el caso de que se alineen dos secuencias para conseguir la máxima identidad. Entre las modificaciones ejemplares se incluyen las inserciones, 5 delecciones y sustituciones de monómeros. Los enzimas modificados (es decir, los catalizadores basados en proteínas o en ácidos nucleicos) de la invención han sido creados, o se crean opcionalmente, mediante diversos métodos generadores de diversidad. Aunque puede utilizarse esencialmente cualquier método para producir un enzima modificado, entre determinadas técnicas ejemplares se incluyen la recombinación (por ejemplo mediante recombinación recursiva, recombinación sintética o similar) de dos o más ácidos nucleicos codificantes de uno o más 10 enzimas parentales, o mediante la mutación de uno o más ácidos nucleicos que codifican enzimas, por ejemplo utilizando mutagénesis de conjunto recursiva, mutagénesis por inserción de casete, mutagénesis aleatoria, mutagénesis *in vivo*, mutagénesis sitio-dirigida, o similares. Un ácido nucleico codificante de un enzima parental típicamente incluye un gen que, mediante los mecanismos de transcripción y traducción, produce una secuencia de aminoácidos correspondiente a un enzima parental, por ejemplo una forma nativa del enzima. Los enzimas 15 modificados también incluyen enzimas quiméricos que presentan secuencias componente identificables (por ejemplo dominios estructurales y/o funcionales, etc.) derivadas de dos o más progenitores. También se encuentran incluidos en la definición de enzimas modificados aquellos que comprenden modificaciones químicas (por ejemplo grupos sustituyentes unidos, grupos sustituyentes alterados, etc.) respecto a una secuencia de referencia. De manera similar a los enzimas, los ADNzimas y los ribozimas también pueden comprender modificaciones similares.

Un "marcaje" se refiere a una fracción unida (covalente o no covalentemente) o capaz de unirse a una molécula, proporcionando dicha fracción, o siendo capaz de proporcionar, información sobre la molécula (por ejemplo información descriptiva, identificativa, etc., sobre la molécula). Entre los marcajes ejemplares se incluyen los marcajes fluorescentes, los marcajes débilmente fluorescentes, los marcajes no fluorescentes, los marcajes 20 colorimétricos, los marcajes quimioluminiscentes, los marcajes bioluminiscentes, los marcajes radioactivos, los grupos modificadores de la masa, los anticuerpos, los antígenos, la biotina, los haptenos y los enzimas (incluyendo, por ejemplo, peroxidasa, fosfatasa, etc.).

Un "soporte sólido" se refiere a un material sólido que puede derivatizarse o unirse de otro modo con una fracción química, tal como un cebador de ácidos nucleicos, un ácido nucleico molde, o similar. Entre los soportes sólidos ejemplares se incluyen una placa, una perla, una microperla, una fibra, una cerda, un peine, un chip de hibridación, una membrana, un cristal simple, una capa cerámica, una monocapa autoensamblante, y similares.

La expresión "en solución" se refiere a una condición de reacción en la que por lo menos los reactivos no se unen a un soporte sólido. Por ejemplo, entre determinaciones reacciones de extensión de la invención se incluyen la incubación de ácidos nucleicos molde, cebadores de ácidos nucleicos, nucleótidos terminadores 2', nucleótidos extensibles y biocatalizadores que incorporan nucleótidos conjuntamente en solución.

El término "corte" se refiere a un procedimiento de liberación de un material o compuesto respecto de otro compuesto o material o respecto de un soporte sólido, por ejemplo para permitir el análisis del compuesto mediante métodos en fase solución. Ver, por ejemplo, Wells *et al.*, "Cleavage and Analysis of Material from Single Resin Beads", J. Org. Chem. 63:6430-6431, 1998.

Un "carácter", utilizado en referencia a un carácter de una cadena de caracteres, se refiere a una subunidad de la cadena. En una realización preferente, el carácter de una cadena de caracteres codifica una subunidad de una molécula biológica codificada. De esta manera, por ejemplo en el caso de que la molécula biológica codificada sea un polinucleótido u oligonucleótido, un carácter de la cadena codifica un único nucleótido.

Una "cadena de caracteres" representa cualquier entidad capaz de almacenar información de secuencia (por ejemplo la estructura de subunidades de una molécula biológica, tal como la secuencia de nucleótidos de un ácido nucleico, etc.). En una realización, la cadena de caracteres puede ser una secuencia simple de caracteres (letras, números u otros símbolos) o puede ser una representación numérica de dicha información en forma tangible o intangible (por ejemplo electrónica, magnética, etc.). La cadena de caracteres no es necesario que sea "lineal", sino que también puede existir en varias otras formas, por ejemplo de lista encadenada u otra matriz no lineal (por ejemplo utilizada como un código para generar una matriz lineal de caracteres) o similar. Las cadenas de caracteres preferentemente son aquellas que codifican cadenas polinucleótidas, directa o indirectamente, incluyendo cualesquiera cadenas encriptadas, o imágenes, o disposiciones de objetos que pueden transformarse sin ambigüedad en cadenas de caracteres que representan secuencias de monómeros o multímeros en polinucleótidos, o similares (constituidos de monómeros naturales o artificiales).

El término "resolución" se refiere a la identificación de una o más propiedades de por lo menos determinados miembros de una población dada. En algunas realizaciones de la invención, por ejemplo, los biocatalizadores que incorporan nucleótidos producen múltiples cebadores de ácidos nucleicos extendidos diferentes, los cuales se resuelven de manera que por lo menos una parte de una secuencia de bases de un ácido nucleico molde sea determinable a partir de los cebadores de ácidos nucleicos extendidos resueltos. A título ilustrativo adicional, una población de cebadores de ácidos nucleicos extendidos se resuelve opcionalmente mediante la determinación de las

masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos individuales en la población. En algunas realizaciones, los cebadores de ácidos nucleicos extendidos marcados se resuelven mediante la separación de los cebadores de ácidos nucleicos extendidos en una población y la detección de las señales detectables producidas por estos marcajes.

La expresión "espectrometría de iones en fase gaseosa" se refiere a la utilización de un espectrómetro de iones en fase gaseosa para detectar iones en fase gaseosa. Los espectrómetros de iones en fase gaseosa incluyen una fuente de iones que proporciona iones en fase gaseosa. Entre los espectrómetros de iones en fase gaseosa se incluyen los espectrómetros de masas, los dispositivos de medición de corrientes iónicas totales, los espectrómetros de movilidad iónica, y similares.

Un "espectrómetro de masas" es un instrumento analítico que puede utilizarse para determinar los pesos moleculares de diversas sustancias, tales como productos de una reacción catalizada enzimáticamente. Típicamente, un espectrómetro de masas comprende cuatro partes: un puerto de introducción de muestra, una fuente de ionización, un analizador de masas y un detector. Se introduce opcionalmente una muestra mediante diversos tipos de puertos de introducción, por ejemplo una sonda sólida, cromatografía de gases (CG) o cromatografía de líquidos (CL) en fase gaseosa, líquida o sólida. A continuación, la muestra típicamente se ioniza en la fuente de ionización para formar uno o más iones. Los iones resultantes se introducen y son manipulados en el analizador de masas (por ejemplo un analizador de masas de tiempo de vuelo ("time of flight", TOF), etc.). Los iones supervivientes se detectan a partir de las proporciones entre masa y carga. En una realización, el espectrómetro de masas bombardea la sustancia investigada con un láser o haz de electrones y registra cuantitativamente el resultado en forma de espectro de fragmentos iónicos positivos y negativos. La separación de los fragmentos iónicos se realiza basándose en la proporción entre masa y carga de los iones. En el caso de que todos los iones presenten una sola carga, dicha separación se basa esencialmente en la masa. Un espectrómetro de masas con cuadrupolo utiliza cuatro polos eléctricos para el analizador de masas. Estas técnicas se describen en mayor detalle en muchos textos, entre ellos, por ejemplo, Dawson, *Quadrupole Mass Spectrometry and its Applications*, Springer Verlag, 1995. En un sistema de espectrometría de masas por electropulverización, la ionización se produce con un campo eléctrico que se utiliza para generar gotas cargadas y posteriormente iones analito mediante evaporación iónica. Ver Cole, "Electrospray Ionization Mass Spectrometry" John Wiley and Sons, Inc., 1997.

Una "mezcla" se refiere a una combinación de dos o más componentes diferentes. Una "mezcla de reacción" se refiere a una mezcla que comprende moléculas que pueden participar y/o facilitar una reacción dada. Por ejemplo, una "mezcla de reacción de secuenciación de ADN" se refiere a una mezcla de reacción que comprende componentes necesarios para una reacción de secuenciación de ADN. De esta manera, una mezcla de reacción de secuenciación de ADN resulta adecuada para la utilización en un método de secuenciación de ADN para determinar la secuencia de ácidos nucleicos de un ácido nucleico molde o diana, aunque la mezcla de reacción inicialmente puede encontrarse incompleta, de manera que el inicio de la reacción de secuenciación es controlado por el usuario. De esta manera, la reacción puede iniciarse tras la adición de un componente final, tal como el enzima, proporcionando una mezcla de reacción de secuenciación de ADN completa. Típicamente, una reacción de secuenciación de ADN contiene un tampón, adecuado para la actividad de polimerización, nucleótidos extensibles y por lo menos un nucleótido terminador 2'. La mezcla de reacción también puede contener un cebador de ácidos nucleicos adecuado para la extensión en un ácido nucleico molde realizada por un enzima polimerasa. El cebador de ácidos nucleicos o uno de los nucleótidos generalmente se marca con una fracción detectable, tal como un marcaje fluorescente. Generalmente, la reacción es una mezcla que comprende cuatro nucleótidos extensibles y por lo menos un nucleótido terminador 2'. Típicamente, la polimerasa es una ADN polimerasa termoestable (por ejemplo una ADN polimerasa CS5 G46E E678G, una ADN polimerasa CS6 G46E E678G, una ADN polimerasa Taq E615G, una ADN polimerasa Δ ZO5R, una ADN polimerasa CS5 G46E L329A E678G, etc.) y el nucleótido terminador 2' es un nucleótido 2'-monofosfato-3'-hidroxil-5'-trifosfato.

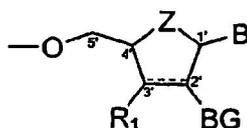
II. NUCLEÓTIDOS TERMINADORES 2'

La presente invención se refiere de manera general a métodos de marcaje terminal y/o de bloqueo de la extensión dependiente de molde de ácidos nucleicos utilizando nucleótidos terminadores 2', que típicamente incluyen un grupo hidroxilo en una posición 3' de un anillo sacárido intacto (por ejemplo anillos sacáridos de pentosa) o anillos de análogos de sacárido (por ejemplo anillos carbocíclicos, etc.) y un grupo bloqueante (por ejemplo un grupo bloqueante cargado negativamente, un grupo bloqueante voluminoso y/o similares) en una posición 2' de la fracción sacárida. Los biocatalizadores que incorporan nucleótidos de la invención comprenden la capacidad de extender, por ejemplo, cebadores de ácidos nucleicos con nucleótidos terminadores 2' de una manera dirigida por un molde. En determinadas realizaciones, los biocatalizadores que incorporan nucleótidos extienden ácidos nucleicos independientemente de un ácido nucleico molde, tal como en el caso de ácidos nucleicos que se marcan terminalmente utilizando los nucleótidos terminadores 2' indicados en la presente memoria. Tras la incorporación de un nucleótido terminador 2' en el extremo terminal de, por ejemplo, un cebador de ácidos nucleicos extendido, el ácido nucleico típicamente deja de ser extensible por un biocatalizador que incorpora nucleótidos de la invención. Además, inesperadamente un cebador de ácidos nucleicos extendido que comprende un nucleótido terminador 2' es generalmente resistente a la actividad enzimática correctora de errores (por ejemplo una actividad de exonucleasa 3'-5' de una ADN polimerasa con corrección de errores, etc.). En consecuencia, un biocatalizador que

incorpora nucleótidos utilizado en un método de la invención opcionalmente incluye una actividad exonucleasa 3'-5', por ejemplo para mejorar la fidelidad de la secuencia respecto a enfoques que utilizan catalizadores que presentan actividades correctoras de errores reducidas o nulas. En determinadas realizaciones de la invención, los métodos de secuenciación utilizan una ADN polimerasa que no presenta una mutación de F por Y en la hélice O del enzima o que, de otra manera, no presenta una mutación que incrementa la incorporación de desoxinucleótidos 3' por parte del enzima.

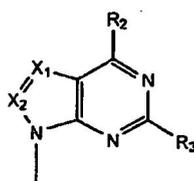
A título ilustrativo, las figuras 1A-D ilustran esquemáticamente nucleótidos terminadores 2' según determinadas realizaciones de la invención. En particular, la figura 1A muestra esquemáticamente un nucleósido terminador adenosina tetrafosfato; la figura 1B ilustra esquemáticamente un nucleósido terminador guanosina tetrafosfato; la figura 1C ilustra esquemáticamente un nucleósido terminador uridina tetrafosfato, y la figura 1D muestra esquemáticamente un nucleósido terminador citidina tetrafosfato.

Un nucleótido terminador 2' según la presente invención generalmente incluye la fórmula:

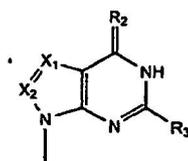


en la que R_1 es H, OH, un grupo hidrofílico, o un grupo hidrofóbico; B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico (con o sin heteroátomos exocíclicos) o por lo menos un grupo arilo, o combinaciones de los mismos; BG es un grupo bloqueante; Z es O o CH_2 y --- representa un enlace sencillo o doble. En determinadas realizaciones, un nucleótido de la invención se encuentra marcado. Además, un nucleótido terminador 2' generalmente comprende 1, 2, 3 ó más grupos fosfato unidos en la posición 5'. En una realización de la invención, por ejemplo, el nucleótido comprende un nucleósido 2'-monofosfato-3'-hidroxil-5'-trifosfato.

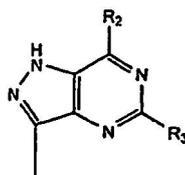
Los nucleótidos terminadores 2' de la invención opcionalmente incluyen esencialmente cualquier anillo heterocíclico o grupo arilo (es decir, como la base o grupo B). De acuerdo con lo anteriormente expuesto, no se realiza ningún intento en la presente memoria por describir la totalidad de los posibles grupos que pueden utilizarse. Sin embargo, a título ilustrativo, los grupos B en los que las bases se aparean con otro ácido nucleico, por ejemplo mediante un enlace de hidrógeno o mediante un mecanismo de apilamiento de bases, se incluyen en la posición 1' de la fracción sacárida de los nucleótidos y nucleótidos en determinadas realizaciones de la invención. A título ilustrativo adicional de aspectos de la invención, se proporcionan posteriormente determinados grupos B representativos. En algunas realizaciones, por ejemplo, B comprende la fórmula:



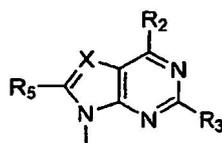
en la que X_1 y X_2 se seleccionan independientemente de entre CH y N; R_2 es H, OH o NR_4R_5 ; R_3 es H, OH o NR_6R_7 ; R_4 y R_5 se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo y un grupo ariloxi, y combinaciones de los mismos, y R_6 y R_7 se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo y un grupo ariloxi, y combinaciones de los mismos. En otras realizaciones, B comprende la fórmula:



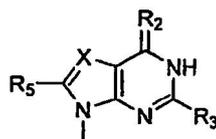
en la que X_1 y X_2 se seleccionan independientemente de entre CH y N; R_2 es O o S; R_3 es H, OH o NR_4R_5 ; y R_4 y R_5 se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo y un grupo ariloxi, y combinaciones de los mismos. En algunas realizaciones, B comprende la fórmula:



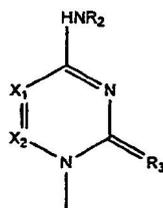
5 en la que R_2 es H, OH o NR_4R_5 ; R_3 es H, OH o NR_6R_7 ; R_4 y R_5 se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo y un grupo ariloxi, y combinaciones de los mismos, y R_6 y R_7 se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo y un grupo ariloxi, y combinaciones de los mismos. En algunas realizaciones, B comprende la fórmula:



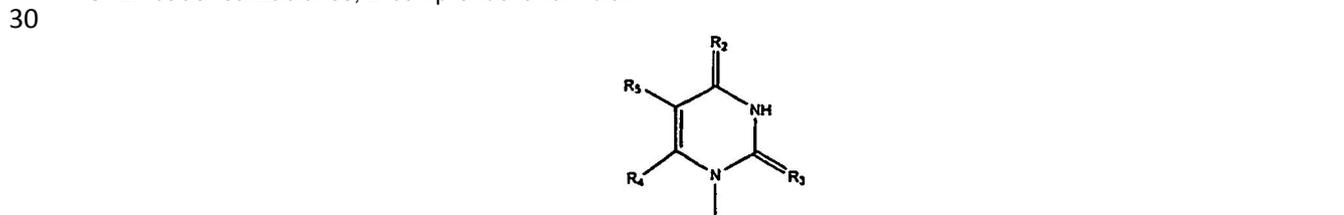
10 en la que X es CH o N; R_2 y R_3 se seleccionan independientemente de entre H, OH y NHR_4 ; R_4 es H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo o un grupo ariloxi, o combinaciones de los mismos, y R_5 es OH, NH_2 , SH, un grupo halo, un grupo éter, un grupo tioéter, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo alquilamina, un grupo alquenilamina o un grupo alquinilamina, o combinaciones de los mismos. En otras realizaciones, B comprende la fórmula:



20 en la que X es CH o N; R_2 es O o S; R_3 es H, OH o NHR_4 ; R_4 es H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo o un grupo ariloxi, o combinaciones de los mismos, y R_5 es OH, NH_2 , SH, un grupo halo, un grupo éter, un grupo tioéter, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo alquilamina, un grupo alquenilamina o un grupo alquinilamina, o combinaciones de los mismos. En determinadas realizaciones, B comprende la fórmula:

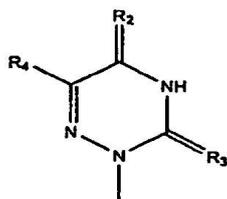


25 en la que X_1 y X_2 se seleccionan independientemente de entre CH y N; R_2 es H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo o un grupo ariloxi, o combinaciones de los mismos, y R_3 es O o S. En otras realizaciones, B comprende la fórmula:

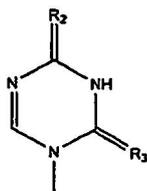


30 en la que R_2 y R_3 se seleccionan independientemente de entre O y S; y R_4 y R_5 se seleccionan independientemente de entre H, NH_2 , SH, OH, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi y un grupo halo, y combinaciones de los mismos. En algunas realizaciones, B comprende la fórmula:

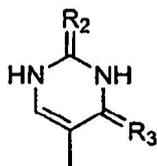
35



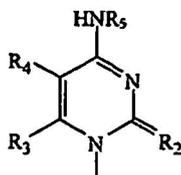
- 5 en la que R_2 y R_3 se seleccionan independientemente de entre O y S; y R_4 es H, NH_2 , SH, OH, un grupo alquilo, un grupo alquenoilo, un grupo alquinoilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi o un grupo halo, o combinaciones de los mismos. En otras realizaciones, B comprende la fórmula:



- 10 en la que R_2 y R_3 se seleccionan independientemente de entre O y S. En algunas realizaciones, B comprende la fórmula:



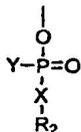
- 15 en la que R_2 y R_3 se seleccionan independientemente de entre O y S. En otras realizaciones, B comprende la fórmula:



- 20 en la que R_2 es O o S; R_3 y R_4 se seleccionan independientemente de entre H, NH_2 , SH, OH, COOH, COOCH_3 , $\text{COOCH}_2\text{CH}_3$, CHO, NO_2 , CN, un grupo alquilo, un grupo alquenoilo, un grupo alquinoilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi, un grupo halo y combinaciones de los mismos, y R_5 es un grupo alquilo, un grupo alcoxi, un grupo alquenoilo, un grupo alquenoilo, un grupo alquinoilo, un grupo alquinoxio, un grupo arilo, un grupo ariloxi, un grupo bencilo, un grupo benciloxi, o combinaciones de los mismos.

