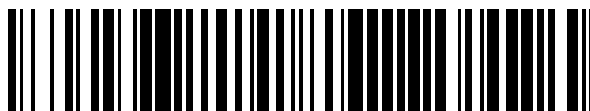


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 929 047**

51 Int. Cl.:

C07K 14/405 (2006.01)

C12N 9/22 (2006.01)

C12N 15/10 (2006.01)

C12N 15/113 (2010.01)

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.06.2017** **PCT/US2017/039133**

87 Fecha y número de publicación internacional: **28.12.2017** **WO17223528**

96 Fecha de presentación y número de la solicitud europea: **23.06.2017** **E 17816348 (1)**

97 Fecha y número de publicación de la concesión europea: **10.08.2022** **EP 3475295**

54 Título: **Transportador de nucleósido trifosfato novedoso y usos del mismo**

30 Prioridad:

24.06.2016 US 201662354650 P

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

24.11.2022

73 Titular/es:

THE SCRIPPS RESEARCH INSTITUTE (100.0%)
10550 North Torrey Pines Road
La Jolla, CA 92037, US

72 Inventor/es:

ROMESBERG, FLOYD E. y
ZHANG, YORKE

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 929 047 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

Transportador de nucleósido trifosfato novedoso y usos del mismo

5 **DECLARACIÓN CON RESPECTO A LA INVESTIGACIÓN CON FONDOS FEDERALES**

La invención desvelada en el presente documento se realizó, al menos en parte, con el apoyo del gobierno de EE.UU. bajo la subvención n.º GM060005 concedida por The National Institutes of Health (NIH). Por consiguiente, el gobierno de EE.UU. posee determinados derechos sobre la invención.

10

LISTADO DE SECUENCIAS

La presente solicitud contiene un Listado de secuencias que se ha presentado electrónicamente en formato ASCII. Dicha copia en ASCII, creada el 22 de junio de 2017, se denomina 46085-707_601_SL.txt y tiene un tamaño de 101.123 bytes.

15

Antecedentes de la divulgación

Los oligonucleótidos y sus aplicaciones han revolucionado la biotecnología. Sin embargo, los oligonucleótidos que incluyen tanto ADN como ARN incluyen solo los cuatro nucleótidos naturales de adenosina (A), guanosina (G), citosina (C), timina (T) para el ADN, y los cuatro nucleótidos naturales de adenosina (A), guanosina (G), citosina (C) y uridina (U) para el ARN, lo que restringe significativamente las funciones y aplicaciones potenciales de los oligonucleótidos.

20

La capacidad de sintetizar/amplificar oligonucleótidos (ADN o ARN) de forma específica de la secuencia con polimerasas, por ejemplo, mediante PCR o sistemas de amplificación isotérmicos (por ejemplo, transcripción con ARN polimerasa T7), ha revolucionado la biotecnología. Además de todas las aplicaciones potenciales en nanotecnología, esto ha permitido una amplia gama de nuevas tecnologías, tales como la evolución *in vitro* a través de *SELEX* (evolución sistemática de ligandos por enriquecimiento exponencial, por sus siglas en inglés) de aptámeros y enzimas de ARN y ADN. Véase, por ejemplo, Oliphant AR, Brandl CJ y Struhl K (1989), Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 proteins, *Mol. Cell Biol.*, 9:2944-2949; Tuerk C y Gold L (1990), Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, 249:505-510; Ellington AD y Szostak JW (1990), *In vitro* selection of RNA molecules that bind specific ligands, *Nature*, 346:818-822.

25

30

En algunos aspectos, estas aplicaciones están restringidas por la limitada diversidad química/física presente en el alfabeto genético natural (los cuatro nucleótidos naturales A, C, G y T en el ADN, y los cuatro nucleótidos naturales A, C, G y U en el ARN).

35

El número de acceso de la base de datos UniParc. UPI0001D42ADE desvela una secuencia de aminoácidos truncada en el extremo N del transportador PtNTT2. El documento WO 2015/157555 A2 desvela un transportador de nucleótido trifosfato expresado de forma recombinante que importa de manera eficiente los nucleótidos trifosfato no naturales a las células.

40

Sumario de la divulgación

En el presente documento, la invención es tal como se describe en las reivindicaciones.

45

En el presente documento, en determinadas realizaciones, se desvela un transportador de nucleósido trifosfato aislado y modificado de *Phaeodactylum tricornutum* (PtNTT2) que comprende una delección, en donde el transportador de nucleósido trifosfato aislado y modificado se obtiene de una célula genomanipulada. En algunas realizaciones, la delección es una delección terminal o una delección interna. En algunas realizaciones, la delección es una delección terminal. En algunas realizaciones, la delección es una delección interna. En algunas realizaciones, la delección terminal es una delección del extremo N, una delección del extremo C o una delección de ambos extremos. En algunas realizaciones, la delección terminal es una delección del extremo N. En algunos aspectos de la divulgación, la delección comprende aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos en el extremo N. En algunas realizaciones, el transportador de nucleósido trifosfato aislado y modificado comprende una delección de aproximadamente 66 residuos de aminoácidos en el extremo N. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4. En algunas realizaciones, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 4. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido

55

60

65

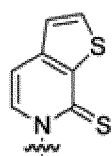
trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 8. En algunas realizaciones, la célula genomanipulada comprende una célula procariota. En algunas realizaciones, la célula genomanipulada es *E. coli*.

En el presente documento, en determinadas realizaciones, se desvela una molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato aislado y modificado descrito anteriormente.

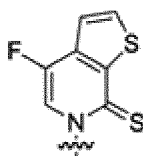
En el presente documento, en determinadas realizaciones, se desvela el uso de un transportador de nucleósido trifosfato modificado descrito anteriormente para la incorporación de un trifosfato no natural durante la síntesis de una molécula de ácido nucleico.

En el presente documento, en determinadas realizaciones, se desvela una célula genomanipulada que comprende una primera molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2). En algunas realizaciones, el ácido nucleico del transportador de nucleósido trifosfato modificado se incorpora en la secuencia genómica de la célula genomanipulada. En algunas realizaciones, la célula genomanipulada comprende un plásmido que comprende el transportador de nucleósido trifosfato modificado. En algunas realizaciones, el transportador de nucleósido trifosfato modificado es un transportador de nucleósido trifosfato de codones optimizados de *Phaeodactylum tricornutum*. En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección. En algunas realizaciones, la delección es una delección terminal o una delección interna. En algunas realizaciones, la delección es un truncamiento del extremo N, un truncamiento del extremo C o un truncamiento de ambos extremos. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos en el extremo N. En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 66 residuos de aminoácidos en el extremo N. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4. En algunas realizaciones, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 4. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 8. En algunas realizaciones, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor seleccionado de un promotor de *E. coli* o un promotor de fagos. En algunas realizaciones, el promotor se selecciona de P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_A , P_{tac} o P_{N25} . En algunas realizaciones, el transportador de nucleósido trifosfato modificado está bajo el control del promotor P_{lacUV5} . En algunas realizaciones, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor de un operón *lac*. En algunas realizaciones, el transportador de nucleósido trifosfato modificado está codificado dentro de un plásmido pSC. En algunas realizaciones, el transportador de nucleósido trifosfato modificado reduce el tiempo de duplicación de la célula genomanipulada. En algunas realizaciones, el transportador de nucleósido trifosfato modificado permite una retención de pares de bases no naturales de aproximadamente el 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 99 % o más. En algunas realizaciones, la célula genomanipulada comprende además una segunda molécula de ácido nucleico que codifica un polipéptido Cas9 o variantes del mismo, una tercera molécula de ácido nucleico que codifica un ARN guía único (ARNgu) que comprende un armazón ARNcr-ARNtracr; y una cuarta molécula de ácido nucleico que comprende un nucleótido no natural. En algunas realizaciones, la segunda molécula de ácido nucleico, la tercera molécula de ácido nucleico y la cuarta molécula de ácido nucleico están codificadas en uno o más plásmidos. En algunas realizaciones, el ARNgu codificado por la tercera molécula de ácido nucleico comprende un motivo diana que reconoce una modificación en la posición del nucleótido no natural dentro de la cuarta molécula de ácido nucleico. En algunas realizaciones, la modificación en la posición del nucleótido no natural dentro de la tercera molécula de ácido nucleico genera una tercera molécula de ácido nucleico modificada. En algunas realizaciones, la modificación es una sustitución. En algunas realizaciones, la modificación es una delección. En algunas realizaciones, la modificación es una inserción. En algunas realizaciones, el ARNgu codificado por la tercera molécula de ácido nucleico comprende además un elemento de reconocimiento de motivo adyacente de protoespaciador (PAM, por sus siglas en inglés). En algunas realizaciones, el elemento PAM es adyacente al extremo 3' del motivo diana. En algunas realizaciones, la combinación del polipéptido Cas9 o variantes del mismo y ARNgu modula la replicación de la cuarta molécula de ácido nucleico modificada. En algunas realizaciones, la combinación del polipéptido Cas9 o variantes del mismo, ARNgu y el transportador de nucleósido trifosfato modificado modula la replicación de la cuarta molécula de ácido nucleico modificada. En algunas realizaciones, la combinación del polipéptido Cas9 o variantes del mismo, ARNgu y el transportador de nucleósido trifosfato modificado disminuye la velocidad de replicación de la cuarta molécula de ácido nucleico modificada en aproximadamente un 80 %, 85 %, 95 %, 99 % o más. En algunas realizaciones, la producción

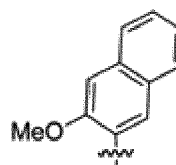
de la cuarta molécula de ácido nucleico en la célula modificada aumenta en aproximadamente un 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o más. En algunas realizaciones, el polipéptido Cas9 o variantes del mismo generan una rotura bicatenaria. En algunas realizaciones, el polipéptido Cas9 es un Cas9 de tipo silvestre. En algunas realizaciones, el nucleótido no natural comprende una base no natural seleccionada del grupo que consiste en 2-aminoadenin-9-ilo, 2-aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, 2-propilo y derivados de alquilo de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-desazaguanina, 3-desazaadenina, 4-tio-uracilo, 4-tio-timina, uracil-5-ilo, hipoxantin-9-ilo (I), 5-metil-citosina, 5-hidroximetil-citosina, xantina, hipoxantina, 5-bromo y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5-sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxixiurea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo y 5-yodouracilo, derivados de 6-alquilo de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azocitosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-desazaguanina, 7-desazaguanosina, 7-desaza-adenina, 7-desaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y adeninas y guaninas sustituidas por 8-hidroxilo; N4-etilcitosina, purinas sustituidas por N-2, purinas sustituidas por N-6, purinas sustituidas por O-6, aquellas que aumentan la estabilidad de la formación de dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos de tamaño expandido, ácidos nucleicos fluorados, pirimidinas tricíclicas, fenoxazina citidina([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazin-2(3H)-ona), pinzas G, fenoxazina citidina (9-(2-aminoetoxy)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido [3',2':4,5]pirrolo [2,3-d]pirimidin-2-ona), 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metio-N6-isopenteniladenina, ácido uracil-5-oxiacético, wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico del ácido uracil-5-oxacético, ácido uracil-5-oxacético, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza por un heterociclo. En algunas realizaciones, la base no natural se selecciona del grupo que consiste en



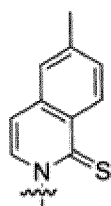
dTPT3



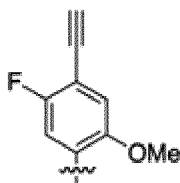
dFTPT3



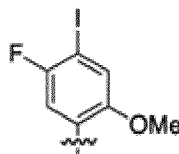
dNaM



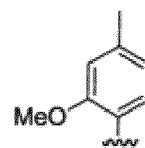
d5SICS



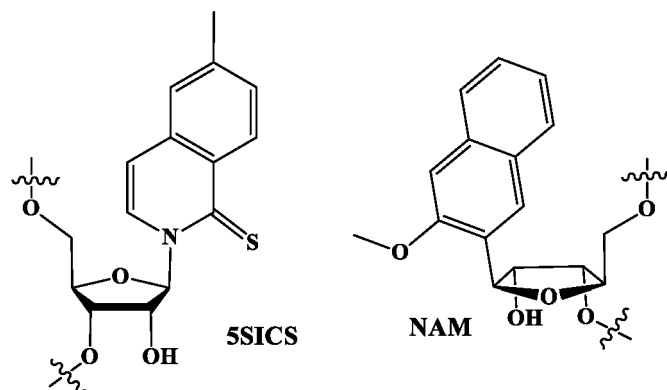
dFEMO



dFIMO



dMMO2



En algunas realizaciones, el nucleótido no natural comprende además un resto de azúcar no natural. En algunas realizaciones, el resto de azúcar no natural se selecciona del grupo que consiste en una modificación en la posición 2': OH; alquilo inferior sustituido, alcarilo, aralquilo, O-alcarilo u O-aralquilo, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F; O-alquilo, S-alquilo, N-alquilo; O-alqueno, S-alqueno, N-alqueno; O-alquino, S-alquino, N-alquino; O-alquil-O-alquilo, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, en donde el alquilo, alqueno y alquino pueden ser alquilo C₁-C₁₀, alqueno C₂-C₁₀, alquino C₂-C₁₀ sustituidos o sin sustituir, -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-ONH₂ y -O(CH₂)_nON[(CH₂)_nCH₃]₂, en donde n y m son de 1 a aproximadamente 10; y/o una modificación en la posición 5': 5'-vinilo, 5'-metilo (R o S), una modificación en la posición 4', 4'-S, heterocicloalquilo, heterocicloalcarilo, aminoalquilamino, polialquilamino, sililo sustituido, un grupo de escisión de ARN, un grupo indicador, un intercalante, un grupo para mejorar las propiedades farmacocinéticas de un oligonucleótido, o un grupo para mejorar las propiedades farmacodinámicas de un oligonucleótido y cualquier combinación de los mismos. En algunas realizaciones, el nucleótido no natural comprende además una cadena principal no natural. En algunas realizaciones, la cadena principal no natural se selecciona del grupo que consiste en fosforotioato, fosforotioato quiral, fosforoditioato, fosfotriéster, aminoalquilfosfotriéster, fosfonatos C₁-C₁₀, fosfonato de 3'-alqueno, fosfonatos quirales, fosfinatos, fosforoamidatos, fosforamidato de 3'-amino, aminoalquilfosforamidatos, tionofosforamidatos, tionoalquilfosfonatos, tionoalquilfosfotriésteres y boranofosfatos. En algunas realizaciones, el ARN_{gu} tiene menos de aproximadamente un 20 %, 15 %, 10 %, 5 %, 3 %, 1 % o menos de velocidad de unión inespecífica. En algunas realizaciones, la célula genomanipulada comprende además una molécula de ácido nucleico adicional que codifica un ARN guía único (ARN_{gu}) adicional que comprende un armazón de ARN_{cr}-ARN_{tracr}. En algunas realizaciones, la célula genomanipulada es un organismo semisintético.

En el presente documento, en determinadas realizaciones, se desvela un método *in vivo* para aumentar la producción de una molécula de ácido nucleico que contiene un nucleótido no natural que comprende una célula genomanipulada descrita anteriormente.

En el presente documento, en determinadas realizaciones, se desvela una molécula de ácido nucleico que contiene un nucleótido no natural producido por una célula genomanipulada descrita anteriormente.

En el presente documento, en determinadas realizaciones, se desvela un plásmido aislado y purificado que comprende una molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PNTT2); y una región promotora seleccionada de un plásmido pSC o locus lacZYA. En algunas realizaciones, el transportador de nucleósido trifosfato modificado es un transportador de nucleósido trifosfato de codones optimizados de *Phaeodactylum tricornutum*. En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección. En algunas realizaciones, la delección es una delección terminal o una delección interna. En algunas realizaciones, la delección es un truncamiento del extremo N, un truncamiento del extremo C o un truncamiento de ambos extremos. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos en el extremo N. En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 66 residuos de aminoácidos en el extremo N. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4. En algunas realizaciones, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 4. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 8. En algunos aspectos de la

divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 8. En algunas realizaciones, la región promotora se selecciona de P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_{λ} , P_{tac} o P_{N25} . En algunas realizaciones, la región promotora se selecciona de P_{lacI} , P_{bla} o P_{lac} . En algunas realizaciones, el plásmido es un plásmido procariota.

5 En el presente documento, en determinadas realizaciones, se desvela un método *in vivo* para aumentar la producción de una molécula de ácido nucleico que contiene un nucleótido no natural que comprende incubar una célula con un plásmido aislado y purificado descrito anteriormente.

10 En el presente documento, en determinadas realizaciones, se desvela un kit que comprende un transportador de nucleósido trifosfato aislado y modificado descrito anteriormente.

En el presente documento, en determinadas realizaciones, se desvela un kit que comprende una célula genomanipulada descrita anteriormente.

15 En el presente documento, en determinadas realizaciones, se desvela un kit que comprende un plásmido aislado y purificado descrito anteriormente.

Breve descripción de los dibujos

20 Diversos aspectos de la divulgación se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de las características y ventajas de la presente divulgación haciendo referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en la que se utilizan los principios de la divulgación y los dibujos adjuntos de los cuales:

25 La figura 1A-figura 1B ilustran las UBP y la optimización del transportador. La figura 1A muestra la estructura química de las UBP **dNaM-d5SICS** y **dNaM-dTPT3** en comparación con el par de bases dC-dG natural. La figura 1B muestra la comparación de la aptitud y la captación de [α - ^{32} P]-dATP en DM1 y las diversas cepas construidas: pCDF y *PtNTT2*(1-575) inducible (gris); pSC y *PtNTT2*(66-575) constitutivo (azul); *PtNTT2*(66-575) integrado y constitutivo (verde). Los triángulos de color blanco representan las cepas de control correspondientes sin *PtNTT2*. Los plásmidos pCDF están en C41(DE3) de *E. coli*; los plásmidos pSC y los integrantes están en BL21(DE3) de *E. coli*. Ninguna de las cepas de *PtNTT2* tiene codones optimizados para la expresión basada en plásmidos ni codones optimizados para la expresión cromosómica, a menos que se indique lo contrario. u.r.d. = unidades relativas de descomposición. Las barras de error representan la d.e. de la media, $n = 3$ cultivos cultivados y ensayados en paralelo; las barras de error en algunos puntos de datos son más pequeñas que su marcador.

35 La figura 2A-figura 2B ilustran una mayor retención de UBP resultante de la optimización del transportador y de UBP. La figura 2A muestra las retenciones de UBP de los plásmidos pUCX1, pUCX2 y pBRX2 en las cepas DM1 e YZ3. Las barras de error representan la d.e. de la media, $n = 4$ transformaciones para pUCX1 y pUCX2, $n = 3$ para DM1 pBRX2 y $n = 5$ para YZ3 pBRX2. La figura 2B muestra las retenciones de UBP de las variantes pUCX2, en donde el UBP está flanqueado por todas las combinaciones posibles de nucleótidos naturales (NXN, en donde N = G, C, A, o T y X = **NaM**), en la cepa YZ3 cultivada en medios complementados con **dNaMTP** y **d5SICSTP** (barras de color gris) o **dNaMTP** y **dTPT3TP** (barras de color negro).

45 La figura 3A-figura 3C ilustran el sistema de edición basado en Cas9. La figura 3A ilustra el modelo de inmunidad mediada por Cas9 con respecto a la pérdida de UBP. La figura 3B muestra retenciones de UBP de variantes pUCX2 en la cepa YZ2 con un plásmido pCas9 que expresa un ARN_g no diana (gris) o un ARN_g específico (negro). Las barras de error representan la d.e. de la media, $n = 3$ transformaciones para todas las secuencias excepto CXA y CXG específicos, en donde $n = 5$. La figura 3C muestra retenciones de UBP de los plásmidos pAIO en la cepa YZ3 (gris), que no expresa Cas9, o en la cepa YZ4 (negro) con expresión de Cas9. En la figura 3B y la figura 3C, se indican los nucleótidos que flanquean inmediatamente X = **NaM**, así como la distancia al PAM. (N) representa el nucleótido N en el ARN_g que se dirige a una mutación de sustitución del UBP; todos los plásmidos pCas9 y pAIO también expresan un ARN_g dirigido a la mutación de delección. Las barras de error representan la d.e. de la media, $n \geq 3$ colonias; véase la figura 11 para conocer los valores exactos de n , las secuencias y las concentraciones de IPTG usadas para inducir Cas9 en YZ4.

55 La figura 4 muestra la retención simultánea de dos UBP durante el crecimiento prolongado. Las cepas YZ3 e YZ4 se transformaron con pAIO2X y se colocaron en placas en medios sólidos que contenían **dNaMTP** y **dTPT3TP**, con o sin IPTG para inducir Cas9. Se inocularon colonias individuales en medios líquidos de la misma composición y los cultivos se cultivaron a una DO_{600} de ~ 2 (punto 1). Posteriormente, los cultivos se diluyeron 30.000 veces y se volvieron a cultivar a una DO_{600} de ~ 2 (punto 2), y a continuación, este proceso de dilución-recrecimiento se repitió dos veces más (puntos 3 y 4). Como control sin inmunidad, la cepa YZ3 se cultivó en ausencia de IPTG y se indican dos cultivos representativos en color gris. La cepa YZ4 se cultivó en presencia de cantidades variables de IPTG y los promedios de los cultivos se indican en color verde (0 μ M, $n = 5$), azul (20 μ M, $n = 5$) y rojo (40 μ M, $n = 4$). Las retenciones del UBP en *gfp* y *serT* se indican con líneas continuas o discontinuas, respectivamente. Después del cuarto crecimiento, dos de los cultivos de YZ4 desarrollados con IPTG 20 μ M se subcultivaron en medios sólidos de la misma composición.

Se inocularon tres colonias seleccionadas al azar de cada placa ($n = 6$ en total) en medios líquidos de la misma composición, y cada uno de los seis cultivos se cultivó a una DO_{600} de ~ 1 (punto 5), se diluyó 300.000 veces en medios que contenían IPTG 0, 20 y 40 μM , y volvió a crecer a una DO_{600} de ~ 1 (punto 6). Este proceso de dilución-recrecimiento se repitió posteriormente (punto 7). Los plásmidos pAIO2X se aislaron en cada uno de los puntos numerados y se analizaron para determinar la retención de UBP. Las duplicaciones celulares se estiman a partir de una DO_{600} (véase la sección Métodos) y no tuvieron en cuenta el crecimiento en medios sólidos (lo que las convierte en una subestimación del crecimiento real). Las barras de error representan la d.e. de la media.

La figura 5A-figura 5F ilustran la captación de dATP y el crecimiento de células que expresan PtNTT2 en función de la concentración del inductor (IPTG) o la fuerza del promotor, la estirpe de la cepa y la presencia de secuencias señal del extremo N. La figura 5A y la figura 5D muestran la captación de $[\alpha\text{-}^{32}P]\text{-dATP}$. Las barras de error representan la d.e. de la media, $n = 3$ cultivos. u.r.d. = unidades relativas de descomposición, lo que corresponde al número total de recuentos radiactivos por minuto normalizados con respecto a la DO_{600} promedio en la ventana de 1 h de captación, estando la captación de C41(DE3) pCDF-1b PtNTT2(1-575) (es decir, DM1) inducida con IPTG 1000 μM establecido en 1. La delección de las secuencias señal del extremo N reduce drásticamente la actividad de captación en C41(DE3), pero la actividad se puede restaurar con niveles más altos de expresión en BL21(DE3). La figura 5B muestra las curvas de crecimiento de las cepas C41(DE3). La inducción de PtNTT2(1-575) es tóxica. La figura 5C muestra las curvas de crecimiento de las cepas BL21(DE3). La inducción de T7 RNAP en BL21(DE3) es tóxica (véanse los rastros del vector vacío), lo que enmascara el efecto de eliminar las secuencias señal del extremo N de PtNTT2 en el crecimiento celular. La figura 5E muestra las curvas de crecimiento de las cepas transportadoras basadas en plásmidos. Las cepas expresan PtNTT2(66-575) sin codones optimizados (co) a menos que se indique lo contrario. La figura 5F muestra las curvas de crecimiento de cepas transportadoras cromosómicamente integradas. Las cepas expresan PtNTT2(66-575) con codones optimizados a menos que se indique lo contrario. La cepa YZ4 también contiene un gen Cas9 integrado cromosómicamente.

La figura 6 muestra mapas de plásmidos. Los promotores y terminadores se representan con características de color blanco y gris, respectivamente. * representa el derivado del origen pMB 1 de pUC19, que contiene una mutación que aumenta su número de copias. Los plásmidos que contienen un UBP generalmente se indican con la secuencia TK1 (naranja), pero como se describe en el texto y se ha indicado anteriormente, las variantes pUCX2 y pAIO con otras secuencias que contienen UBP también posicionan el UBP en el locus aproximado que se muestra con TK1 anteriormente. El ARNgu (N) representa el ARN guía que reconoce una mutación de sustitución natural del UBP, siendo N el nucleótido presente en el ARN guía. El ARNgu (Δ) representa el ARN guía que reconoce una delección de un solo par de bases del UBP; este y su promotor y terminador asociados (indicados por $\pm$) solo están presentes en determinados experimentos. Los genes *serT* y *gfp* no tienen promotores.

La figura 7A-figura 7B ilustra geles de ensayo de desplazamiento de biotina. La figura 7A muestra el esquema de ensayo de desplazamiento de biotina y los geles representativos de la figura 2b. El plásmido de entrada se refiere al producto de ligadura usado para transformar SSO. * representa una banda cuya movilidad no cambia en ausencia de estreptavidina (datos no mostrados) y no aparece en ninguna muestra de cultivos derivados de clones (datos no mostrados). Es probable que la banda corresponda a un plásmido completamente natural derivado del cebado no específico durante la PCR usada para generar la inserción para la ligadura, y está presente en cantidades muy pequeñas en el plásmido de entrada, pero se enriquece durante la replicación *in vivo* por competencia contra secuencias UBP desafiantes. Dichas bandas no se incluyen en el cálculo de la retención. La figura 7B ilustra geles representativos de la figura 4. Cada carril (excluyendo los controles de oligonucleótidos) corresponde a una muestra de plásmido pAIO2X aislada de un cultivo YZ4 derivado de clones, desarrollado con la concentración de IPTG indicada, después de una duplicación estimada de 108 células en cultivo líquido (punto 7 en la figura 4). Cada muestra de plásmido se divide y se analiza en reacciones paralelas de desplazamiento de biotina que ensayan el contenido de UBP en los locus *gfp* y *serT* (cebadores de color rojo y azul, respectivamente). Las muestras de 80 μM no se muestran en el gráfico de la figura 4.