- 25 Los grupos bloqueantes (GB) utilizados en la posición 2' de la fracción sacárida también incluyen diversas realizaciones. En algunas realizaciones, por ejemplo, el GB es un grupo cargado negativamente y/o un grupo voluminoso. A título ilustrativo adicional, el GB se selecciona opcionalmente de entre, por ejemplo, CN, NO_2 , N_3 , un grupo sililo, un grupo halo, un grupo alcohol, un grupo éter, un grupo aldehído, un grupo ácido, un grupo éster, un grupo amino y combinaciones de los mismos. Más concretamente, el GB comprende opcionalmente la fórmula:

30



- 35 en la que X es O, S, NR_3 , CR_3R_4 , o SiR_3R_4 ; Y es $\text{CR}_5\text{R}_6\text{R}_7$, $\text{SiR}_5\text{R}_6\text{R}_7$, OR_5 , SR_5 , o NHR_5 ; R_2 es H, OH, NHR_8 , SR_8 , un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo, un grupo alquinoilo, un grupo alcoxi, o combinaciones de los mismos, y R_3 , R_4 , R_5 , R_6 , R_7 y R_8 se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo, un grupo alquinoilo o combinaciones de los mismos. La figura 2A ilustra esquemáticamente un nucleótido que comprende un grupo bloqueante que presenta dicha fórmula. A título ilustrativo adicional, el GB comprende opcionalmente la fórmula:



5 en la que X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃, o NHR₃; R₂ es H, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueno, un grupo alquino, un grupo alcoxi, o combinaciones de los mismos, y R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueno, un grupo alquino o combinaciones de los mismos. La figura 2B ilustra esquemáticamente un nucleótido terminador 2' que comprende un grupo bloqueante que presenta dicha fórmula.

10 Los nucleótidos terminadores 2', los nucleótidos extensibles, los cebadores de ácidos nucleicos (por ejemplo cebadores de ácidos nucleicos extendidos) y/o otros ácidos nucleicos utilizados según los métodos de la invención comprenden opcionalmente por lo menos un marcaje. Por ejemplo, el marcaje se une opcionalmente a, por ejemplo, un anillo homocíclico, un anillo heterocíclico o un grupo arilo del nucleótido terminador 2' (por ejemplo mediante C⁵ de una pirimidina, N⁴ de citidina, N⁷ de una purina, N⁶ de adenosina, C⁸ de una purina u otro sitio de unión conocido de la técnica), por ejemplo mediante una amida, éster, tioéster, éter, tioéter, enlace carbono-carbono u otro tipo de enlace covalente. Además, o alternativamente, el marcaje se une a una fracción sacárida (por ejemplo un sacárido ribosa, etc.) o un análogo de la misma (por ejemplo un anillo carbocíclico, etc.) de un nucleótido terminador 2' y/o un grupo fosfato de un nucleótido terminador 2', tal como mediante un enlace covalente que es un enlace amida, éster, tioéster, éter, tioéter, carbono-carbono u otro enlace. Los enlaces covalentes típicamente se forman en reacciones entre grupos electrofílicos y nucleofílicos de marcajes y nucleótidos de la invención. En determinadas realizaciones, los marcajes y nucleótidos se conjugan directamente entre sí (por ejemplo mediante enlaces carbono-carbono sencillos, dobles, triples o aromáticos, o mediante enlaces carbono-nitrógeno, nitrógeno-nitrógeno, carbono-oxígeno, carbono-azufre, fósforo-oxígeno, fósforo-nitrógeno, etc.). Opcionalmente, un conector une el marcaje al nucleótido terminador 2'. Puede utilizarse una amplia diversidad de conectores o adaptarse para la utilización en la conjugación de marcajes y nucleótidos. En la presente memoria se hace referencia a determinadas ilustraciones no limitativas de dichos conectores.

15 A título ilustrativo adicional, las figuras 3A-C ilustran esquemáticamente tetrafosfatos marcados con pigmento según determinadas realizaciones de la invención. En particular, la figura 3A muestra esquemáticamente un pigmento informador unido a una base de un nucleótido terminador 2'; la figura 3B ilustra esquemáticamente un pigmento informador unido a un grupo bloqueante de un nucleótido terminador 2', y la figura 3C muestra esquemáticamente un pigmento informador unido a una fracción sacárida de un nucleótido terminador 2'. Las figuras 4A y B muestran esquemáticamente nucleótidos tetrafosfato marcados según determinadas realizaciones de la invención. Más concretamente, las figuras 4A y B muestran esquemáticamente marcajes unidos mediante conectores a bases de los nucleósidos tetrafosfato, en los que R se selecciona de entre el grupo que consiste de: H, OH, un grupo alquilo, un grupo arilo, un grupo alquilo ramificado, un grupo alquilarilo ramificado, un grupo alqueno y un grupo alquino. Además, la figura 5 ilustra esquemáticamente un marcaje unido a un nucleósido tetrafosfato mediante un conector según una realización de la invención. Las figuras 6A-D muestran esquemáticamente diversos nucleótidos terminadores 2' que presentan unidos pigmentos fluorescentes según determinadas realizaciones de la invención. En particular, la figura 6A muestra esquemáticamente una adenosina tetrafosfato marcada con R6G; la figura 6B ilustra esquemáticamente una guanosina tetrafosfato marcada con R110; la figura 6C ilustra esquemáticamente una uridina tetrafosfato marcada con TAMRA, y la figura 6D muestra esquemáticamente una citidina tetrafosfato marcada con ROX.

20 Se utiliza opcionalmente esencialmente cualquier marcaje para marcar los nucleótidos y nucleósidos de la invención. En algunas realizaciones, por ejemplo, el marcaje comprende un pigmento fluorescente (por ejemplo un pigmento rodamina (por ejemplo R6G, R110, TAMRA, ROX, etc.), un pigmento fluoresceína (por ejemplo JOE, VIC, TET, HEX, FAM, etc.), un pigmento halofluoresceína, un pigmento cianina (por ejemplo CY3, CY3.5, CY5, CY5.5, etc.), un pigmento BODIPY[®] (por ejemplo FL, 530/550, TR, TMR, etc.), un pigmento ALEXA FLUOR[®] (por ejemplo 488, 532, 546, 568, 594, 555, 653, 647, 660, 680, etc.), un pigmento diclororrodamina, un pigmento de transferencia energética (por ejemplo pigmentos BIGDYE[™] v1, pigmentos BIGDYE[™] v2, pigmentos BIGDYE[™] v3, etc.), pigmentos Lucifer (por ejemplo amarillo Lucifer, etc.), CASCADE BLUE[®], verde Oregon, y similares. Se proporcionan detalles adicionales referentes a los pigmentos fluorescentes en, por ejemplo, Haugland, Molecular Probes Handbook of Fluorescent Probes and Research Products, novena edición (2003) y actualizaciones del mismo. Los pigmentos fluorescentes generalmente se encuentran fácilmente disponibles de diversos proveedores comerciales, incluyendo, por ejemplo, Molecular Probes, Inc. (Eugene, OR), Amersham Biosciences Corp. (Piscataway, NJ), Applied Biosystems (Foster City, CA), etc. Entre otros marcajes se incluyen, por ejemplo, biotina, marcajes débilmente fluorescentes (Yin *et al.*, Appl Environ Microbiol. 69(7):3938, 2003; Babendure *et al.*, Anal. Biochem. 317(1):1, 2003, y Jankowiak *et al.*, Chem Res Toxicol. 16(3):304, 2003), marcajes no fluorescentes, marcajes colorimétricos, marcajes quimioluminiscentes (Wilson *et al.*, Analyst 128(5):480, 2003; y Roda *et al.*, Luminescence 18(2):72, 2003), marcajes Raman, marcajes electroquímicos, marcajes bioluminiscentes (Kitayama *et al.*, Photochem Photobiol. 77(3): 333, 2003; Arakawa *et al.*, Anal. Biochem. 314(2):206, 2003, y Maeda, J. Pharm. Biomed. Anal. 30(6):1725, 2003), y un reactivo de marcaje α-metil-PEG tal como se indica en, por ejemplo, la solicitud provisional de patente US n° 60/428,484, presentada el 22 de noviembre de 2002.

En determinadas realizaciones, el marcaje comprende un isótopo radioactivo, tal como ^3H , ^{14}C , ^{22}Na , ^{32}P , ^{33}P , ^{35}S , ^{42}K , ^{45}Ca , ^{59}Fe , ^{125}I , ^{203}Hg o similar. A título de ejemplo adicional, el marcaje opcionalmente incluye además por lo menos un grupo modificador de la masa. Por ejemplo, el grupo modificador de la masa se selecciona opcionalmente de entre, por ejemplo, deuterio, F, Cl, Br, I, S, N_3 , XY, CH_3 , SPO_4 , BH_3 , SiY_3 , $\text{Si}(\text{CH}_3)_3$, $\text{Si}(\text{CH}_3)_2(\text{C}_2\text{H}_5)$, $\text{Si}(\text{CH}_3)(\text{C}_2\text{H}_5)_2$, $\text{Si}(\text{C}_2\text{H}_5)_3$, $(\text{CH}_2)_n\text{CH}_3$, $(\text{CH}_2)_n\text{NY}_2$, CH_2CONY_2 , $(\text{CH}_2)_n\text{OH}$, CH_2F , CHF_2 , CF_3 , y un grupo fosforotioato, en el que X es O, NH, NY, S, NHC(S), $\text{OCO}(\text{CH})_n\text{COO}$, $\text{NHCO}(\text{CH}_2)_n\text{COO}$, OSO_2O , $\text{OCO}(\text{CH}_2)_n$, $\text{NHC}(\text{S})\text{NH}$, $\text{OCO}(\text{CH}_2)_n\text{S}$, $\text{OCO}(\text{CH}_2)\text{S}$, $\text{NC}_4\text{O}_2\text{H}_2\text{S}$, $\text{OPO}(\text{O-alkyl})$, o $\text{OP}(\text{O-alkilo})$; n es un número entero de entre 1 y 20, ambos inclusive, e Y es H, deuterio, un grupo alquilo, un grupo alcoxi, un grupo arilo, un grupo polioximetileno, un grupo polioximetileno monoalquilado, un grupo polietilén-imina, un grupo poliamida, un grupo poliéster, un grupo sililo alquilado, un heterooligo, un poliaminoácido, un grupo heterooligo/poliaminoácido o un grupo polietilenglicol. Se proporcionan detalles adicionales referentes al marcaje de ácidos nucleicos y al análisis de secuencias en, por ejemplo, Sterky *et al.*, "Sequence analysis of genes and genomes", J. Biotech. 76(2000):1, 2000; Sensen (Ed.), Biotechnology, volumen 5B, Genomics and Bioinformatics, John Wiley & Sons, Inc., 2001; y Sensen (Ed.), Essentials of Genomics and Bioinformatics, John Wiley & Sons, Inc., 2002.

Se encuentra disponible una gran diversidad de conectores para unir marcajes a ácidos nucleicos y resultarán evidentes para el experto en la materia. Un conector generalmente presenta una estructura que es estérica y electrónicamente adecuada para la incorporación en un ácido nucleico. Entre los conectores opcionalmente se incluyen, por ejemplo, éter, tioéter, carboxamida, sulfonamida, urea, uretano, hidrazina u otras fracciones. A título ilustrativo adicional, los conectores generalmente incluyen entre aproximadamente uno y aproximadamente 25 átomos no hidrógeno seleccionados de entre, por ejemplo, C, N, O, P, Si, S, etc., y comprenden esencialmente cualquier combinación de, por ejemplo, enlaces éter, tioéter, amina, éster, carboxamida, sulfonamida, hidrazida y aromáticos o heteroaromáticos. En algunas realizaciones, por ejemplo, un conector comprende una combinación de enlaces sencillos carbono-carbono y enlaces carboxamida o tioéter. Aunque se utilizan opcionalmente segmentos lineales de conectores más largos, el segmento lineal más largo típicamente contiene entre aproximadamente 3 y aproximadamente 15 átomos no de hidrógeno, incluyendo uno o más heteroátomos.

Entre los ejemplos no limitativos de fracciones conectoras se incluyen grupos sustituidos (por ejemplo funcionalizados) o no sustituidos, tales como conectores imidazol/biotina, grupos polimetileno, grupos arileno, grupos alquilarileno, grupos arilentalquilo, grupos ariltio, grupos amidoalquilo, grupos alquinilalquilo, grupos alquenilalquilo, grupos alquilo, grupos alcoxi, grupos tio, grupos aminoalquilo, fosfatos derivatizados con morfina, ácidos péptido-nucleicos (por ejemplo N-(2-aminoetil)glicina, etc.), y similares. Algunos de estos conectores y otros se describen en mayor detalle en, por ejemplo, la patente US nº 6.339.392, de Haugland *et al.*; en la patente US nº 5.047.519, de Hobbs, Jr. *et al.*; en la patente US nº 4.711.958, de Iizuka *et al.*; en la patente US nº 5.175.269, de Stavrianopoulos; en la patente US nº 4.711.955, de Ward *et al.*; en la patente US nº 5.241.060, de Engelhardt *et al.*; en la patente US nº 5.328.824, de Ward *et al.*; y en la publicación de patente US nº 2002/0151711, de Khan *et al.* Se proporcionan detalles adicionales referentes al marcaje de ácidos nucleicos y conectores en, por ejemplo, Hermanson, Bioconjugate Techniques, Elsevier Science, 1996. En determinadas realizaciones, son conectores adecuados comprenden fracciones fotocortables, tales como fracciones 2-nitrobencilo, fracciones 2-nitrobencilo α -sustituidos (por ejemplo fracciones 1-(2-nitrofenil)etilo), fracciones 3,5-dimetoxibencilo, ácido tiohidroxámico, fracciones 7-nitroindolina, fracciones 9-fenilxantilo, fracciones benzoína, fracciones hidroxifenacilo, fracciones NHS-ASA, y similares. Los conectores fotocortables se describen en mayor detalle en, por ejemplo, la publicación de patente US nº 2003/0099972, de Olejnik *et al.* En algunas realizaciones, entre los conectores se incluyen metales, tales como átomos de platino. Estos se describen en mayor detalle en, por ejemplo, la patente US nº 5.714.327, de Houthoff *et al.* Se encuentran disponibles comercialmente varios conectores de longitud variable, disponibles de diversos proveedores, incluyendo, por ejemplo, Qiagen Operon Technologies, Inc. (Alameda, CA), BD Biosciences Clontech (Palo Alto, CA) y Molecular BioSciences (Boulder, CO). Los nucleótidos terminadores 2' también se describen en, por ejemplo, la solicitud provisional de patente US nº 60/519.661, titulada "Synthesis and compositions of 2'-terminator nucleotides", presentada el 12 de noviembre de 2003, de Gelfand *et al.*

III. MÉTODOS DE MARCAJE Y SECUENCIACIÓN

En determinados aspectos, la invención proporciona métodos para extender ácidos nucleicos (por ejemplo oligonucleótidos o similares), por ejemplo para marcar el extremo de ácidos nucleicos para la utilización como sondas, entre otras aplicaciones. Entre estos métodos típicamente se incluyen incubar ácidos nucleicos que deben extenderse con biocatalizadores que incorporan nucleótidos (por ejemplo transferasas terminales, polinucleótido fosforilasas, etc.) y nucleótidos terminadores 2' marcados. En algunas realizaciones, los biocatalizadores que incorporan nucleótidos extienden los ácidos nucleicos para producir ácidos nucleicos extendidos mediante la incorporación de nucleótidos terminadores 2' marcados en los extremos 3'-terminales de los ácidos nucleicos, por ejemplo de una manera independiente de molde. En el caso de que los ácidos nucleicos extendidos se utilizan como sondas, los métodos típicamente incluyen además la hibridación de los ácidos nucleicos extendidos con ácidos nucleicos diana y la detección de señales detectables producidas por los marcajes, detectando de esta manera los ácidos nucleicos diana.