La figura 8A-figura 8B muestra una caracterización adicional de la propagación de UBP. La figura 8A muestra las curvas de crecimiento de los experimentos que se muestran en la figura 2a. YZ3 y DM1 (inducidas con IPTG 1 mM) se transformaron con los plásmidos que contenían UBP indicados, o sus correspondientes controles totalmente naturales, y se cultivaron en medios que contenían dNaMTP y d5SICSTP. Cada línea representa una transformación y el crecimiento posterior en cultivo líquido. El eje x representa el tiempo dedicado al cultivo líquido, excluyendo 1 h de recuperación después de la electroporación (véase la sección Métodos). Las curvas de crecimiento terminan en la DO_{600} a la que se recogieron las células para el aislamiento plasmídico y el análisis de la retención de UBP. Es probable que el escalonamiento de las curvas a lo largo del eje x para las réplicas dentro de una combinación determinada de cepa y plásmido se deba a una variabilidad menor en las frecuencias de transformación entre transformaciones (y, por lo tanto, a las diferencias en el número de células inoculadas en cada cultivo), mientras que las diferencias en la pendiente entre las curvas indican diferencias en la aptitud. El crecimiento de YZ3 es comparable entre los tres plásmidos que contienen UBP (y entre cada plásmido que contiene UBP con su respectivo control natural), mientras que el crecimiento de DM1 se ve afectado por los plásmidos que contienen UBP, especialmente para pUCX1 y pUCX2. La figura 8B muestra las retenciones de variantes *gfp* pUCX2 propagadas en YZ3 por transformación, colocación en placas en medios sólidos, aislamiento de colonias individuales y posterior inoculación y crecimiento en medios líquidos, en comparación con las retenciones de plásmidos propagados por la transformación y el crecimiento de YZ3 en medios

líquidos solamente. Las células se colocaron en placas a partir de las mismas transformaciones descritas en la figura 2b. Los medios sólidos y líquidos contenían dNaMTP y dTPT3TP. Las células se recogieron a $DO_{600} \sim 1$. Se inocularon cinco colonias para cada una de las variantes pUCX2 indicadas, pero algunas colonias no crecieron (indicado por un espacio en blanco en la tabla). Las retenciones para muestras aisladas de transformantes cultivados únicamente en medios líquidos se ensayaron a partir de las mismas muestras que se muestran en la figura 2b, pero se ensayaron y se normalizaron con respecto a un control de oligonucleótidos en paralelo con las muestras de transformantes en placas para facilitar las comparaciones en la retención. Véase la sección Métodos para obtener detalles adicionales sobre la normalización de la retención de UBP. Para las muestras con un cambio cercano a cero, no se puede determinar si el UBP se perdió por completo *in vivo* o si la muestra provino de una colonia que se transformó con un plásmido completamente natural (algo de lo cual surge durante la construcción del plásmido, específicamente durante la PCR usada para generar la inserción que contiene UBP).

La figura 9A-figura 9B ilustran el efecto de dNaM-dTPT3 sobre la escisión de ADN mediada por Cas9 *in vitro*. Se ensayó la escisión *in vitro* mediada por Cas9 para seis sustratos de ADN, en donde el tercer par de bases cadena arriba del PAM era uno de los cuatro pares de bases naturales o el UBP (en ambos contextos de hebra). Los cuatro ARN_g que son complementarios a cada molde natural se prepararon mediante transcripción *in vitro* con T7 RNAP. Para tener en cuenta las diferencias en la actividad de ARN_g y/o variaciones menores en la preparación, se muestra entre paréntesis un porcentaje relativo de escisión máxima para cada ARN_g frente a los seis sustratos de ADN. Los valores representan medias ± 1 d.e. ($n = 3$ réplicas técnicas). En varios casos, la presencia de un nucleótido no natural redujo significativamente la escisión en comparación con el ADN complementario al ARN_g. Estos datos sugieren que Cas9 programado con uno o más ARN_g complementarios a una o más de las secuencias naturales degradaría preferentemente el ADN que había perdido el UBP. La figura 9B desvela las SEQ ID NO 201 y 202, respectivamente, en orden de aparición.

La figura 10A-figura 10D muestra la inmunidad mediada por Cas9 con respecto a la pérdida de UBP en TK1. La figura 10A muestra secuencias de ARN_g usadas para mejorar la retención del UBP (SEQ ID NO 203, 204, 204, 205 y 206, respectivamente, en orden de aparición). La figura 10B muestra que la retención de UBP para pUCX2 TK1 se potencia al dirigir Cas9 a la mutación principal (dTPT3→dA). A medida que las células continúan creciendo en ausencia de ARN_g correctos dirigidos a las mutaciones, la retención de UBP disminuye. Las barras de error representan la d.e. de la media, $n = 3$ transformaciones. En la figura 10A y la figura 10B, hEGFP es un ARN_g no diana. La figura 10C muestra el cromatograma de secuenciación de Sanger que ilustra la mutación de dNaM con respecto a dT en ausencia de un ARN_g para dirigir la actividad de la nucleasa Cas9 (SEQ ID NO 207 y 208, respectivamente, en orden de aparición). La figura 10D muestra el cromatograma de secuenciación de Sanger que ilustra que la pérdida de retención en presencia de Cas9 y un ARN_g dirigido (TK1-A) se debe al crecimiento de células con plásmidos que poseen una mutación de un solo par de bases. Las especies que contenían UBP se agotaron antes de la secuenciación. La figura 10D desvela las SEQ ID NO 207 y 208, respectivamente, en orden de aparición.

La figura 11 muestra secuencias Cas9 NXN. Anteriormente se muestran 22 nt de cada secuencia que contiene UBP examinada en la figura 3. X = dNaM, Y = dTPT3 (SEQ ID NO 209-224, respectivamente, en orden de aparición). La secuencia del ARN_g que se dirige a la mutación de sustitución del UBP (N) es la secuencia de 18 nt 5' con respecto a NGG PAM, estando X o Y reemplazado por el nucleótido natural indicado. La secuencia del ARN_g que se dirige a la mutación de delección de UBP (Δ) es la secuencia de 19 nt 5' con respecto a NGG PAM sin X o Y. Se realizaron experimentos con YZ3 sin IPTG. Las retenciones que se muestran en la figura 3C se promedian a partir de los valores y el número de colonias que se indican aquí.

La figura 12 muestra las curvas de crecimiento de YZ4 que replica pAIO2X. Curvas de crecimiento para la primera dilución-recrecimiento (punto 2) en la figura 4. Las curvas terminan a la DO_{600} a la que se recogieron los cultivos tanto para el aislamiento plasmídico como para la dilución para el siguiente recrecimiento. Los tiempos de duplicación se calculan a partir de los puntos temporales recopilados entre DO_{600} 0,1-1,0 para cada curva y se promedian para cada cepa y/o condición de IPTG.

Descripción detallada de la divulgación

Los nucleósidos son moléculas hidrófilas que requieren proteínas de transporte para la permeación de las membranas celulares. Los transportadores de nucleósidos (NT) son un grupo de proteínas de transporte de membrana que facilitan el cruce de los nucleósidos a través de las membranas celulares y las vesículas. En algunos casos, hay dos tipos de transportadores de nucleósidos, transportadores de nucleósidos de concentración que impulsan un proceso de concentración por gradiente electroquímico, y transportadores de nucleósidos de equilibrio que impulsan un proceso bidireccional de equilibrio por gradiente químico. En algunos casos, un transportador de nucleósidos abarca además un transportador de nucleósido trifosfato.

Los nucleósidos naturales comprenden adenina, guanina, timina, uracilo y citosina; y son reconocidos por los transportadores de nucleótidos para la permeación de las membranas celulares. Los nucleósidos no naturales, en algunos casos, no son reconocidos por transportadores de nucleótidos endógenos o son reconocidos pero la eficiencia del transporte es baja.

En algunas realizaciones, en el presente documento se describen transportadores de nucleótidos modificados que reconocen y facilitan el transporte de ácidos nucleicos no naturales a una célula. En algunos casos, el transportador de nucleótidos modificado aumenta la importación de ácidos nucleicos no naturales a una célula en relación con un transportador de nucleótidos endógeno. En algunos casos, el transportador de nucleótidos modificado aumenta la retención de ácidos nucleicos no naturales dentro de una célula. En casos adicionales, el transportador de nucleótidos modificado minimiza la toxicidad debido a su expresión y, opcionalmente, mejora el tiempo de duplicación celular y la aptitud en relación con una célula en ausencia del transportador.

Transportadores de nucleósido trifosfato

En determinadas realizaciones, en el presente documento se describen transportadores de nucleósidos trifosfatos modificados para transportar ácidos nucleicos no naturales a una célula. En algunos casos, el transportador de nucleósido trifosfato modificado es de *Phaeodactylum tricornutum* (PtNTT2). En algunos casos, el transportador de nucleósido trifosfato modificado comprende además una delección. En algunos casos, la delección es una delección terminal (por ejemplo, una delección del extremo N o una delección del extremo C) o es una delección interna.

En algunas realizaciones, en el presente documento se describe un transportador de nucleósido trifosfato aislado y modificado de *Phaeodactylum tricornutum* (PtNTT2) que comprende una delección. En algunos aspectos de la divulgación, la delección comprende aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70, 80, 90 o más residuos de aminoácidos. En algunos aspectos de la divulgación, la delección comprende aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 5 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 10 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 15 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 20 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 22 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 25 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 30 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 40 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 44 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 50 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 60 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 66 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección de aproximadamente 70 o más residuos de aminoácidos.

En algunas realizaciones, en el presente documento se describe un transportador de nucleósido trifosfato aislado y modificado de *Phaeodactylum tricornutum* (PtNTT2) que comprende una delección del extremo N. En algunos casos, la delección del extremo N comprende aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70, 80, 90 o más residuos de aminoácidos. En algunos casos, la delección del extremo N comprende aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 5 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 10 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 15 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 20 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 22 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato aislado y modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 25 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 30 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 40 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) comprende una delección del extremo N de aproximadamente 44 o más residuos de aminoácidos. En algunos casos, el transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum*

modificado que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia un ácido nucleico natural y/o el ácido nucleico no natural sin el azúcar modificado. En algunas realizaciones, el transportador de nucleósido trifosfato modificado tiene una especificidad por un ácido nucleico no natural que comprende una base modificada que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia un ácido nucleico natural y/o el ácido nucleico no natural sin la base modificada. En algunas realizaciones, el transportador de nucleósido trifosfato modificado tiene una especificidad por un ácido nucleico no natural que comprende un trifosfato que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia un ácido nucleico que comprende un trifosfato y/o el ácido nucleico no natural sin el trifosfato. Por ejemplo, un transportador de nucleósido trifosfato modificado puede tener una especificidad por un ácido nucleico no natural que comprende un trifosfato que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia el ácido nucleico no natural con un difosfato o monofosfato, o sin fosfato, o una combinación de los mismos.

En algunas realizaciones, un transportador de nucleósido trifosfato modificado descrito en el presente documento tiene una especificidad por un ácido nucleico no natural y una especificidad por un ácido nucleico natural de al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia el ácido nucleico natural. En algunas realizaciones, el transportador de nucleósido trifosfato modificado tiene una especificidad por un ácido nucleico no natural que comprende un azúcar modificado y una especificidad por un ácido nucleico natural que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia el ácido nucleico natural. En algunas realizaciones, el transportador de nucleósido trifosfato modificado tiene una especificidad por un ácido nucleico no natural que comprende una base modificada y una especificidad por un ácido nucleico natural que es al menos aproximadamente un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad del transportador de nucleósido trifosfato de tipo silvestre hacia el ácido nucleico natural.

En algunas realizaciones, una secuencia de un transportador de nucleósido trifosfato modificado se modifica adicionalmente para mejorar la expresión y la actividad celular. En algunos casos, el uso de codones se modifica para introducir sitios de pausa ribosómicos para ralentizar la traducción y mejorar el direccionamiento del polipéptido transportador de nucleósido trifosfato modificado a translocónes de membrana (Fluman, *et al.*, "mRNA-programmed translation pauses in the targeting of *E. coli* membrane proteins", *eLife* 2014; 3:e03440). En algunos casos, la modificación de una o más hélices transmembrana, por ejemplo, la modificación de una primera hélice transmembrana y/o la generación de un transportador quimérico que comprende una primera hélice transmembrana de una proteína diferente (por ejemplo, un transportador relacionado) puede mejorar la expresión y las actividades celulares (Marshall, *et al.*, "A link between integral membrane protein expression and simulated integration efficiency", *Cell Reports*, 16(8): 2169-2177 (2016)). En algunos casos, se incorpora un péptido señal endógeno, modificado o heterólogo en la secuencia de un transportador de nucleósido trifosfato modificado para mejorar la expresión y la actividad celular. En algunos casos, el péptido señal se une opcionalmente en marco con la secuencia del transportador de nucleósido trifosfato modificado a través de un enlazador. En algunos casos, el enlazador es un enlazador no escindible. En otros casos, el enlazador es un enlazador escindible. Los péptidos señal de ejemplo se ilustran en la Tabla 3. En algunos casos, un péptido señal de la Tabla 3, opcionalmente unido a un enlazador, se incorpora a la secuencia de un transportador de nucleósido trifosfato modificado descrito en el presente documento.

En algunas realizaciones, la expresión del transportador de nucleósido trifosfato modificado se ajusta a través de la modificación del sitio de unión al ribosoma para modular la velocidad de síntesis del polipéptido transportador de nucleósido trifosfato modificado. Véanse, por ejemplo, Howard, *et al.*, "Automated design of synthetic ribosome binding sites to control protein expression", *Nature Biotechnology* 27: 946-950 (2009); y Mutalik, *et al.*, "Precise and reliable gene expression via standard transcription and translation initiation elements", *Nature Methods* 10: 354-360 (2013).

En algunas realizaciones, la expresión del transportador de nucleósido trifosfato modificado se modula mediante la unión de una etiqueta de degradación ajustable. En algunos casos, una etiqueta de degradación ajustable comprende una pequeña secuencia de aminoácidos que, cuando se fusiona con una proteína diana, marca la proteína para su degradación por una proteasa análoga en una célula bacteriana. La etiqueta de degradación ajustable y los pares de proteasas análogos de ejemplo y incluyen, pero sin limitación, *ssrA* de *E. coli* (*ec-ssrA*)/Lon de *E. coli* (*ec-Lon*), y *ssrA* de *Mesoplasma florum* (*mf-ssrA*)/Lon de *Mesoplasma florum* (*mf-Lon*). En algunos casos, la etiqueta de degradación ajustable comprende una etiqueta modificada que altera la dinámica de expresión y/o degradación con respecto a una etiqueta de degradación no modificada. En algunos casos, una etiqueta de degradación ajustable contemplada en el presente documento comprende una etiqueta de degradación descrita en la Publicación de Patente PCT WO2014/160025A2. En algunos casos, una etiqueta de degradación ajustable contemplada en el presente documento comprende una etiqueta de degradación descrita en Cameron, *et al.*, "Tunable protein degradation in bacteria", *Nature Biotechnology* 32: 1276-1281 (2014).

En algunas realizaciones, la expresión del transportador de nucleósido trifosfato modificado se modula mediante la disponibilidad de una molécula endógena o exógena (por ejemplo, nucleótido trifosfato no natural o aminoácido no natural) durante la traducción. En algunos casos, la expresión del transportador de nucleósido trifosfato modificado se correlaciona con el número de copias de los codones raros, en los que la velocidad de lectura ribosómica de un codón raro modula la traducción del transportador. Véase, por ejemplo, Wang, *et al.*, "An engineered rare codon device for optimization of metabolic pathways", Scientific Reports 6:20608 (2016).

En algunos casos, un transportador de nucleósido trifosfato modificado se caracteriza de acuerdo con su velocidad de disociación de un sustrato de ácido nucleico. En algunas realizaciones, un transportador de nucleósido trifosfato modificado tiene una velocidad de disociación relativamente baja para uno o más ácidos nucleicos naturales y no naturales. En algunas realizaciones, un transportador de nucleósido trifosfato modificado tiene una velocidad de disociación relativamente alta para uno o más ácidos nucleicos naturales y no naturales. La velocidad de disociación es una actividad de un transportador de nucleósido trifosfato modificado y aislado que se puede ajustar para afinar las velocidades de reacción en los métodos expuestos en el presente documento.

Los transportadores de nucleósido trifosfato modificados de fuentes nativas o variantes de los mismos pueden cribarse usando un ensayo que detecta la importación de un ácido nucleico no natural que tenga una estructura particular. En un ejemplo, los transportadores de nucleósido trifosfato modificados se pueden cribar en cuanto a la capacidad de importar un ácido nucleico no natural o UBP; por ejemplo, d5SICSTP, dNaMTP o d5SICSTP-dNaMTP UBP. Se puede usar un NTT, por ejemplo, un transportador heterólogo, que muestra una propiedad modificada para el ácido nucleico no natural en comparación con el transportador de tipo silvestre. Por ejemplo, la propiedad modificada puede ser, por ejemplo, K_m , k_{cat} , $V_{m\acute{a}x}$, la importación de NTT en presencia de un ácido nucleico no natural (o de un nucleótido de origen natural), la longitud de lectura de molde promedio por una célula con el transportador de nucleósido trifosfato modificado en presencia de un ácido nucleico no natural, la especificidad del transportador por un ácido nucleico no natural, la velocidad de unión de un ácido nucleico no natural o la velocidad de liberación del producto, o cualquier combinación de los mismos. En una realización, la propiedad modificada es una reducción de K_m para un ácido nucleico no natural y/o un aumento de k_{cat}/K_m o $V_{m\acute{a}x}/K_m$ para un ácido nucleico no natural. De manera similar, el transportador de nucleósido trifosfato modificado tiene opcionalmente una mayor velocidad de unión de un ácido nucleico no natural, una mayor velocidad de liberación de producto y/o una mayor velocidad de importación de células, en comparación con un transportador de tipo silvestre.

Al mismo tiempo, un transportador de nucleósido trifosfato modificado puede importar ácidos nucleicos naturales, por ejemplo, A, C, G y T, a la célula. Por ejemplo, un transportador de nucleósido trifosfato modificado muestra opcionalmente una actividad de importación específica para un ácido nucleico natural que es al menos aproximadamente un 5 % tan alta (por ejemplo, 5 %, 10 %, 25 %, 50 %, 75 %, 100 % o más), como un transportador de tipo silvestre correspondiente. Opcionalmente, el transportador de nucleósido trifosfato modificado muestra una k_{cat}/K_m o $V_{m\acute{a}x}/K_m$ para un nucleótido de origen natural que es al menos aproximadamente un 5 % tan alta (por ejemplo, aproximadamente un 5 %, 10 %, 25 %, 50 %, 75 % o un 100 % o más) como en el NTT de tipo silvestre.

Los transportadores de nucleósido trifosfato modificados usados en el presente documento que pueden tener la capacidad de importar un ácido nucleico no natural de una estructura particular también se pueden producir usando un enfoque de evolución dirigida. Se puede usar un ensayo de síntesis de ácido nucleico para cribar variantes de transportadores que tengan especificidad por cualquiera de diversos ácidos nucleicos no naturales. Por ejemplo, las variantes de transportadores se pueden cribar para determinar la capacidad de importar un ácido nucleico no natural o UBP; por ejemplo, d5SICSTP, dNaMTP o d5SICSTP-dNaMTP UBP en ácidos nucleicos. En algunas realizaciones, tal ensayo es un ensayo *in vitro*, por ejemplo, que usa una variante de transportador recombinante. En algunas realizaciones, tal ensayo es un ensayo *in vivo*, por ejemplo, que expresa una variante de transportador en una célula. Dichas técnicas de evolución dirigida se pueden usar para cribar variantes de cualquier transportador adecuado para determinar la actividad hacia cualquiera de los ácidos nucleicos no naturales expuestos en el presente documento.

Células genomanipuladas

En algunas realizaciones, en el presente documento se describe una célula genomanipulada que comprende una molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato modificado. En algunos casos, la molécula de ácido nucleico codifica un transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2). En algunos casos, el ácido nucleico del transportador de nucleósido trifosfato modificado se incorpora en la secuencia genómica de la célula genomanipulada.

La célula genomanipulada puede ser cualquier procariota adecuado. En algunos casos, la célula genomanipulada es una bacteria gramnegativa. En otros casos, la célula genomanipulada es una bacteria grampositiva. Las bacterias de ejemplo incluyen, pero sin limitación, bacterias *Bacillus* (por ejemplo, *B. subtilis*, *B. megaterium*), bacterias *Acinetobacter*, bacterias *Norcardia*, bacterias *Xanthobacter*, bacterias *Escherichia* (por ejemplo, *E. coli* (por ejemplo, las cepas DH10B, Stbl2, DH5-alfa, DB3, DB3.1), DB4, DB5, JDP682 y ccdA-over (por ejemplo, la Solicitud de EE.UU. N.º 09/518.188))), bacterias *Streptomyces*, bacterias *Erwinia*, bacterias *Klebsiella*, bacterias *Serratia* (por ejemplo, *S. marcessans*), bacterias *Pseudomonas* (por ejemplo, *P. aeruginosa*), bacterias *Salmonella* (por ejemplo, *S.*

typhimurium, *S. typhi*), bacterias *Megasphaera* (por ejemplo, *Megasphaera elsdenii*). Las bacterias también incluyen, pero sin limitación, bacterias fotosintéticas (por ejemplo, bacterias verdes no de azufre (por ejemplo, bacterias *Chloroflexus* (por ejemplo, *C. aurantiacus*), bacterias *Chloronema* (por ejemplo, *C. gigateum*)), bacterias verdes del azufre (por ejemplo, bacterias *Chlorobium* (por ejemplo, *C. limicola*), bacterias *Pelodictyon* (por ejemplo, *P. luteolum*), bacterias púrpuras del azufre (por ejemplo, bacterias *Chromatium* (por ejemplo, *C. okenii*) y bacterias púrpuras no del azufre (por ejemplo, bacterias *Rhodospirillum* (por ejemplo, *R. rubrum*), bacterias *Rhodobacter* (por ejemplo, *R. sphaeroides*, *R. capsulatus*) y bacterias *Rhodomicrobium* (por ejemplo, *R. vanellii*)).

En algunos casos, la célula genomanipulada comprende un plásmido que comprende el transportador de nucleósido trifosfato modificado. En algunos casos, el transportador de nucleósido trifosfato modificado es un transportador de nucleósido trifosfato de codones optimizados de *Phaeodactylum tricornutum*.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección. En algunos casos, la delección es una delección terminal (por ejemplo, una delección del extremo N o del extremo C). En otros casos, la delección es una delección interna.

Como se ha descrito anteriormente, el transportador de nucleósido trifosfato modificado desvelado en el presente documento comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos aspectos de la divulgación, la delección es una delección del extremo N, y el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos en el extremo N. En algunos casos, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 66 residuos de aminoácidos en el extremo N.

En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4, 6 u 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4. En algunos casos, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 4. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 8.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor. En algunos casos, el promotor se obtiene de una fuente de *E. coli*. En otros casos, el promotor se obtiene de una fuente de fago. Los promotores de ejemplo incluyen, pero sin limitación, P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_{λ} , P_{tac} o $PN25$. En algunos casos, el promotor reemplaza al operón lac . En algunos casos, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor seleccionado de P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_{λ} , P_{tac} o $PN25$. En algunos casos, el transportador de nucleósido trifosfato modificado está bajo el control del promotor P_{lacUV5} .

En algunos casos, el transportador de nucleósido trifosfato modificado está codificado dentro de un plásmido pSC.

En algunas realizaciones, la célula genomanipulada comprende además una segunda molécula de ácido nucleico que codifica un polipéptido Cas9 o variantes del mismo, una tercera molécula de ácido nucleico que codifica un ARN guía único (ARNgu) que comprende un armazón ARNcr-ARNtracr; y una cuarta molécula de ácido nucleico que comprende un nucleótido no natural.

El sistema CRISPR/Cas implica (1) una integración de regiones cortas de material genético que son homólogas a una molécula de ácido nucleico de interés que comprende un nucleótido no natural, denominadas "espaciadores", en matrices agrupadas en el genoma huésped, (2) la expresión de ARN guía cortos (ARNcr) de los espaciadores, (3) la unión de los ARNcr a porciones específicas de la molécula de ácido nucleico de interés denominadas protoespaciadores, y (4) la degradación de protoespaciadores por nucleasas asociadas a CRISPR (Cas). En algunos casos, se ha descrito un sistema CRISPR tipo II en la bacteria *Streptococcus pyogenes*, en el que Cas9 y dos pequeños ARN no codificantes (pre-ARNcr y ARNtracr (ARN CRISPR transactivador)) actúan en conjunto para dirigirse a y degradar una molécula de ácido nucleico de interés en una manera específica de secuencia (Jinek *et al.*, "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity", Science 337(6096):816-821 (agosto de 2012, epub 28 de junio de 2012)).

En algunos casos, un sistema CRISPR/Cas utiliza un polipéptido Cas9 o una variante del mismo. Cas9 es una nucleasa bicatenaria con dos sitios de corte activos, uno para cada cadena de la doble hélice. En algunos casos, el polipéptido Cas9 o variantes del mismo generan una rotura bicatenaria. En algunos casos, el polipéptido Cas9 es un

Cas9 de tipo silvestre. En algunos casos, el polipéptido Cas9 es un Cas9 optimizado para la expresión en una célula genomanipulada descrita en el presente documento.

En algunos casos, los dos ARN no codificantes se fusionan adicionalmente en un ARN guía único (ARNgu). En algunos casos, el ARNgu comprende un motivo diana que reconoce una modificación en la posición del nucleótido no natural dentro de una molécula de ácido nucleico de interés. En algunas realizaciones, la modificación es una sustitución, inserción o delección. En algunos casos, el ARNgu comprende un motivo diana que reconoce una sustitución en la posición del nucleótido no natural dentro de una molécula de ácido nucleico de interés. En algunos casos, el ARNgu comprende un motivo diana que reconoce una delección en la posición del nucleótido no natural dentro de una molécula de ácido nucleico de interés. En algunos casos, el ARNgu comprende un motivo diana que reconoce una inserción en la posición del nucleótido no natural dentro de una molécula de ácido nucleico de interés.

En algunos casos, el motivo diana tiene una longitud de entre 10 y 30 nucleótidos. En algunos casos, el motivo diana tiene una longitud de entre 15 y 30 nucleótidos. En algunos casos, el motivo diana tiene aproximadamente 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleótidos de longitud. En algunos casos, el motivo diana tiene una longitud de aproximadamente 15, 16, 17, 18, 19, 20, 21 o 22 nucleótidos.

En algunos casos, el ARNgu comprende además un elemento de reconocimiento de motivo adyacente de protoespaciador (PAM). En algunos casos, el PAM se ubica junto al extremo 3' del motivo diana. En algunos casos, un nucleótido dentro del motivo diana que forma el emparejamiento de bases de Watson-Crick con la modificación en la posición no natural del nucleótido dentro de la molécula de ácido nucleico de interés se ubica entre 3 y 22, entre 5 y 20, entre 5 y 18, entre 5 y 15, entre 5 y 12, o entre 5 y 10 nucleótidos desde el extremo 5' de PAM. En algunos casos, un nucleótido dentro del motivo diana que forma el emparejamiento de bases de Watson-Crick con la modificación en la posición no natural del nucleótido dentro de la molécula de ácido nucleico de interés se ubica aproximadamente a 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 nucleótidos desde el extremo 5' de PAM.

En algunos casos, la segunda molécula de ácido nucleico, la tercera molécula de ácido nucleico y la cuarta molécula de ácido nucleico están codificadas en uno o más plásmidos. En algunos casos, el ARNgu codificado por la tercera molécula de ácido nucleico comprende un motivo diana que reconoce una modificación en la posición del nucleótido no natural dentro de la cuarta molécula de ácido nucleico. En algunos casos, la modificación en la posición del nucleótido no natural dentro de la tercera molécula de ácido nucleico genera una tercera molécula de ácido nucleico modificada. En algunos casos, la modificación es una sustitución, una delección o una inserción. En algunos casos, el ARNgu codificado por la tercera molécula de ácido nucleico comprende además un elemento de reconocimiento de motivo adyacente de protoespaciador (PAM, por sus siglas en inglés). En algunos casos, el elemento PAM es adyacente al extremo 3' del motivo diana. En algunos casos, la combinación del polipéptido Cas9 o variantes del mismo y ARNgu modula la replicación de la cuarta molécula de ácido nucleico modificada. En algunos casos, la combinación del polipéptido Cas9 o variantes del mismo, ARNgu y el transportador de nucleósido trifosfato modificado modula la replicación de la cuarta molécula de ácido nucleico modificada.

En algunos casos, la célula genomanipulada comprende además una molécula de ácido nucleico adicional que codifica un ARN guía único (ARNgu) adicional que comprende un armazón de ARNcr-ARNtracr.

En algunos casos, la combinación del polipéptido Cas9 o variantes del mismo, ARNgu y el transportador de nucleósido trifosfato modificado disminuye la velocidad de replicación de la cuarta molécula de ácido nucleico modificada en aproximadamente un 80 %, 85 %, 95 %, 99 % o más. En algunos casos, la producción de la cuarta molécula de ácido nucleico en la célula modificada aumenta en aproximadamente un 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o más.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado se expresa en una cepa huésped modificada (por ejemplo, la célula genomanipulada), optimizada para la expresión y actividad del transportador de nucleósido trifosfato modificado y/o la captación general de nucleósidos trifosfato, naturales o no naturales. Por ejemplo, la expresión de porinas de membrana externa, incluyendo, pero sin limitación, OmpA, OmpF, OmpC, puede modificarse, o el transportador de nucleósido trifosfato modificado puede expresarse en la célula huésped (por ejemplo, la célula genomanipulada) que también expresa una porina de membrana externa heteróloga. Como alternativa, la célula huésped (por ejemplo, la célula genomanipulada) puede permeabilizarse (químicamente o por medios genéticos) para mejorar la captación de nucleósidos trifosfato. En algunas realizaciones, la célula huésped (por ejemplo, la célula modificada) puede contener delecciones de proteínas secretadas endógenamente no esenciales para mejorar la capacidad de la maquinaria de secreción del huésped para la expresión del transportador de nucleósido trifosfato modificado.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado reduce el tiempo de duplicación de la célula huésped (por ejemplo, la célula genomanipulada).

En algunos casos, el transportador de nucleósido trifosfato modificado permite una retención de pares de bases no naturales de aproximadamente el 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 99 % o más.

Plásmidos que codifican un transportador de nucleósido trifosfato modificado

En algunas realizaciones, en el presente documento también se describe un plásmido aislado y purificado que comprende una molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2); y una región promotora seleccionada de un plásmido pSC o locus lacZYA.

En algunos casos, el transportador de nucleósido trifosfato modificado es un transportador de nucleósido trifosfato de codones optimizados de *Phaeodactylum tricornutum*.

En algunos casos, el transportador de nucleósido trifosfato modificado comprende una delección. En algunos casos, la delección es una delección terminal o una delección interna. En algunos casos, la delección es un truncamiento del extremo N, un truncamiento del extremo C o un truncamiento de ambos extremos.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos. En algunos casos, la delección es una delección del extremo N. En algunos casos, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 5, 10, 15, 20, 22, 25, 30, 40, 44, 50, 60, 66, 70 o más residuos de aminoácidos en el extremo N. En algunos casos, el transportador de nucleósido trifosfato modificado comprende una delección de aproximadamente 66 residuos de aminoácidos en el extremo N.

En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4, 6 u 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 4. En algunos casos, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 4. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 6. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con el SEQ ID NO: 8. En algunos aspectos de la divulgación, el transportador de nucleósido trifosfato aislado y modificado comprende un 100 % de identidad de secuencia con el SEQ ID NO: 8.

En algunas realizaciones, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor. Los promotores de ejemplo incluyen, pero sin limitación, P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_{λ} , P_{tac} o P_{N25} . En algunos casos, el promotor reemplaza al operón lac. En algunos casos, el transportador de nucleósido trifosfato modificado está bajo el control de un promotor seleccionado de P_{bla} , P_{lac} , P_{lacUV5} , P_{H207} , P_{λ} , P_{tac} o P_{N25} . En algunos casos, el transportador de nucleósido trifosfato modificado está bajo el control del promotor P_{lacUV5} .

En algunos casos, el transportador de nucleósido trifosfato modificado está codificado dentro de un plásmido pSC.

En algunas realizaciones, en el presente documento también se desvela un método *in vivo* para aumentar la producción de una molécula de ácido nucleico que contiene un nucleótido no natural que comprende incubar una célula con un plásmido aislado y purificado descrito anteriormente.

Moléculas de ácido nucleico

En algunas realizaciones, un ácido nucleico (por ejemplo, también denominado en el presente documento molécula de ácido nucleico de interés) es de cualquier fuente o composición, tal como ADN, ADNc, ADNg (ADN genómico), ARN, ARNc (ARN inhibidor corto), ARNi, ARNt, ARNm o ARNr (ARN ribosómico), por ejemplo, y está en cualquier forma (por ejemplo, lineal, circular, superenrollado, monocatenario, bicatenario y similares). En algunas realizaciones, los ácidos nucleicos comprenden nucleótidos, nucleósidos o polinucleótidos. En algunos casos, los ácidos nucleicos comprenden ácidos nucleicos naturales y no naturales. En algunos casos, un ácido nucleico también comprende ácidos nucleicos no naturales, tales como análogos de ADN o ARN (por ejemplo, que contienen análogos de bases, análogos de azúcar y/o una cadena principal no nativa y similares). Se entiende que la expresión "ácido nucleico" no se refiere ni infiere una longitud específica de la cadena polinucleotídica, por lo que los polinucleótidos y oligonucleótidos también se incluyen en la definición. Los nucleótidos naturales de ejemplo incluyen, sin limitación, ATP, UTP, CTP, GTP, ADP, UDP, CDP, GDP, AMP, UMP, CMP, GMP, dATP, dTTP, dCTP, dGTP, dADP, dTDP, dCDP, dGDP, dAMP, dTMP, dCMP y dGMP. Los desoxirribonucleótidos naturales de ejemplo incluyen dATP, dTTP, dCTP, dGTP, dADP, dTDP, dCDP, dGDP, dAMP, dTMP, dCMP y dGMP. Los ribonucleótidos naturales de ejemplo incluyen ATP, UTP, CTP, GTP, ADP, UDP, CDP, GDP, AMP, UMP, CMP y GMP. Para el ARN, la base de uracilo es uridina. Un ácido nucleico a veces es un vector, plásmido, fagémido, secuencia de replicación autónoma (ARS, por sus siglas en inglés), centrómero, cromosoma artificial, cromosoma artificial de levadura (por ejemplo, YAC) u otro ácido nucleico capaz de replicarse o replicarse en una célula huésped. En algunos casos, un ácido nucleico no natural es un análogo de ácido nucleico. En casos adicionales, un ácido nucleico no natural es de una fuente extracelular. En

otros casos, un ácido nucleico no natural está disponible para el espacio intracelular de un organismo proporcionado en el presente documento, por ejemplo, un organismo modificado genéticamente.

Ácidos nucleicos no naturales

5

Un análogo de nucleótido, o nucleótido no natural, comprende un nucleótido que contiene algún tipo de modificación de los restos de base, azúcar o fosfato. En algunas realizaciones, una modificación comprende una modificación química. En algunos casos, las modificaciones se producen en el grupo 3'OH o 5'OH, en la cadena principal, en el componente de azúcar o en la base de nucleótidos. Las modificaciones, en algunos casos, incluyen opcionalmente moléculas enlazadoras de origen no natural y/o enlaces cruzados intercatenarios o intracatenarios. En un aspecto, el ácido nucleico modificado comprende la modificación de uno o más del grupo 3'OH o 5'OH, la cadena principal, el componente de azúcar o la base de nucleótidos y/o la adición de moléculas enlazadoras de origen no natural. En un aspecto, una cadena principal modificada comprende una cadena principal a la cadena principal de fosfodiéster. En un aspecto, un azúcar modificado comprende un azúcar distinto de la desoxirribosa (en el ADN modificado) o distinto de la ribosa (ARN modificado). En un aspecto, una base modificada comprende una base distinta de adenina, guanina, citosina o timina (en el ADN modificado) o una base distinta de adenina, guanina, citosina o uracilo (en el ARN modificado).

10

15

20

En algunas realizaciones, el ácido nucleico comprende al menos una base modificada. En algunos casos, el ácido nucleico comprende 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 o más bases modificadas. En algunos casos, las modificaciones del resto base incluyen modificaciones naturales y sintéticas de A, C, G y T/U, así como diferentes bases de purina o pirimidina. En algunas realizaciones, una modificación es una forma modificada de adenina, guanina citosina o timina (en el ADN modificado) o una forma modificada de adenina, guanina citosina o uracilo (ARN modificado).

25

30

35

40

45

50

55

60

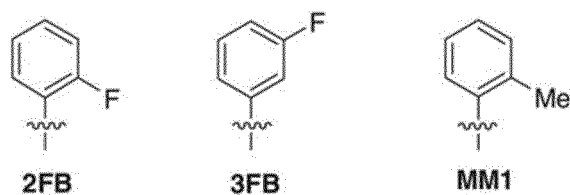
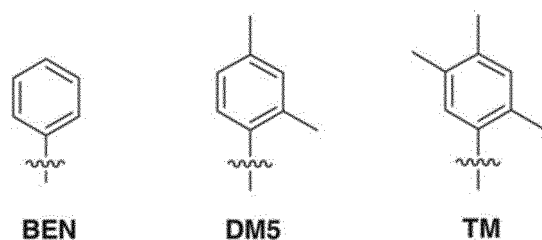
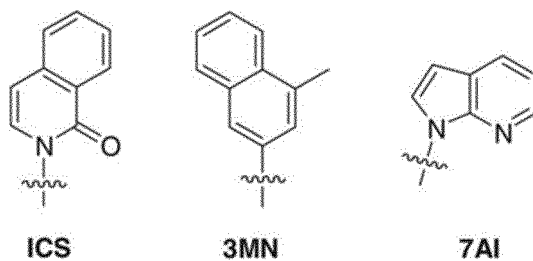
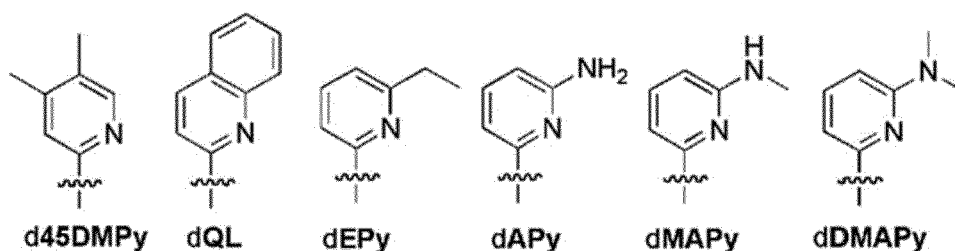
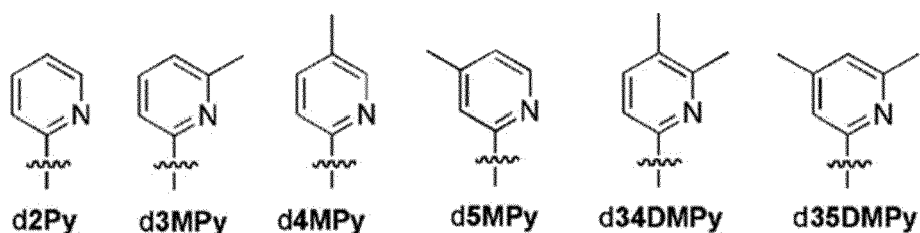
Una base modificada de un ácido nucleico no natural incluye, pero sin limitación, uracil-5-ilo, hipoxantin-9-ilo (I), 2-aminoadenin-9-ilo, 5-metilcitosina (5-me-C), 5-hidroximetil-citosina, xantina, hipoxantina, 2-aminoadenina, 6-metilo y otros derivados de alquilo de adenina y guanina, 2-propilo y otros derivados de alquilo de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo y citosina, 5-propinil uracilo y citosina, 6-azo uracilo, citosina y timina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo, 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas y guaninas 8-sustituidas, 5-halo, particularmente, 5-bromo, 5-trifluorometilo y otros uracilos y citosinas 5-sustituidos, 7-metilguanina y 7-metiladenina, 8-azaguanina y 8-azaadenina, 7-desazaguanina y 7-desazaadenina y 3-desazaguanina y 3-desazaadenina. Determinados ácidos nucleicos no naturales, tales como pirimidinas 5-sustituidas, 6-azapirimidinas y purinas sustituidas por N-2, purinas sustituidas por N-6, purinas sustituidas por O-6, 2-aminopropiladenina, 5-propiluracilo, 5-propinilcitosina, 5-metilcitosina, aquellas que aumentan la estabilidad de la formación de dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos de tamaño expandido, ácidos nucleicos fluorados, pirimidinas 5-sustituidas, 6-azapirimidinas y purinas sustituidas por N-2, N-6 y O-6, incluyendo 2-aminopropiladenina, 5-propiniluracilo y 5-propinilcitosina. 5-metilcitosina (5-me-C), 5-hidroximetil-citosina, xantina, hipoxantina, 2-aminoadenina, 6-metilo, otros derivados de alquilo de adenina y guanina, 2-propilo y otros derivados de alquilo de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo, 5-halocitosina, 5-propinilo ($-C\equiv C-Cl^{1/4}$) uracilo, 5-propinil citosina, otros derivados de alquilo de ácidos nucleico de pirimidina, 6-azo uracilo, 6-azocitosina, 6-azotimina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo, 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas y guaninas 8-sustituidas, 5-halo particularmente 5-bromo, 5-trifluorometilo, otros uracilos y citosinas 5-sustituidas, 7-metilguanina, 7-metiladenina, 2-F-adenina, 2-amino-adenina, 8-azaguanina, 8-azaadenina, 7-desazaguanina, 7-desazaadenina, 3-desazaguanina, 3-desazaadenina, pirimidinas tricíclicas, fenoxazina citidina([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazin-2(3H)-ona), pinzas G, fenoxazina citidina (por ejemplo, 9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido[3',2':4,5]pirrolo[2,3-d]pirimidin-2-ona), aquellos en los que la base de purina o pirimidina se reemplaza por otros heterociclos, 7-desaza-adenina, 7-desazaguanosina, 2-aminopiridina, 2-piridona, , azacitosina, 5-bromocitosina, bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxiiurea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo y 5-yodouracilo, 2-amino-adenina, 6-tioguanina, 2-tio-timina, 4-tiotimina, 5-propinil-uracilo, 4-tio-uracilo, N4-etilcitosina, 7-desazaguanina, 7-desaza-8-azaguanina, 5-hidroxicitosina, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina, y los descritos en las Patentes de EE.UU. N.º 3.687.808; 4.845.205; 4.910.300; 4.948.882; 5.093.232; 5.130.302; 5.134.066; 5.175.273; 5.367.066; 5.432.272; 5.457.187; 5.459.255; 5.484.908; 5.502.177; 5.525.711; 5.552.540; 5.587.469; 5.594.121; 5.596.091; 5.614.617; 5.645.985; 5.681.941; 5.750.692; 5.763.588; 5.830.653 y 6.005.096; WO 99/62923; Kandimalla *et al.*, (2001) Bioorg. Med. Chem. 9:807-813; The Concise Encyclopedia of Polymer Science and Engineering, Kroschwitz, J.I., Ed., John Wiley & Sons, 1990, 858-859; Englisch *et al.*, Angewandte Chemie, Edición Internacional, 1991, 30, 613; y Sanghvi, Capítulo 15, Antisense Research and Applications, Crooke and Lebleu Eds., CRC Press, 1993, 273-288. Pueden encontrarse modificaciones de bases adicionales, por ejemplo, en la Patente de EE.UU. N.º 3.687.808; Englisch *et al.*, Angewandte Chemie, Edición Internacional, 1991, 30, 613; y Sanghvi, Capítulo 15, Antisense Research and Applications, páginas 289-302, Crooke y Lebleu ed., CRC Press, 1993.

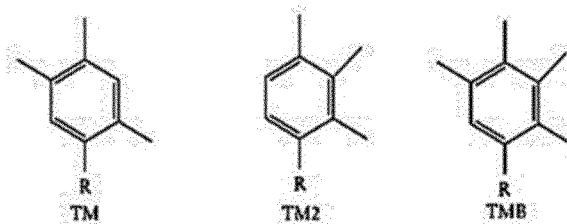
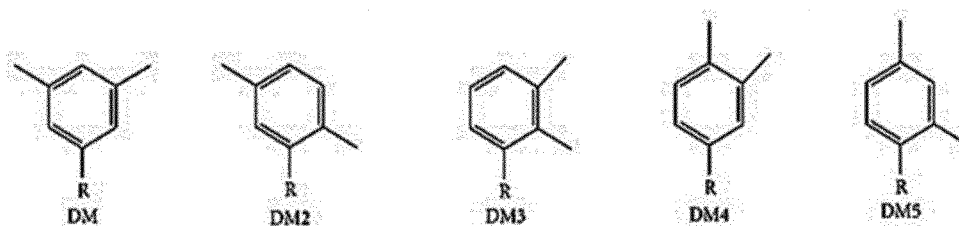
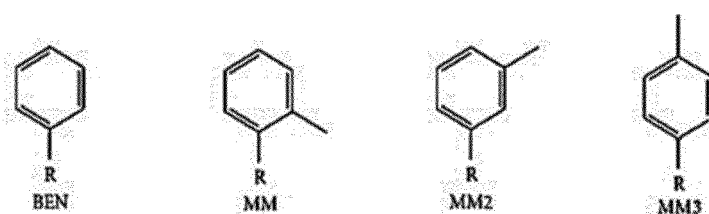
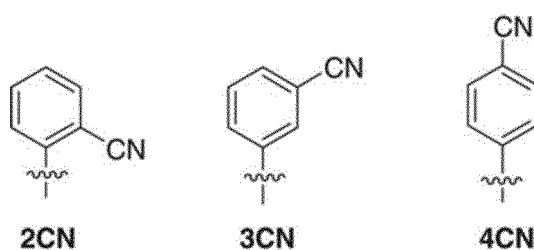
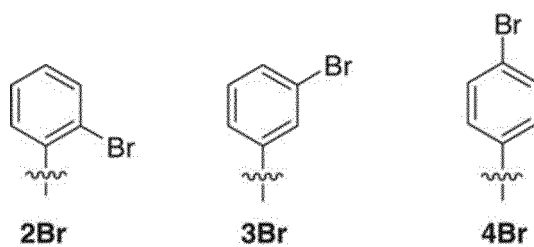
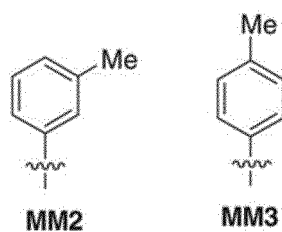
65

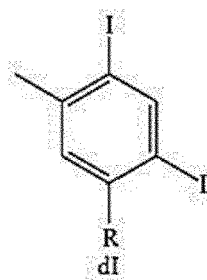
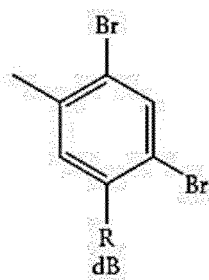
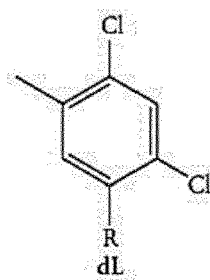
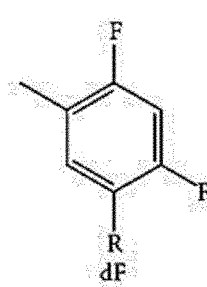
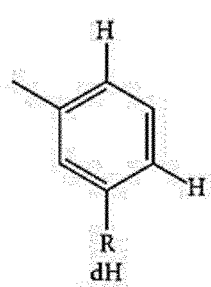
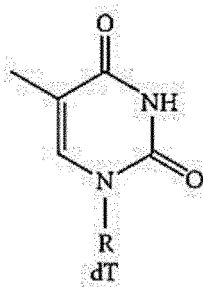
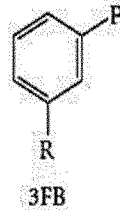
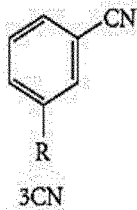
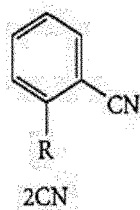
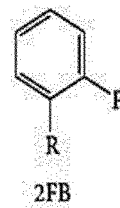
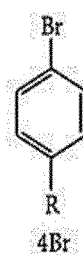
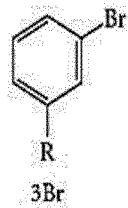
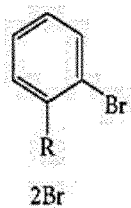
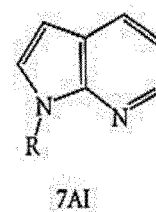
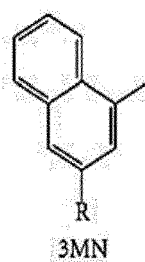
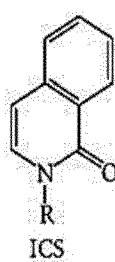
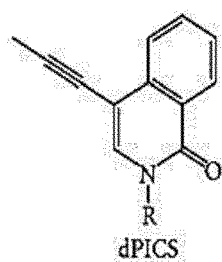
Los ácidos nucleicos no naturales que comprenden diversas bases heterocíclicas y diversos restos de azúcar (y análogos de azúcar) están disponibles en la materia y, en algunos casos, el ácido nucleico incluye una o varias bases heterocíclicas distintas de los cinco componentes de bases principales de los ácidos nucleicos de origen natural. Por

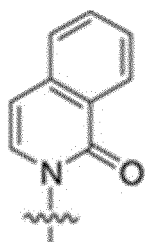
ejemplo, la base heterocíclica incluye, en algunos casos, uracil-5-ilo, citosin-5-ilo, adenin-7-ilo, adenin-8-ilo, guanin-7-ilo, guanin-8-ilo, 4-aminopirrolo [2.3-d] pirimidin-5-ilo, 2-amino-4-oxopirrolo [2, 3-d] pirimidin-5-ilo, 2-amino-4-oxopirrolo [2.3-d] pirimidin-3-ilo, en donde las purinas se unen al resto de azúcar del ácido nucleico a través de la posición 9, las pirimidinas a través de la posición 1, las pirrolopirimidinas a través de la posición 7 y las pirazolopirimidinas a través de la posición 1.

En algunas realizaciones, una base modificada de un ácido nucleico no natural se representa a continuación, en donde la línea ondulada identifica un punto de unión a la (desoxi)ribosa o ribosa.