En algunas realizaciones, los métodos de la invención incluyen incubar un ácido nucleico molde con por lo menos un biocatalizador que incorpora nucleótidos, por lo menos un nucleótido terminador 2' y por lo menos un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico molde. El biocatalizador que incorpora nucleótidos extiende el cebador de ácidos nucleicos para producir por lo menos un cebador de ácidos nucleicos extendido, incorporando el nucleótido terminador 2' en un extremo terminal del cebador de ácidos nucleicos extendido.

Los métodos de secuenciación de la invención típicamente también incluyen incubar el ácido nucleico molde con por lo menos un nucleótido extensible (por ejemplo un ribonucleótido, un desoxirribonucleótido y/o similar), que opcionalmente se encuentra marcado. El marcaje de ácidos nucleicos se ha descrito en mayor detalle anteriormente. Aunque se utilizan opcionalmente otras proporciones molares, los nucleótidos terminadores 2' y los nucleótidos extensibles típicamente se encuentran presentes en una proporción molar de 1:1 ó inferior. Los cebadores de ácidos nucleicos extendidos producidos mediante los métodos de la invención típicamente son complementarios a una subsecuencia del ácido nucleico molde o complementarios a una secuencia de longitud completa del ácido nucleico molde.

Los métodos de la invención también incluyen generalmente la incubación, por ejemplo del ácido nucleico molde, con por lo menos una pirofosfatasa (por ejemplo una pirofosfatasa termoestable). Se ha demostrado que la pirofosfatasa mejora los resultados de secuenciación utilizando tanto polimerasas mesófilas como ADN polimerasa termoestable mediante la reducción del nivel de pirofosforólisis a medida que se acumulan los productos de extensión. En algunas realizaciones, la pirofosfatasa no se encuentra incluida en la secuenciación del ADN o en otras mezclas de reacción. Más concretamente, la utilización de algunos de los enzimas indicados o a los que se hace referencia en la presente memoria elimina la necesidad del gasto adicional de adición de un segundo enzima a la mezcla de reacción de secuenciación.

En la práctica de la presente invención, se utilizan opcionalmente muchas técnicas convencionales de la biología molecular y de ADN recombinante. Estas técnicas son bien conocidas y se explican en, por ejemplo, *Current Protocols in Molecular Biology*, volúmenes I, II y III, 1997 (editor: F.M. Ausubel); Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, tercera edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001; Berger y Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, volumen 152, Academic Press, Inc., San Diego, CA (Berger), *DNA Cloning: A Practical Approach*, volúmenes I y II, 1985 (editor: D.N. Glover); *Oligonucleotide Synthesis*, 1984 (editor: M.L. Gait); *Nucleic Acid Hybridization*, 1985 (Hames y Higgins); *Transcription and Translation*, 1984 (editores: Hames y Higgins); *Animal Cell Culture*, 1986 (editor: R. I. Freshney); *Immobilized Cells and Enzymes*, 1986 (IRL Press); Perbal, 1984, *A Practical Guide to Molecular Cloning; the series, Methods in Enzymology* (Academic Press, Inc.); *Gene Transfer Vectors for Mammalian Cells*, 1987 (editores: J.H. Miller y M.P. Calos, Cold Spring Harbor Laboratory); *Methods in Enzymology*, volúmenes 154 y 155 (editores Wu y Grossman, y Wu, respectivamente).

Entre los ácidos nucleicos molde que pueden secuenciarse según los métodos descritos en la presente memoria se incluyen secuencias de ácido desoxirribonucleico (ADN) o de ácido ribonucleico (ARN). Estas secuencias pueden obtenerse de fuentes biológicas, recombinantes o de otras fuentes artificiales, o pueden purificarse a partir de fuentes naturales, incluyendo células, tejidos o fuentes ambientales. Entre otros tipos de moléculas que pueden secuenciarse se incluyen ácido nucleico-poliámidas (APN) (Nielsen *et al.*, *Science* 254:1497, 1991) o cualquier secuencia de bases unida mediante un esqueleto químico que puede formar pares de bases o hibridarse con una estructura química complementaria.

Entre las bases de ADN, ARN y APN se incluyen purinas, pirimidinas y derivados y modificaciones de purina y de pirimidina, las cuales se unen linealmente a un esqueleto químico. Son estructuras de esqueleto químico comunes, la desoxirribosa-fosfato, la ribosa-fosfato y la poliámidas. Las purinas de tanto el ADN como el ARN son la adenina (A) y la guanina (G). Entre otras cuya existencia se conoce se incluye la xantina, la hipoxantina, la 2-diaminopurina y la 1-diaminopurina y otras bases más modificadas. Las pirimidinas son la citosina (C), que es común tanto al ADN como al ARN, el uracilo (U), presente predominantemente en el ARN, y la timidina (T), que se encuentra casi

- exclusivamente en el ADN. Entre algunas de las pirimidinas más atípicas se incluyen la metilcitosina, la hidroximetilcitosina, el metiluracilo, el hidroximetiluracilo, el dihidroxipentiluracilo y otras modificaciones de bases. Estas bases interactúan de una manera complementaria formando pares de bases, incluyendo, por ejemplo, guanina con citosina y adenina con timidina. La presente invención se refiere además al apareamiento de bases no tradicional, tal como el apareamiento de bases de Hoogsteen, que ha sido identificado en determinadas moléculas de ARNt y que se conjetura que existe en un triple hélice. Los ácidos nucleicos se han descrito en mayor detalle anteriormente, incluyendo en la sección de definiciones.
- Los ácidos nucleicos molde se purifican opcionalmente, por ejemplo para eliminar sustancias que podrían resultar perjudiciales (por ejemplo toxinas), peligrosas (por ejemplo infecciosas) o que podrían interferir con la reacción de hibridación o con la sensibilidad de dicha reacción (por ejemplo metales, sales, proteínas, lípidos). La purificación puede implicar técnicas tales como la extracción química con sales, cloroformo o fenol; la sedimentación, la centrifugación, la cromatografía u otras técnicas conocidas por el experto ordinario en la materia.
- En el caso de que se disponga de cantidades suficientes de ácidos nucleicos molde y los ácidos nucleicos sean suficientemente puros o puedan purificarse de manera que cualesquiera sustancias que pudiesen interferir con la hibridación resulten eliminadas, los ácidos nucleicos molde podrán secuenciarse directamente. Es decir, la información de la secuencia puede obtenerse sin crear copias complementarias u homólogas de una secuencia diana. Sin embargo, los ácidos nucleicos molde también pueden amplificarse para incrementar el número de copias del molde utilizando, por ejemplo, reacciones en cadena de la polimerasa (PCR) u otras técnicas de amplificación. Opcionalmente también se utiliza un protocolo de amplificación de ácidos nucleicos para incrementar el número de copias del cebador de ácidos nucleicos utilizado en los métodos de la invención. La amplificación de ácidos nucleicos generalmente implica la desnaturalización del ADN molde mediante calentamiento en presencia de un gran exceso molar de cada uno de entre dos o más cebadores oligonucleótidos y cuatro dNTP (dGTP, dCTP, dATP, dTTP). La mezcla de reacción se enfría hasta una temperatura que permita hibridar el cebador oligonucleótido con secuencias diana, seguido de la extensión de los cebadores hibridados con ADN polimerasa. El ciclo de desnaturalización, hibridación y síntesis de ADN, la parte principal de la amplificación por PCR, se repite muchas veces para generar grandes cantidades de producto, que puede identificarse fácilmente.
- Aunque la PCR es un método fiable de amplificación de secuencias de molde, también pueden utilizarse algunos otros procedimientos, entre ellos, por ejemplo, la reacción en cadena de la ligasa, la replicación de secuencias autosostenida, la amplificación mediante la replicasa ϕ , la reacción en cadena de la polimerasa ligada a la reacción en cadena de la ligasa, la reacción en cadena de la ligasa con huecos, la detección de la cadena de la ligasa, la amplificación por círculo rodante y la amplificación por desplazamiento de cadena. El principio de la reacción en cadena de la ligasa se basa en parte en la ligación de dos cebadores oligonucleótidos sintéticos contiguos que se hibridan de manera única con una cadena del ADN o ARN diana. En el caso de que se encuentre presente la diana, los dos oligonucleótidos pueden ser unidos covalentemente por una ligasa. También se proporciona un segundo par de cebadores, casi totalmente complementarios al primer par de cebadores. El molde y los cuatro cebadores se introducen en un termociclador con una ligasa termoestable. A medida que se eleva y se reduce la temperatura, los oligonucleótidos se renaturalizan en posición inmediatamente contigua entre sí sobre el molde y se ligan. El producto ligado de una reacción sirve como molde para la siguiente ronda de ligación. La presencia de diana se manifiesta como un fragmento de ADN con una longitud igual a la suma de los dos oligonucleótidos contiguos.
- Se encuentran ejemplos de técnicas suficientes para dirigir expertos en la materia a través de los métodos de amplificación *in vitro*, incluyendo la reacción en cadena de la polimerasa (PCR), por ejemplo para la amplificación de ácidos nucleicos molde en una muestra o cebadores de ácidos nucleicos, diseño de cebadores de ácidos nucleicos, etc., en Berger, Sambrook y Ausubel, así como en Mullis *et al.*, patente US nº 4.683.202, 1987; PCR Protocols A Guide to Methods and Applications (Innis *et al.*, editores), Academic Press Inc., San Diego, CA, 1990 (Innis); Amheim y Levinson (1 de octubre de 1990), C&EN 36-47, The Journal of NIH Research 3:81-94, 1991; Kwok *et al.*, Proc. Natl. Acad. Sci. USA 86, 1173, 1989; Guatelli *et al.*, Proc. Natl. Acad. Sci. USA 87:1874, 1990; Lomell *et al.*, J. Clin. Chem 35:1826, 1989; Landegren *et al.*, Science 241:1077-1080, 1988; Van Brunt, Biotechnology 8:291-294, 1990; Wu y Wallace, Gene 4:560, 1989; Barringer *et al.*, Gene 89:117, 1990, y Sooknanan y Malek, Biotechnology 13:563, 1995.
- Los ácidos nucleicos molde opcionalmente se fragmentan en una pluralidad de fragmentos utilizando enfoques físicos, químicos o enzimáticos para crear un juego de fragmentos de longitud uniforme o relativamente uniforme. Por ejemplo, las secuencias se cortan enzimáticamente utilizando, por ejemplo, nucleasas tales como las ADNasas o ARNasas (nucleasa de judía mungo, nucleasa microcócica, ADNasa I, ARNasa I, ARNasa T1), endonucleasas de restricción de tipo I ó II, u otras endonucleasas específicas de sitio o no específicas. Los tamaños de los fragmentos de ácidos nucleicos típicamente se encuentran comprendidos entre aproximadamente 5 y aproximadamente 1.000 nucleótidos de longitud, más típicamente entre aproximadamente 10 y aproximadamente 200 nucleótidos, y todavía más típicamente entre aproximadamente 12 y aproximadamente 100 nucleótidos de longitud. Los tamaños del orden de aproximadamente 5, 10, 12, 15, 18, 20, 24, 26, 30 y 35 resultan útiles para llevar a cabo análisis a pequeña escala de regiones cortas de un ácido nucleico molde, mientras que los tamaños de fragmentos del orden de 25, 50, 75, 125, 150, 175, 200 y 250 nucleótidos y mayores típicamente resultan útiles para analizar rápidamente secuencias diana de mayor tamaño.

- Los cebadores de ácidos nucleicos, los ácidos nucleicos de molde y/o otros ácidos nucleicos opcionalmente se sintetizan químicamente, por ejemplo según el método de la fosforamidita triéster en fase sólida descrito por Beaucage y Caruthers, *Tetrahedron Letts.* 22(20):1859-1862, 1981, u otra técnica de síntesis conocida de la técnica, por ejemplo utilizando un sintetizador automático, tal como describen Needham-VanDevanter *et al.*, *Nucleic Acids Res.* 12:6159-6168, 1984. Se encuentran disponible comercialmente una amplia diversidad de equipos para la síntesis automática de oligonucleótidos. También se utilizan opcionalmente enfoques de síntesis multinucleótido (por ejemplo la síntesis trinucleótido).
- Además, puede obtenerse esencialmente cualquier ácido nucleico (y virtualmente cualquier ácido nucleico marcado, estándar o no estándar) mediante pedido estándar o específico, de cualquiera de entre una diversidad de proveedores comerciales, tales como Midland Certified Reagent Company (Midland, TX), The Great American Gene Company (Ramona, CA), ExpressGen Inc. (Chicago, IL), Operon Technologies Inc. (Alameda, CA) y muchos otros.
- La hibridación entre bases complementarias de ADN, ARN, APN o combinaciones de ADN, ARN y APN, se produce bajo una amplia diversidad de condiciones que varían en temperatura, concentración salina, fuerza electrostática, composición del tampón y similares. Los ejemplos de dichas condiciones y métodos para aplicarlas se describen en, por ejemplo, Hames y Higgins, *supra*. La hibridación generalmente tiene lugar entre aproximadamente 0°C y aproximadamente 70°C, durante periodos de entre aproximadamente un minuto y aproximadamente una hora, dependiendo de la naturaleza de la secuencia que debe hibridarse y de su longitud. Sin embargo, se reconoce que las hibridaciones pueden producirse en segundos o en horas, dependiendo de las condiciones de la reacción. A título ilustrativo, las condiciones típicas de hibridación para una mezcla de dos 20-meros es llevar la mezcla a 68°C, seguido del enfriamiento hasta la temperatura ambiente (22°C) durante cinco minutos o a temperaturas muy bajas, tales como 2°C en 2 microlitros. La hibridación entre ácidos nucleicos puede facilitarse utilizando tampones tales como Tris-EDTA (TE), Tris-HCl y HEPES, soluciones salinas (por ejemplo NaCl, KCl, CaCl₂) u otras soluciones acuosas, reactivos y productos químicos. Entre los ejemplos de dichos reactivos se incluyen proteínas de unión de cadena sencilla, tales como proteína Rec A, proteína del gen 32 de T4, proteína de unión de cadena sencilla de *E. coli* y proteínas de unión del surco mayor o menor de ácidos nucleicos. Entre otros ejemplos de dichos reactivos y productos químicos se incluyen iones divalentes, iones polivalentes y sustancias intercalantes, tales como bromuro de etidio, actinomicina D, psoralén y angelicina.
- En otras realizaciones de la invención, se incuban ácidos nucleicos molde con el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' y el cebador de ácidos nucleicos en solución. En otras realizaciones, el ácido nucleico de molde o el cebador de ácidos nucleicos se une (por ejemplo covalente o no covalentemente) a un soporte sólido. Entre los ejemplos de soporte sólidos que pueden utilizarse se incluyen un plástico, una cerámica, un metal, una resina, un gel y una membrana. Entre los tipos útiles de soportes sólidos se incluyen placas, perlas, microperlas, fibras, cerdas, chips de hibridación, membranas, cristales simples, cerámicas y monocapas autoensamblantes.
- Los ácidos nucleicos pueden unirse al soporte sólido mediante unión covalente, tal como mediante conjugación con un agente de acoplamiento o mediante unión no covalente, tal como interacción electrostática, enlace de hidrógeno o acoplamiento de anticuerpo-antígeno, o mediante combinaciones de los mismos. Entre los agentes de acoplamiento típicos se incluyen biotina/avidina, biotina/estreptavidina, proteína A de *Staphylococcus aureus*/fragmento Fc de anticuerpo IgG y quimeras de estreptavidina/proteína A (Sano *et al.*, *Bio/Technology* 9:1378, 1991), o derivados o combinaciones de dichos agentes. Los ácidos nucleicos pueden unirse al soporte sólido mediante un enlace fotocortable, un enlace electrostático, un enlace disulfuro, un enlace peptídico, un enlace diéster o una combinación de dichos enlaces. Los ácidos nucleicos además se unen opcionalmente a soportes sólidos mediante un enlace selectivamente liberable, tal como 4,4'-dimetoxitriilo o un derivado del mismo. Entre los derivados que se ha encontrado que resultan útiles se incluyen el ácido 3 ó 4-[bis(4-metoxifenil)]-metilbenzoico, el ácido N-succinimidil-3 ó 4-[bis(4-metoxifenil)]-metilbenzoico, el ácido N-succinimidil-3 ó 4-[bis(4-metoxifenil)]-hidroximetilbenzoico, el ácido N-succinimidil-3 ó 4-[bis(4-metoxifenil)]-clorometilbenzoico y sales de dichos ácidos.
- Además, los ácidos nucleicos se unen opcionalmente a soportes sólidos mediante fracciones espaciadoras entre los ácidos nucleicos y el soporte sólido. Entre los espaciadores útiles se incluyen un agente de acoplamiento, tal como se ha indicado anteriormente para la unión de parejas de acoplamiento alternativas o adicionales, o para convertir la unión al soporte sólido en cortable.
- Las uniones cortables pueden crearse mediante la unión de fracciones químicas cortables entre los ácidos nucleicos y el soporte sólido, incluyendo, por ejemplo, un oligopéptido, oligonucleótido, oligopoliamida, oligoacrilamida, oligoetileno, glicerol, cadenas alquilo de entre aproximadamente 6 y 20 átomos de carbono, y combinaciones de los mismos. Estas fracciones pueden cortarse con, por ejemplo, agentes químicos añadidos, radiación electromagnética o enzimas. Entre las uniones ejemplares cortables con enzimas se incluyen enlaces peptídicos que pueden ser cortados por proteasas y enlaces fosfodiéster que pueden ser cortados por nucleasas.
- Algunos agentes químicos, tales como el β-mercaptoetanol, el ditiotreitól (DTT) y otros agentes reductores cortan los enlaces disulfuro. Entre otros agentes que pueden resultar útiles se incluyen agentes oxidantes, agentes hidratantes

y otros compuestos selectivamente activos. La radiación electromagnética, tal como la luz ultravioleta, infrarroja y visible, corta los enlaces fotocortables. Las uniones también pueden ser reversibles, por ejemplo utilizando calor o tratamiento enzimático, o uniones químicas o magnéticas reversibles. La liberación y nueva unión pueden llevarse a cabo utilizando, por ejemplo, campos magnéticos o eléctricos.

Los biocatalizadores que incorporan nucleótidos utilizados en los métodos descritos en la presente memoria típicamente comprenden enzimas, tales como polimerasas, transferasas terminales, transcriptasas inversas, telomerasas, polinucleótido fosforilasas, y similares. Por ejemplo, la polimerasa opcionalmente no presenta una mutación de F por Y en la hélice O del enzima o, de otra manera, no presenta una mutación que incrementa la incorporación de desoxinucleótidos 3' por parte del enzima. Opcionalmente, el enzima comprende una actividad de exonucleasa 3' a 5' y/o es un enzima termoestable. El enzima típicamente se deriva de un organismo, tal como *Thermus antranikianii*, *Thermus aquaticus*, *Thermus caldophilus*, *Thermus chliarophilus*, *Thermus filiformis*, *Thermus flavus*, *Thermus igniterrae*, *Thermus lacteus*, *Thermus oshimai*, *Thermus ruber*, *Thermus rubens*, *Thermus scotoductus*, *Thermus silvanus*, *Thermus species Z05*, *Thermus species sps 17*, *Thermus thermophilus*, *Thermotoga maritima*, *Thermotoga neapolitana*, *Thermosipho africanus*, *Anaerocellum thermophilum*, *Bacillus caldotenax*, *Bacillus stearothermophilus*, o similar.

En algunas realizaciones, el enzima ha sido modificado. Entre los enzimas modificados ejemplares se incluyen, por ejemplo, una ADN polimerasa CS5 G46E E678G, una ADN polimerasa CS6 G46C E678G, una polimerasa Δ ZO5R, una ADN polimerasa CS5 G46E L329A E678G, y similares. Los enzimas modificados de la invención generalmente comprenden una capacidad incrementada de incorporar nucleótidos terminadores 2' respecto a un enzima no modificado. A título ilustrativo adicional, los enzimas modificados de la invención típicamente comprenden mutaciones que incrementan la incorporación de ribonucleótidos, que incrementan la incorporación de análogos 2'-modificados de ribonucleótidos y/o que reducen o eliminan la actividad de exonucleasa 5' a 3', por ejemplo respecto a un enzima que no presenta una o más de dichas mutaciones. Se proporcionan detalles adicionales referentes a los biocatalizadores que incorporan nucleótidos que resultan útiles en la práctica de los métodos de la presente invención en, por ejemplo, la patente US nº 5.939.292, titulada "Thermostable DNA polymerases having reduced discrimination against ribo-NTPs", publicada el 17 de agosto de 1999, de Gelfand *et al.*; en la patente US nº 4.889.818, titulada "Purified thermostable enzyme", publicada el 26 de diciembre de 1989, de Gelfand *et al.*; en la patente US nº 5.374.553, titulada "DNA encoding a thermostable nucleic acid polymerase enzyme from Thermotoga maritima", publicada el 20 de diciembre de 1994, de Gelfand *et al.*; en la patente US nº 5.445.170, titulada "Mutated thermostable nucleic acid polymerase enzyme from Thermus species Z05", publicada el 3 de octubre de 1995, de Abramson *et al.*; en la patente US nº 5.466.591, titulada "5' to 3' exonuclease mutations of thermostable DNA polymerases", publicada el 14 de noviembre de 1995, de Abramson *et al.*; en la patente US nº 5.624.833, titulada "Purified thermostable nucleic acid polymerase enzyme from Thermotoga maritima", publicada el 29 de abril de 1997, de Gelfand *et al.*; en la patente US nº 5.674.738, titulada "DNA encoding thermostable nucleic acid polymerase from Thermus species Z05", publicada el 7 de octubre de 1997, de Abramson *et al.*; en la patente US nº 5.789.224, titulada "Recombinant expression vectors and purification methods for Thermus thermophilus DNA polymerase", publicada el 4 de agosto de 1998, de Gelfand *et al.*; en la patente US nº 5.795.762, titulada "5' to 3' exonuclease mutations of thermostable DNA polymerases", publicada el 18 de agosto de 1998, de Abramson *et al.*; en la solicitud publicada de patente US nº 2002/0012970, titulada "High temperature reverse transcription using mutant DNA polymerases", publicada el 31 de enero de 2002, de Smith *et al.*; y en la solicitud de patente US nº 10/401.403, presentada el 26 de marzo de 2003, cada una las cuales se incorporan como referencia.

La producción de enzimas modificados de, por ejemplo, eficiencia incrementada para incorporar nucleótidos terminadores 2' puede conseguirse mediante diversos procedimientos, incluyendo, por ejemplo, la mutagénesis sitio-dirigida. Ver, por ejemplo, Sambrook *et al.*, *supra*. Más concretamente, la mutagénesis sitio-dirigida se consigue de manera general mediante mutagénesis específica de sitio dirigida por cebadores. Esta técnica típicamente se lleva a cabo utilizando un cebador oligonucleótido sintético que es complementario a un ADN fágico de cadena sencilla que debe mutagenizarse, excepto por un apareamiento incorrecto limitado que representa la mutación deseada. Brevemente, se utiliza el oligonucleótido sintético como cebador para dirigir la síntesis de una cadena complementaria al plásmido o fago, y el ADN de doble cadena resultante se transforma en una bacteria huésped que puede alojar un fago. Las bacterias resultantes pueden someterse a ensayo mediante, por ejemplo, análisis de secuencias de ADN o hibridación de sondas con el fin de identificar aquellas placas que portan la secuencia génica mutada que se desea. A título ilustrativo adicional, también pueden utilizarse muchos otros enfoques para modificar los ácidos nucleicos, tales como los métodos de "PCR recombinante" (ver, por ejemplo, Innis *et al.*, *supra*).

Los biocatalizadores que incorporan nucleótidos típicamente producen múltiples cebadores de ácidos nucleicos extendidos diferentes y los métodos generalmente incluyen la resolución de múltiples cebadores de ácidos nucleicos extendidos diferentes, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico de molde es determinable a partir de los cebadores de ácidos nucleicos extendidos que se han resuelto. Por ejemplo, los cebadores de ácidos nucleicos extendidos opcionalmente se resuelven mediante la determinación de las masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos. En determinadas realizaciones, los cebadores de ácidos nucleicos extendidos comprenden además marcajes y los cebadores de ácidos nucleicos extendidos se resuelven mediante separación de los cebadores de ácidos nucleicos extendidos marcados unos de otros y la detección (por ejemplo espectrofotométricamente, etc.) de las señales detectables

producidas por los marcajes. A título ilustrativo, los cebadores de ácidos nucleicos extendidos marcados se separan mediante por lo menos una técnica de separación, tal como electroforesis, cromatografía y espectrometría de iones en fase gaseosa y/o similar.

5 IV. MEZCLAS DE REACCIÓN

La invención proporciona además mezclas de reacción que comprenden por lo menos un nucleótido terminador 2' marcado, tal como se indica en la presente memoria (por ejemplo un nucleósido 2'-monofosfato-3'-hidroxil-5'-trifosfato marcado, etc.) y por lo menos un biocatalizador que incorpora nucleótidos tal como se indica en la presente memoria. En algunas realizaciones, la mezcla de reacción incluye además por lo menos una pirofosfatasa (por ejemplo una pirofosfatasa termoestable). La mezcla de reacción opcionalmente incluye además uno o más nucleótidos extensibles (por ejemplo ribonucleótidos, desoxirribonucleótidos y/o similares). Opcionalmente se marca por lo menos uno de los nucleótidos extensibles. El marcaje se ha descrito en mayor detalle anteriormente. Típicamente, el nucleótido terminador 2' y los nucleótidos extensibles se encuentran presentes en una proporción molar de 1:1 ó inferior. En determinadas realizaciones, la mezcla de reacción incluye además un ácido nucleico de molde y un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico de molde. Opcionalmente, el ácido nucleico de molde o el cebador de ácidos nucleicos se une (por ejemplo covalente o no covalentemente) a un soporte sólido. En algunas de dichas realizaciones, el cebador comprende un marcaje (por ejemplo pigmentos fluorescentes, isótopos radioactivos, grupo modificador de la masa, etc.). Los soportes sólidos y los marcajes se han descrito en mayor detalle anteriormente.

A título ilustrativo adicional, aunque sin limitación de la presente invención, se proporciona un conjunto de condiciones de reacción representativas para la secuenciación de un ADN molde, que resultan útiles en los métodos relacionados con los nucleótidos terminadores 2' descritos en la presente memoria. En particular, los nucleótidos terminadores 2' a los que se hace referencia en dicho conjunto ejemplar de condiciones son los nucleósidos 2'-monofosfato-3'-hidroxil-5'-trifosfato (abreviados como N-tetra-PO₄).

Las extensiones de cebadores de ácidos nucleicos opcionalmente se llevan a cabo en cuatro reacciones separadas. Entre los componentes en común en cada una de las reacciones se incluyen:

30 Tricina 50 µM a pH 8,5,
 KOAc 40 µM,
 Mg(OAc)₂ 4 µM,
 100 µM de cada uno de dATP, dCTP y dTTP,
 35 c7dGTP 150 µM,
 20 ng/µl de ADN molde M13mp18,
 0,5 U/µl de ADN polimerasa CS5 G46E E678G, y
 1,0 U/µml de pirofosfatasa termoestable r*Th*.

40 Entre las reacciones individuales además se incluyen las siguientes:

Reacción A:

45 Volumen de reacción: 10 µl,
 0,1 µM de cebador de ácidos nucleicos FR686N-HEX, y
 A-tetra-PO₄ 3,5 µM

Reacción C:

50 Volumen de reacción: 10 µl,
 0,1 µM de cebador de ácidos nucleicos FR686N-FAM, y
 C-tetra-PO₄ 7,5 µM

Reacción G:

55 Volumen de reacción: 20 µl,
 0,1 µM de cebador de ácidos nucleicos FR686N-TAMRA, y G-tetra-PO₄ 5 µM

Reacción U:

60 Volumen de reacción: 20 µl,
 0,1 µM de cebador de ácidos nucleicos FR686N-ROX, y U-tetra-PO₄ 10 µM

65 Tras el ciclado térmico, las mezclas de reacción opcionalmente se agruparon, se precipitaron con etanol y se resuspendieron en formamida. A continuación, se resolvió la muestra resuspendida, por ejemplo mediante electroforesis, y se analizaron en un secuenciador de ADN (por ejemplo un secuenciador de ADN ABI 377 (Applied

Biosystems, Foster City, CA) o similar).

V. DETECCIÓN DE SEÑALES DETECTABLES

- 5 Los ácidos nucleicos extendidos de la invención pueden detectarse utilizando esencialmente cualquier método de detección. Por ejemplo, la fluorescencia se detecta opcionalmente con detectores o sensores, tales como los tubos fotomultiplicadores (PMT), los dispositivos de carga acoplada (CCD), los CCD intensificados, los fotodiodos, los fotodiodos de avalancha, los sensores ópticos, los detectores de barrido, o similares. Detectores tales como los
- 10 indicados se encuentran disponibles fácilmente de diversos proveedores comerciales, entre ellos, por ejemplo, Applied Biosystems (Foster City, CA). Los sistemas de detección utilizados en la práctica de los métodos de la invención se describen en mayor detalle en, por ejemplo, Skoog *et al.*, Principles of Instrumental Analysis, 5a edición, Harcourt Brace College Publishers, 1998, y Currell, Analytical Instrumentation: Performance Characteristics and Quality, John Wiley & Sons, Inc., 2000.
- 15 En algunas realizaciones de la invención, el método incluye además detectar una masa molecular o peso molecular del cebador de ácidos nucleicos extendido o de un fragmento del mismo. Un genotipo del ácido nucleico molde es determinable típicamente a partir de la masa molecular detectada del cebador de ácidos nucleicos extendido o fragmento del mismo. Por ejemplo, una secuencia específica de ácidos nucleicos típicamente presentará un peso molecular único o relativamente único dependiendo de su tamaño y composición. Ese peso molecular puede
- 20 determinarse, por ejemplo, mediante cromatografía (por ejemplo HPLC), resonancia magnética nuclear (RMN), electroforesis en gel de alta definición, electroforesis capilar (por ejemplo HPCE), espectroscopia o espectrometría de iones en fase gaseosa (por ejemplo espectrometría de masas, etc.). Típicamente, los pesos moleculares se determinan mediante medición de la proporción de masa/carga utilizando la espectrometría de masas.
- 25 La espectrometría de masas de biopolímeros tales como los ácidos nucleicos puede llevarse a cabo utilizando una diversidad de técnicas (ver, por ejemplo, las patentes US nº 4.442.354, nº 4.931.639, nº 5.002.868, nº 5.130.538, nº 5.135.870 y nº 5.174.962). Las dificultades asociadas a la volatilización de las moléculas de elevado peso molecular, tales como el ADN y el ARN, se han superado, por lo menos en parte, con avances en las técnicas, procedimientos y diseño electrónico. Además, sólo resultan necesarias cantidades reducidas de muestra para el análisis, siendo la
- 30 muestra típica una mezcla de 10 fragmentos aproximadamente. Las cantidades comprendidas entre aproximadamente 0,1 femtomoles y aproximadamente 1,0 nanomol, preferentemente entre aproximadamente 1,0 femtomol y aproximadamente 1.000 femtomoles, y más preferentemente entre aproximadamente 10 femtomoles y aproximadamente 100 femtomoles típicamente resultan suficientes para el análisis. Estas cantidades pueden aplicarse fácilmente sobre las posiciones individuales de una superficie adecuada o unirse a un soporte.
- 35 Entre las técnicas ejemplares que pueden utilizarse para volatilizar un ácido nucleico se incluyen el bombardeo con átomos rápidos, la desorción de plasma, la desorción/ionización por láser asistida por matriz, la electropulverización, la liberación fotoquímica, la liberación eléctrica, la liberación de gotas, la ionización por resonancia y combinaciones de estas técnicas.
- 40 En la ionización electrohidrodinámica, termopulverización, aeropulverización y electropulverización, el ácido nucleico se disuelve en un solvente y se inyecta con ayuda de calor, aire o electricidad directamente en la cámara de ionización. En el caso de que el método de ionización incluya un haz lumínico, un haz de partículas o una descarga eléctrica, la muestra puede unirse a una superficie e introducirse en la cámara de ionización. En dichas situaciones,
- 45 puede unirse una pluralidad de muestras a una única superficie o a múltiples superficies e introducirse simultáneamente en la cámara de ionización y todavía analizarse individualmente. El sector apropiado de la superficie que contiene el ácido nucleico deseado puede desplazarse hasta encontrarse en el camino del haz ionizante. Tras pulsar el haz e ionizar las moléculas unidades a la superficie, se desplaza un sector diferente de la superficie para que se encuentre bajo el camino del haz y se analiza una segunda muestra, con la misma molécula u
- 50 otra diferente, sin recargar el aparato. También pueden introducirse múltiples muestras en regiones eléctricamente aisladas de una superficie. Típicamente se conectan diferentes sectores de un soporte sólido, tal como un chip, a una fuente eléctrica y se ionizan individualmente. La superficie a la que se une la muestra puede conformarse para la máxima eficiencia del método de ionización utilizado. Para la ionización de campo y la desorción de campo, una clavija o borde afilado es un soporte sólido eficiente, y para el bombardeo de partículas y la ionización por láser, una
- 55 superficie plana.
- Un objetivo de la ionización para la espectrometría de masas es producir una molécula completa con una carga. Opcionalmente, se utiliza la desorción/ionización por láser asistida por matriz (MALDI) (ver, por ejemplo, Sauer *et al.*, "Facile method for automated genotyping of single nucleotide polymorphisms by mass spectrometry", Nucleic Acids Res. 30(5):e22, 2002) o la espectrometría de masas por electropulverización (ES) para determinar el peso molecular y, de esta manera, información sobre la secuencia de los ácidos nucleicos de molde. El experto ordinario en la materia reconocerá que puede utilizarse una diversidad de métodos que resultan apropiados para moléculas de gran tamaño, tales como ácidos nucleicos. Típicamente, un ácido nucleico se disuelve en un solvente y se inyecta en la
- 60 cámara de ionización mediante ionización electrohidrodinámica, termopulverización, aeropulverización o electropulverización. Los ácidos nucleicos también pueden unirse a una superficie e ionizarse con un haz de partículas o de luz. Entre las partículas que se han utilizado con éxito se incluyen el plasma (desorción de plasma),
- 65

los iones (bombardeo de iones rápidos) o átomos (bombardeo de átomos rápidos). También se han producido iones mediante la aplicación rápida de energía láser (desorción por láser) y energía eléctrica (desorción de campo).

5 En el análisis de espectrometría de masas, la muestra se ioniza brevemente con un pulso de haces láser o mediante pulverización inducida por un campo eléctrico. Los iones se aceleran en un campo eléctrico y se envían a alta
10 velocidad a la parte del analizador del espectrómetro. La velocidad del ión acelerado es directamente proporcional a la carga (z) e inversamente proporcional a la masa (m) del ión. La masa de la molécula puede deducirse a partir de las características de vuelo de su ión. Para iones pequeños, el detector típico presenta un campo magnético, que funciona restringiendo el flujo de iones a un camino circular. Los radios de los caminos de las partículas de igual
15 carga en un campo magnético uniforme son directamente proporcionales a las masas. Es decir, una partícula más pesada con la misma carga que una partícula más ligera presentará un radio de vuelo más grande en un campo magnético. Se considera generalmente que no resulta práctico medir las características de vuelo de iones grandes, tales como ácidos nucleicos, en un campo magnético porque la proporción relativamente alta de masa a carga (m/z) requiere un imán de tamaño o potencia poco habituales. Para superar esta limitación, el método de electropulverización, por ejemplo, puede aplicar consistentemente múltiples iones a una molécula. Las cargas múltiples sobre un ácido nucleico reducirán la proporción de masa a carga, permitiendo que un analizador de cuadrupolo convencional detecte especies de hasta 100.000 daltons.