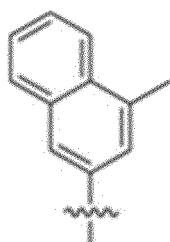




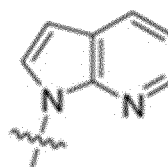




ICS



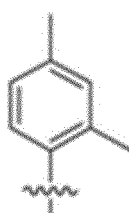
3MN



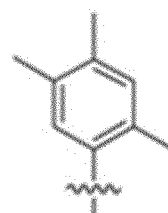
7AI



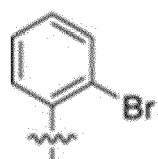
BEN



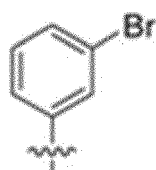
DM5



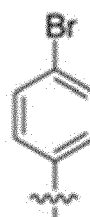
TM



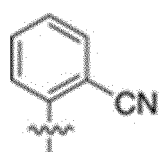
2Br



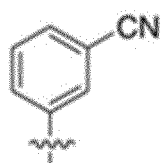
3Br



4Br



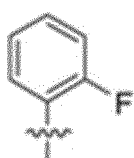
2CN



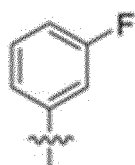
3CN



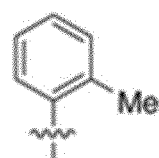
4CN



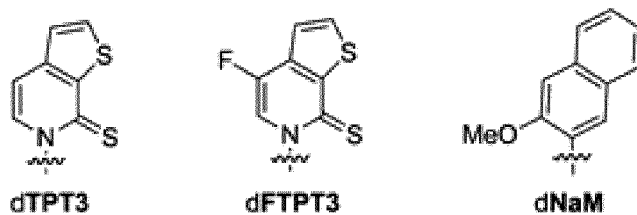
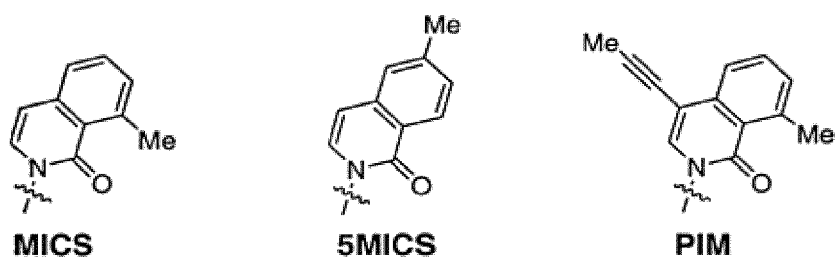
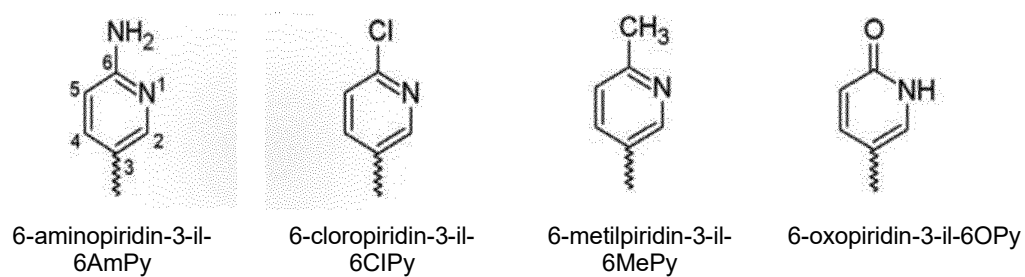
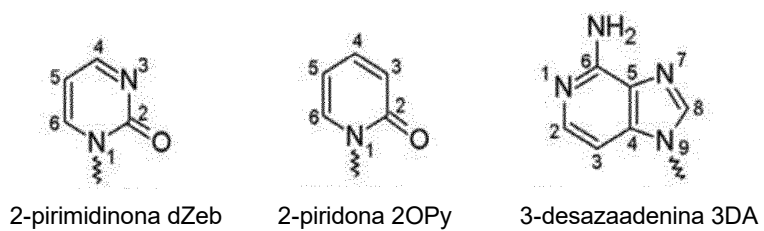
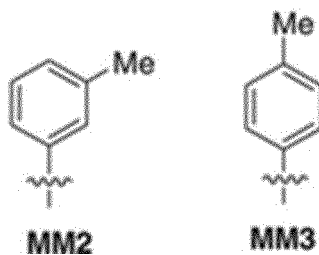
2FB

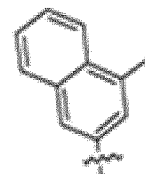
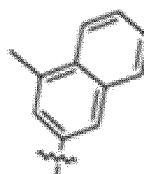
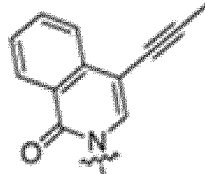
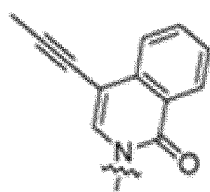
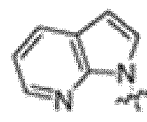
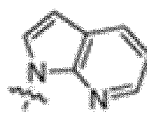
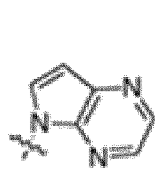
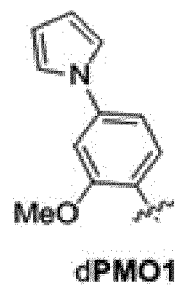
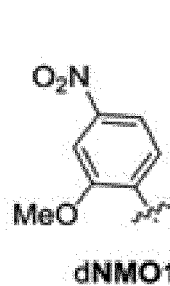
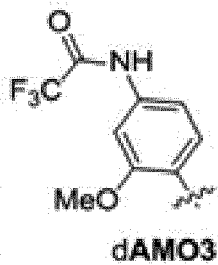
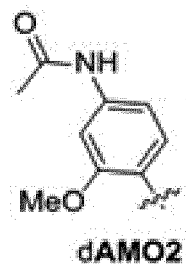
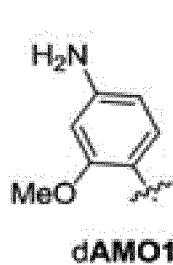
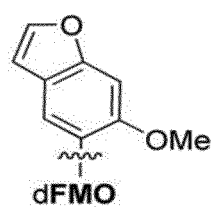
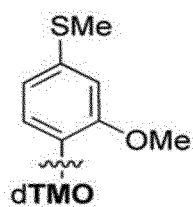
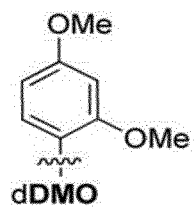
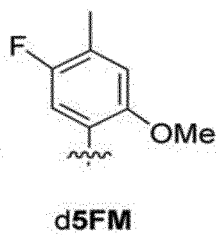
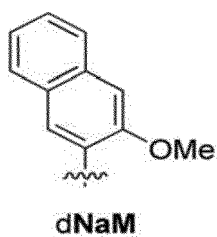
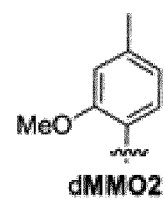
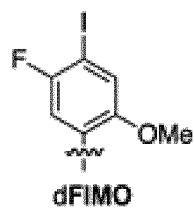
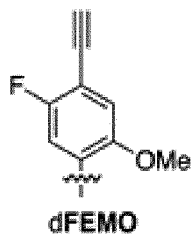
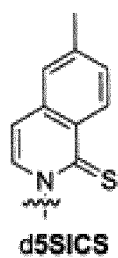


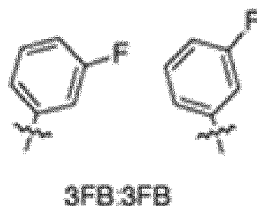
3FB



MM1







En algunas realizaciones, los análogos de nucleótidos también se modifican en el resto de fosfato. Los restos de fosfato modificados incluyen, pero sin limitación, aquellos con modificación en el enlace entre dos nucleótidos y contiene, por ejemplo, un fosforotioato, fosforotioato quiral, fosforoditioato, fosfotriéster, aminoalquilfosfotriéster, metilo y otros alquil fosfonatos incluyendo fosfonatos de 3'-alquileo y fosfonatos quirales, fosfinatos, fosforamidatos que incluyen 3'-amino fosforamidato y aminoalquilfosforamidatos, tionofosforamidatos, tionoalquilfosfonatos, tionoalquilfosfotriésteres y boranofosfonatos. Se entiende que estos enlaces fosfato o fosfato modificado entre dos nucleótidos son a través de un enlace 3'-5' o un enlace 2'-5', y el enlace contiene polaridad invertida tal como 3'-5' a 5'-3' o 2'-5' a 5'-2'. También se incluyen diversas sales, sales mixtas y formas de ácidos libres. Numerosas patentes de Estados Unidos enseñan cómo hacer y usar los nucleótidos que contienen fosfatos modificados e incluyen, pero sin limitación, 3.687.808; 4.469.863; 4.476.301; 5.023.243; 5.177.196; 5.188.897; 5.264.423; 5.276.019; 5.278.302; 5.286.717; 5.321.131; 5.399.676; 5.405.939; 5.453.496; 5.455.233; 5.466.677; 5.476.925; 5.519.126; 5.536.821; 5.541.306; 5.550.111; 5.563.253; 5.571.799; 5.587.361; y 5.625.050.

En algunas realizaciones, los ácidos nucleicos no naturales incluyen 2',3'-didesoxi-2',3'-dideshidro-nucleósidos (documento PCT/US2002/006460), derivados de ADN y ARN 5'-sustituidos (documento PCT/US2011/033961; Saha *et al.*, J. Org Chem., 1995, 60, 788-789; Wang *et al.*, Bioorganic & Medicinal Chemistry Letters, 1999, 9, 885-890; y Mikhailov *et al.*, Nucleosides & Nucleotides, 1991, 10(1-3), 339-343; Leonid *et al.*, 1995, 14(3-5), 901-905; y Eppacher *et al.*, Helvetica Chimica Acta, 2004, 87, 3004-3020; PCT/JP2000/004720; PCT/JP2003/002342; PCT/JP2004/013216; PCT/JP2005/020435; PCT/JP2006/315479; PCT/JP2006/324484; PCT/JP2009/056718; PCT/JP2010/067560), o monómeros 5'-sustituidos preparados como el monofosfato con bases modificadas (Wang *et al.*, Nucleosides Nucleotides & Nucleic Acids, 2004, 23 (1 y 2), 317-337).

En algunas realizaciones, los ácidos nucleicos no naturales incluyen modificaciones en la posición 5' y la posición 2' del anillo de azúcar (documento PCT/US94/02993), tales como nucleósidos 2'-O-protegidos 5'-CH₂-sustituidos (Wu *et al.*, Helvetica Chimica Acta, 2000, 83, 1127-1143 y Wu *et al.*, Bioconjugate Chem. 1999, 10, 921-924). En algunos casos, los ácidos nucleicos no naturales incluyen dímeros de nucleósidos unidos a amida que se han preparado para su incorporación en oligonucleótidos en los que el nucleósido unido en 3' en el dímero (5' a 3') comprende un 2'-OCH₃ y un 5'-(S)-CH₃ (Mesmaeker *et al.*, Synlett, 1997, 1287-1290). Los ácidos nucleicos no naturales pueden incluir nucleósidos modificados con 5'-CH₂ (u O) 2'-sustituidos (documento PCT/US92/01020). Los ácidos nucleicos no naturales pueden incluir monómeros y dímeros de ADN y ARN de 5'-metilfosfonato (Bohringer *et al.*, Tet. Lett., 1993, 34, 2723-2726; Collingwood *et al.*, Synlett, 1995, 7, 703-705; y Hutter *et al.*, Helvetica Chimica Acta, 2002, 85, 2777-2806). Los ácidos nucleicos no naturales pueden incluir monómeros de 5'-fosfonato que tienen una sustitución en 2' (documento US2006/0074035) y otros monómeros de 5'-fosfato modificados (documento WO1997/35869). Los ácidos nucleicos no naturales pueden incluir monómeros de metilfosfonato 5'-modificados (documentos EP614907 y EP629633). Los ácidos nucleicos no naturales pueden incluir análogos de ribonucleósidos 5' o 6'-fosfonato que comprenden un grupo hidroxilo en la posición 5' y/o 6' (Chen *et al.*, Phosphorus, Sulfur and Silicon, 2002, 777, 1783-1786; Jung *et al.*, Bioorg. Med. Chem., 2000, 8, 2501-2509; Gallier *et al.*, Eur. J. Org. Chem., 2007, 925-933; y Hampton *et al.*, J. Med. Chem., 1976, 19(8), 1029-1033). Los ácidos nucleicos no naturales pueden incluir monómeros y dímeros de desoxirribonucleósido 5'-fosfonato que tienen un grupo 5'-fosfato (Nawrot *et al.*, Oligonucleotides, 2006, 16(1), 68-82). Los ácidos nucleicos no naturales pueden incluir nucleósidos que tienen un grupo 6'-fosfonato en donde la posición 5' y/o 6' está sin sustituir o sustituida por un grupo tio-*tert*-butilo (SC(CH₃)₃) (y análogos del mismo); un grupo metilnamino (CH₂NH₂) (y análogos del mismo) o un grupo ciano (CN) (y análogos del mismo) (Fairhurst *et al.*, Synlett, 2001, 4, 467-472; Kappler *et al.*, J. Med. Chem., 1986, 29, 1030-1038; Kappler *et al.*, J. Med. Chem., 1982, 25, 1179-1184; Vruthula *et al.*, J. Med. Chem., 1987, 30, 888-894; Hampton *et al.*, J. Med. Chem., 1976, 19, 1371-1377; Geze *et al.*, J. Am. Chem. Soc., 1983, 105(26), 7638-7640; y Hampton *et al.*, J. Am. Chem. Soc., 1973, 95(13), 4404-4414).

En algunas realizaciones, los ácidos nucleicos no naturales también incluyen modificaciones del resto de azúcar. En algunos casos, los ácidos nucleicos contienen uno o más nucleósidos en donde se ha modificado el grupo de azúcar. Dichos nucleósidos modificados con azúcar pueden impartir una estabilidad de nucleasa mejorada, una afinidad de unión aumentada o alguna otra propiedad biológica beneficiosa. En determinadas realizaciones, los ácidos nucleicos comprenden un resto de anillo de ribofuranosa químicamente modificado. Los ejemplos de anillos de ribofuranosa modificados químicamente incluyen, sin limitación, la adición de grupos sustituyentes (incluyendo los grupos sustituyentes 5' y/o 2'; puente de dos átomos del anillo para formar ácidos nucleicos bicíclicos (BNA); reemplazo del átomo de oxígeno del anillo ribosilo por S, N(R) o C(Ri)(R₂) (R = H, alquilo C₁-C₁₂ o un grupo protector); y combinaciones de los mismos. Se pueden encontrar ejemplos de azúcares modificados químicamente en los

documentos WO2008/101157, US2005/0130923 y WO2007/134181.

En algunos casos, un ácido nucleico modificado comprende azúcares modificados o análogos de azúcares. Por lo tanto, además de ribosa y desoxirribosa, el resto de azúcar puede ser pentosa, desoxipentosa, hexosa, desoxihexosa, glucosa, arabinosa, xilosa, lixosa o un grupo ciclopentilo "análogo" de azúcar. El azúcar puede estar en forma de piranosilo o furanosilo. El resto de azúcar puede ser el furanósido de ribosa, desoxirribosa, arabinosa o 2'-O-alquilribosa, y el azúcar puede estar unido a las respectivas bases heterocíclicas en configuración anomérica [alfa] o [beta]. Las modificaciones de azúcar incluyen, pero sin limitación, análogos de 2'-alcoxi-ARN, análogos de 2'-amino-ARN, 2'-fluoro-ADN y quimeras 2'-alcoxi o amino-ARN/ADN. Por ejemplo, una modificación de azúcar puede incluir 2'-O-metil-uridina o 2'-O-metil-citidina. Las modificaciones de azúcar incluyen desoxirribonucleósidos sustituidos con 2'-O-alquilo y ribonucleósidos de tipo 2'-O-etilenglicol. Se conoce la preparación de estos azúcares o análogos de azúcar y los respectivos "nucleósidos" en donde dichos azúcares o análogos están unidos a una base heterocíclica (base de ácido nucleico). También se pueden realizar modificaciones del azúcar y combinarlas con otras modificaciones.

Las modificaciones del resto de azúcar incluyen modificaciones naturales de la ribosa y la desoxirribosa, así como modificaciones no naturales. Las modificaciones de azúcar incluyen, pero sin limitación, las siguientes modificaciones en la posición 2': OH; F; O-, S- o N-alquilo; O-, S- o N-alquenilo; O-, S- o N-alquinilo; u O-alquil-O-alquilo, en donde el alquilo, alquenilo y alquinilo pueden ser alquilo C₁ a C₁₀, o alquenilo C₂ a C₁₀ y alquinilo sustituidos o sin sustituir. Las modificaciones de azúcar en 2' también incluyen, pero sin limitación, -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_nONH₂ y -O(CH₂)_nON[(CH₂)_nCH₃]₂, en donde n y m son de 1 a aproximadamente 10.

Otras modificaciones en la posición 2' incluyen, pero sin limitación: alquilo C₁ a C₁₀ inferior, alquilo inferior sustituido, alcarilo, aralquilo, O-alcarilo, O-aralquilo, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, heterocicloalquilo, heterocicloalcarilo, aminoalquilamino, polialquilamino, sililo sustituido, un grupo de escisión de ARN, un grupo indicador, un intercalante, un grupo para mejorar las propiedades farmacocinéticas de un oligonucleótido, o un grupo para mejorar las propiedades farmacocinéticas de un oligonucleótido y otros sustituyentes que tienen propiedades similares. También pueden producirse modificaciones similares en otras posiciones sobre el azúcar, particularmente, la posición 3' del azúcar en el nucleótido del extremo 3' o en oligonucleótidos enlazados en 2'-5' y la posición 5' del nucleótido del extremo 5'. Los azúcares modificados también incluyen aquellos que contienen modificaciones en el oxígeno del anillo puenteado, tal como CH₂ y S. Los análogos de azúcar del nucleótido pueden tener también miméticos de azúcar tales como restos de ciclobutilo en lugar de azúcar pentofuranosilo. Existen numerosas patentes de Estados Unidos que enseñan la preparación de tales estructuras de azúcar modificadas y que detallan y describen varias modificaciones de bases, tales como las Patentes de EE.UU. N.º 4.981.957; 5.118.800; 5.319.080; 5.359.044; 5.393.878; 5.446.137; 5.466.786; 5.514.785; 5.519.134; 5.567.811; 5.576.427; 5.591.722; 5.597.909; 5.610.300; 5.627.053; 5.639.873; 5.646.265; 5.658.873; 5.670.633; 4.845.205; 5.130.302; 5.134.066; 5.175.273; 5.367.066; 5.432.272; 5.457.187; 5.459.255; 5.484.908; 5.502.177; 5.525.711; 5.552.540; 5.587.469; 5.594.121; 5.596.091; 5.614.617; 5.681.941; y 5.700.920.

Los ejemplos de ácidos nucleicos que tienen restos de azúcar modificados incluyen, sin limitación, ácidos nucleicos que comprenden grupos de sustituyentes 5'-vinilo, 5'-metilo (R o S), 4'-S, 2'-F, 2'-OCH₃ y 2'-O(CH₂)₂OCH₃. El sustituyente en la posición 2' también se puede seleccionar de alilo, amino, azido, tio, O-alilo, O-(alquilo C₁-C₁₀), OCF₃, O(CH₂)₂SCH₃, O(CH₂)₂-ON(R_m)(R_n) y O-CH₂-C(=O)-N(R_m)(R_n), en donde cada R_m y R_n es, independientemente, H o alquilo C₁-C₁₀ sustituido o sin sustituir.

En determinadas realizaciones, los ácidos nucleicos descritos en el presente documento incluyen uno o más ácidos nucleicos bicíclicos. En determinadas realizaciones de este tipo, el ácido nucleico bicíclico comprende un puente entre los átomos del anillo de ribosilo 4' y 2'. En determinadas realizaciones, los ácidos nucleicos proporcionados en el presente documento incluyen uno o más ácidos nucleicos bicíclicos, en donde el puente comprende un ácido nucleico bicíclico 4' a 2'. Los ejemplos de dichos ácidos nucleicos bicíclicos 4' a 2' incluyen, pero sin limitación, una de las fórmulas: 4'-(CH₂)₂-O-2' (LNA); 4'-(CH₂)₂-S-2' (ENA); 4'-(CH₂)₂-O-2' (ENA); 4'-CH(CH₃)-O-2' y 4'-CH(CH₂OCH₃)-O-2' y análogos de los mismos (véase, la Patente de EE.UU. N.º 7.399.845); 4'-C(CH₃)(CH₃)-O-2' y análogos de los mismos, (véase los documentos WO2009/006478, WO2008/150729, US2004/0171570, la Patente de EE.UU. N.º 7.427.672, Chattopadhyaya *et al.*, J. Org. Chem., 209, 74, 118-134 y el documento WO2008/154401). Véanse también, por ejemplo: Singh *et al.*, Chem. Commun., 1998, 4, 455-456; Koshkin *et al.*, Tetrahedron, 1998, 54, 3607-3630; Wahlestedt *et al.*, Proc. Natl. Acad. Sci. U.S. A., 2000, 97, 5633-5638; Kumar *et al.*, Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222; Singh *et al.*, J. Org. Chem., 1998, 63, 10035-10039; Srivastava *et al.*, J. Am. Chem. Soc., 2007, 129(26) 8362-8379; Elayadi *et al.*, Curr. Opin. Invers. Drugs, 2001, 2, 558-561; Braasch *et al.*, Chem. Biol., 2001, 8, 1-7; Oram *et al.*, Curr. Opin. Mol. Ther., 2001, 3, 239-243; las Patentes de EE.UU. N.º 4.849.513; 5.015.733; 5.118.800; 5.118.802; 7.053.207; 6.268.490; 6.770.748; 6.794.499; 7.034.133; 6.525.191; 6.670.461; y 7.399.845; las Publicaciones Internacionales N.º WO2004/106356, WO1994/14226, WO2005/021570, WO2007/090071 y WO2007/134181; las Publicaciones de Patente de EE.UU. N.º US2004/0171570, US2007/0287831 y US2008/0039618; las Solicitudes Provisionales de EE.UU. N.º 60/989.574, 61/026.995, 61/026.998, 61/056.564, 61/086.231, 61/097.787 y 61/099.844; y las Solicitudes Internacionales N.º PCT/US2008/064591, PCT US2008/066154, PCT US2008/068922 y PCT/DK98/00393.

En determinadas realizaciones, los ácidos nucleicos comprenden ácidos nucleicos unidos. Los ácidos nucleicos

pueden unirse entre sí usando cualquier enlace entre ácidos nucleicos. Las dos clases principales de grupos de enlace entre ácidos nucleicos se definen por la presencia o ausencia de un átomo de fósforo. Los enlaces entre ácidos nucleicos que contienen fósforo representativos incluyen, pero sin limitación, fosfodiésteres, fosfotriésteres, fosfonatos de metilo, fosforamido y fosforotioatos (P=S). Los grupos de enlace entre ácidos nucleicos que no contienen fósforo representativos incluyen, pero sin limitación, metilenmetilimino (-CH₂-N(CH₃)-O-CH₂-), tiodiéster (-O-C(O)-S-), tionocarbamato (-O-C(O)(NH)-S-); siloxano (-O-Si(H)₂-O-); y N,N*-dimetilhidrazina (-CH₂-N(CH₃)-N(CH₃)). En determinadas realizaciones, los enlaces entre ácidos nucleicos que tienen un átomo quiral se pueden preparar como una mezcla racémica, como enantiómeros separados, por ejemplo, alquilfosfonatos y fosforotioatos. Los ácidos nucleicos no naturales pueden contener una única modificación. Los ácidos nucleicos no naturales pueden contener múltiples modificaciones dentro de uno de los restos o entre diferentes restos.

Las modificaciones de fosfato de la cadena principal con respecto al ácido nucleico incluyen, pero sin limitación, fosfonato de metilo, fosforotioato, fosforamido (puenteado o no puenteado), fosfotriéster, fosforoditioato, fosfoditioato y boranofosfato, y pueden usarse en cualquier combinación. También pueden usarse otros enlaces no fosfato.

En algunas realizaciones, las modificaciones de la cadena principal (por ejemplo, enlaces internucleotídicos de fosfonato de metilo, fosforotioato, fosforoamido y fosforoditioato) pueden conferir actividad inmunomoduladora sobre el ácido nucleico modificado y/o mejorar su estabilidad *in vivo*.

En algunos casos, se une un derivado de fósforo (o grupo de fosfato modificado) al azúcar o resto análogo de azúcar y puede ser un monofosfato, difosfato, trifosfato, alquilfosfonato, fosforotioato, fosforoditioato, fosforamido o similares. Se pueden encontrar polinucleótidos de ejemplo que contienen enlaces fosfato modificados o enlaces no fosfato en Peyrottes *et al.*, 1996, *Nucleic Acids Res.* 24: 1841-1848; Chaturvedi *et al.*, 1996, *Nucleic Acids Res.* 24:2318-2323; y Schultz *et al.*, (1996) *Nucleic Acids Res.* 24:2966-2973; Matteucci, 1997, "Oligonucleotide Analogs: an Overview" en *Oligonucleotides as Therapeutic Agents*, (Chadwick y Cardew, ed.) John Wiley and Sons, Nueva York, NY; Zon, 1993, "Oligonucleoside Phosphorothioates" en *Protocols for Oligonucleotides and Analogs*, Synthesis and Properties, Humana Press, págs. 165-190; Miller *et al.*, 1971, *JACS* 93:6657-6665; Jager *et al.*, 1988, *Biochem.* 27:7247-7246; Nelson *et al.*, 1997, *JOC* 62:7278-7287; la Patente de EE.UU. N.º 5.453.496; y Micklefield, 2001, *Curr. Med. Chem.* 8: 1157-1179.

En algunos casos, la modificación de la cadena principal comprende reemplazar el enlace fosfodiéster con un resto alternativo tal como un grupo aniónico, neutro o catiónico. Los ejemplos de dichas modificaciones incluyen: enlace internucleosídico aniónico; modificación de fosforamido de N3' a P5'; ADN de boranofosfato; prooligonucleótidos; enlaces internucleosídicos neutros tales como fosfonatos de metilo; ADN unido a amida; enlaces metilen(metilimino); enlaces formacetal y tioformacetal; cadenas principales que contienen grupos sulfonilo; oligos morfolino; ácidos nucleicos peptídicos (PNA, por sus siglas en inglés); y oligos de guanidina desoxirribonucleica cargados positivamente (DNG, por sus siglas en inglés) (Micklefield, 2001, *Current Medicinal Chemistry* 8: 1157-1179). Un ácido nucleico modificado puede comprender una cadena principal química o mixta que comprende una o más modificaciones, por ejemplo, una combinación de enlaces fosfato tal como una combinación de enlaces fosfodiéster y fosforotioato.

Los sustitutos del fosfato incluyen, por ejemplo, enlaces internucleosídicos de alquilo o cicloalquilo de cadena corta, enlaces internucleosídicos de heteroátomos mixtos y de alquilo o cicloalquilo o uno o más enlaces internucleosídicos heteroatómicos o heterocíclicos de cadena corta. Estos incluyen aquellos que tienen enlaces morfolino (formados en parte a partir de la parte de azúcar de un nucleósido); cadenas principales de siloxano; sulfuro, cadenas principales de sulfóxido y sulfona; cadenas principales de formaceto y tioformaceto; cadenas principales de metilen formaceto y tioformaceto; cadenas principales que contienen alqueno; cadenas principales de sulfamato; cadenas principales de metilenimino y metilendiazina; cadenas principales de sulfonato y sulfonamida; cadenas principales de amida; y otros que tienen partes componentes mezcladas de N, O, S y CH₂. Numerosas patentes de Estados Unidos desvelan cómo fabricar y usar estos tipos de reemplazos de fosfato e incluyen, pero sin limitación, Patentes de EE.UU. N.º 5.034.506; 5.166.315; 5.185.444; 5.214.134; 5.216.141; 5.235.033; 5.264.562; 5.264.564; 5.405.938; 5.434.257; 5.466.677; 5.470.967; 5.489.677; 5.541.307; 5.561.225; 5.596.086; 5.602.240; 5.610.289; 5.602.240; 5.608.046; 5.610.289; 5.618.704; 5.623.070; 5.663.312; 5.633.360; 5.677.437; y 5.677.439. Se entiende también que en un sustituto de nucleótido se pueden reemplazar los restos de azúcar y de fosfato del nucleótido, mediante, por ejemplo, un enlace de tipo amida (aminoetilglicina) (ANP). Las patentes de EE.UU. N.º 5.539.082; 5.714.331; y 5.719.262 enseñan cómo hacer y usar las moléculas de ANP. Véase también Nielsen *et al.*, *Science*, 1991, 254, 1497-1500. También existe posibilidad de unir otros tipos de moléculas (conjugados) a nucleótidos o análogos de nucleótidos para potenciar, por ejemplo, la captación celular. Los conjugados se pueden enlazar químicamente con el nucleótido o análogos de nucleótidos. Dichos conjugados incluyen, pero sin limitación, restos de lípidos tales como un resto de colesterol (Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA*, 1989, 86, 6553-6556), ácido cólico (Manoharan *et al.*, *Bioorg. Med. Chem. Lett.*, 1994, 4, 1053-1060), un tioéter, por ejemplo, hexil-S-tritilol (Manoharan *et al.*, *Ann. KY. Acad. Sci.*, 1992, 660, 306-309; Manoharan *et al.*, *Bioorg. Med. Chem. Lett.*, 1993, 3, 2765-2770), un tiocolesterol (Oberhauser *et al.*, *Nucl. Acids Res.*, 1992, 20, 533-538), una cadena alifática, *por ejemplo*, residuos de dodecandiol o undecilo (Saison-Behmoaras *et al.*, *EMSOJ*, 1991, 10, 1111-1118; Kabanov *et al.*, *FEBS Lett.*, 1990, 259, 327-330; Svinarchuk *et al.*, *Biochimie*, 1993, 75, 49-54), un fosfolípido, por ejemplo, di-hexadecil-rac-glicerol o 1-di-O-hexadecil-rac-glicerol-S-H-fosfonato de trietilamonio (Manoharan *et al.*, *Tetrahedron Lett.*, 1995, 36, 3651-3654; Shea *et al.*, *Nucl. Acids Res.*, 1990, 18, 3777-3783), una poliamina o una cadena de polietilenglicol (Manoharan *et al.*, *Nucleosides & Nucleotides*,

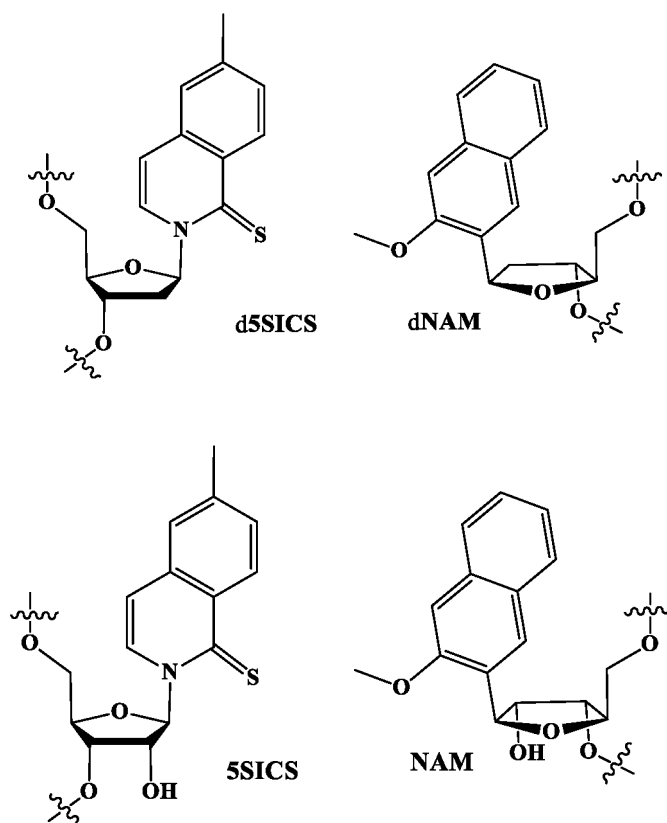
1995, 14, 969-973), o ácido adamantano acético (Manoharan *et al.*, Tetrahedron Lett., 1995, 36, 3651-3654), un resto de palmitilo (Mishra *et al.*, Biochem. Biophys. Acta, 1995, 1264, 229-237) o un resto de octadecilamina o hexilamino-carbonil-oxicolesterol (Crooke *et al.*, J. Pharmacol. Exp. Ther., 1996, 277, 923-937). Numerosas patentes de Estados Unidos enseñan la preparación de dichos conjugados e incluyen, pero sin limitación, las Patentes de EE.UU. N.º 4.828.979; 4.948.882; 5.218.105; 5.525.465; 5.541.313; 5.545.730; 5.552.538; 5.578.717; 5.580.731; 5.580.731; 5.591.584; 5.109.124; 5.118.802; 5.138.045; 5.414.077; 5.486.603; 5.512.439; 5.578.718; 5.608.046; 4.587.044; 4.605.735; 4.667.025; 4.762.779; 4.789.737; 4.824.941; 4.835.263; 4.876.335; 4.904.582; 4.958.013; 5.082.830; 5.112.963; 5.214.136; 5.082.830; 5.112.963; 5.214.136; 5.245.022; 5.254.469; 5.258.506; 5.262.536; 5.272.250; 5.292.873; 5.317.098; 5.371.241; 5.391.723; 5.416.203; 5.451.463; 5.510.475; 5.512.667; 5.514.785; 5.565.552; 5.567.810; 5.574.142; 5.585.481; 5.587.371; 5.595.726; 5.597.696; 5.599.923; 5.599.928 y 5.688.941.

Propiedades de emparejamiento de bases de ácidos nucleicos

En algunas realizaciones, un ácido nucleico no natural forma un par de bases con otro ácido nucleico. En algunas realizaciones, un ácido nucleico no natural integrado de manera estable es un ácido nucleico no natural que puede formar un par de bases con otro ácido nucleico, *por ejemplo*, un ácido nucleico natural o no natural. En algunas realizaciones, un ácido nucleico no natural integrado de manera estable es un ácido nucleico no natural que puede formar un par de bases con otro ácido nucleico no natural (par de bases de ácidos nucleicos no naturales (UBP)). Por ejemplo, un primer ácido nucleico no natural puede formar un par de bases con un segundo ácido nucleico no natural. Por ejemplo, un par de nucleótidos trifosfato no naturales que pueden emparejarse por bases cuando se incorporan a ácidos nucleicos incluyen un trifosfato de d5SICS (d5SICSTP) y un trifosfato de dNaM (dNaMTP). Dichos nucleótidos no naturales pueden tener un resto de azúcar ribosa o desoxirribosa. En algunas realizaciones, un ácido nucleico no natural no forma sustancialmente un par de bases con un ácido nucleico natural (A, T, G, C). En algunas realizaciones, un ácido nucleico no natural integrado de manera estable puede formar un par de bases con un ácido nucleico natural.

En algunas realizaciones, un ácido nucleico no natural integrado de manera estable es un ácido nucleico no natural que puede formar un UBP, pero no forma sustancialmente un par de bases con cada uno de los cuatro ácidos nucleicos naturales. En algunas realizaciones, un ácido nucleico no natural integrado de manera estable es un ácido nucleico no natural que puede formar un UBP, pero no forma sustancialmente un par de bases con uno o más ácidos nucleicos naturales. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con A, T y C, pero puede formar un par de bases con G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con A, T y G, pero puede formar un par de bases con C. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con C, G y A, pero puede formar un par de bases con T. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con C, G y T, pero puede formar un par de bases con A. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con A y T, pero puede formar un par de bases con C y G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con A y C, pero puede formar un par de bases con T y G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con A y G, pero puede formar un par de bases con C y T. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con C y T, pero puede formar un par de bases con A y G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con C y G, pero puede formar un par de bases con T y G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con T y G, pero puede formar un par de bases con A y G. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con, G, pero puede formar un par de bases con A, T y C. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con, A, pero puede formar un par de bases con G, T y C. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con, T, pero puede formar un par de bases con G, A y C. Por ejemplo, un ácido nucleico no natural integrado de manera estable puede no formar sustancialmente un par de bases con, C, pero puede formar un par de bases con G, T y A.

A modo de ejemplo, los nucleótidos no naturales capaces de formar un par de bases no naturales (UBP) de ADN o ARN en condiciones *in vivo* incluyen, pero sin limitación, 5SICS, d5SICS, NAM, dNaM y combinaciones de los mismos. En algunas realizaciones, los nucleótidos no naturales incluyen:



5 Polimerasa

Una función particularmente útil de una polimerasa es catalizar la polimerización de una hebra de ácido nucleico usando como molde un ácido nucleico existente. Se describen otras funciones útiles en otras partes del presente documento. Los ejemplos de polimerasas útiles incluyen ADN polimerasas y ARN polimerasas.

La capacidad de mejorar la especificidad, la capacidad de procesamiento u otras características de los ácidos nucleicos no naturales de las polimerasas sería muy deseable en diversos contextos en los que, por ejemplo, se desea la incorporación de ácidos nucleicos no naturales, incluyendo la amplificación, secuenciación, marcado, detección, clonación de ADN y muchas otras. La presente divulgación proporciona polimerasas con propiedades modificadas para ácidos nucleicos no naturales, métodos para preparar dichas polimerasas, métodos para usar dichas polimerasas y muchas otras características que resultarán evidentes tras una revisión completa de lo siguiente.

En algunos casos, en el presente documento la divulgación incluye polimerasas que incorporan ácidos nucleicos no naturales en una copia de molde en crecimiento, por ejemplo, durante la amplificación del ADN. En algunas realizaciones, las polimerasas se pueden modificar de modo que el sitio activo de la polimerasa se modifique para reducir la inhibición de la entrada estérica del ácido nucleico no natural en el sitio activo. En algunas realizaciones, las polimerasas se pueden modificar para proporcionar complementariedad con una o más características no naturales de los ácidos nucleicos no naturales. Dichas polimerasas se pueden expresar o genomanipular en células para incorporar de manera estable un UBP en las células. Por consiguiente, la invención incluye composiciones que incluyen una polimerasa heteróloga o recombinante y métodos de uso de la misma.

Las polimerasas se pueden modificar usando métodos pertenecientes a la ingeniería de proteínas. Por ejemplo, el modelado molecular se puede realizar basándose en estructuras cristalinas para identificar las ubicaciones de las polimerasas en donde se pueden realizar mutaciones para modificar una actividad diana. Un residuo identificado como diana para el reemplazo se puede reemplazar por un residuo seleccionado usando modelado de minimización de energía, modelado de homología y/o sustituciones conservadoras de aminoácidos, tal como se describe en Bordo, *et al.* J Mol Biol 217: 721-729 (1991) y Hayes, *et al.* Proc Natl Acad Sci, USA 99: 15926-15931 (2002).

Puede usarse cualquiera de diversas polimerasas en un método o composición expuesta en el presente documento incluyendo, por ejemplo, enzimas basadas en proteínas aisladas de sistemas biológicos y variantes funcionales de las mismas. Se entenderá que la referencia a una polimerasa particular, tal como las que se ilustran a continuación, incluye variantes funcionales de la misma a menos que se indique de otro modo. En algunas realizaciones, una polimerasa es una polimerasa de tipo silvestre. En algunas realizaciones, una polimerasa es una polimerasa modificada o mutante.

También se pueden usar polimerasas con características para mejorar la entrada de ácidos nucleicos no naturales en las regiones del sitio activo y para coordinarse con nucleótidos no naturales en la región del sitio activo. En algunas realizaciones, una polimerasa modificada tiene un sitio de unión de nucleótidos modificado.

En algunas realizaciones, una polimerasa modificada tiene una especificidad por un ácido nucleico no natural que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia el ácido nucleico no natural. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural que comprende un azúcar modificado que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia un ácido nucleico no natural y/o el ácido nucleico no natural sin el azúcar modificado. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural que comprende una base modificada que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia un ácido nucleico natural y/o el ácido nucleico no natural sin la base modificada. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural que comprende un trifosfato que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia un ácido nucleico que comprende un trifosfato y/o el ácido nucleico no natural sin el trifosfato. Por ejemplo, una polimerasa modificada o de tipo silvestre puede tener una especificidad por un ácido nucleico no natural que comprende un trifosfato que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia el ácido nucleico no natural con un difosfato o monofosfato, o sin fosfato, o una combinación de los mismos.

En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad relajada por un ácido nucleico no natural. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural y una especificidad por un ácido nucleico natural que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia el ácido nucleico natural. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural que comprende un azúcar modificado y una especificidad por un ácido nucleico natural que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia el ácido nucleico natural. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una especificidad por un ácido nucleico no natural que comprende una base modificada y una especificidad por un ácido nucleico natural que es al menos aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa de tipo silvestre hacia el ácido nucleico natural.

La ausencia de actividad de exonucleasa puede ser una característica de tipo silvestre o una característica impartida por una variante o polimerasa genomanipulada. Por ejemplo, un fragmento exo minus Klenow es una versión mutada del fragmento Klenow que carece de actividad de exonucleasa de corrección de errores 3' a 5'.

El método de la invención puede usarse para expandir el rango de sustrato de cualquier ADN polimerasa que carezca de una actividad de corrección de errores de exonucleasa 3' a 5' intrínseca o en donde se ha deshabilitado una actividad de corrección de errores de exonucleasa 3' a 5', por ejemplo, a través de mutación. Los ejemplos de ADN polimerasas incluyen polA, polB (véase, por ejemplo, Parrel & Loeb, Nature Struc Biol 2001) polC, polD, polY, polX y transcriptasas inversas (RT) pero preferentemente son polimerasas procesivas de alta fidelidad (documento PCT/GB2004/004643). En algunas realizaciones, una polimerasa modificada o de tipo silvestre carece sustancialmente de actividad de exonucleasa de corrección de errores 3' a 5'. En algunas realizaciones, una polimerasa modificada o de tipo silvestre carece sustancialmente de actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico no natural. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una actividad de exonucleasa de corrección de errores 3' a 5'. En algunas realizaciones, una polimerasa modificada o de tipo silvestre tiene una actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico natural y carece sustancialmente de actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico no natural.

En algunas realizaciones, una polimerasa modificada tiene una actividad de exonucleasa de corrección de errores 3' a 5' que es al menos aproximadamente el 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la actividad de exonucleasa de corrección de errores de la polimerasa de tipo silvestre. En algunas realizaciones, una polimerasa modificada tiene una actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico no natural que es al menos aproximadamente el 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la actividad de exonucleasa de corrección de errores de la polimerasa de tipo silvestre con respecto a un ácido nucleico natural. En algunas realizaciones, una polimerasa modificada tiene una actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico no natural y una actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico natural que es al menos aproximadamente el 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la actividad de exonucleasa de corrección de errores de la polimerasa de tipo silvestre con

respecto a un ácido nucleico natural. En algunas realizaciones, una polimerasa modificada tiene una actividad de exonucleasa de corrección de errores 3' a 5' para un ácido nucleico natural que es al menos aproximadamente el 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la actividad de exonucleasa de corrección de errores de la polimerasa de tipo silvestre con respecto al ácido nucleico natural.

En algunas realizaciones, las polimerasas se caracterizan según su velocidad de disociación de los ácidos nucleicos. En algunas realizaciones, una polimerasa tiene una velocidad de disociación relativamente baja para uno o más ácidos nucleicos naturales y no naturales. En algunas realizaciones, una polimerasa tiene una velocidad de disociación relativamente alta para uno o más ácidos nucleicos naturales y no naturales. La velocidad de disociación es una actividad de una polimerasa que puede ajustarse para sintonizar las velocidades de reacción en los métodos expuestos en el presente documento.

En algunas realizaciones, las polimerasas se caracterizan según su fidelidad cuando se usan con un ácido nucleico natural y/o no natural particular o colecciones de ácido nucleico natural y/o no natural. La fidelidad generalmente se refiere a la precisión con la que una polimerasa incorpora los ácidos nucleicos correctos en una cadena de ácido nucleico en crecimiento al fabricar una copia de un molde de ácido nucleico. La fidelidad de la ADN polimerasa se puede medir como la relación de incorporaciones de ácidos nucleicos naturales y no naturales correctas con respecto a incorrectas cuando los ácidos nucleico naturales y no naturales están presentes, por ejemplo, a concentraciones equivalentes, para competir por la síntesis de la hebra en el mismo sitio en el complejo binario de ácido nucleico polimerasa-hebra-molde. La fidelidad de la ADN polimerasa se puede calcular como la relación de (k_{cat}/K_m) para el ácido nucleico natural y no natural y (k_{cat}/K_m) para el ácido nucleico natural y no natural incorrecto; en donde k_{cat} y K_m son parámetros de Michaelis-Menten en la cinética enzimática en estado estacionario (Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, 2ª ed., pág. 350, W. H. Freeman & Co., Nueva York). En algunas realizaciones, una polimerasa tiene un valor de fidelidad de al menos aproximadamente 100, 1000, 10.000, 100.000 o 1×10^6 , con o sin actividad de corrección de errores.

En algunas realizaciones, las polimerasas de fuentes nativas o variantes de las mismas se criban usando un ensayo que detecta la incorporación de un ácido nucleico no natural que tenga una estructura particular. En un ejemplo, las polimerasas se pueden cribar para determinar la capacidad de incorporar un ácido nucleico no natural o UBP; por ejemplo, d5SICSTP, dNaMTP o d5SICSTP-dNaMTP UBP. Puede usarse una polimerasa, por ejemplo, una polimerasa heteróloga, que muestra una propiedad modificada para el ácido nucleico no natural en comparación con la polimerasa de tipo silvestre. Por ejemplo, la propiedad modificada puede ser, por ejemplo, K_m , k_{cat} , $V_{máx}$, la capacidad de procesamiento de polimerasa en presencia de un ácido nucleico no natural (o de un nucleótido de origen natural), la longitud de lectura promedio del molde por la polimerasa en presencia de un ácido nucleico no natural, la especificidad de la polimerasa por un ácido nucleico no natural, la velocidad de unión de un ácido nucleico no natural, la velocidad de liberación de producto

(pirofosfato, trifosfato, etc.), la velocidad de ramificación o cualquier combinación de los mismos. En una realización, la propiedad modificada es una reducción de K_m para un ácido nucleico no natural y/o un aumento de k_{cat}/K_m o $V_{máx}/K_m$ para un ácido nucleico no natural. De manera similar, la polimerasa tiene opcionalmente una velocidad aumentada de unión de un ácido nucleico no natural, una velocidad aumentada de liberación de producto y/o una velocidad de ramificación disminuida, en comparación con una polimerasa de tipo silvestre.

Al mismo tiempo, una polimerasa puede incorporar ácidos nucleicos naturales, por ejemplo, A, C, G y T, en una copia de ácido nucleico en crecimiento. Por ejemplo, una polimerasa muestra opcionalmente una actividad específica para un ácido nucleico natural que es al menos aproximadamente un 5 % tan alta (por ejemplo, 5 %, 10 %, 25 %, 50 %, 75 %, 100 % o más), como en una polimerasa de tipo silvestre correspondiente y una capacidad de procesamiento con ácidos naturales en presencia de un molde que es al menos un 5 % tan alta (por ejemplo, 5 %, 10 %, 25 %, 50 %, 75 %, 100 % o más) como en la polimerasa de tipo silvestre en presencia del ácido nucleico natural. Opcionalmente, la polimerasa presenta una k_{cat}/K_m o $V_{máx}/K_m$ para un nucleótido de origen natural que es al menos aproximadamente un 5 % tan alta (por ejemplo, aproximadamente un 5 %, 10 %, 25 %, 50 %, 75 % o 100 % o más) como en la polimerasa de tipo silvestre.

Las polimerasas usadas en el presente documento que pueden tener la capacidad de incorporar un ácido nucleico no natural de una estructura particular también se pueden producir usando un enfoque de evolución dirigida. Puede usarse un ensayo de síntesis de ácido nucleico para cribar variantes de polimerasa que tengan especificidad por cualquiera de diversos ácidos nucleicos no naturales. Por ejemplo, las variantes de polimerasa se pueden cribar para determinar la capacidad de incorporar un ácido nucleico no natural o UBP; por ejemplo, d5SICSTP, dNaMTP o d5SICSTP-dNaMTP UBP en ácidos nucleicos. En algunas realizaciones, tal ensayo es un ensayo *in vitro*, por ejemplo, que usa una variante de polimerasa recombinante. En algunas realizaciones, tal ensayo es un ensayo *in vivo*, por ejemplo, que expresa una variante de polimerasa en una célula. Dichas técnicas de evolución dirigida se pueden usar para cribar variantes de cualquier polimerasa adecuada para la actividad hacia cualquiera de los ácidos nucleicos no naturales expuestos en el presente documento.

Las polimerasas modificadas de las composiciones descritas pueden ser opcionalmente una ADN polimerasa de tipo $\Phi 29$ modificada y/o recombinante. Opcionalmente, la polimerasa puede ser una polimerasa modificada y/o

recombinante Φ 29, B103, GA-1, PZA, Φ 15, BS32, M2Y, Nf, G1, Cp-1, PRD1, PZE, SF5, Cp-5, Cp-7, PR4, PR5, PR722 o L17 polimerasa.

Las polimerasas de ácido nucleico generalmente útiles en la invención incluyen ADN polimerasas, ARN polimerasas, transcriptasas inversas y formas mutantes o alteradas de las mismas. Las ADN polimerasas y sus propiedades se describen en detalle en, entre otros lugares, DNA Replication 2ª edición, Kornberg y Baker, W. H. Freeman, Nueva York, N. Y. (1991). Las ADN polimerasas convencionales conocidas útiles en la invención incluyen, pero sin limitación, ADN polimerasa de *Pyrococcus furiosus* (Pfu) (Lundberg *et al.*, 1991, Gene, 108: 1, Stratagene), ADN polimerasa de *Pyrococcus woesei* (Pwo) (Hinnisdaels *et al.*, 1996, Biotechniques, 20:186-8, Boehringer Mannheim), ADN polimerasa de *Thermus thermophilus* (Tth) (Myers y Gelfand 1991, Biochemistry 30:7661), ADN polimerasa de *Bacillus stearothermophilus* (Stenesh y McGowan, 1977, Biochim Biophys Acta 475:32), ADN polimerasa de *Thermococcus litoralis* (Tli) (también denominada ADN polimerasa Vent™, Cariello *et al.*, 1991, Polynucleotides Res, 19: 4193, New England Biolabs), ADN polimerasa 9°Nm™ (New England Biolabs), fragmento Stoffel, Thermo Sequenase® (Amersham Pharmacia Biotech Reino Unido), Terminator™ (New England Biolabs), ADN polimerasa de *Thermotoga maritima* (Tma) (Diaz y Sabino, 1998 Braz J Med. Res, 31 : 1239), ADN polimerasa de *Thermus aquaticus* (Taq) (Chien *et al.*, 1976, J. Bacteriol, 127: 1550), DNA polymerase, ADN polimerasa de *Pyrococcus kodakaraensis* KOD (Takagi *et al.*, 1997, Appl. Environ. Microbiol. 63:4504), ADN polimerasa JDF-3 (de *thermococcus sp.* JDF-3, Solicitud de Patente WO 0132887), ADN polimerasa de *Pyrococcus* GB-D (PGB-D) (también denominada ADN polimerasa Deep Vent™, Juncosa-Ginesta *et al.*, 1994, Biotechniques, 16:820, New England Biolabs), ADN polimerasa UITma (del termófilo *Thermotoga maritima*; Diaz y Sabino, 1998 Braz J. Med. Res, 31 :1239; PE Applied Biosystems), ADN polimerasa Tgo (de *thermococcus gorgonarius*, Roche Molecular Biochemicals), ADN polimerasa I de *E. coli* (Lecomte y Doubleday, 1983, Polynucleotides Res. 11 :7505), ADN polimerasa T7 (Nordstrom *et al.*, 1981, J Biol. Chem. 256:3112) y ADN polimerasa II DP11/DP2 de arqueas (Cann *et al.*, 1998, Proc. Natl. Acad. Sci. USA 95:14250). Se contemplan tanto polimerasas mesófilas como las polimerasas termófilas. Las ADN polimerasas termófilas incluyen, pero sin limitación, ThermoSequenase®, 9°Nm™, Terminator™, Taq, Tne, Tma, Pfu, Tfl, Tth, Tli, fragmento Stoffel, ADN polimerasa Vent™ y Deep Vent™, ADN polimerasa KOD, Tgo, JDF-3 y mutantes, variantes y derivados de las mismas. También se contempla una polimerasa que es un mutante deficiente en exonucleasa 3'. Las transcriptasas inversas útiles en la invención incluyen, pero sin limitación, transcriptasas inversas de VIH, HTLV-I, HTLV-II, FeLV, FIV, SIV, AMV, MMTV, MoMuLV y otros retrovirus (véanse Levin, Cell 88:5-8 (1997); Verma, Biochim Biophys Acta. 473:1-38 (1977); Wu *et al.*, CRC Crit Rev Biochem. 3:289-347(1975)). Los ejemplos adicionales de polimerasas incluyen, pero sin limitación, ADN polimerasa 9°N, ADN polimerasa Taq, ADN polimerasa Phusion®, ADN polimerasa Pfu, ADN polimerasa RB69, ADN polimerasa KOD y ADN polimerasa VentR® Gardner *et al.* (2004) "Comparative Kinetics of Nucleotide Analog Incorporation by Vent DNA Polymerase (J. Biol. Chem., 279(12), 11834-11842; Gardner y Jack "Determinants of nucleotide sugar recognition in an archaeon DNA polymerase" Nucleic Acids Research, 27(12) 2545-2553). Las polimerasas aisladas de organismos no termófilos pueden ser inactivables por calor. Son ejemplos las ADN polimerasas de fagos. Se entenderá que las polimerasas de cualquiera de diversas fuentes pueden modificarse para aumentar o disminuir su tolerancia a condiciones de alta temperatura. En algunas realizaciones, una polimerasa puede ser termófila. En algunas realizaciones, una polimerasa termófila puede ser inactivable por calor. Las polimerasas termófilas son normalmente útiles para condiciones de alta temperatura o en condiciones de termociclado tales como las empleadas para las técnicas de reacción en cadena de la polimerasa (PCR).