20 Los iones ácidos nucleicos generados mediante desorción/ionización por láser asistida por matriz únicamente presentan una carga unitaria y debido a su gran masa, generalmente utilizan el análisis mediante un analizador de masas por tiempo de vuelo (TOF). Los analizadores de tiempo de vuelo típicamente son tubos largos con un detector en un extremo. Durante el funcionamiento de un analizador TOF, se ioniza brevemente una muestra y se acelera por el tubo. Tras la detección, se calcula el tiempo necesario para recorrer el tubo del detector. La masa de la molécula puede calcularse a partir del tiempo de vuelo. Los analizadores de masas TOF no utilizan típicamente un
25 campo magnético y pueden detectar iones de carga unitaria con una masa de hasta 100.000 daltons. Para una resolución mejor, el espectrómetro de masas de tiempo de vuelo puede incluir un reflectrón, una región en el extremo del tubo de vuelo que acelera negativamente los iones. Las partículas en movimiento que entran en la región del reflectrón, que contiene un campo de polaridad opuesta al campo acelerante, se retardan hasta una velocidad cero y después se aceleran inversamente con la misma velocidad pero en la dirección opuesta. Durante la
30 utilización de un analizador de masas con reflectrón, el detector se sitúa en el mismo lado del tubo de vuelo que la fuente de iones con el fin de detectar los iones de retorno y efectivamente se duplica la longitud efectiva del tubo de vuelo y el poder de resolución. El cálculo de la proporción de masa a carga a partir de los datos de tiempos de vuelo considera el tiempo transcurrido en el reflectrón.

35 Los iones con la misma proporción de carga a masa típicamente abandonarán los aceleradores de iones con un abanico de energías debido a que las regiones de ionización de un espectrómetro de masas no son una fuente puntual. Los iones generados más lejos del tubo de vuelo pasan más tiempo en el campo acelerante y entran en el tubo de vuelo a una velocidad más elevada. De esta manera, los iones de una única especie de molécula llegarán al detector en tiempos diferentes. En el análisis de masas por tiempo de vuelo, un mayor tiempo en el tubo de vuelo en teoría proporciona más sensibilidad, pero debido a las diferentes velocidades de los iones, también puede incrementarse el ruido (el fondo). Un reflectrón, a parte de duplicar efectivamente la longitud efectiva del tubo de
40 vuelo, puede reducir el error e incrementar la sensibilidad al reducir la dispersión de los tiempos de entrada en el detector de una única especie de ión. Un ión con una velocidad más alta entrará en el reflectrón a una velocidad más alta y permanecerá en la región del reflectrón durante más tiempo que un ión a menor velocidad. En el caso de que los voltajes del electrodo del reflectrón se dispongan apropiadamente, la contribución máxima en anchura de la distribución inicial de velocidades puede corregirse en gran medida a la altura del plano del detector. La corrección proporcionada por el reflectrón conduce a una resolución incrementada de las masas para todos los iones estables (es decir, aquellos que no se disocian durante el vuelo) en el espectro.

50 Aunque un reflectrón de campo lineal funciona adecuadamente reduciendo el ruido e incrementando la sensibilidad, los reflectrones con potencias de campo más complejas presentan capacidades de corrección superiores y pueden utilizarse varios reflectrones complejos. El reflectrón de doble estadio presenta una primera región con un campo eléctrico más débil y una segunda región con un campo eléctrico más fuerte. El reflectrón de campo cuadrático y el de campo curvado presentan un campo eléctrico que se incrementa como función de la distancia. Estas funciones, tal como indica su nombre, pueden ser una función exponencial cuadrática o compleja. Los reflectrones de doble
55 estadio, de campo cuadrático y de campo curvado, aunque más sofisticados, también resultan más precisos que el reflectrón lineal.

60 La detección de iones en un espectrómetro de masas típicamente se lleva a cabo utilizando detectores de electrones. Para detectar los iones de masa elevada producidos por el espectrómetro de masas, los iones se convierten en electrones o en iones de masa reducida en un electrodo de conversión. Estos electrones o iones de masa reducida seguidamente se utilizan para iniciar la multiplicación en cascada de electrones en un multiplicador de electrones y se amplifican adicionalmente con un amplificador lineal rápido. Las señales del análisis múltiple de una sola muestra se combinan para mejorar la proporción de señal a ruido y las formas de pico, lo que también
65 incrementa la exactitud de la determinación de las masas.

Pueden detectarse múltiples iones primarios directamente mediante la utilización de la resonancia de ciclotrón iónico y el análisis de Fourier. Lo anterior resulta útil para el análisis de una escalera de secuenciación completa inmovilizada sobre una superficie. En este método, se ioniza una pluralidad de muestras simultáneamente y los iones se capturan en una celda con un campo magnético potente. Un campo de RF excita la población de iones en órbitas ciclotrónicas. Debido a que las frecuencias de las órbitas son una función de la masa, se obtiene una señal de salida que representa el espectro de las masas de los iones. Esta salida de datos se analiza con un ordenador utilizando un análisis de Fourier, que reduce la señal combinada a sus frecuencias componentes y proporciona de esta manera una medición de las masas iónicas presentes en la muestra de iones. La resonancia ciclotrónica de iones y el análisis de Fourier pueden determinar las masas de todos los ácidos nucleicos en una muestra. La aplicación de este método resulta especialmente útil para una escalera de secuenciación.

Los datos de la espectrometría de masas, obtenidos individualmente o en paralelo (multiplexados), pueden determinar la masa molecular de una muestra de ácidos nucleicos. La masa molecular, en combinación con la secuencia conocida de la muestra, puede analizarse para determinar la longitud de la secuencia. Debido a que diferentes bases presentan diferente peso molecular, la salida de un espectrómetro de masas de alta resolución en combinación con la secuencia conocida y la historia de reacción de la muestra, determinarán la secuencia y la longitud del ácido nucleico analizado. En la espectroscopía de masas de una escalera de secuenciación, generalmente se conoce la secuencia de bases de los cebadores. A partir de una secuencia conocida de una longitud determinada, puede deducirse la base añadida de una secuencia una base más larga mediante la comparación de las masas de las dos moléculas. Se continúa este procedimiento hasta determinar la secuencia completa de una escalera de secuenciación.

VI. SISTEMAS

En otro aspecto, la invención se refiere a un sistema para extender un ácido nucleico. El sistema incluye: (a) por lo menos un recipiente que comprende un nucleótido terminador 2' marcado. Típicamente, el sistema comprende una pluralidad de recipientes, por ejemplo para llevar a cabo múltiples reacciones de extensión en paralelo. El sistema incluye además: (b) por lo menos un modulador térmico (por ejemplo un dispositivo de termociclado, etc.) operablemente conectado al recipiente para modular la temperatura en el mismo, y/o (c) por lo menos un componente de transferencia de líquidos (por ejemplo un pipeteador automático, etc.) que transfiere líquido hacia y/o desde el recipiente. Los dispositivos de termociclado, algunos de los cuales se realizan en dispositivos de microfluidos, y diversos dispositivos de transferencia de líquidos, adecuados o adaptables para la utilización en los sistemas de la invención son generalmente conocidos de la técnica. El sistema además incluye opcionalmente por lo menos un detector operablemente conectado al recipiente para la detección de señales detectables producidas en el recipiente. El sistema típicamente incluye además por lo menos un controlador operablemente conectado al modulador térmico para llevar a cabo la modulación de la temperatura en el recipiente y/o al componente de transferencia de líquidos para llevar a cabo la transferencia de líquido hacia y/o desde el recipiente.

Entre los sistemas de la invención se incluyen diversas realizaciones. Por ejemplo, los componentes de detección que se estructuran para detectar las señales detectables producidas en otro componente del sistema (por ejemplo en un recipiente de reacción, etc.) o en una posición próxima al mismo. Los detectores de señales adecuados que se utilizan opcionalmente, o que se adaptan para su utilización, en estos sistemas detectan, por ejemplo, fluorescencia, fosforescencia, radioactividad, masa, concentración, pH, carga, absorbancia, índice de refracción, luminiscencia, temperatura, magnetismo o similar. Los detectores opcionalmente realizan un seguimiento de una señal o de una pluralidad de señales anteriormente y/o posteriormente a la realización de, por ejemplo, una etapa de ensayo dada. Por ejemplo, el detector opcionalmente realiza un seguimiento de una pluralidad de señales ópticas, que corresponden en posición a los resultados "en tiempo real". Entre los detectores o sensores ejemplares se incluyen tubos fotomultiplicadores, matrices de CCD, sensores ópticos, sensores de temperatura, sensores de presión, sensores de pH, sensores de conductividad, detectores de barrido, o similares. Cada uno de ellos, así como otros tipos de sensores se incorporan opcionalmente con facilidad en los sistemas indicados en la presente memoria. Opcionalmente, entre los sistemas de la presente invención se incluyen múltiples detectores.

Puede utilizarse o adaptarse para la utilización en los sistemas de la invención esencialmente cualquier componente analítico. Entre determinados componentes analíticos ejemplares que se utilizan opcionalmente en dichos sistemas se incluyen, por ejemplo, una columna de cromatografía líquida, una columna de electroforesis en gel, una columna de electrocromatografía, un detector de dispersión lumínica por resonancia, un espectroscopio de emisión, un espectroscopio de fluorescencia, un espectroscopio de fosforescencia, un espectroscopio de luminiscencia, un espectrofotómetro, un fotómetro, un calorímetro, un espectrómetro de masas, un espectrómetro de resonancia magnética nuclear, un espectrómetro de resonancia paramagnética de electrones, un espectroscopio de resonancia de espín electrónico, un turbidímetro, un nefelómetro, un espectroscopio de Raman, un refractómetro, un interferómetro, un analizador de difracción de rayos X, un analizador de difracción de electrones, un polarímetro, un analizador de dispersión de la rotación óptica, un espectrómetro de dicroísmo circular, un potenciómetro, un cronopotenciómetro, un culombímetro, un amperímetro, un conductímetro, un gravímetro, un gravímetro térmico, un titrímetro, un calorímetro de barrido diferencial, un analizador de activación radioactiva, un analizador de dilución isotópica radioactiva, o similares. También se utilizan diversos componentes sintéticos, o se adaptan para su utilización, en los sistemas de la invención, incluyendo, por ejemplo, sintetizadores automáticos de ácidos nucleicos.

Los componentes analíticos y sintéticos que se incluyen opcionalmente en los sistemas de la invención se describen en mayor detalle en, por ejemplo, Skoog *et al.*, Principles of Instrumental Analysis, 5a edición, Harcourt Brace College Publishers, 1998, y Currell, Analytical Instrumentation: Performance Characteristics and Quality, John Wiley & Sons, Inc., 2000.

5 Entre los sistemas de la invención típicamente también se incluyen controladores que se conectan operablemente a uno o más componentes (por ejemplo componentes analíticos, componentes sintéticos, moduladores térmicos, componentes de transferencia de líquidos, detectores, etc.) del sistema para controlar el funcionamiento de los componentes. Más concretamente, generalmente se incluyen controladores a modo de componentes separados o
10 incorporados del sistema que se utilizan para, por ejemplo, recibir datos de los detectores, para producir y/o regular la temperatura en los recipientes, para producir y/o regular el flujo de líquido hacia o desde recipientes seleccionados, o similares. Los controladores y/o otros componentes del sistema se acoplan opcionalmente a un procesador apropiadamente programado, ordenador, dispositivo digital u otro dispositivo informático (por ejemplo incluyendo un conversor analógico a digital o digital a analógico, según resulte necesario), que funciona ordenando
15 el funcionamiento de estos instrumentos de acuerdo con instrucciones preprogramadas o introducidas por el usuario, recibiendo datos e información de dichos instrumentos, e interpretando, manipulando y ofreciendo esta información al usuario. Los controladores adecuados son generalmente conocidos de la técnica y se encuentran disponibles de diversos proveedores comerciales.

20 Cualquier controlador u ordenador incluye opcionalmente un monitor que con frecuencia es una pantalla de tubo de rayos catódicos ("CTR"), una pantalla plana (por ejemplo una pantalla de matriz activa de cristal líquido, una pantalla de cristal líquido, etc.) u otros. Los circuitos de ordenador con frecuencia se incluyen en una caja, que incluye numerosos chips de circuito integrado, tales como un microprocesador, memoria, circuitos de interfaz y otros. La caja opcionalmente también incluye una unidad de disco duro, una unidad de disquetera, un disco extraíble de alta
25 capacidad, tal como un CD-ROM reescribible y otros elementos periféricos comunes. Los dispositivos de entrada, tales como un teclado o un ratón opcionalmente permiten la introducción de información por parte del usuario. Estos componentes se ilustran en mayor detalle posteriormente.

El ordenador típicamente incluye software apropiado para recibir instrucciones del usuario, en forma de entrada por
30 parte del usuario en un conjunto de campos de parámetros, por ejemplo en un GUI, o en forma de instrucciones preprogramadas, por ejemplo preprogramadas para una diversidad de diferentes operaciones específicas. A continuación, el software convierte dichas instrucciones en un lenguaje apropiado para ordenar la activación de uno o más controladores para que lleven a cabo la operación deseada. El ordenador seguidamente recibe los datos de, por ejemplo, sensores/detectores incluidos en el sistema e interpreta los datos, los proporciona en un formato
35 comprensible para el usuario o utiliza los datos para iniciar instrucciones adicionales para el controlador, según la programación, por ejemplo el control de los reguladores de flujo de líquido en respuesta a datos de peso de líquido recibidos de balanzas de peso o similares.

A título ilustrativo, algunas realizaciones de la invención proporcionan ordenadores y/o medios legibles por
40 ordenador que comprenden datos que comprenden por lo menos un carácter correspondiente a por lo menos un nucleótido terminador 2' tal como se indica en la presente memoria. Típicamente, los datos comprende una pluralidad de cadenas de caracteres correspondientes a una pluralidad de secuencias de ácidos nucleicos.

El ordenador puede ser, por ejemplo, un PC (Intel x86 o DOSTM compatible con chip Pentium, OS2TM, WINDOWSTM,
45 WINDOWS NTTM, WINDOWS95TM, WINDOWS98TM, WINDOWS 2000TM, WINDOWS XPTM, una máquina basada en Linux, un MACINTOSHTM, un Power PC o una máquina basada en UNIX (por ejemplo una estación de trabajo SUNTM) u otro ordenador común disponible comercialmente conocido por el experto en la materia. La aplicaciones de sobremesa estándares, tales como el software de procesamiento de textos (por ejemplo Microsoft WordTM o Corel WordPerfectTM) y software de bases de datos (por ejemplo software de hojas de cálculo, tal como Microsoft ExcelTM,
50 Corel Quattro ProTM, o programas de bases de datos, tales como Microsoft AccessTM o ParadoxTM) pueden adaptarse a la presente invención. El software para llevar a cabo, por ejemplo, el control de moduladores de la temperatura y reguladores del flujo de líquidos opcionalmente es construido por el experto en la materia utilizando lenguaje de programación estándar, tal como Visual Basic, Fortran, Basic, Java o similar.

55 VII. KITS

La presente invención proporciona además kits para la extensión de ácidos nucleicos. Los kits incluyen: (a) por lo menos un biocatalizador que incorpora nucleótidos tal como se indica en la presente memoria, y (b) por lo menos un nucleótido terminador 2' marcado, tal como se indica en la presente memoria. Por ejemplo, el nucleótido terminador
60 2' opcionalmente incluye por lo menos un marcaje (por ejemplo un isótopo radioactivo, un pigmento fluorescente, un grupo modificador de masa o similar). En algunas realizaciones, el kit incluye además uno o más nucleótidos extensibles y, opcionalmente, por lo menos uno los nucleótidos extensibles comprende un marcaje (por ejemplo un isótopo radioactivo, un pigmento fluorescente, un grupo modificador de la masa, o similar). Opcionalmente, el kit incluye además por lo menos una pirofosfatasa (por ejemplo una pirofosfatasa termoestable, etc.). Típicamente, el kit
65 incluye además: (c) un juego de instrucciones para extender el ácido nucleico utilizando el biocatalizador que incorpora nucleótidos y el nucleótido terminador 2'. Además, el kit opcionalmente incluye además: (d) por lo menos

un recipiente para empaquetar el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' marcado y el juego de instrucciones. En determinadas realizaciones, el kit incluye además un ácido nucleico de molde y un cebador de ácidos nucleicos, en el que el cebador de ácidos nucleicos es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico de molde. Opcionalmente, el ácido nucleico de molde o el cebador de ácidos nucleicos se une a un soporte sólido, por ejemplo tal como se indica en la presente memoria. En algunas de dichas realizaciones, el cebador comprende un marcaje, tal como un isótopo radioactivo, un pigmento fluorescente, un grupo modificador de la masa, o similar.

VIII. EJEMPLO 1: SECUENCIACIÓN CÍCLICA AUTOMÁTICA DE ADN UTILIZANDO UNA ADN POLIMERASA TERMOESTABLE MODIFICADA Y CEBADORES FLUORESCENTES

El presente ejemplo ilustra la aplicación de los nucleótidos terminadores 2' de la invención a la secuenciación cíclica automática de ADN mediante pigmentos-cebadores. En particular, se secuenció ADN molde M13mp18 utilizando ribonucleósidos-2'-monofosfato-5'-trifosfato.

Las reacciones de secuenciación cíclica se llevaron a cabo con ADN polimerasa CS5 G46E E678G (indicada anteriormente) modificada para la incorporación de análogos de ribonucleótidos, cebadores-pigmento y análogos de ribonucleósido-2'-monofosfato-5'-trifosfato. Las reacciones consistían de tricina 50 mM, pH 8,5; KOAc 40 mM, Mg(OAc)₂ 4 mM, 100 mM de cada uno de dATP, dCTP, dTTP, c7dGTP 150 mM, 0,5 unidades/ml de ADN polimerasa CS5 G46E E678G, 1,0 unidad/ml de pirofosfatasa termoestable r*Tth* y 20 ng/ml de molde M13mp18. Se llevaron a cabo cuatro reacciones individuales, una para cada base. Las reacciones para cada una de las bases contenían lo anteriormente indicado más los reactivos siguientes:

Reacciones de adenosina (10 ml):

Adenosina-2'-monofosfato-5'-trifosfato 3,5 μM
Cebador FR686NH3X 0,1 μM

Reacciones de citidina (10 ml):

Citidina-2'-monofosfato-5'-trifosfato 7,5 μM
Cebador FR686NFAM 0,1 μM

Reacciones de guanosina (20 ml):

Guanosina-2'-monofosfato-5'-trifosfato 5 μM
Cebador FR686NTAMRA 0,1 μM

Reacciones de uridina (20 ml):

Uridina-2'-monofosfato-5'-trifosfato 10 μM
Cebador FR686NROX 0,1 μM

En las reacciones de adenosina, la adenosina-2'-monofosfato-5'-trifosfato era pura aproximadamente al 95% (es decir, aproximadamente 5% era adenosina-3'-monofosfato-5'-trifosfato). En las reacciones de citidina, la citidina-2'-monofosfato-5'-trifosfato y la citidina-3'-monofosfato-5'-trifosfato se encontraban presentes en forma de mezcla 50/50. En las reacciones de guanosina, la guanosina-2'-monofosfato-5'-trifosfato era pura aproximadamente al 94% (es decir, aproximadamente 6% era guanosina-3'-monofosfato-5'-trifosfato). En las reacciones de uridina, la uridina-2'-monofosfato-5'-trifosfato era pura al 100%.

Las secuencias de cebadores oligonucleótidos eran las siguientes:

FR686NFAM	FCGCCAGGGTTTTCCAGTE	
	A	
	E = 2'-amino (ribo) C	F = 5' FAM ABD
FR686NHEX	ICGCCAGGGTTTTCCAGTEA	
	E = 2'-amino (ribo) C	I = 5' HEX ABD
FR686NROX	JCGCCAGGGTTTTCCAGTEA	
	E = 2'-amino (ribo) C	J = 6-ROX
FR686NTAMRA	LCGCCAGGGTTTTCCAGTEA	
	E = 2'-amino (ribo) C	L = C6-amino TAMRA

Se introdujo cada una de las cuatro reacciones en un ciclador térmico GeneAmp[®] PCR system 9600 de Perkin-Elmer y se sometieron a 95°C durante 45 segundos y después a 20 ciclos de 95°C durante 15 segundos, 55°C durante 15

segundos, 70°C durante 90 segundos, seguido de 20 ciclos de 95°C durante 15 segundos y 70°C durante 90 segundos. Se agruparon las cuatro reacciones y se precipitaron mediante la adición de 144 ml de etanol al 100% y 6 ml de NaOAc 3 M (pH 5,2) a 4°C durante 15 minutos. Las reacciones agrupadas se microcentrifugaron a 4°C durante 15 minutos para precipitar el ADN y se eliminó el sobrenadante. El pellet se lavó con 350 ml de etanol al 70% frío, se microcentrifugó a 4°C durante 5 minutos, se eliminó el sobrenadante y se secó el pellet de ADN. El ADN precipitado se resuspendió en 10 ml de formamida Hi-Di (Applied Biosystems, Foster City, CA, componente nº 4311320), se calentó a 90°C durante 3 minutos y se dejó sobre hielo. Se cargaron 2 ml de cada muestra en un gel previamente sometido a electroforesis en un gel de 48 cm de acrilamida al 4,25%:bis (29:1), urea 6 M, y se sometió a electroforesis durante 7 horas en un secuenciador de ADN ABI PRISM™ 377 (Applied Biosystems, Foster City, CA).

Los datos se analizaron con el software de análisis de secuenciación 3.4.1 (Applied Biosystems, Foster City, CA) utilizando el archivo de cebadores DP4%Ac{KS}, el lector de nucleótidos semiadaptativo versión 3.3.1b2 y un archivo de matriz específico para los cebadores-pigmento utilizados anteriormente, generados tras el procedimiento en el manual de Applied Biosystems (componente nº 903436). La lectura de nucleótidos automática por el software de análisis fue exacta al 100% para las bases +18 a +739 desde el cebador de secuenciación en comparación con una secuencia de referencia M13mp18. La figura 7 proporciona un perfil espectral de los datos de este análisis de secuencias.

IX. EJEMPLO 2: EXTENSIÓN CÍCLICA DE CEBADORES DE ADN UTILIZANDO UNA ADN POLIMERASA TERMOESTABLE MODIFICADA Y RIBONUCLEÓSIDOS 2'-MONOFOSFATO-5'-TRIFOSFATO MARCADOS CON PIGMENTO

Se llevó a cabo una reacción térmica cíclica de extensión de cebadores con ADN polimerasa CS5 G46E E678G modificada para la incorporación de análogos de ribonucleótidos, cebadores no marcados y uridina-2'-monofosfato-5'-trifosfato marcado con pigmento TAMRA. La reacción de 20 µl consistía de tricina 50 mM, pH 7,5, KOAc 25 mM, Mg(OAc)₂ 2,5 mM, 100 µM de cada uno de dATP, dCTP y dTTP, 150 µM de dITP, 0,5 unidades/µl de ADN polimerasa CS5 G46E E678G, 1,0 unidades/µl de pirofosfatasa inorgánica termoestable *rTth*, 5 ng/µl de molde M13mp18, 0,15 µM de cebador y 0,25 µM de TAMRA-uridina-2'-fosfato-5'-trifosfato.

Se llevó a cabo una reacción de control con ADN polimerasa AmpliTaq, FS, cebador no marcado y ddTTP marcado con pigmento TAMRA. La reacción de 20 µl consistía de Tris 50 mM, pH 9, MgCl₂ 2 mM, 100 µM de cada uno de dATP, dCTP y dTTP, 150 µM de dITP, 0,5 unidades/µl de ADN polimerasa AmpliTaq, FS; 1,0 unidades/µl de pirofosfatasa inorgánica termoestable *rTth*, 5 ng/µl de molde M13mp18, 0,15 µM de cebador FR686N y 0,2 µM de TAMRA-ddTTP.

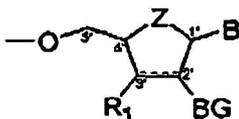
FR686	CGCCAGGGTTTTCCCAGTE
N	A
	E = 2'-amino (ribo) C

Las reacciones se introdujeron en un ciclador térmico GeneAmp® PCR system 9700 de Perkin-Elmer y se sometieron a 96°C durante 20 segundos y después a 25 ciclos de 96°C durante 10 segundos, 50°C durante 5 segundos y 60°C durante 4 minutos. Tras el ciclado, se eliminó de la reacción el terminador no incorporado marcado con pigmento, mediante centrifugación a 700xg durante 2 minutos en una columna Sephadex-G50 (Sigma, componente nº G-50-80). Se calentó la muestra a 95°C durante 3 minutos y se dejó sobre hielo. Las muestras se sometieron a electroforesis en un analizador génico Applied Biosystems 3100 con la aplicación GeneScan con los parámetros del módulo por defecto StdSe50_POP6 utilizando una matriz capilar de 50 cm y polímero POP6.

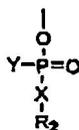
Los datos se analizaron con el software de análisis de fragmentos GeneScan 3.7 de Applied Biosystems. La figura 8 muestra el patrón de fragmentos para los picos T de 77 a 273 bases del cebador FR686N. Más concretamente, la comparación entre el patrón de fragmentos generados con la ADN polimerasa CS5 G46E E678G y TAMRA-uridina-2'-monofosfato-5'-trifosfato (panel B) y el patrón de fragmentos generado con la ADN polimerasa AmpliTaq de control, FS y TAMRA-ddTTP (panel A) reveló un patrón similar de picos.

REIVINDICACIONES

1. Método de extensión de un cebador de ácidos nucleicos, comprendiendo el método: incubar un ácido nucleico molde con por lo menos un biocatalizador que incorpora nucleótidos, por lo menos un nucleótido terminador 2' y por lo menos un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico molde, en el que dicho cebador de ácidos nucleicos comprende ADN, bajo condiciones en las que el biocatalizador que incorpora nucleótidos extiende el cebador de ácidos nucleicos para producir por lo menos un cebador de ácidos nucleicos extendido mediante la incorporación del nucleótido terminador 2' en un extremo terminal del cebador de ácidos nucleicos extendido, en el que el nucleótido terminador 2' comprende la fórmula:



- en la que:
 R₁ es OH,
 B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o combinaciones de los mismos,
 BG es un grupo bloqueante,
 y Z es O o CH₂, y ---- representa un enlace sencillo o doble,
 en el que BG comprende la fórmula:



- en la que X es O, S, NR₃, CR₃R₄, o SiR₃R₄;
 Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅, o NHR₅;
 R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y
 R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo o combinaciones de los mismos, o



- en la que:
 X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃, o NHR₃;
 R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y
 R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, o combinaciones de los mismos.

2. Método según la reivindicación 1, en el que el nucleótido terminador 2' comprende un nucleósido 2'-monofosfato-3'-hidroxil-5'-trifosfato o un grupo bloqueante cargado negativamente.

3. Método según la reivindicación 1, en el que el biocatalizador que incorpora nucleótidos comprende un enzima seleccionado de entre el grupo que consiste de: una polimerasa, una transferasa terminal, una transcriptasa inversa, una polinucleótido fosforilasa y una telomerasa.

4. Método según la reivindicación 3, en el que el enzima comprende una mutación que incrementa la incorporación de ribonucleótidos.

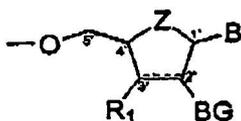
5. Método según la reivindicación 3, en el que el enzima comprende una mutación que reduce o elimina la actividad de exonucleasa 5' a 3'.

6. Método según la reivindicación 3, en el que el enzima comprende una mutación que incrementa la incorporación de análogos 2'-modificados de ribonucleótidos.

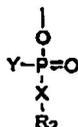
7. Método según la reivindicación 1, que comprende además la detección de la totalidad o de una parte del cebador

de ácidos nucleicos extendido o de un fragmento del mismo.

- 5 8. Método según la reivindicación 1, en el que por lo menos uno de entre el nucleótido terminador 2', el cebador de ácidos nucleicos extendido y el cebador de ácidos nucleicos comprende por lo menos un marcaje.
9. Método según la reivindicación 8, en el que se une un marcaje a uno de entre una base heterocíclica del nucleótido terminador 2', una fracción sacárida del nucleótido terminador 2' y un grupo fosfato del nucleótido terminador 2'.
- 10 10. Método según la reivindicación 8, en el que un conector une un marcaje al nucleótido terminador 2'.
11. Método según la reivindicación 8, que comprende además la detección de una señal detectable producida por el marcaje.
- 15 12. Método según la reivindicación 8, en el que el marcaje comprende un pigmento fluorescente, un marcaje débilmente fluorescente, un marcaje no fluorescente, un marcaje colorimétrico, un marcaje quimioluminiscente, un marcaje bioluminiscente, un isótopo radioactivo, un anticuerpo, un antígeno, biotina, un hapteno o un enzima.
- 20 13. Método según la reivindicación 1, en el que el ácido nucleico de molde o el cebador de ácidos nucleicos se une a un soporte sólido.
14. Método según la reivindicación 1, que comprende además incubar el ácido nucleico de molde con por lo menos un nucleótido extensible.
- 25 15. Método según la reivindicación 14, que comprende incubar el ácido nucleico de molde con por lo menos una pirofosfatasa que minimiza la pirofosforolisis.
16. Método según la reivindicación 14, en el que los nucleótidos terminadores 2' y los nucleótidos extensibles se encuentran presentes en una proporción molar de 1:1 ó inferior.
- 30 17. Método según la reivindicación 14, en el que el biocatalizador que incorpora nucleótidos produce múltiples cebadores de ácidos nucleicos extendidos diferentes y el método comprende la resolución de múltiples cebadores de ácidos nucleicos extendidos diferentes, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico de molde es determinable a partir del cebador de ácidos nucleicos extendido que se ha resuelto.
- 35 18. Método según la reivindicación 17, en el que los cebadores de ácidos nucleicos extendidos se resuelven mediante la determinación de por lo menos una de las masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos.
- 40 19. Método según la reivindicación 17, en el que los cebadores de ácidos nucleicos extendidos comprenden además marcajes y los cebadores de ácidos nucleicos extendidos se resuelven mediante separación de los cebadores de ácidos nucleicos extendidos marcados unos de otros y la detección de las señales detectables producidas por los marcajes.
- 45 20. Método de extensión de un ácido nucleico, comprendiendo el método: incubar por lo menos un ácido nucleico con por lo menos un biocatalizador que incorpora nucleótidos y por lo menos un nucleótido terminador 2', de manera que el biocatalizador que incorpora nucleótidos extiende el ácido nucleico, produciendo por lo menos un ácido nucleico extendido mediante la incorporación del nucleótido terminador 2' marcado en un extremo terminal del ácido nucleico, en el que el nucleótido terminador 2' marcado comprende la fórmula:
- 50



- 55 en la que: R₁ es OH, B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o combinaciones de los mismos, BG es un grupo bloqueante, y Z es O o CH₂, y ---- representa un enlace sencillo o doble, en el que BG comprende la fórmula:



en la que:

X es O, S, NR₃, CR₃R₄ ó SiR₃R₄,

Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅ ó NHR₅,

5 R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y

R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo, un grupo alquinilo, o combinaciones de los mismos, o



10 en el que:
X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃ ó NHR₃;
R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo, un grupo alquinilo,
15 un grupo alcoxi, o combinaciones de los mismos, y
R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenoilo o combinaciones de los mismos.

20 21. Método según la reivindicación 20, en el que el biocatalizador que incorpora nucleótidos comprende un enzima seleccionado de entre el grupo que consiste de: una transferasa terminal y una polinucleótido fosforilasa.

22. Método según la reivindicación 20, en el que el nucleótido terminador 2' marcado comprende un nucleósido 2'-monofosfato-3'-hidroxil-5'-trifosfato.

25 23. Método según la reivindicación 20, que comprende además hibridar el ácido nucleico extendido con otro ácido nucleico y detectar una señal detectable producida por el marcaje.

30 24. Método según la reivindicación 20, en el que el ácido nucleico comprende un cebador de ácidos nucleicos que es por lo menos parcialmente complementario respecto a por lo menos una subsecuencia de un ácido nucleico de molde, y el método comprende incubar el ácido nucleico de molde con el biocatalizador que incorpora nucleótidos, el nucleótido terminador 2' marcado y el cebador de ácidos nucleicos.

35 25. Método según la reivindicación 24, en el que el biocatalizador que incorpora nucleótidos comprende un enzima seleccionado de entre el grupo que consiste de: una polimerasa, una transferasa terminal, una transcriptasa inversa, una polinucleótido fosforilasa y una telomerasa.

26. Método según la reivindicación 24, que comprende además incubar el ácido nucleico de molde con por lo menos un nucleótido extensible.

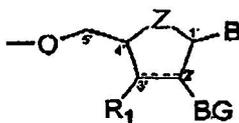
40 27. Método según la reivindicación 26, en el que el biocatalizador que incorpora nucleótidos produce múltiples cebadores de ácidos nucleicos extendidos diferentes y el método comprende resolver múltiples cebadores de ácidos nucleicos extendidos diferentes, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico de molde es determinable a partir del cebador de ácidos nucleicos extendido que se ha resuelto.

45 28. Método según la reivindicación 27, en el que los cebadores de ácidos nucleicos extendidos se resuelven mediante la determinación de por lo menos una de las masas moleculares, tamaños y/o propiedades de carga de los cebadores de ácidos nucleicos extendidos.

50 29. Método según la reivindicación 27, en el que el cebador de ácidos nucleicos extendidos se resuelven mediante la separación de los cebadores de ácidos nucleicos extendidos unos respecto a otros y la detección de las señales detectables producidas por los marcajes.

55 30. Método de inhibición de la extensión posterior de un ácido nucleico extendido, comprendiendo el método: poner en contacto por lo menos un ácido nucleico con por lo menos un biocatalizador que incorpora nucleótidos y por lo menos un nucleósido o nucleótido terminador 2' ó una sal farmacéuticamente aceptable del mismo, en el que el ácido nucleico comprende ADN y en el que el nucleósido o nucleótido terminador 2', o sal farmacéuticamente aceptable del mismo, no es extensible por el biocatalizador que incorpora nucleótidos, de manera que el biocatalizador que incorpora nucleótidos extiende el ácido nucleico, produciendo por lo menos un ácido nucleico extendido mediante la incorporación del nucleósido o nucleótido terminador 2' marcado, o la sal farmacéuticamente
60 aceptable del mismo, en un extremo terminal del ácido nucleico, inhibiendo de esta manera la extensión posterior del ácido nucleico extendido.

en el que el nucleótido terminador 2' comprende la fórmula:



en la que:

R₁ es OH,

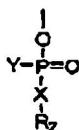
5 B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o combinaciones de los mismos,

BG es un grupo bloqueante,

y Z es O o CH₂, y ---- representa un enlace sencillo o doble,

en el que BG comprende la fórmula:

10



en la que:

X es O, S, NR₃, CR₃R₄ ó SiR₃R₄;

15 Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅ ó NHR₅,

R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y

R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, o combinaciones de los mismos, o

20



en la que:

X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃, o NHR₃,

25 R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi, o combinaciones de los mismos, y

R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo o combinaciones de los mismos.