En algunas realizaciones, la polimerasa comprende Φ 29, B103, GA-1, PZA, Φ 15, BS32, M2Y, Nf, G1, Cp-1, PRD1, PZE, SF5, Cp-5, Cp-7, PR4, PR5, PR722, L17, ThermoSequenase®, 9°Nm™, Terminator™ ADN polimerasa, Tne, Tma, Tfl, Tth, Tli, fragmento Stoffel, ADN polimerasa Vent™ y Deep Vent™, ADN polimerasa KOD, Tgo, JDF-3, Pfu, Taq, ADN polimerasa T7, ARN polimerasa T7, PGB-D, ADN polimerasa UITma, ADN polimerasa I de *E. coli*, ADN polimerasa II de *E. coli*, ADN polimerasa II DP11/DP2 de arqueas, ADN polimerasa 9°N, ADN polimerasa Taq, ADN polimerasa Phusion®, ADN polimerasa Pfu, ARN polimerasa SP6, ADN polimerasa RB69, transcriptasa inversa del virus de la mieloblastosis aviar (AMV), transcriptasa inversa del virus de la leucemia murina de Moloney (MMLV), transcriptasa inversa SuperScript® II y transcriptasa inversa SuperScript® III.

En algunas realizaciones, la polimerasa es ADN polimerasa 1-fragmento de Klenow, polimerasa Vent, ADN polimerasa Phusion®, ADN polimerasa KOD, polimerasa Taq, ADN polimerasa T7, ARN polimerasa T7, ADN polimerasa Terminator™, polimerasa POLB, ARN polimerasa SP6, ADN polimerasa I de *E. coli*, ADN polimerasa II de *E. coli*, transcriptasa inversa del virus de la mieloblastosis aviar (AMV), transcriptasa inversa del virus de la leucemia murina de Moloney (MMLV), transcriptasa inversa SuperScript® II o transcriptasa inversa SuperScript® III.

Además, dichas polimerasas se pueden usar para aplicaciones de secuenciación y/o amplificación de ADN, incluyendo aplicaciones en tiempo real, por ejemplo, en el contexto de amplificación o secuenciación que incluyen la incorporación de residuos de ácidos nucleicos no naturales en el ADN por la polimerasa. En otras realizaciones, el ácido nucleico no natural que se incorpora puede ser el mismo que un residuo natural, por ejemplo, cuando un marcador u otro resto del ácido nucleico no natural se elimina por la acción de la polimerasa durante la incorporación, o el ácido nucleico no natural puede tener una o más características que lo distinguen de un ácido nucleico natural.

Reactivos de ácidos nucleicos y herramientas

Un reactivo de ácido nucleico para su uso con un método, célula o microorganismo genomanipulado descrito en el

presente documento comprende uno o más ORF. Un ORF puede proceder de cualquier fuente adecuada, a veces de ADN genómico, ARNm, ARN de transcripción inversa o ADN complementario (ADNc) o una biblioteca de ácidos nucleicos que comprende uno o más de los anteriores, y es de cualquier especie de organismo que contenga una secuencia de ácido nucleico de interés, proteína de interés o actividad de interés. Los ejemplos no limitantes de organismos a partir de los cuales se puede obtener un ORF incluyen bacterias, levaduras, hongos, seres humanos, insectos, nematodos, bovinos, equinos, caninos, felinos, ratas o ratones, por ejemplo. En algunas realizaciones, se aísla o se purifica un reactivo de ácido nucleico u otro reactivo descrito en el presente documento.

Un reactivo de ácido nucleico a veces comprende una secuencia de nucleótidos adyacente a un ORF que se traduce junto con el ORF y codifica una etiqueta de aminoácido. La secuencia de nucleótidos que codifica la etiqueta se localiza en 3' y/o 5' de un ORF en el reactivo de ácido nucleico, codificando así una etiqueta en el extremo C o N de la proteína o péptido codificado por el ORF. Se puede utilizar cualquier etiqueta que no anule la transcripción y/o traducción *in vitro* y puede seleccionarse apropiadamente por el experto. Las etiquetas pueden facilitar el aislamiento y/o la purificación del producto de ORF deseado de los medios de cultivo o fermentación.

Un ácido nucleico o reactivo de ácido nucleico puede comprender determinados elementos, por ejemplo, elementos reguladores, a menudo seleccionados de acuerdo con el uso pretendido del ácido nucleico. Cualquiera de los siguientes elementos puede incluirse o excluirse de un reactivo de ácido nucleico. Un reactivo de ácido nucleico, por ejemplo, puede incluir uno o más o todos los siguientes elementos nucleotídicos: uno o más elementos promotores, una o más regiones no traducidas 5' (5'UTR), una o más regiones en las que puede insertarse una secuencia de nucleótidos diana (un "elemento de inserción"), una o más secuencias de nucleótidos diana, una o más regiones no traducidas 3' (3'UTR) y uno o más elementos de selección. Se puede proporcionar un reactivo de ácido nucleico con uno o más de tales elementos y se pueden insertar otros elementos en el ácido nucleico antes de que el ácido nucleico se introduzca en el organismo deseado. En algunas realizaciones, un reactivo de ácido nucleico proporcionado comprende un promotor, 5'UTR, 3'UTR opcional y uno o más elementos de inserción mediante los cuales se inserta (es decir, se clona) una secuencia de nucleótidos diana en el reactivo de ácido nucleotídico. En determinadas realizaciones, un reactivo de ácido nucleico proporcionado comprende un promotor, uno o más elementos de inserción y una 3'UTR opcional, y se inserta una 5'UTR/secuencia de nucleótidos diana con una 3'UTR opcional. Los elementos pueden disponerse en cualquier orden adecuado para la expresión en el sistema de expresión elegido (por ejemplo, expresión en un organismo elegido, o expresión en un sistema acelular, por ejemplo), y en algunas realizaciones, un reactivo de ácido nucleico comprende los siguientes elementos en la dirección 5' a 3': (1) un elemento promotor, 5'UTR y uno o más elementos de inserción; (2) un elemento promotor, 5'UTR y secuencia de nucleótidos diana; (3) un elemento promotor, 5'UTR, uno o más elementos de inserción y 3'UTR; y (4) un elemento promotor, 5'UTR, secuencia de nucleótidos diana y 3'UTR.

Los reactivos de ácido nucleico, por ejemplo, casetes de expresión y/o vectores de expresión, pueden incluir diversos elementos reguladores, incluyendo promotores, potenciadores, secuencias de inicio de la traducción, secuencias de terminación de la transcripción y otros elementos. Un "promotor" es generalmente una secuencia o secuencias de ADN que funcionan cuando se encuentran en una ubicación relativamente fija con respecto al sitio de inicio de la transcripción. Por ejemplo, el promotor puede estar cadena arriba del segmento de ácido nucleico del transportador de nucleótidos trifosfato. Un "promotor" contiene los elementos centrales necesarios para la interacción de bases de la ARN polimerasa y los factores de transcripción y puede contener elementos cadena arriba y elementos de respuesta. "Potenciador" generalmente se refiere a una secuencia de ADN que funciona sin una distancia fija desde el sitio de inicio de la transcripción y puede estar en 5' o 3' con respecto a la unidad de transcripción. Además, los potenciadores pueden estar dentro de un intrón, así como dentro de la secuencia codificante. Normalmente, tienen entre 10 y 300 pb de longitud y funcionan en cis. Los potenciadores tienen la función de aumentar la transcripción de los promotores cercanos. Los potenciadores, como los promotores, a menudo también contienen elementos de respuesta que median en la regulación de la transcripción. Los potenciadores a menudo determinan la regulación de la expresión.

Como se ha indicado anteriormente, los reactivos de ácido nucleico también pueden comprender una o más 5' UTR y una o más 3' UTR. Por ejemplo, los vectores de expresión usados en células huésped procariotas (por ejemplo, virus, bacteria) pueden contener secuencias que señalan la terminación de la transcripción que puede afectar a la expresión del ARNm. Estas regiones se pueden transcribir como segmentos poliadenilados en la parte no traducida del ARNm que codifica la proteína del factor tisular. Las regiones no traducidas 3' también incluyen sitios de terminación de la transcripción. En algunas realizaciones preferidas, una unidad de transcripción comprende una región de poliadenilación. Un beneficio de esta región es que aumenta la probabilidad de que la unidad transcrita sea procesada y transportada como ARNm. La identificación y el uso de señales de poliadenilación en construcciones de expresión están bien establecidos. En algunas realizaciones preferidas, se pueden usar señales de poliadenilación homólogas en las construcciones transgénicas.

Una 5' UTR puede comprender uno o más elementos endógenos con respecto a la secuencia de nucleótidos de la que se origina, y algunas veces incluye uno o más elementos exógenos. Una 5' UTR puede originarse a partir de cualquier ácido nucleico adecuado, tal como ADN genómico, ADN plasmídico, ARN o ARNm, por ejemplo, de cualquier organismo adecuado (por ejemplo, virus, bacteria, levadura, hongos, planta, insecto o mamífero). El experto puede seleccionar los elementos apropiados para la 5' UTR basándose en el sistema de expresión elegido (por ejemplo, expresión en un organismo elegido o expresión en un sistema libre de células, por ejemplo). Una 5' UTR a veces

comprende uno o más de los siguientes elementos conocidos por el experto: secuencias potenciadoras (por ejemplo, transcripcionales o traduccionales), sitio de inicio de la transcripción, sitio de unión del factor de transcripción, sitio de regulación de la traducción, sitio de inicio de la traducción, sitio de unión del factor de traducción, sitio de unión de proteínas accesorias, sitios de unión del agente de regulación de la retroalimentación, caja Pribnow, caja TATA, elemento -35, caja E (elemento de unión de hélice-bucle-hélice), sitio de unión a ribosomas, replicón, sitio interno de entrada a ribosomas (IRES, por sus siglas en inglés), elemento silenciador y similares. En algunas realizaciones, un elemento promotor puede aislarse de manera que todos los elementos 5' UTR necesarios para la regulación condicional adecuada estén contenidos en el fragmento del elemento promotor, o dentro de una subsecuencia funcional de un fragmento del elemento promotor.

Una 5' UTR en el reactivo de ácido nucleico puede comprender una secuencia de nucleótidos potenciadora de la traducción. Una secuencia de nucleótidos potenciadora de la traducción a menudo se encuentra entre el promotor y la secuencia de nucleótidos diana en un reactivo de ácido nucleico. Una secuencia potenciadora de la traducción a menudo se une a un ribosoma, a veces es una secuencia de ribonucleótidos de unión a ARNr 18S (es decir, una secuencia de unión al ribosoma 40S) y, a veces, es una secuencia interna de entrada al ribosoma (IRES). Un IRES generalmente forma un armazón de ARN con estructuras terciarias de ARN colocadas con precisión que entran en contacto con una subunidad ribosómica 40S a través de una serie de interacciones intermoleculares específicas. Se conocen ejemplos de secuencias potenciadoras ribosómicas y se pueden identificar por el experto (por ejemplo, Mignone *et al.*, *Nucleic Acids Research* 33: D141-D146 (2005); Paulous *et al.*, *Nucleic Acids Research* 31: 722-733 (2003); Akbergenov *et al.*, *Nucleic Acids Research* 32: 239-247 (2004); Mignone *et al.*, *Genome Biology* 3(3): reviews0004.1-0001.10 (2002); Gallie, *Nucleic Acids Research* 30: 3401-3411 (2002); Shaloiiko *et al.*, DOI: 10.1002/bit.20267; y Gallie *et al.*, *Nucleic Acids Research* 15: 3257-3273 (1987)).

Una secuencia potenciadora de la traducción a veces es una secuencia eucariota, tal como una secuencia consenso de Kozak u otra secuencia (por ejemplo, secuencia de pólipos hidroides, número de acceso de GenBank U07128). Una secuencia potenciadora de la traducción a veces es una secuencia procariota, tal como una secuencia consenso Shine-Dalgarno. En determinadas realizaciones, la secuencia potenciadora de la traducción es una secuencia de nucleótidos viral. Una secuencia potenciadora de la traducción a veces es de una 5' UTR de un virus vegetal, tal como el virus del mosaico del tabaco (TMV), el virus del mosaico de la alfalfa (AMV); el virus del grabado del tabaco (ETV); el virus Y de la patata (PVY); el virus del mosaico del nabo (poty) y el virus del mosaico transmitido por semillas de guisantes, por ejemplo. En determinadas realizaciones, se incluye una secuencia omega de aproximadamente 67 bases de longitud del TMV en el reactivo de ácido nucleico como una secuencia potenciadora de la traducción (por ejemplo, desprovista de nucleótidos de guanosina e incluye una región central poli (CAA) de 25 nucleótidos de longitud).

Una 3' UTR puede comprender uno o más elementos endógenos con respecto a la secuencia de nucleótidos de la que se origina, y algunas veces incluye uno o más elementos exógenos. Una 3' UTR puede originarse a partir de cualquier ácido nucleico adecuado, tal como ADN genómico, ADN plasmídico, ARN o ARNm, por ejemplo, de cualquier organismo adecuado (por ejemplo, un virus, bacteria, levadura, hongos, planta, insecto o mamífero). El experto puede seleccionar elementos apropiados para la 3' UTR basándose en el sistema de expresión elegido (por ejemplo, expresión en un organismo elegido, por ejemplo). Una 3' UTR a veces comprende uno o más de los siguientes elementos conocidos por el experto: sitio de regulación de la transcripción, sitio de inicio de la transcripción, sitio de terminación de la transcripción, sitio de unión del factor de transcripción, sitio de regulación de la traducción, sitio de terminación de la traducción, sitio de inicio de la traducción, sitio de unión del factor de traducción, sitio de unión a ribosomas, replicón, elemento potenciador, elemento silenciador y cola de poliadenosina. Una 3' UTR a menudo incluye una cola de poliadenosina y algunas veces no, y si hay una cola de poliadenosina presente, se pueden añadir o eliminar uno o más restos de adenosina (por ejemplo, se pueden añadir o restar aproximadamente 5, aproximadamente 10, aproximadamente 15, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 35, aproximadamente 40, aproximadamente 45 o aproximadamente 50 restos de adenosina).

En algunas realizaciones, se usa una modificación de una 5' UTR y/o una 3' UTR para alterar (por ejemplo, aumentar, añadir, disminuir o eliminar sustancialmente) la actividad de un promotor. La alteración de la actividad promotora puede, a su vez, alterar la actividad de un péptido, polipéptido o proteína (por ejemplo, la actividad enzimática, por ejemplo), mediante un cambio en la transcripción de la secuencia de nucleótidos de interés a partir de un elemento promotor unido operativamente que comprende la 5' o 3' UTR modificada. Por ejemplo, un microorganismo se puede genomanipular mediante modificación genética para expresar un reactivo de ácido nucleico que comprende una 5' o 3' UTR modificada que puede añadir una actividad novedosa (por ejemplo, una actividad que normalmente no se encuentra en el organismo huésped) o aumentar la expresión de una actividad existente aumentando la transcripción de un promotor homólogo o heterólogo unido operativamente a una secuencia de nucleótidos de interés (por ejemplo, secuencia de nucleótidos homóloga o heteróloga de interés), en determinadas realizaciones. En algunas realizaciones, un microorganismo puede genomanipularse mediante modificación genética para expresar un reactivo de ácido nucleico que comprende una 5' o 3' UTR modificada que puede disminuir la expresión de una actividad al disminuir o eliminar sustancialmente la transcripción de un promotor homólogo o heterólogo unido operativamente a una secuencia de nucleótidos de interés, en determinadas realizaciones.

La expresión de un transportador de nucleótido trifosfato a partir de un casete de expresión o vector de expresión

puede controlarse mediante cualquier promotor capaz de expresarse en células procariotas. Normalmente se requiere un elemento promotor para la síntesis de ADN y/o la síntesis de ARN. Un elemento promotor a menudo comprende una región de ADN que puede facilitar la transcripción de un gen particular, proporcionando un sitio de inicio para la síntesis de ARN correspondiente a un gen. Los promotores generalmente están ubicados cerca de los genes que regulan, están ubicados cadena arriba del gen (por ejemplo, 5' del gen) y están en la misma hebra de ADN que la hebra codificante del gen, en algunas realizaciones. En algunas realizaciones, un elemento promotor puede aislarse de un gen u organismo e insertarse en conexión funcional con una secuencia de polinucleótidos para permitir la expresión alterada y/o regulada. Un promotor no nativo (por ejemplo, un promotor que normalmente no está asociado con una secuencia de ácido nucleico determinada) usado para la expresión de un ácido nucleico, a menudo se denomina promotor heterólogo. En determinadas realizaciones, se puede insertar un promotor heterólogo y/o una 5'UTR en conexión funcional con un polinucleótido que codifica un polipéptido que tiene una actividad deseada como se describe en el presente documento. Las expresiones "unido operativamente" y "en conexión funcional con", como se usan en el presente documento con respecto a los promotores, se refieren a una relación entre una secuencia codificante y un elemento promotor. El promotor está unido operativamente o en conexión funcional con la secuencia codificante cuando la expresión de la secuencia codificante a través de la transcripción se regula o se controla por el elemento promotor. Las expresiones "unido operativamente" y "en conexión funcional con" se utilizan indistintamente en el presente documento con respecto a los elementos promotores.

Un promotor a menudo interactúa con una ARN polimerasa. Una polimerasa es una enzima que cataliza la síntesis de ácidos nucleicos usando un reactivo de ácido nucleico preexistente. Cuando el molde es un molde de ADN, se transcribe una molécula de ARN antes de que se sintetice la proteína. Las enzimas que tienen actividad polimerasa adecuadas para su uso en los presentes métodos incluyen cualquier polimerasa que sea activa en el sistema elegido con el molde elegido para sintetizar la proteína. En algunas realizaciones, un promotor (por ejemplo, un promotor heterólogo) también denominado en el presente documento como elemento promotor, se puede unir operativamente a una secuencia de nucleótidos o un marco de lectura abierto (ORF). La transcripción del elemento promotor puede catalizar la síntesis de un ARN correspondiente a la secuencia de nucleótidos o secuencia de ORF unida operativamente al promotor, lo que a su vez conduce a la síntesis de un péptido, polipéptido o proteína deseados.

Los elementos promotores a veces muestran capacidad de respuesta al control regulador. A veces, los elementos promotores también pueden regularse por un agente selectivo. Es decir, la transcripción de elementos promotores a veces se puede activar, desactivar, regular positivamente o regular negativamente, en respuesta a un cambio en las condiciones o señales ambientales, nutricionales o internas (por ejemplo, promotores inducibles por calor, promotores regulados por luz, promotores regulados por retroalimentación, promotores influenciados por hormonas, promotores específicos de tejido, promotores influenciados por oxígeno y el pH, promotores que responden a agentes selectivos (por ejemplo, kanamicina) y similares, por ejemplo). Los promotores influenciados por señales ambientales, nutricionales o internas frecuentemente están influenciados por una señal (directa o indirecta) que se une al promotor o cerca de este y aumenta o disminuye la expresión de la secuencia diana bajo ciertas condiciones.

Los ejemplos no limitantes de agentes selectivos o reguladores que influyen en la transcripción de un elemento promotor utilizado en las realizaciones descritas en la presente invención incluyen, sin limitación, (1) segmentos de ácido nucleico que codifican productos que proporcionan resistencia contra compuestos de otro modo tóxicos (por ejemplo, antibióticos); (2) segmentos de ácido nucleico que codifican productos que por lo demás están ausentes en la célula receptora (por ejemplo, productos esenciales, genes de ARNt, marcadores auxotróficos); (3) segmentos de ácido nucleico que codifican productos que suprimen la actividad de un producto génico; (4) segmentos de ácido nucleico que codifican productos que pueden identificarse fácilmente (por ejemplo, marcadores fenotípicos tales como antibióticos (por ejemplo, β -lactamasa), β -galactosidasa, proteína verde fluorescente (GFP), proteína amarilla fluorescente (YFP), proteína roja fluorescente (RFP), proteína cian fluorescente (CFP) y proteínas de la superficie celular); (5) segmentos de ácido nucleico que se unen a productos que de otro modo son perjudiciales para la supervivencia y/o la función celular; (6) segmentos de ácido nucleico que de otro modo inhiben la actividad de cualquiera de los segmentos de ácido nucleico descritos en los N.º 1-5 anteriores (por ejemplo, oligonucleótidos no codificantes); (7) segmentos de ácido nucleico que se unen a productos que modifican un sustrato (por ejemplo, endonucleasas de restricción); (8) segmentos de ácido nucleico que pueden usarse para aislar o identificar una molécula deseada (por ejemplo, sitios de unión a proteínas específicos); (9) segmentos de ácido nucleico que codifican una secuencia de nucleótidos específica que de otro modo puede no ser funcional (por ejemplo, para la amplificación por PCR de subpoblaciones de moléculas); (10) segmentos de ácido nucleico que, cuando están ausentes, confieren directa o indirectamente resistencia o sensibilidad a compuestos particulares; (11) segmentos de ácido nucleico que codifican productos que son tóxicos o convierten un compuesto relativamente no tóxico en un compuesto tóxico (por ejemplo, timidina cinasa del herpes simple, citosina desaminasa) en las células receptoras; (12) segmentos de ácido nucleico que inhiben la replicación, la división o la heredabilidad de las moléculas de ácido nucleico que los contienen; y/o (13) segmentos de ácido nucleico que codifican funciones de replicación condicional, por ejemplo, replicación en determinados huéspedes o cepas de células huésped o bajo ciertas condiciones ambientales (por ejemplo, temperatura, condiciones nutricionales y similares). En algunas realizaciones, el agente regulador o selectivo se puede añadir para cambiar las condiciones de crecimiento existentes a las que está sujeto el organismo (por ejemplo, crecimiento en cultivo líquido, crecimiento en un fermentador, crecimiento en placas de nutrientes sólidos y similares, por ejemplo).

En algunas realizaciones, la regulación de un elemento promotor puede usarse para alterar (por ejemplo, aumentar, añadir, disminuir o eliminar sustancialmente) la actividad de un péptido, polipéptido o proteína (por ejemplo, la actividad enzimática, por ejemplo). Por ejemplo, un microorganismo se puede genomanipular mediante modificación genética para expresar un reactivo de ácido nucleico que puede añadir una actividad novedosa (por ejemplo, una actividad que normalmente no se encuentra en el organismo huésped) o aumentar la expresión de una actividad existente aumentando la transcripción de un promotor homólogo o heterólogo unido operativamente a una secuencia de nucleótidos de interés (por ejemplo, secuencia de nucleótidos homóloga o heteróloga de interés), en determinadas realizaciones. En algunas realizaciones, un microorganismo puede genomanipularse mediante modificación genética para expresar un reactivo de ácido nucleico que puede disminuir la expresión de una actividad al disminuir o eliminar sustancialmente la transcripción de un promotor homólogo o heterólogo unido operativamente a una secuencia de nucleótidos de interés, en determinadas realizaciones.

Los ácidos nucleicos que codifican proteínas heterólogas, por ejemplo, transportadores de nucleótidos trifosfato, pueden insertarse o emplearse con cualquier sistema de expresión adecuado. En algunas realizaciones, un reactivo de ácido nucleico a veces se integra de manera estable en el cromosoma del organismo huésped, o un reactivo de ácido nucleico puede ser una delección de una porción del cromosoma huésped, en determinadas realizaciones (por ejemplo, organismos genéticamente modificados, en donde la alteración del genoma huésped confiere la capacidad de mantener selectiva o preferencialmente el organismo deseado portador de la modificación genética). Dichos reactivos de ácido nucleico (por ejemplo, ácidos nucleicos u organismos modificados genéticamente cuyo genoma alterado confiere un rasgo seleccionable al organismo) pueden seleccionarse por su capacidad para guiar la producción de una proteína o molécula de ácido nucleico deseada. Cuando se desee, el reactivo de ácido nucleico se puede alterar de modo que los codones codifiquen (i) el mismo aminoácido, usando un ARNt diferente al especificado en la secuencia nativa, o (ii) un aminoácido diferente al normal, incluyendo aminoácidos no convencionales o no naturales (incluyendo aminoácidos marcados de forma detectable).

La expresión recombinante se logra de manera útil usando un casete de expresión que puede formar parte de un vector, tal como un plásmido. Un vector puede incluir un promotor unido operativamente a un ácido nucleico que codifica un transportador de nucleótidos trifosfato. Un vector también puede incluir otros elementos necesarios para la transcripción y la traducción como se describe en el presente documento. Un casete de expresión, un vector de expresión y las secuencias en un casete o vector pueden ser heterólogos con respecto a la célula con la que se ponen en contacto los nucleótidos no naturales. Por ejemplo, una secuencia transportadora de nucleótidos trifosfato puede ser heteróloga con respecto a la célula.

Se puede producir diversos vectores de expresión procariotas adecuados para transportar, codificar y/o expresar transportadores de nucleótidos trifosfato. Dichos vectores de expresión incluyen, por ejemplo, pET, pET3d, pCR2.1, pBAD, pUC y vectores de levadura. Los vectores pueden usarse, por ejemplo, en diversas situaciones *in vivo* e *in vitro*. Los ejemplos no limitantes de promotores procariotas que pueden usarse incluyen promotores SP6, T7, T5, *tac*, *bla*, *trp*, *gal*, *lac* o maltosa. Los vectores virales que pueden emplearse incluyen los relacionados con lentivirus, adenovirus, virus adenoasociado, virus del herpes, virus variolovacunal, virus de la polio, virus del SIDA, virus trófico neuronal, Sindbis y otros virus. También son útiles las familias virales que comparten las propiedades de estos virus que las hacen adecuadas para su uso como vectores. Los vectores retrovirales que pueden emplearse incluyen los descritos en Verma, American Society for Microbiology, págs. 229-232, Washington, (1985). Por ejemplo, dichos vectores retrovirales pueden incluir el virus de la leucemia murina de Maloney, MMLV, y otros retrovirus que expresan propiedades deseables. Normalmente, los vectores virales contienen genes tempranos no estructurales, genes tardíos estructurales, un transcrito de ARN polimerasa III, repeticiones terminales invertidas necesarias para la replicación y encapsidación, y promotores para controlar la transcripción y replicación del genoma viral. Cuando se genomanipulan como vectores, los virus normalmente tienen uno o más de los genes tempranos eliminados y se inserta un gen o casete génico/promotor en el genoma viral en lugar del ácido nucleico viral eliminado.