30 31. Método según la reivindicación 30, en el que el ácido nucleico comprende ADN microbiano.

32. Método según la reivindicación 31, en el que el ADN microbiano, el biocatalizador que incorpora nucleótidos y el nucleósido o nucleótido terminador 2', o la sal farmacéuticamente aceptable del mismo, se pone en contacto en un huésped infectado por un microbio que comprende el ADN microbiano.

35

33. Método de secuenciación de un ácido nucleico diana, comprendiendo el método:

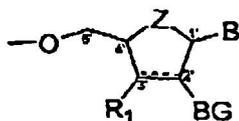
(a) incubar por lo menos un ácido nucleico diana con una o más polimerasas, uno o más nucleótidos terminadores 2', uno o más nucleótidos extensibles, y uno o más cebadores de ácidos nucleicos que son complementarios a por lo menos una subsecuencia del ácido nucleico diana, en el que las polimerasas extienden los cebadores de ácidos nucleicos para producir productos de extensión de cebador que incorporan los nucleótidos terminadores 2' en los extremos 3'-terminales de los productos de extensión de cebador, y

40

(b) identificar los nucleótidos terminadores 2' en los productos de extensión de cebador, de manera que por lo menos una parte de una secuencia de bases del ácido nucleico diana sea determinable a partir de los nucleótidos terminadores 2' identificados.

45

en el que el nucleótido terminador 2' comprende la fórmula:

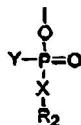


en la que:

R₁ es OH,

50 B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o

en el que BG comprende la fórmula:



5 en la que:

X es O, S, NR₃, CR₃R₄, o SiR₃R₄,

Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅, o NHR₅,

R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi, o combinaciones de los mismos, y

10 R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo o combinaciones de los mismos, o



15 en la que:

X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃ ó NHR₃,

R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi, o combinaciones de los mismos, y

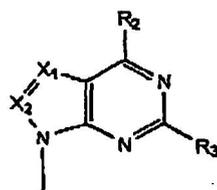
20 R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo o combinaciones de los mismos.

41. Mezcla de reacción según la reivindicación 40, en la que el nucleótido terminador 2' comprende un nucleósido 2'-monofosfato-3'-hidroxilo.

25 42. Mezcla de reacción según la reivindicación 40, que comprende además por lo menos una pirofosfatasa.

43. Mezcla de reacción según la reivindicación 40, en la que B comprende una fórmula seleccionada de entre el grupo que consiste de:

30 a)



en la que:

35 X₁ y X₂ se seleccionan independientemente de entre CH y N,

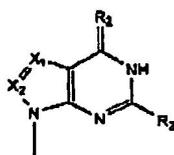
R₂ es H, OH ó NR₄R₅,

R₃ es H, OH ó NR₆R₇,

R₄ y R₅ se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi y combinaciones de los mismos, y

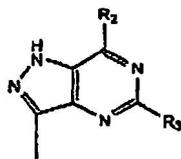
40 R₆ y R₇ se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi y combinaciones de los mismos,

b)



45 en la que: X₁ y X₂ se seleccionan independientemente de entre CH y N, R₂ es O ó S, R₃ es H, OH ó NR₄R₅, y R₄ y R₅ se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi y combinaciones de los mismos,

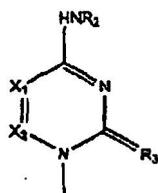
c)



5 en la que: R₂ es H, OH ó NR₄R₅, R₃ es H, OH ó NR₆R₇, R₄ y R₅ se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi y combinaciones de los mismos, y

R₆ y R₇ se seleccionan independientemente de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi y combinaciones de los mismos,

10 d)



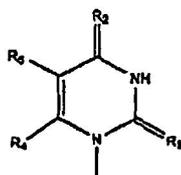
15 en la que:

X₁ y X₂ se seleccionan independientemente de entre CH y N,

R₂ se selecciona de entre H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, o combinaciones de los mismos, y

R₃ es O o S.

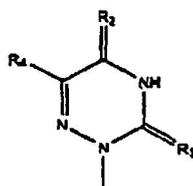
e)



20 en la que:

R₂ y R₃ se seleccionan independientemente de entre O y S, y

25 R₄ y R₅ se seleccionan independientemente de entre H, NH₂, SH, OH, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi, un grupo halo, y combinaciones de los mismos,

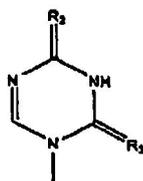


30 en la que:

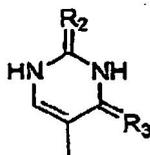
R₂ y R₃ se seleccionan independientemente de entre O y S, y

R₄ es H, NH₂, SH, OH, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi, un grupo halo, o combinaciones de los mismos, en el que R₂ y R₃ se seleccionan independientemente de entre O y S,

35 g)

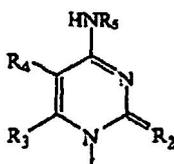


en la que R_2 y R_3 se seleccionan independientemente de entre O y S,
h)



5

en la que R_2 y R_3 se seleccionan independientemente de entre O y S,
i)



10

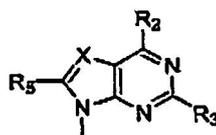
en la que:

R_2 es O ó S, y

R_3 y R_4 se seleccionan independientemente de entre H, NH_2 , SH, OH, COOH, COOCH₃, COOCH₂CH₃, CHO, NO₂, CN, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi, un grupo alcoxi, un grupo halo, o combinaciones de los mismos, y

R_5 es un grupo alquilo, un grupo alcoxi, un grupo alquenilo, un grupo alquenoxi, un grupo alquinilo, un grupo alquinoxio, un grupo arilo, un grupo ariloxi, un grupo bencilo, un grupo benciloxi o combinaciones de los mismos,
j)

20



en la que:

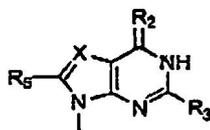
X es CH o N, R_2 y R_3 se seleccionan independientemente de entre H, OH y NHR_4 ,

R_4 es H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi o combinaciones de los mismos, y

R_5 es OH, NH_2 , SH, un grupo halo, un grupo éter, un grupo tioéter, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo alquilamina, un grupo alquenilamina, un grupo alquinilamina o combinaciones de los mismos, y

30

k)



en la que:

X es CH o N, R_2 es O ó S,

R_3 es H, OH ó NHR_4 ,

R_4 es H, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo bencilo, un grupo arilo, un grupo ariloxi o combinaciones de los mismos, y

R_5 es OH, NH_2 , SH, un grupo halo, un grupo éter, un grupo tioéter, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo alquilamina, un grupo alquenilamina, un grupo alquinilamina o combinaciones de los mismos.

40

44. Mezcla de reacción según la reivindicación 40, en la que se une el marcaje a uno de entre una base heterocíclica del nucleótido terminador 2', una fracción sacárida del nucleótido terminador 2' y un grupo fosfato del nucleótido terminador 2'.

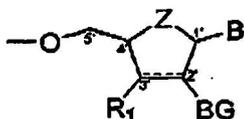
45

45. Mezcla de reacción según la reivindicación 40, en la que un conector une un marcaje al nucleótido terminador 2'.

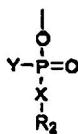
46. Mezcla de reacción según la reivindicación 40, en la que el marcaje comprende un pigmento fluorescente, un

marcaje débilmente fluorescente, un marcaje no fluorescente, un marcaje colorimétrico, un marcaje quimioluminiscente, un marcaje bioluminiscente, un isótopo radioactivo, un anticuerpo, un antígeno, biotina, un hapteno o un enzima.

- 5 47. Mezcla de reacción según la reivindicación 40, que comprende además por lo menos uno o más nucleótidos extensibles.
48. Mezcla de reacción según la reivindicación 47, en la que se marca por lo menos uno de los nucleótidos extensibles.
- 10 49. Mezcla de reacción según la reivindicación 47, en la que el nucleótido terminador 2' y los nucleótidos extensibles se encuentran presentes en una proporción molar de 1:1 ó inferior.
- 15 50. Mezcla de reacción según la reivindicación 40, en la que el biocatalizador que incorpora nucleótidos comprende un enzima seleccionado de entre el grupo que consiste de: una polimerasa, una transferasa terminal, una transcriptasa inversa, una polinucleótido fosforilasa y una telomerasa.
- 20 51. Mezcla de reacción según la reivindicación 50, en la que el enzima comprende una mutación que incrementa la incorporación de ribonucleótidos.
52. Mezcla de reacción según la reivindicación 50, en la que el enzima comprende una mutación que reduce o elimina la actividad de exonucleasa 5' a 3'.
- 25 53. Mezcla de reacción según la reivindicación 50, en la que el enzima comprende una mutación que incrementa la incorporación de análogos 2'-modificados de ribonucleótidos.
54. Mezcla de reacción según la reivindicación 40, que comprende además un ácido nucleico de molde y un cebador de ácidos nucleicos que es por lo menos parcialmente complementario a por lo menos una subsecuencia del ácido nucleico de molde.
- 30 55. Mezcla de reacción según la reivindicación 54, en la que el ácido nucleico de molde o el cebador de ácidos nucleicos se une a un soporte sólido.
56. Mezcla de reacción según la reivindicación 54, en la que el cebador comprende un marcaje.
- 35 57. Mezcla de reacción según la reivindicación 56, en la que el marcaje comprende un pigmento fluorescente, un marcaje débilmente fluorescente, un marcaje no fluorescente, un marcaje colorimétrico, un marcaje quimioluminiscente, un marcaje bioluminiscente, un isótopo radioactivo, un anticuerpo, un antígeno, biotina, un hapteno o un enzima.
- 40 58. Kit para extender un ácido nucleico, que comprende:
 (a) por lo menos un biocatalizador que incorpora nucleótidos, y
 (b) por lo menos un nucleótido terminador 2' marcado,
- 45 en el que el nucleótido terminador 2' marcado comprende la fórmula:



- 50 en la que:
 R₁ es OH,
 B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o combinaciones de los mismos,
 BG es un grupo bloqueante,
- 55 y Z es O ó CH₂, y $\overset{\text{---}}{\text{---}}$ representa un enlace sencillo o doble,
 en el que BG comprende la fórmula:
- en la que:



- 5 X es O, S, NR₃, CR₃R₄ ó SiR₃R₄,
 Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅ ó NHR₅,
 R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueniilo, un grupo alquinilo, un grupo alcoxi, o combinaciones de los mismos, y
 R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueniilo, un grupo alquinilo o combinaciones de los mismos, o



- 10 en la que:
 X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃, o NHR₃,
 R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueniilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y
 R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alqueniilo, un grupo alquinilo o combinaciones de los mismos.

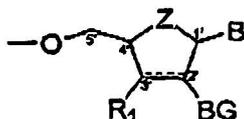
59. Kit según la reivindicación 58, en el que el nucleótido terminador 2' comprende un nucleósido 2'-monofosfato-3'-hidroxilo.
60. Kit según la reivindicación 58, que comprende además uno o más nucleótidos extensibles.
61. Kit según la reivindicación 60, en el que por lo menos uno de los nucleótidos extensibles comprende un marcaje.
62. Kit según la reivindicación 61, en el que el marcaje comprende un isótopo radioactivo, un pigmento fluorescente o un grupo modificador de la masa.
63. Kit según la reivindicación 58, que comprende además por lo menos una pirofosfatasa.
64. Kit según la reivindicación 58, que comprende además:
 (c) un juego de instrucciones para extender el ácido nucleico utilizando el biocatalizador que incorpora nucleótidos y el nucleótido terminador 2' marcado.
65. Kit según la reivindicación 58, en el que el biocatalizador que incorpora nucleótidos comprende un enzima seleccionado de entre el grupo que consiste de: una polimerasa, una transferasa terminal, una transcriptasa inversa, una polinucleótido fosforilasa y una telomerasa.
66. Kit según la reivindicación 65, en el que el enzima comprende una mutación que incrementa la incorporación de ribonucleótidos.
67. Kit según la reivindicación 65, en el que el enzima comprende una mutación que reduce o elimina la actividad de exonucleasa 5' a 3'.
68. Kit según la reivindicación 65, en el que el enzima comprende una mutación que incrementa la incorporación de análogos 2'-modificados de ribonucleótidos.
69. Kit según la reivindicación 65, en el que el enzima comprende una actividad de exonucleasa 3' a 5'.
70. Kit según la reivindicación 58, en el que el marcaje comprende un isótopo radioactivo, un pigmento fluorescente o un grupo modificador de la masa.
71. Kit según la reivindicación 58, en el que el ácido nucleico comprende un cebador de ácidos nucleicos y el kit comprende además un ácido nucleico molde y el cebador de ácidos nucleicos, en el que el cebador de ácidos nucleicos es complementario respecto a por lo menos una subsecuencia del ácido nucleico molde.
72. Kit según la reivindicación 71, en el que el ácido nucleico molde o el cebador de ácidos nucleicos se une a un soporte sólido.

73. Kit según la reivindicación 71, en el que el cebador de ácidos nucleicos comprende un marcaje.

74. Kit según la reivindicación 73, en el que el marcaje comprende un isótopo radioactivo, un pigmento fluorescente o un grupo modificador de la masa.

5
75. Sistema para extender un ácido nucleico, que comprende:
(a) por lo menos un recipiente que comprende un nucleótido terminador 2' marcado, y por lo menos uno de entre:
10 (b) por lo menos un modulador térmico operablemente conectado al recipiente para modular la temperatura en el mismo, y
(c) por lo menos un componente de transferencia de líquidos que transfiere líquidos hacia y/o desde el recipiente,

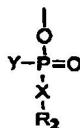
15 en el que el nucleótido terminador 2' marcado comprende la fórmula:



20 en la que:

R₁ es OH, B es por lo menos un anillo homocíclico, por lo menos un anillo heterocíclico, por lo menos un grupo arilo, o combinaciones de los mismos, BG es un grupo bloqueante,

25 y Z es O ó CH₂, y  representa un enlace sencillo o doble, en el que BG comprende la fórmula:



30 en la que:

X es O, S, NR₃, CR₃R₄ ó SiR₃R₄;

Y es CR₅R₆R₇, SiR₅R₆R₇, OR₅, SR₅ ó NHR₅,

35 R₂ es H, OH, NHR₈, SR₈, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y

R₃, R₄, R₅, R₆, R₇ y R₈ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, o combinaciones de los mismos, o

en la que:



40 X es CR₃R₄R₅, SiR₃R₄R₅, OR₃, SR₃, o NHR₃;

R₂ es H, OH, NHR₆, SR₆, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo, un grupo alcoxi o combinaciones de los mismos, y

R₃, R₄, R₅ y R₆ se seleccionan independientemente de entre H, un grupo alquilo, un grupo bencilo, un grupo arilo, un grupo alquenilo, un grupo alquinilo o combinaciones de los mismos.

76. Sistema según la reivindicación 75, que comprende además por lo menos un detector operablemente conectado al recipiente para detectar señales detectables producidas en el recipiente.

77. Sistema según la reivindicación 75, que comprende además por lo menos un controlador operablemente conectado al modulador térmico para llevar a cabo la modulación de la temperatura en el recipiente y/o al componente de transferencia de líquidos para llevar a cabo la transferencia de líquido hacia y/o desde el recipiente.