50 Clonación

Se puede utilizar cualquier estrategia de clonación conveniente conocida en la técnica para incorporar un elemento, tal como un ORF, en un reactivo de ácido nucleico. Se pueden utilizar métodos conocidos para insertar un elemento en el molde independientemente de un elemento de inserción, tal como (1) escindir el molde en uno o más sitios de enzimas de restricción existentes y ligar un elemento de interés y (2) añadir sitios de enzimas de restricción al molde hibridando cebadores oligonucleotídicos que incluyen uno o más sitios de enzimas de restricción adecuados y amplificando mediante reacción en cadena de la polimerasa (descrita con mayor detalle en el presente documento). Otras estrategias de clonación aprovechan uno o más sitios de inserción presentes o insertados en el reactivo de ácido nucleico, tal como un sitio de hibridación de cebador oligonucleotídico para PCR, por ejemplo, y otros descritos en el presente documento. En algunas realizaciones, se puede combinar una estrategia de clonación con manipulación genética tal como recombinación (por ejemplo, recombinación de un reactivo de ácido nucleico con una secuencia de ácido nucleico de interés en el genoma del organismo a modificar, como se describe adicionalmente en el presente documento). En algunas realizaciones, el uno o más ORF clonados pueden producir (directa o indirectamente) transportadores de nucleótidos trifosfato y/o polimerasas modificados o de tipo silvestre, mediante genomanipulación de un microorganismo con uno o más ORF de interés, cuyo microorganismo comprende actividades alteradas de la actividad del transportador de nucleótidos trifosfato o la actividad de polimerasa.

Un ácido nucleico se puede escindir específicamente poniendo en contacto el ácido nucleico con uno o más agentes de escisión específicos. Los agentes de escisión específicos a menudo se escinden específicamente de acuerdo con una secuencia de nucleótidos particular en un sitio particular. Los ejemplos de agentes de escisión específicos de enzimas incluyen, sin limitación, endonucleasas (por ejemplo, DNasa (por ejemplo, DNasa I, II); RNasa (por ejemplo, RNasa E, F, H, P); la enzima Cleavase™; ADN polimerasa Taq; ADN polimerasa I de *E. coli*; endonucleasas FEN-1 murinas; endonucleasas de restricción de tipo I, II o III tales como Acc I, Afl III, Alu I, Alw44 I, Apa I, Asn I, Ava I, Ava II, BamH I, Ban II, Bcl I, Bgl I, Bgl II, Bln I, Bsal, Bsm I, BsmBI, BssH II, BstE II, Cfo I, Cla I, Dde I, Dpn I, Dra I, EclX I, EcoR I, EcoR II, EcoR V, Hae II, Hae III, Hind II, Hind III, Hpa I, Hpa II, Kpn I, Ksp I, Mlu I, MluN I, Msp I, Nci I, Nco I, Nde I, Nde II, Nhe I, Not I, Nru I, Nsi I, Pst I, Pvu I, Pvu II, Rsa I, Sac I, Sal I, Sau3A I, Sca I, ScrF I, Sfi I, Sma I, Spe I, Sph I, Ssp I, Stu I, Sty I, Swa I, Taq I, Xba I, Xho I); glicosilasas (por ejemplo, uracil-ADN glicosilasa (UDG), 3-metiladenina ADN glicosilasa, 3-metiladenina ADN glicosilasa II, pirimidina hidrato-ADN glicosilasa, FaPy-ADN glicosilasa, ADN glicosilasa con emparejamiento incorrecto a timina, hipoxantina-ADN glicosilasa, 5-Hidroximetiluracil ADN glicosilasa (HmUDG), 5-Hidroximetilcitosina ADN glicosilasa o 1,N6-eteno-adenina ADN glicosilasa); exonucleasas (por ejemplo, exonucleasa III); ribozimas y DNazimas. La muestra de ácido nucleico puede tratarse con un agente químico o sintetizarse usando nucleótidos modificados, y el ácido nucleico modificado puede escindirse. En ejemplos no limitantes, el ácido nucleico de muestra puede tratarse con (i) agentes alquilantes tales como metilnitrosourea que generan varias bases alquiladas, incluyendo N3-metiladenina y N3-metilguanina, que se reconocen y se escinden por alquil purina ADN glicosilasa; (ii) bisulfito de sodio, que provoca la desaminación de residuos de citosina en el ADN para formar residuos de uracilo que pueden escindirse por uracil N-glicosilasa; y (iii) un agente químico que convierte la guanina en su forma oxidada, 8-hidroxiguanina, que puede escindirse mediante formamidopirimidina ADN N-glicosilasa. Los ejemplos de procesos de escisión química incluyen, sin limitación, alquilación (por ejemplo, alquilación de ácido nucleico modificado con fosforotioato); escisión de la labilidad ácida del ácido nucleico que contiene P3'-N5'-fosforamidoato; y tratamiento de ácido nucleico con tetróxido de osmio y piperidina.

En algunas realizaciones, el reactivo de ácido nucleico incluye uno o más sitios de inserción de recombinasa. Un sitio de inserción de recombinasa es una secuencia de reconocimiento en una molécula de ácido nucleico que participa en una reacción de integración/recombinación mediante proteínas de recombinación. Por ejemplo, el sitio de recombinación para la recombinasa Cre es loxP, que es una secuencia de 34 pares de bases compuesta por dos repeticiones invertidas de 13 pares de bases (que sirven como sitios de unión de recombinasa) que flanquean una secuencia central de 8 pares de bases (por ejemplo, Sauer, Curr. Opin. Biotech. 5:521-527 (1994)). Otros ejemplos de sitios de recombinación incluyen secuencias attB, attP, attL y attR y mutantes, fragmentos, variantes y derivados de las mismas, que se reconocen por la proteína de recombinación λ Int y por el factor de integración del huésped (IHF) de proteínas auxiliares, FIS y escisionasa (Xis) (por ejemplo, las Patentes de EE.UU. N.º 5.888.732; 6.143.557; 6.171.861; 6.270.969; 6.277.608; y 6.720.140; las Sol. de Patente de EE.UU. N.º 09/517.466 y 09/732.914; la Publicación de Patente de EE.UU. N.º US2002/0007051; y Landy, Curr. Opin. Biotech. 3:699-707 (1993)).

Los ejemplos de ácidos nucleicos de clonación de recombinasa se encuentran en los sistemas Gateway® (Invitrogen, California), que incluyen al menos un sitio de recombinación para clonar moléculas de ácido nucleico deseadas *in vivo* o *in vitro*. En algunas realizaciones, el sistema utiliza vectores que contienen al menos dos sitios de recombinación de sitio específico diferentes, a menudo basados en el sistema lambda del bacteriófago (por ejemplo, att1 y att2), y se mutan de los sitios de tipo silvestre (att0). Cada sitio mutado tiene una especificidad única para el sitio att de su compañero afín (es decir, el sitio de recombinación del compañero de unión) del mismo tipo (por ejemplo, attB1 con attP1 o attL1 con attR1) y no reaccionará de forma cruzada con los sitios de recombinación del otro tipo mutante o con el sitio att0 de tipo silvestre. Las diferentes especificidades de sitio permiten la clonación o el enlace direccional de las moléculas deseadas, proporcionando así la orientación deseada de las moléculas clonadas. Los fragmentos de ácido nucleico flanqueados por sitios de recombinación se clonan y se subclonan usando el sistema Gateway® reemplazando un marcador seleccionable (por ejemplo, ccdB) flanqueado por sitios att en la molécula de plásmido receptora, a veces denominado vector de destino. A continuación, se seleccionan los clones deseados mediante la transformación de una cepa huésped sensible a ccdB y la selección positiva de un marcador en la molécula receptora. Se pueden usar estrategias similares para la selección negativa (por ejemplo, el uso de genes tóxicos) en otros organismos tales como timidina cinasa (TK) en mamíferos e insectos.

Un reactivo de ácido nucleico a veces contiene uno o más elementos de origen de replicación (ORI, por sus siglas en inglés). En algunas realizaciones, un molde comprende dos o más ORI, en donde uno funciona de manera eficiente en un organismo (por ejemplo, una bacteria) y otro funciona de manera eficiente en otro organismo (por ejemplo, un eucariota, como una levadura, por ejemplo). En algunas realizaciones, un ORI puede funcionar eficazmente en una especie (por ejemplo, *S. cerevisiae*, por ejemplo) y otro ORI puede funcionar eficazmente en una especie diferente (por ejemplo, *S. pombe*, por ejemplo). Un reactivo de ácido nucleico también incluye a veces uno o más sitios de regulación de la transcripción.

Un reactivo de ácido nucleico, por ejemplo, un casete de expresión o un vector, puede incluir una secuencia de ácido nucleico que codifica un producto marcador. Se usa un producto marcador para determinar si un gen se ha entregado a la célula y si una vez entregado se está expresando. Los genes marcadores de ejemplo incluyen el gen lacZ de *E. coli* que codifica la β -galactosidasa y la proteína verde fluorescente. En algunas realizaciones, el marcador puede ser un marcador seleccionable. Cuando dichos marcadores seleccionables se transfieren con éxito a una célula huésped,

la célula huésped transformada puede sobrevivir si se pone bajo presión selectiva. Hay dos categorías distintas de regímenes selectivos que se usan ampliamente. La primera categoría se basa en el metabolismo de una célula y el uso de una línea celular mutante que carece de la capacidad de crecer independientemente de un medio complementado. La segunda categoría es la selección dominante que se refiere a un esquema de selección usado en cualquier tipo de célula y no requiere el uso de una línea celular mutante. Normalmente, estos esquemas usan un fármaco para detener el crecimiento de una célula huésped. Aquellas células que tienen un gen nuevo expresarían una proteína que transmite resistencia al fármaco y sobrevivirían a la selección. Los ejemplos de dicha selección dominante usan los fármacos neomicina (Southern *et al.*, J. Molec. Appl. Genet. 1: 327 (1982)), ácido micofenólico, (Mulligan *et al.*, Science 209: 1422 (1980)) o higromicina, (Sugden, *et al.*, Mol. Cell. Biol. 5: 410-413 (1985)).

Un reactivo de ácido nucleico puede incluir uno o más elementos de selección (por ejemplo, elementos para la selección de la presencia del reactivo de ácido nucleico y no para la activación de un elemento promotor que puede regularse selectivamente). Los elementos de selección a menudo se utilizan usando procesos conocidos para determinar si un reactivo de ácido nucleico está incluido en una célula. En algunas realizaciones, un reactivo de ácido nucleico incluye dos o más elementos de selección, en donde uno funciona de manera eficiente en un organismo y otro funciona de manera eficiente en otro organismo. Los ejemplos de elementos de selección incluyen, pero sin limitación, (1) segmentos de ácido nucleico que codifican productos que proporcionan resistencia contra compuestos de otro modo tóxicos (por ejemplo, antibióticos); (2) segmentos de ácido nucleico que codifican productos que por lo demás están ausentes en la célula receptora (por ejemplo, productos esenciales, genes de ARNt, marcadores auxotróficos); (3) segmentos de ácido nucleico que codifican productos que suprimen la actividad de un producto génico; (4) segmentos de ácido nucleico que codifican productos que pueden identificarse fácilmente (por ejemplo, marcadores fenotípicos tales como antibióticos (por ejemplo, β -lactamasa), β -galactosidasa, proteína verde fluorescente (GFP), proteína amarilla fluorescente (YFP), proteína roja fluorescente (RFP), proteína cian fluorescente (CFP) y proteínas de la superficie celular); (5) segmentos de ácido nucleico que se unen a productos que de otro modo son perjudiciales para la supervivencia y/o la función celular; (6) segmentos de ácido nucleico que de otro modo inhiben la actividad de cualquiera de los segmentos de ácido nucleico descritos en los N.º 1-5 anteriores (por ejemplo, oligonucleótidos no codificantes); (7) segmentos de ácido nucleico que se unen a productos que modifican un sustrato (por ejemplo, endonucleasas de restricción); (8) segmentos de ácido nucleico que pueden usarse para aislar o identificar una molécula deseada (por ejemplo, sitios de unión a proteínas específicos); (9) segmentos de ácido nucleico que codifican una secuencia de nucleótidos específica que de otro modo puede no ser funcional (por ejemplo, para la amplificación por PCR de subpoblaciones de moléculas); (10) segmentos de ácido nucleico que, cuando están ausentes, confieren directa o indirectamente resistencia o sensibilidad a compuestos particulares; (11) segmentos de ácido nucleico que codifican productos que son tóxicos o convierten un compuesto relativamente no tóxico en un compuesto tóxico (por ejemplo, timidina cinasa del herpes simple, citosina desaminasa) en las células receptoras; (12) segmentos de ácido nucleico que inhiben la replicación, la división o la heredabilidad de las moléculas de ácido nucleico que los contienen; y/o (13) segmentos de ácido nucleico que codifican funciones de replicación condicional, por ejemplo, replicación en determinados huéspedes o cepas de células huésped o bajo ciertas condiciones ambientales (por ejemplo, temperatura, condiciones nutricionales y similares).

Un reactivo de ácido nucleico puede ser de cualquier forma útil para la transcripción y/o traducción *in vivo*. Un ácido nucleico a veces es un plásmido, tal como un plásmido superenrollado, a veces es un cromosoma artificial de levadura (por ejemplo, YAC), a veces es un ácido nucleico lineal (por ejemplo, un ácido nucleico lineal producido por PCR o por digestión de restricción), a veces es monocatenario y, a veces, bicatenario. A veces, un reactivo de ácido nucleico se prepara mediante un proceso de amplificación, tal como un proceso de reacción en cadena de la polimerasa (PCR) o un proceso de amplificación mediado por transcripción (TMA). En la TMA, se usan dos enzimas en una reacción isotérmica para producir productos de amplificación detectados por emisión de luz (por ejemplo, Biochemistry, 25 de junio de 1996;35(25):8429-38). Se conocen procesos de PCR estándar (por ejemplo, Patentes de EE.UU. N.º 4.683.202; 4.683.195; 4.965.188; y 5.656.493), y generalmente se realizan en ciclos. Cada ciclo incluye la desnaturalización por calor, en la que se disocian los ácidos nucleicos híbridos; enfriamiento, en el que se hibridan los oligonucleótidos cebadores; y extensión de los oligonucleótidos por una polimerasa (es decir, polimerasa Taq). Un ejemplo de un proceso cíclico de PCR es tratar la muestra a 95 °C durante 5 minutos; repetir cuarenta y cinco ciclos de 95 °C durante 1 minuto, 59 °C durante 1 minuto, 10 segundos y 72 °C durante 1 minuto 30 segundos; y a continuación tratar la muestra a 72 °C durante 5 minutos. Con frecuencia se realizan múltiples ciclos usando un termociclador disponible comercialmente. Los productos de amplificación por PCR a veces se almacenan durante un tiempo a una temperatura inferior (por ejemplo, a 4 °C) y a veces se congelan (por ejemplo, a -20 °C) antes del análisis.

Kits/artículos de fabricación

En el presente documento, en determinadas realizaciones, se desvelan kits y artículos de fabricación para su uso con uno o más métodos descritos en el presente documento. Dichos kits incluyen un portador, envase o recipiente que está compartimentado para recibir uno o más recipientes tales como viales, tubos y similares, comprendiendo cada uno de los recipientes uno de los elementos separados que se usarán en un método descrito en el presente documento. Los recipientes adecuados incluyen, por ejemplo, frascos, viales, jeringas y tubos de ensayo. En una realización, los recipientes se forman a partir de diversos materiales, tales como vidrio o plástico.

En algunas realizaciones, un kit incluye un material de envasado adecuado para albergar el contenido del kit. En

algunos casos, el material de envasado se construye mediante métodos bien conocidos, preferentemente para proporcionar un entorno estéril y libre de contaminantes. Los materiales de envasado empleados en el presente documento pueden incluir, por ejemplo, los que se utilizan habitualmente en kits comerciales vendidos para su uso con sistemas de secuenciación de ácidos nucleicos. Los materiales de envasado de ejemplo incluyen, sin limitación, vidrio, plástico, papel, láminas y similares, capaces de contener dentro de límites fijos un componente expuesto en el presente documento.

El material de envasado puede incluir una etiqueta que indique un uso particular de los componentes. El uso del kit que se indica en la etiqueta puede ser uno o más de los métodos expuestos en el presente documento según sea apropiado para la combinación particular de componentes presentes en el kit. Por ejemplo, una etiqueta puede indicar que el kit es útil para un método de síntesis de un polinucleótido o para un método de determinación de la secuencia de un ácido nucleico.

Las instrucciones para el uso de los reactivos o componentes envasados también se pueden incluir en un kit. Las instrucciones incluirán normalmente una expresión tangible que describa los parámetros de reacción, tales como las cantidades relativas de componentes del kit y muestra a mezclar, períodos de tiempo de mantenimiento para mezclas de reactivo/muestra, temperatura, condiciones de tampón y similares.

Se entenderá que no todos los componentes necesarios para una reacción particular necesitan estar presentes en un kit particular. Más bien, se pueden proporcionar uno o más componentes adicionales de otras fuentes. Las instrucciones proporcionadas con un kit pueden identificar el uno o más componentes adicionales que se proporcionarán y dónde se pueden obtener.

En algunas realizaciones, se proporciona un kit que es útil para incorporar de manera estable un ácido nucleico no natural en un ácido nucleico celular, por ejemplo, usando los métodos proporcionados por la presente invención para preparar células genomanipuladas. En una realización, un kit descrito en el presente documento incluye una célula manipulada genéticamente y uno o más ácidos nucleicos no naturales. En otro aspecto de la divulgación, un kit descrito en el presente documento incluye un plásmido aislado y purificado que comprende una secuencia seleccionada de las SEQ ID NO: 1-9. En un aspecto adicional de la divulgación, un kit descrito en el presente documento incluye un plásmido aislado y purificado que comprende una secuencia de las SEQ ID NO: 2, 3, 5 o 7.

En aspectos adicionales de la divulgación, el kit descrito en el presente documento proporciona una célula y una molécula de ácido nucleico que contiene un gen heterólogo para su introducción en la célula para proporcionar así una célula manipulada genéticamente, tales como vectores de expresión que comprenden el ácido nucleico de cualquiera del aspecto de la divulgación descrita anteriormente en este párrafo.

Terminología específica

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto en la materia a la que pertenece la materia objeto reivindicada. Debe entenderse que la descripción general anterior y la siguiente descripción detallada son solo ilustrativas y explicativas y no son restrictivas de ninguna materia objeto reivindicada. En la presente solicitud, el uso del singular incluye el plural a menos que se indique específicamente lo contrario. Cabe señalar que, como se usa en la memoria descriptiva y en las reivindicaciones adjuntas, las formas en singular "un", "una" y "el/la" incluyen referencias en plural a menos que el contexto dicte claramente lo contrario. En la presente solicitud, el uso de "o" significa "y/o" a menos que se indique de otro modo. Además, el uso de la expresión "que incluye", así como otras formas, tales como "incluyen", "incluye", e "incluido/a", no es limitante.

Como se usa en el presente documento, los intervalos y cantidades pueden expresarse como "aproximadamente" un valor o intervalo concreto. Aproximadamente también incluye la cantidad exacta. Por lo tanto, "aproximadamente 5 µl" significa "aproximadamente 5 µl" y también "5 µl". Generalmente, el término "aproximadamente" incluye una cantidad que pueda esperarse que se encuentre dentro del error experimental.

Los encabezados de sección usados en el presente documento son solo para fines organizativos y no deben interpretarse como una limitación de la materia objeto descrita.

Ejemplos

Estos ejemplos se proporcionan únicamente con propósitos ilustrativos y no limitan el ámbito de las reivindicaciones proporcionadas en el presente documento.

Ejemplo 1

Todos los organismos naturales almacenan información genética en un alfabeto genético de cuatro letras y dos pares de bases. En algunos casos, se generó una forma mutante de *Escherichia coli*, se cultivó en presencia de los nucleósidos trifosfato no naturales dNaMTP y d5SICSTP, y se les proporcionó los medios para importarlos a través

de la expresión de un transportador de nucleósido trifosfato transportado por plásmido, replicas de ADN que contenían un solo UBP dNaM-d5SICS. En algunos casos, el organismo creció mal y no pudo almacenar indefinidamente la información no natural, que es un requisito previo para la verdadera vida semisintética. A continuación se describe un transportador genomanipulado, junto con un UBP químicamente optimizado, para generar un organismo semisintético (SSO, por sus siglas en inglés).

Métodos

A menos que se indique lo contrario, los cultivos bacterianos líquidos se cultivaron en 2x YT (16 g/l de peptona de caseína, 10 g/l de extracto de levadura, 5 g/l de NaCl) complementado con fosfato de potasio (50 mM, pH 7), denominado de aquí en adelante como "medio", y se incubaron a 37 °C en una placa de fondo plano de 48 pocillos (CELLSTAR, Greiner Bio-One) con agitación a 200 rpm. Se preparó medio de crecimiento sólido con agar al 2 %. Se usaron antibióticos, según correspondiera, en las siguientes concentraciones: carbenicilina, 100 µg/ml; estreptomicina, 50 µg/ml; kanamicina, 50 µg/ml; zeocina, 50 µg/ml; cloranfenicol, 33 µg/ml para plásmidos, 5 µg/ml para integrantes cromosómicos. Todos los agentes selectivos se adquirieron comercialmente. El crecimiento celular, indicado como DO₆₀₀, se midió usando un lector multietiqueta Perkin Elmer Envision 2103 con un filtro de 590/20 nm.

A menos que se indique lo contrario, todos los reactivos de biología molecular se obtuvieron en New England Biolabs (NEB) y se usaron según los protocolos del fabricante. Las PCR para la clonación y la construcción de cepas se realizaron con ADN polimerasa Q5. El termociclado se realizó con un termociclador PTC-200 (MJ Research), a excepción de las PCR usadas para generar inserciones Golden Gate que contienen UBP y las PCR usadas en el ensayo de desplazamiento de biotina, que se realizaron con un termociclador en tiempo real CFX-Connect (Bio-Rad) para monitorizar la amplificación del producto con SYBR Green (Thermo Fisher). Cuando fue necesario, los cebadores se fosforilaron usando polinucleótido cinasa T4. Los plásmidos linealizados por PCR se trataron con DpnI para eliminar el molde de plásmido y las ligaduras se realizaron con ADN ligasa T4. Los plásmidos ensamblados por PCR y Golden Gate se purificaron mediante columna giratoria (DNA Clean and Concentrator-5, Zymo Research). Los fragmentos de ADN aislados mediante electroforesis en gel de agarosa se purificaron usando el kit de recuperación de ADN en gel Zymoclean (Zymo Research). Las PCR de colonias se realizaron con ADN polimerasa Taq. Los fragmentos de ADN natural y los plásmidos se cuantificaron mediante A_{260/280} usando un NanoDrop 2000 (Thermo Fisher) o un Infinite M200 Pro (Tecan). Los fragmentos de ADN y los plásmidos que contienen uno o más UBP, que normalmente eran <20 ng/µl, se cuantificaron con el kit de ensayo Qubit dsDNA HS (Thermo Fisher).

Las secuencias de todos los oligonucleótidos de ADN usados en este estudio se proporcionan en la Tabla 2. Los oligonucleótidos naturales se adquirieron en IDT (San Diego, California, EE.UU.) con purificación y desalinización estándar. La síntesis génica de las secuencias de genes *PtNTT2* y GFP de codones optimizados se realizó por GeneArt Gene Synthesis (Thermo Fisher) y GenScript, respectivamente, y se proporcionó amablemente por Synthorx. La secuenciación se realizó por Eton Biosciences (San Diego, California, EE.UU.) o Genewiz (San Diego, California, EE.UU.). Los plásmidos se aislaron usando kits miniprep comerciales (QIAprep, Qiagen o ZR Plasmid Miniprep Classic, Zymo Research).

Se adquirió [α-³²P]-dATP (3000 Ci/mmol, 10 mCi/ml) en PerkinElmer (Shelton, Connecticut, EE.UU.). Se sintetizaron trifosfatos de dNaM, d5SICS, dTPT3 y dMMO2^{bio} como se describe en Li *et al.*, "Natural-like replication of an unnatural base pair for the expansion of the genetic alphabet and biotechnology applications", J. Am. Chem. Soc. 136, 825-829 (2014); o se proporcionaron amablemente por Synthorx (San Diego, California, EE.UU.). El oligonucleótido TK1 que contenía dNaM se describió en Malyshev, *et al.*, "A semi-synthetic organism with an expanded genetic alphabet", Nature, 509, 385-388 (2014). Todos los demás oligonucleótidos no naturales que contenían dNaM se sintetizaron por Biosearch Technologies (Petaluma, California, EE.UU.) con purificación mediante cartucho de fase inversa.

La cepa de *E. coli* C41(DE3) se proporcionó amablemente por J. P. Audia (University of South Alabama, EE.UU.). pKIKOarsBKM fue un regalo de Lars Nielsen y Claudia Vickers (plásmido de Addgene n.º 46766). pRS426 se proporcionó amablemente por Richard Kolodner (University of California San Diego, EE.UU.).

Construcción de los plásmidos *PtNTT2*

La construcción de pCDF-1b-*PtNTT2* se describió por Malyshev, *et al.*, "A semi-synthetic organism with an expanded genetic alphabet", Nature, 509, 385-388 (2014). Para crear pCDF-1b-*PtNTT2*(66-575) se usaron los cebadores fosforilados YZ552 y pCDF-1b-fwd para linealizar pCDF-1b-*PtNTT2* mediante PCR y el producto resultante se ligó intramolecularmente. Los plásmidos de clones individuales se aislaron y se confirmaron secuenciando el gen *PtNTT2* usando los cebadores T7 seq y T7 term seq.

Para crear los plásmidos pSC-P(_{lacI}, bla, lac, lacUV5)*PtNTT2*(66-575)-T₀, se usaron los cebadores fosforilados YZ581 y YZ576 para amplificar el gen *PtNTT2*(66-575) y su secuencia de unión ribosomal correspondiente y terminador, de una versión de pCDF-1b-*PtNTT2*(66-575) que reemplaza el terminador T7 por un terminador λ T₀. Esta inserción se ligó al plásmido pHSG576 linealizado con los cebadores DM002 e YZ580. Se verificó un único clon del plásmido pSC-*PtNTT2*(66-575)-T₀ resultante mediante la secuenciación del gen *PtNTT2* usando los cebadores DM052 e YZ50. A continuación, se linealizó pSC-*PtNTT2*(66-575)-T₀ con los cebadores YZ580 e YZ581 y se ligó a un dúplex de cebador fosforilado

correspondiente al promotor P_{lac} , P_{bla} , P_{lac} o P_{lacUV5} (YZ584/YZ585, YZ582/YZ583, YZ599/YZ600 y YZ595/YZ596, respectivamente) para producir los plásmidos pSC- $P_{(lac, bla, lac, lacUV5)}$ PtNTT2(66-575)- T_0 . La orientación correcta del promotor y las secuencias del gen del promotor se confirmaron de nuevo mediante secuenciación usando los cebadores DM052 e YZ50. Se generó pSC- P_{bla} PtNTT2(66-575 co)- T_0 de manera análoga a pSC- P_{bla} PtNTT2(66-575)- T_0 usando una versión de terminador λ T_0 de pCDF-1b-PtNTT2(66-575) que contenía una secuencia de PtNTT2 de codones optimizados (véase la Tabla 2).