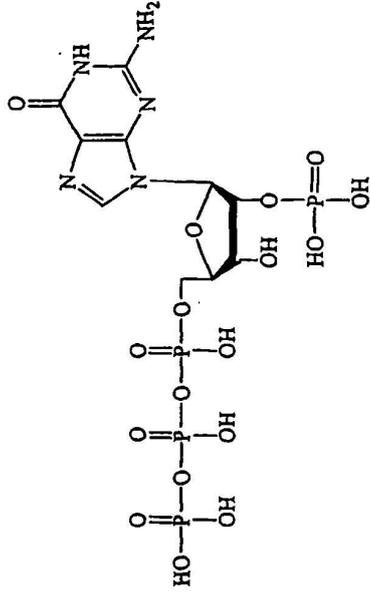


Fig. 1B

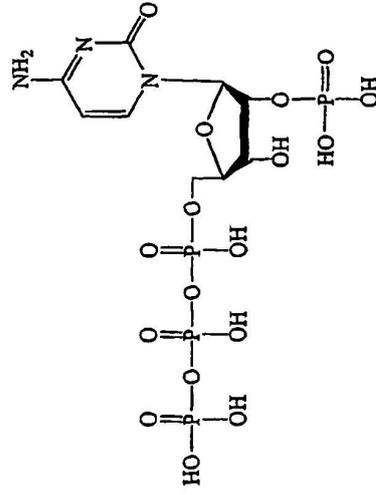


Fig. 1D

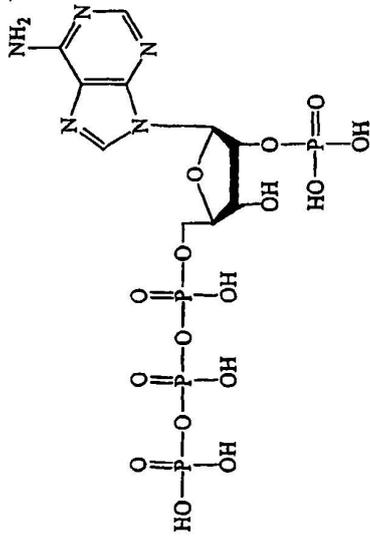


Fig. 1A

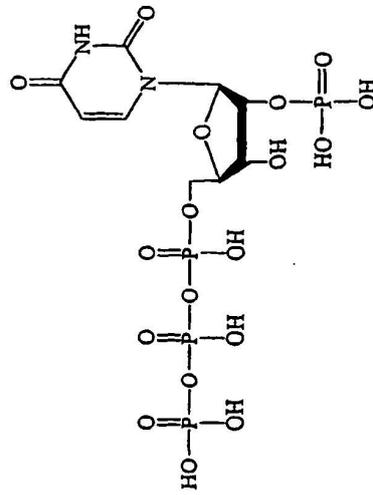


Fig. 1C

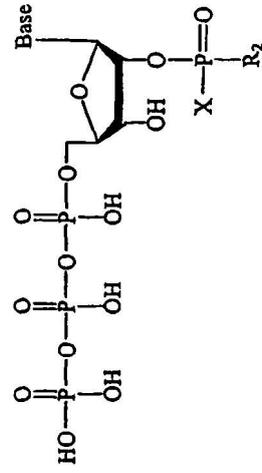


Fig. 2B

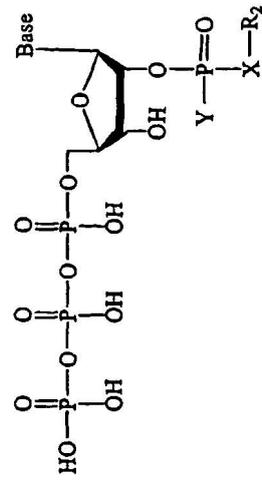


Fig. 2A

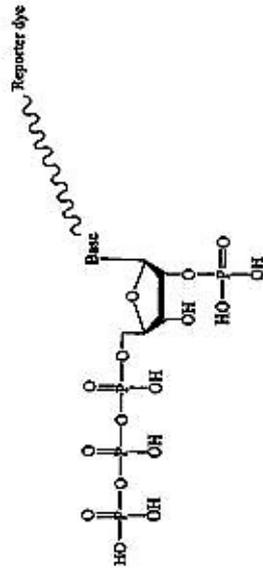


Fig. 3A

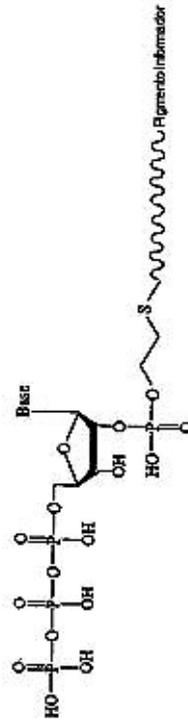


Fig. 3B

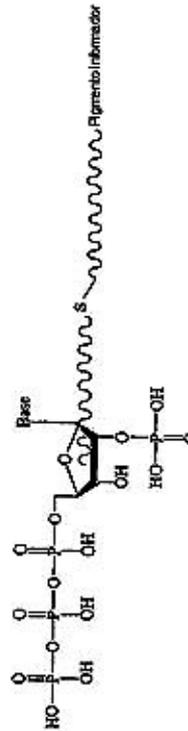
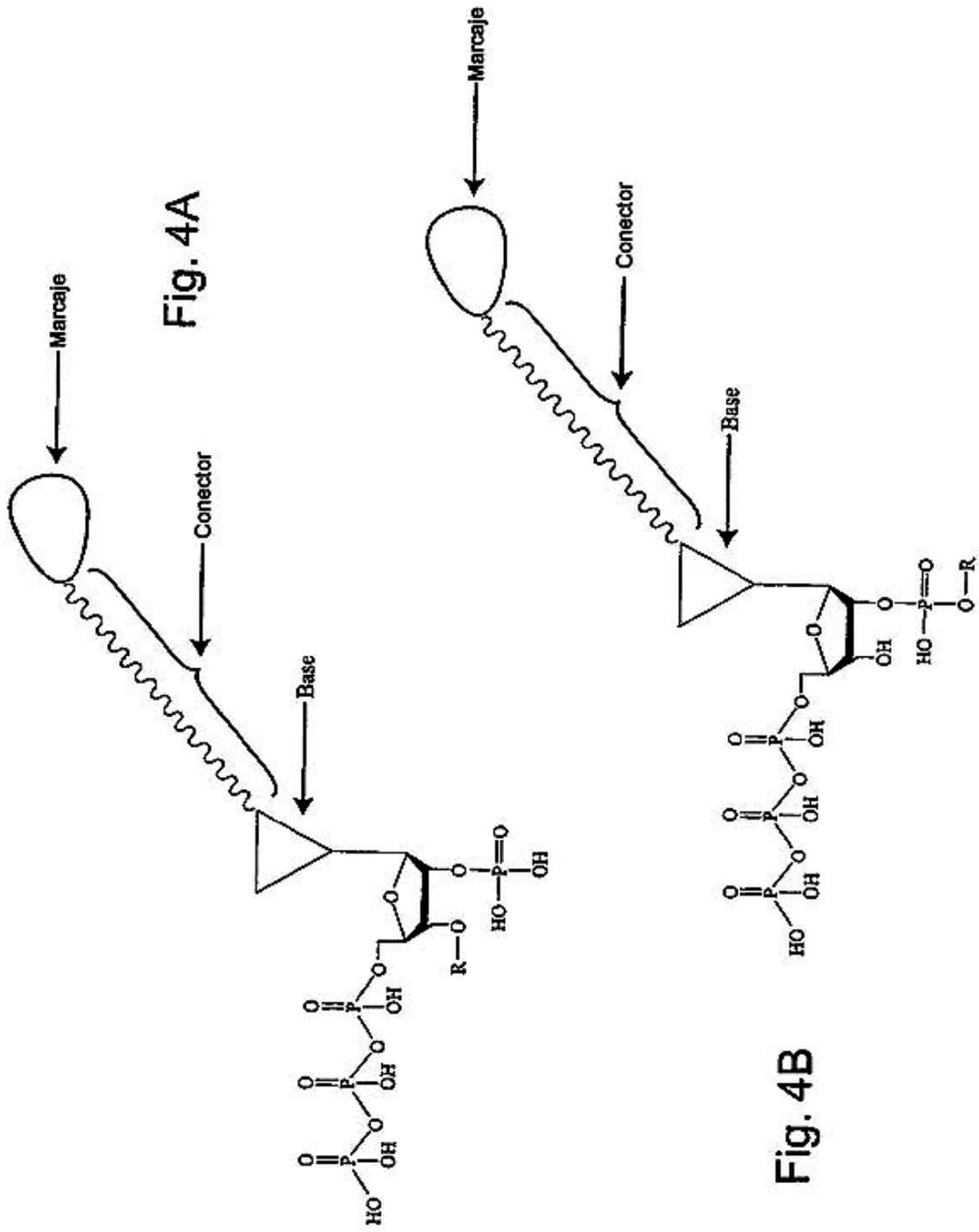


Fig. 3C



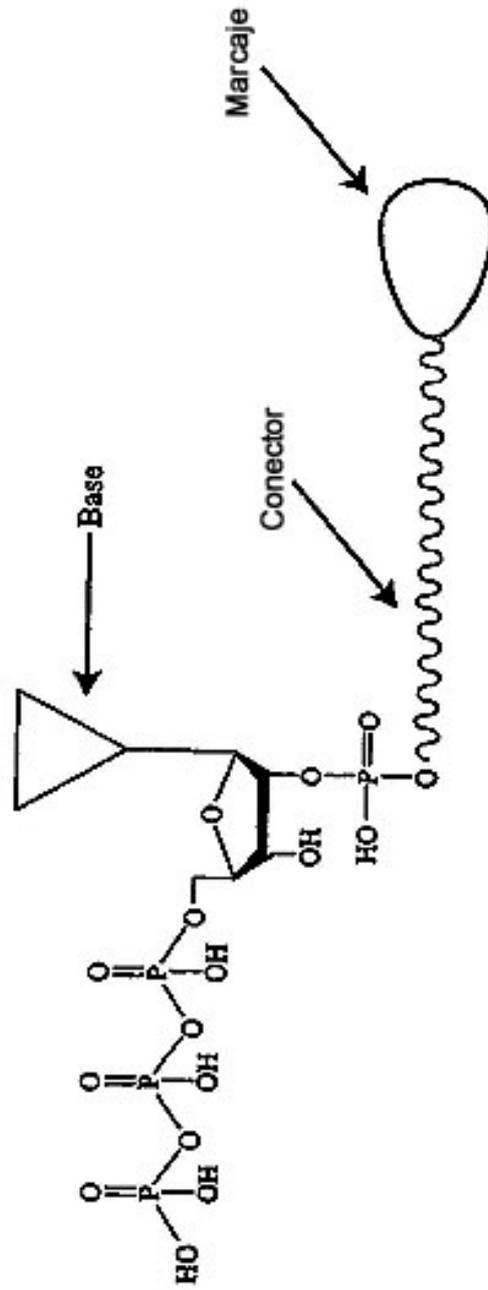


Fig. 5

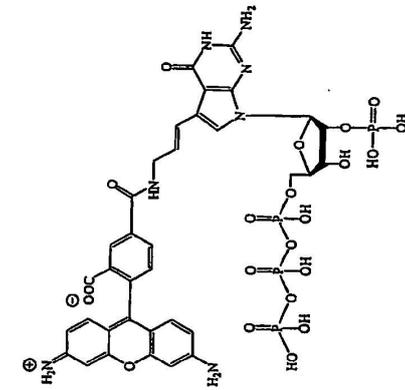


Fig. 6B

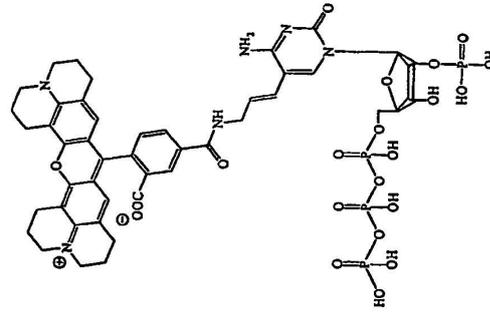


Fig. 6D

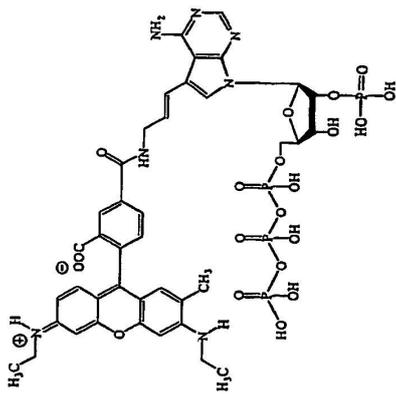


Fig. 6A

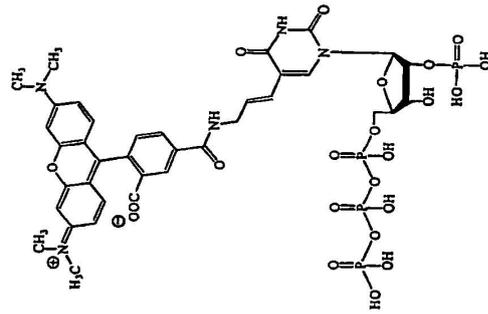


Fig. 6C

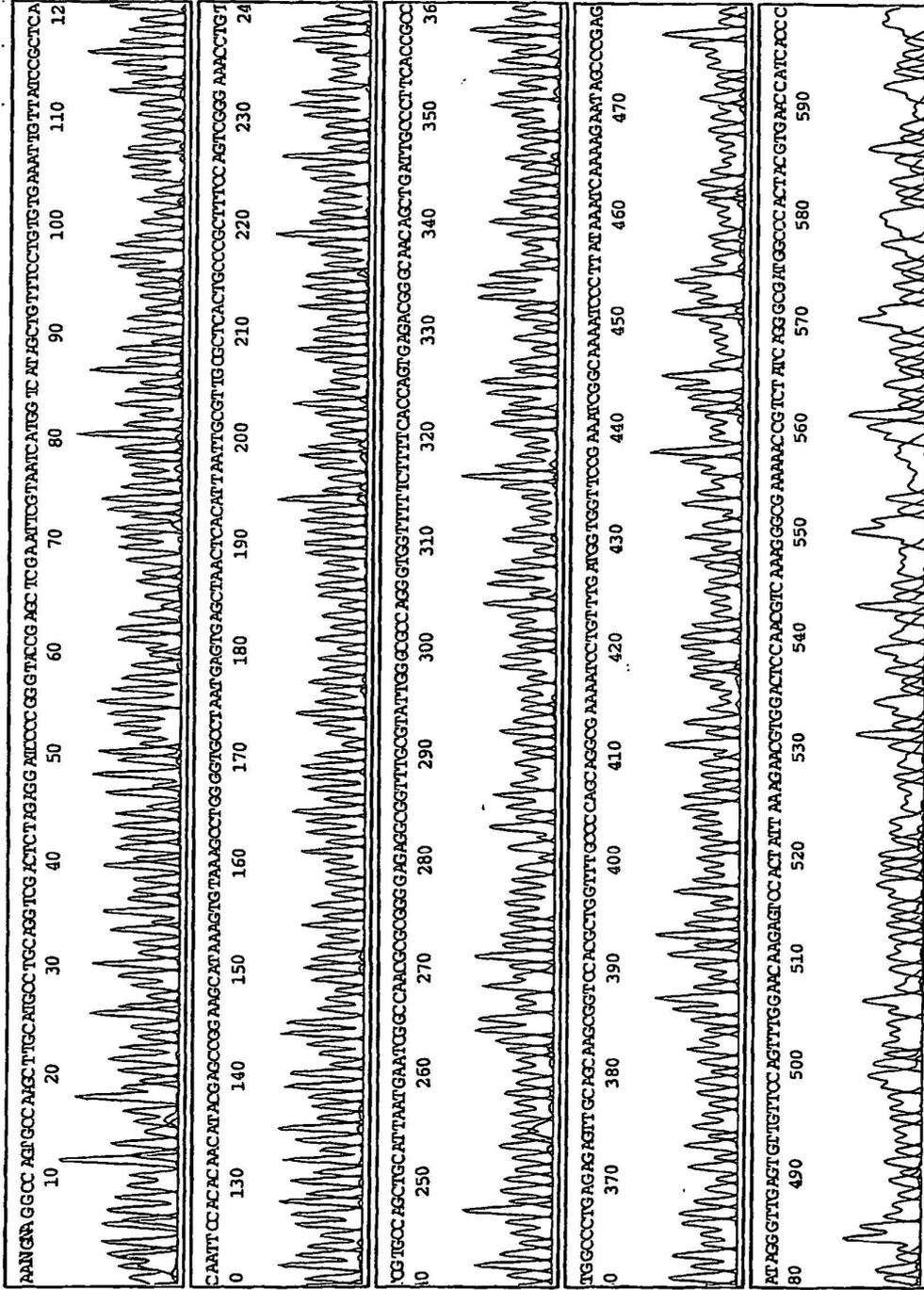


Fig. 7

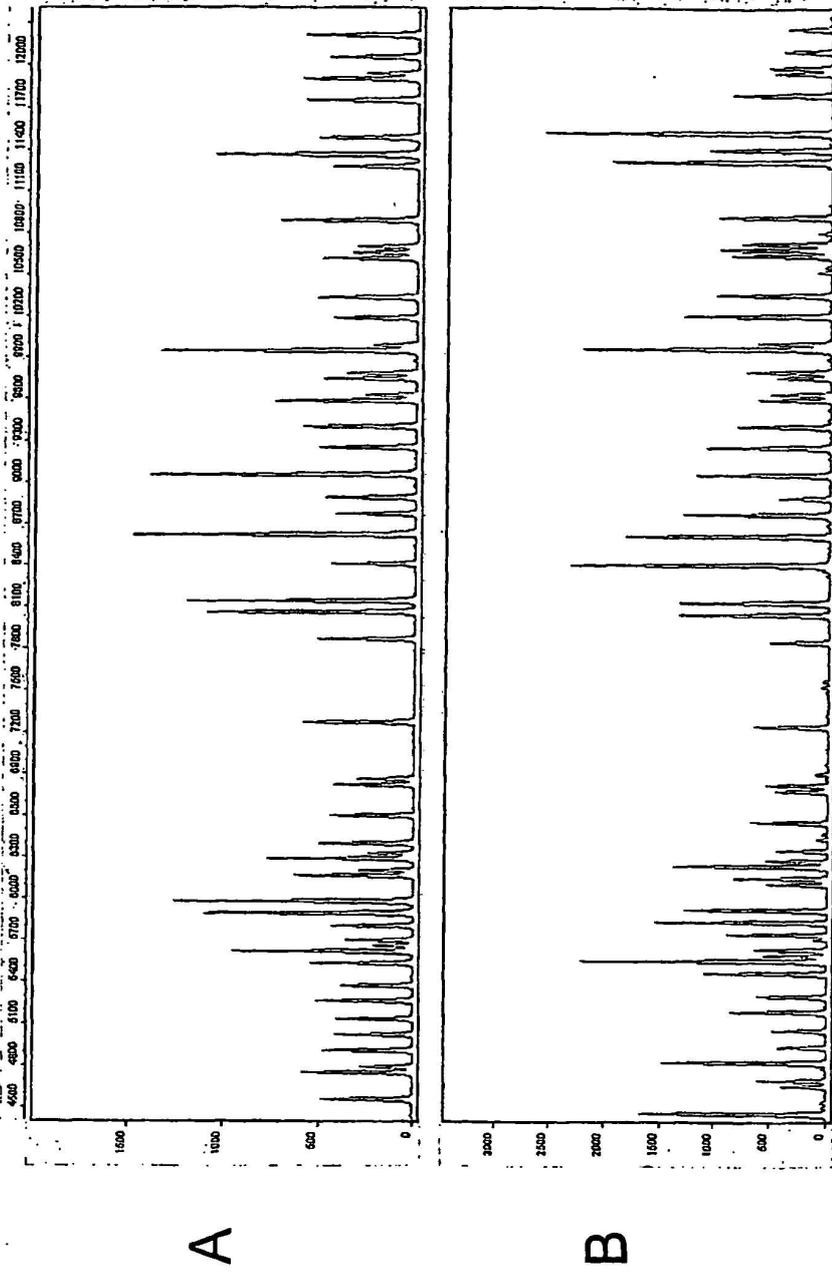


Fig. 8