Construcción de las cepas de PtNTT2 y Cas9

El casete de expresión PtNTT2(66-575) y su marcador de resistencia al cloranfenicol Cm^R en los plásmidos pSC tiene ~2,8 kb, un tamaño que es prohibitivo para la integración cromosómica con los pequeños tramos de homología (~50 pb) que se pueden introducir mediante cebadores durante la PCR, como se hace tradicionalmente en la recombinación. Se usó la recombinación homóloga en *S. cerevisiae* para construir una serie de plásmidos molde de integración con el casete de expresión PtNTT2(66-575) y el Cm^R flanqueado por ~1 kb de la secuencia 5' con respecto a *lacZ* y ~1 kb de secuencia 3' con respecto a 246 pb cadena abajo del codón iniciador *lacA*. El locus *lacZYA* se eligió para que la integración del transportador también desactivará la lactosa permeasa *lacY*, creando así una cepa BL21 (DE3) que permite la entrada celular uniforme de IPTG y, por lo tanto, la inducción homogénea y finamente valorable de promotores que contienen operadores *lac*.

Para crear los plásmidos molde de integración, pRS426 se digirió con PvuI-HF y el fragmento de plásmido de 3810 pb resultante se aisló mediante electroforesis y purificación en gel de agarosa. A continuación, se repararon los huecos de este fragmento en la cepa BY4741 de *S. cerevisiae* a través de la transformación química mediada por acetato de litio del fragmento plasmídico y los productos de PCR de las siguientes combinaciones de cebador/cebador/molde: YZ7/YZ12/pBR322, YZ613/YZ580/ADN genómico de *E. coli*, YZ614/615/ADN genómico de *E. coli* y DM052/YZ612/pSC- P_{lacUV5} (66-575)- T_0 . El plásmido resultante, 426.*lac*-ZYA:: P_{lacUV5} PtNTT2(66-575)- T_0 Cm^R , se aisló (Zymoprep Yeast Plasmid Miniprep, Zymo Research), se digirió con PvuI-HF y XbaI (para reducir la estirpe durante la integración, ya que el plásmido lanzadera pRS426 también contiene un origen pMB1 de *E. coli*) y se usó como molde para generar un fragmento de integración lineal mediante PCR con los cebadores YZ616 e YZ617. La integración de este fragmento en BL21(DE3) para generar la cepa YZ2 se realizó usando pKD46 como se describe en Datsenko, *et al.*, "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products", PNAS 97, 6640-6645 (2000). Los integrantes se confirmaron mediante PCR de colonias de las uniones 5' y 3' usando los cebadores YZ618 y YZ587 (producto de 1601 pb) y YZ69 e YZ619 (producto de 1402 pb), respectivamente, detección de la delección de *lacZ* mediante crecimiento en placas que contenían X-gal (80 µg/ml) e IPTG (100 µM), y PCR y secuenciación del transportador con los cebadores DM053 e YZ50.

El plásmido 426.*lac*ZYA:: Cm^R se generó a partir de la linealización de 426.*lac*ZYA:: P_{lacUV5} PtNTT2(66-575)- T_0 Cm^R usando los cebadores fosforilados YZ580 y pCDF-1b-rev, y la posterior ligadura intramolecular. A continuación, se integró 426.*lac*ZYA:: Cm^R en BL21(DE3) para crear una cepa de control isogénica sin transportador para los ensayos de captación de dATP.

Para crear los plásmidos 426.*lac*ZYA:: $P_{(bla, lac, lacUV5)}$ PtNTT2(66-575 co)- T_0 Cm^R , el plásmido 426.*lac*-ZYA::PtNTT2(66-575)- T_0 Cm^R (un plásmido sin promotor generado de forma análoga a 426.*lac*-ZYA:: P_{lacUV5} PtNTT2(66-575)- T_0 Cm^R usando pSC-PtNTT2(66-575)- T_0), denominado en lo sucesivo 426.trunc, se digirió con PvuI-HF y AvrII, y el fragmento plasmídico de 5969 pb resultante se aisló mediante electroforesis y purificación en gel de agarosa, y se reparó el hueco usando productos de PCR de las siguientes combinaciones de cebador/cebador/molde: **1**, YZ12/YZ580/426.trunc; **2**, DM053/YZ610/pSC- $P_{(bla, lac, lacUV5)}$ PtNTT2(66-575)- T_0 ; **3**, YZ581/YZ50/pSC- P_{bla} PtNTT2(66-575 co)- T_0 . Los plásmidos 426.*lac*ZYA:: $P_{(tac, N25, \lambda, H207)}$ PtNTT2(66-575 co)- T_0 Cm^R se generaron de forma análoga con la excepción de que los fragmentos **2** se reemplazaron por fragmentos correspondientes a los promotores P_{lac} , P_{N25} , P_{λ} y P_{H207} , que se generaron por hibridación y extensión de los pares de cebadores YZ703/YZ704, YZ707/YZ708, YZ709/YZ710 y YZ711/YZ712, respectivamente, con el fragmento de Klenow. A continuación, los plásmidos 426.*lac*ZYA:: $P_{(bla, lac, lacUV5, tac, N25, \lambda, H207)}$ PtNTT2(66-575 co)- T_0 Cm^R se usaron para integrar el transportador en BL21(DE3) usando los cebadores YZ616 y YZ617, y la recombinación, como se ha descrito anteriormente. La cepa YZ3 representa BL21(DE3) integrada con *lac*ZYA:: P_{lacUV5} PtNTT2(66-575 co)- T_0 Cm^R .

Para crear la cepa YZ4, el fragmento de 4362 pb de pCas9-Multi digerido con SpeI y AvrII se ligó en pKIKOarsBKM digerido con SpeI y el plásmido resultante, pKIKOarsB:: P_{lacO} -Cas9- T_{rMB} Km^R , se usó como molde para generar un fragmento de integración lineal a través de PCR con los cebadores YZ720 e YZ721. A continuación, el fragmento se integró en BL21(DE3) como se ha descrito anteriormente y se confirmó mediante PCR de colonias con los cebadores YZ720 e YZ721 y la secuenciación del producto con los cebadores TG1-TG6. P_{lacUV5} PtNTT2(66-575 co)- T_0 Cm^R se integró posteriormente en su cepa, como se ha descrito anteriormente, para generar la cepa YZ4.

Ensayo de captación de dATP

Los ensayos de captación radiactiva se realizaron como se describe en Haferkamp, *et al.*, "Tapping the nucleotide pool of the host: novel nucleotide carrier proteins of Protochlamydia amoebophila", Mol. Microbiol. 60, 1534-1545 (2006) con las siguientes modificaciones: las cepas C41(DE3) y BL21(DE3) que llevan transportadores basados en plásmidos

y sus controles plasmídicos vacíos apropiados, así como integrantes del transportador cromosómico BL21(DE3) y su transportador isogénico apropiado sin control, se cultivaron durante una noche con los antibióticos apropiados (estreptomycin para plásmidos pCDF y cloranfenicol para plásmidos pSC e integrantes) en 500 µl de medio. Los cultivos se diluyeron a una DO_{600} de 0,02 en 500 µl de medio fresco, se cultivaron durante 2,5 h, se indujeron con IPTG (0-1 mM, solo cepas pCDF) o se cultivaron (todas las demás cepas) durante 1 h, y se incubaron con dATP enriquecido con $[\alpha\text{-}^{32}\text{P}]\text{-dATP}$ (concentración final = 250 µM (0,5 µCi/ml)) durante ~1 h. Este esquema experimental es análogo al protocolo utilizado para preparar las células para la transformación con plásmidos que contienen UBP, simulando 1 h de incubación de dATP 1 h de recuperación en presencia de trifosfatos no naturales después de la electroporación. Se cultivó en paralelo una placa duplicada de 48 pocillos sin $[\alpha\text{-}^{32}\text{P}]\text{-dATP}$ para controlar el crecimiento.

Después de la incubación con dATP, se recogieron 200 µl de cada cultivo a través de una placa de filtro de fibra de vidrio de 0,65 µm de 96 pocillos (MultiScreen, EMD Millipore) al vacío y se lavó con fosfato de potasio frío (3 x 200 µl, 50 mM, pH 7) y ddH₂O fría (1 x 200 µl). Los filtros se eliminaron de la placa y se expusieron durante una noche a un tamiz de fósforo de almacenamiento (BAS-IP MS, GE Healthcare Life Sciences), del que posteriormente se tomaron imágenes con un escáner láser de superficie plana (Typhoon 9410, GE Healthcare Life Sciences). La imagen resultante se cuantificó mediante análisis densitométrico usando Image Studio Lite (LI-COR). Las intensidades de las imágenes sin procesar de cada muestra se normalizaron con respecto al período de tiempo y la DO_{600} promedio durante la incubación con dATP (es decir, se normalizaron según una estimación del área bajo la curva de crecimiento correspondiente a la ventana de captación), seguido de la resta de las señales normalizadas de los controles sin transportador negativos apropiados.

Los tiempos de duplicación para las cepas cultivadas en el ensayo de captación de dATP se calcularon duplicando el tiempo como $(t_2 - t_1) / \log_2(DO_{600,2} / DO_{600,1})$, con un promedio de tres intervalos de tiempo de -30 min correspondientes aproximadamente a 30 min antes de la captación de dATP y 60 min durante la captación de dATP.

Construcción de plásmidos de destino Golden Gate para pUCX1, pUCX2 y pBRX2

Aunque el UBP se clonó en plásmidos a través de la clonación por extensión con polimerasa circular (CPEC)^{4,27}, el método da como resultado un plásmido con doble muesca que no se puede tratar con exonucleasa T5 para degradar el plásmido lineal no incorporado y las inserciones y, por lo tanto, hace difícil cuantificar con precisión el rendimiento de la reacción de clonación y controlar la cantidad de plásmido de entrada usado para transformar células durante un experimento de replicación *in vivo*. Además, el plásmido lineal no incorporado y las inserciones de una reacción CPEC también pueden moldear reacciones de PCR con los cebadores usados en el ensayo de desplazamiento de biotina y, por lo tanto, los ensayos de desplazamiento de biotina en productos CPEC no reflejan realmente el contenido de UBP de los plásmidos que en realidad se transforman en células. Para eludir estas complicaciones, el UBP se incorporó en plásmidos usando el ensamblaje Golden Gate.

Para crear pUCX1 GG y pUCX2 GG, los plásmidos de destino Golden Gate para pUCX1 y pUCX2, respectivamente, pUC19 se linealizó con cebadores fosforilados pUC19-lin-fwd y pUC19-lin-rev, y el producto resultante se ligó intramolecularmente para eliminar la secuencia TK1 natural de 75 nt. A continuación, el plásmido resultante se linealizó con los cebadores fosforilados YZ51 e YZ52, y el producto resultante se ligó intramolecularmente para mutar el sitio de reconocimiento Bsal en el marcador de resistencia a la ampicilina Amp^R. A continuación, este plásmido se linealizó con los cebadores pUC19-lin-fwd y pUC19-lin-rev (para pUCX1), o los cebadores YZ95 y YZ96 (para pUCX2), y se ligó a una inserción generada a partir de la PCR con los cebadores fosforilados YZ93 e YZ94 y el molde pCas9-Multi, para introducir dos sitios de reconocimiento Bsal (para la clonación por el ensamblaje Golden Gate) y un marcador de resistencia a la zeocina (un casete de relleno usado para diferenciar entre plásmidos con o sin inserción) en pUC19.

Para crear pBRX2 GG, el plásmido de destino Golden Gate para pBRX2, el fragmento de 2934 pb de pBR322 digerido con Aval y EcoRI-HF se rellenó terminalmente con el fragmento de Klenow y se ligó intramolecularmente para eliminar el casete de resistencia a tetraciclina. El sitio de reconocimiento Bsal dentro de Amp^R se mutó como se ha descrito anteriormente. A continuación, el plásmido se linealizó con los cebadores YZ95 e YZ96 y se ligó al casete Bsal-zeo^R-Bsal como se ha descrito anteriormente. Por lo tanto, pBRX2 es un análogo de menor número de copias de pUCX2.

Ensamblaje Golden Gate de plásmidos que contienen UBP

Se generaron plásmidos que contenían uno o más UBP mediante ensamblaje Golden Gate. Las inserciones que contenían el UBP se generaron mediante PCR de oligonucleótidos sintetizados químicamente que contenían dNaM, usando dTPT3TP y dNaMTP, y cebadores que introducen sitios de reconocimiento Bsal terminales que, cuando se digieren, producen salientes compatibles con un plásmido de destino apropiado; véase la Tabla 2 para obtener una lista completa de cebadores, moldes y sus correspondientes plásmidos de destino Golden Gate. Los oligonucleótidos molde (0,025 ng por 50 µl de reacción) se amplificaron por PCR usando concentraciones de reactivo y equipo en las siguientes condiciones de termociclado (tiempos indicados como mm:ss): [96 °C 1:00 120 x [96 °C 0:15 | 60 °C 0:15 | 68 °C 4:00]].

Para ensamblar los plásmidos que contenían UBP, el plásmido de destino (200-400 ng), una o más inserciones de

PCR (relación molar de inserción:plásmido de 3:1), la ADN ligasa T4 (200 U), Bsal-HF (20 U) y ATP (1 mM) se combinaron en tampón 1x NEB CutSmart (volumen final 30 µl) y se sometieron a termociclado en las siguientes condiciones: [37 °C 20:00 | 40 x [37 °C 5:00 | 16 °C 5:00 | 22 °C 2:30] 37 °C 20:00 | 55 °C 15:00 | 80 °C 30:00]. Después de la reacción Golden Gate, se añadieron exonucleasa T5 (10 U) y más cantidad de Bsal-HF (20 U), y la reacción se incubó (37 °C, 1 h) para digerir el plásmido no incorporado y los fragmentos de inserción. Los plásmidos ensamblados se cuantificaron por Qubit.

Construcción de los plásmidos de destino Golden Gate para pCas9 y pAIO

Para crear pCas9-Multi, el plásmido de destino Golden Gate para la clonación de casetes de ARN_g junto con Cas9, pDZAZ²⁹ y un gen Cas9 amplificado por PCR (Cebadores JL126 y JL128, plásmido molde de Addgene n.º 41815) se digirieron con KpnI y XbaI, y se ligaron para crear pCas9(-). Este plásmido y un casete de GFPT2-ARN_g amplificado por PCR (plásmido molde de Addgene n.º 41820, que contiene la secuencia de ARN_g; el promotor ProK y el terminador se introdujeron mediante PCR) se digirieron con Sall y se ligaron para crear pCas9-GFPT2. A continuación, este plásmido se linealizó con los cebadores BL557 y BL558 (para eliminar los sitios de reconocimiento BsmBI dentro de Cas9) y se circularizó a través del ensamblaje Gibson. A continuación, el plásmido resultante se linealizó con los cebadores BL559 y BL560 (para reintroducir dos sitios BsmBI en la cadena principal del plásmido) y se circularizó a través de ensamblaje Gibson para producir pCas9-Multi, que se confirmó mediante secuenciación con los cebadores TG1-TG6. La digestión de pCas9-Multi con BsmBI da como resultado un plásmido linealizado con salientes que permiten la clonación simultánea de uno o más ARN_g por ensamblaje Golden Gate (véase la sección a continuación).

Para crear pAIO-Multi, pCas9-Multi se linealizó con los cebadores BL731 y BL732 (para eliminar Cas9 e introducir sitios de reconocimiento Bsal para la clonación de UBP), se fosforiló, se ligó intramolecularmente y se confirmó mediante secuenciación con el cebador BL450. La digestión de pAIO-Multi con Bsal da como resultado un plásmido linealizado con salientes idénticos a los producidos por la digestión con Bsal del plásmido de destino pUCX2 y, por lo tanto, las inserciones generadas por PCR para clonar el UBP en pUCX2 también se pueden usar para clonar UBP en pAIO-Multi y sus derivados. Después de la clonación de los casetes de ARN_g en pAIO-Multi (véase la siguiente sección a continuación), el protocolo de ensamblaje Golden Gate para la clonación en un UBP fue idéntico al descrito anteriormente para pUCX2, con la excepción de que se usó el producto de pAIO-Multi (con ARN_g) amplificado con BL731 y BL732 en lugar del propio plásmido.

Clonación de ARN_g en pCas9 y pAIO

Se clonaron casetes de ARN_g duales en pCas9-Multi o pAIO-Multi a través del ensamblaje Golden Gate. Para generar el primer casete de ARN_g de cada par, pCas9-GFPT2 (1 ng) se amplificó por PCR con los cebadores 1^{er} ARN_g GG (200 nM) y BL562 (200 nM), y la ADN polimerasa OneTaq, en las siguientes condiciones de termociclado: [30 x [94 °C 0:30 | 152 °C 0:15 | 68 °C 0:30]]. Los productos de PCR se purificaron mediante electroforesis y purificación en gel de agarosa. El cebador 1^{er} ARN_g GG es un cebador de 70 nt que posee (de 5' a 3') un sitio de restricción BsmBI, 10 nt de homología con el promotor ProK, una guía variable (espaciador) de 18 nt complementaria a una mutación de UBP y 25 nt de homología con el armazón de ARN_g no variable. Para generar el segundo casete de ARN_g, pCas9-GFPT2 se amplificó por PCR con los cebadores BL563 y 2^o ARN_g Rev, y los cebadores BL566 y 2^o ARN_g Fwd, y los dos productos resultantes se combinaron y se amplificaron mediante PCR de extensión superpuesta usando los cebadores BL563 y BL566, seguido de electroforesis y purificación en gel de agarosa.

Para ensamblar los plásmidos guía, pCas9-Multi (40 ng) o pAIO-Multi (20 ng), el ADN purificado del primer casete de ARN_g (4,5 ng) y del segundo casete de ARN_g (8 ng), la ADN ligasa T4 (200 U), BsmBI (5 U) y ATP (1 mM) se combinaron en tampón de reacción 1x NEB CutSmart (volumen final 20 µl) y se sometieron a termociclado en las siguientes condiciones: [5 x [37 °C 6:00 | 16 °C 8:00] 15 x [55 °C 6:00 | 16 °C 8:00]]. Los plásmidos ensamblados se transformaron en células electrocompetentes para la posterior secuenciación y prueba.

Para ensamblar pCas9-TK1-A, un plásmido que contiene solo un casete de ARN_g, pCas9-GFPT2 se amplificó con los cebadores BL566 y BL567, y el producto resultante se ligó en pCas9-Multi mediante ensamblaje Golden Gate como se ha descrito anteriormente.

Para ensamblar pCas9-hEGFP, un plásmido que contiene un casete de ARN_g no diana para experimentos de TK1, los cebadores BL514 y BL515 se hibridaron y se ligaron, mediante ensamblaje Gibson, en pCas9-GFPT2 linealizado con los cebadores BL464 y BL465.

Construcción de pAIO2X

pAIO2X GG, el plásmido de destino Golden Gate para pAIO2X, se obtiene de tres plásmidos, usando inserciones generadas por PCR y múltiples etapas de clonación mediante ligadura y digestión con enzimas de restricción. Las inserciones de pSYN36, que contiene una *superfolder* con codones optimizados *gfp* con un sitio de entrada Golden Gate para la clonación en secuencias que corresponden a los nucleótidos 409-483 de *gfp*, y pET-22b-ESerGG, que contiene un gen *serT* de *E. coli* con un sitio de entrada Golden Gate para la clonación en secuencias que corresponden a los nucleótidos 10-65 de *serT*, se clonaron en la guía dual de pAIO BsmBI, una versión de pAIO-Multi que contiene

dos casetes de ARN_g, estando las secuencias guía de direccionamiento (espaciador) reemplazadas por dos pares ortogonales de sitios de reconocimiento BsmBI que permiten guiar la clonación usando dúplex de cebadores recocidos.

Para crear pAIO2X-GFP151/Eser-69 GG, los dúplex de cebadores hibridados de YZ310/YZ316 y YZ359/YZ360 se ligaron en pAIO2X GG usando los mismos reactivos del ensamblaje Golden Gate y las mismas condiciones de termociclado usadas para la clonación de UBP, con la excepción de que BsaI se reemplazó por BsmBI, cada dúplex de cebador se usó en una relación molar de inserción:plásmido de 50:1 con 30 fmol de plásmido de destino, y la reacción se escaló en un tercio a 10 µl. Después del ensamblaje, la reacción no se digirió con enzimas adicionales ni se purificó, y se transformó directamente en DH5α de *E. coli* químicamente competente. Después del aislamiento de clones de plásmidos individuales y la confirmación de las guías mediante secuenciación con el cebador BL450, los UBP se clonaron en el plásmido mediante ensamblaje Golden Gate con BsaI, como se describe en la sección, Ensamblaje Golden Gate de plásmidos que contienen UBP.

Ensayo de escisión de Cas9 *in vitro*

Para generar los sustratos de ADN para los ensayos de escisión de Cas9 *in vitro*, los moldes BL408, BL409, BL410, BL487, BL488 y BL489 (1 ng por 50 µl de reacción) se amplificaron mediante PCR con los cebadores BL415 (400 nM) y BL416 (400 nM), y ADN polimerasa OneTaq en tampón de reacción estándar OneTaq 1x complementado con dNaMTP (100 µM), dTPT3TP (100 µM) y MgCl₂ (1,5 mM), en las siguientes condiciones de termociclado: [25 x [95 °C 0:15 | 56 °C 0:15 | 68 °C 1:30]].

Para generar los moldes de ADN para la transcripción *in vitro* de ARN_g, los moldes BL318, BL484, BL485 y BL486 (1 ng por 50 µl de reacción), que contienen el promotor T7 y una secuencia espaciadora de ARN CRISPR (ARN_{cr}), se amplificaron por PCR con los cebadores BL472 (200 nM) y BL473 (200 nM), y ADN polimerasa OneTaq en tampón de reacción estándar OneTaq 1x complementado con MgCl₂ (6 mM), en las siguientes condiciones de termociclado: [20 x [95 °C 0:15 | 60 °C 0:15 | 68 °C 1:30]]. A continuación, el ADN de esta primera reacción de PCR (0,5 µl) se transfirió a una segunda reacción de PCR (100 µl) que contenía los cebadores BL472 (400 nM), BL439 (500 nM) y BL440 (600 nM) y se termocicló en las siguientes condiciones: [4 x [95 °C 0:15 | 68 °C 0:15 | 68 °C 1:30] 20 x [95 °C 0:15 | 60 °C 0:15 | 68 °C 1:30]]. Se realizó la transcripción *in vitro* de los productos de PCR con ARN polimerasa T7 y los ARN_g transcritos se purificaron mediante PAGE, escisión de banda y extracción (37 °C, durante una noche) en una solución acuosa de NaCl (200 mM) y EDTA (1 mM, pH 7), seguido de concentración y purificación por precipitación con etanol.

Para las reacciones de escisión *in vitro*, la nucleasa Cas9 (125 nM) se incubó con cada ARN_g transcrito (125 nM) en tampón de reacción de nucleasa Cas9 1x durante 5 min, a continuación se añadió sustrato de ADN y la reacción se incubó (37 °C, 10 minutos). La reacción se interrumpió con tampón de carga SDS-PAGE (Tris-HCl 62 mM, SDS al 2,5 %, azul de bromofenol al 0,002 %, β-mercaptoetanol 0,7 M y glicerol al 10 %), se desnaturalizó con calor (95 °C, 10 min) y a continuación se cargó en un gel SDS-PAGE. Las bandas de escisión resultantes se cuantificaron mediante análisis densitométrico usando ImageJ³¹. Para cada ARN_g, las eficiencias de escisión sin procesar se dividieron por la escisión máxima observada para ese ARN_g en el conjunto de los seis sustratos de ADN, para tener en cuenta las diferencias en la actividad del ARN_g y/o variaciones menores en la preparación. Los experimentos se realizaron por triplicado técnico y los promedios representan un promedio de tres reacciones de escisión *in vitro* realizadas en paralelo.

Experimentos de replicación de plásmidos *in vivo*

Se prepararon células electrocompetentes YZ3 mediante crecimiento durante una noche en ~5 ml de medio complementado con cloranfenicol, dilución a una DO₆₀₀ de 0,02 en el mismo medio (volúmenes variables, ~10 ml de medio por transformación) y crecimiento a una DO₆₀₀ de ~0,3-0,4. A continuación, las células se enfriaron rápidamente en un baño de agua helada con agitación, se sedimentaron (2500 x g, 10 min) y se lavaron dos veces con un volumen de cultivo de ddH₂O enfriada con hielo. A continuación, las células electrocompetentes se resuspendieron en ddH₂O enfriada con hielo (50 µl por transformación), se mezclaron con un plásmido con ensamblaje Golden Gate (~1 µl, ~1 ng) que contenía el UBP y se transfirieron a una cubeta de electroporación de 0,2 cm de separación previamente enfriada. Las células se sometieron a electroporación (Gene Pulser II, Bio-Rad) según las recomendaciones del fabricante (voltaje 25 kV, condensador 2,5 µF, resistencia 200 Ω) y a continuación se diluyeron inmediatamente con 950 µl de medio precalentado complementado con cloranfenicol. A continuación, una alícuota (10-40 µl) de esta dilución se diluyó inmediatamente 5 veces con el mismo medio precalentado, pero se complementó adicionalmente con dNaMTP (250 µM) y d5SICSTP (250 µM). Las muestras se incubaron (37 °C, 1 h) y a continuación se usó ~15 % de la muestra para inocular medio (volumen final 250-300 µl) complementado con cloranfenicol, carbenicilina, dNaMTP (250 µM) y d5SICSTP (250 µM). A continuación, se controló el crecimiento de las células, se recogieron a la densidad (DO₆₀₀) indicada en el texto principal y se sometieron a aislamiento plasmídico. Las diluciones de la mezcla de recuperación también se extendieron sobre medios sólidos con cloranfenicol y carbenicilina para determinar las eficiencias de transformación. Los experimentos con dNaMTP (150 µM) y dTPT3TP (37,5 µM) se realizaron de manera análoga.

Los experimentos con DM1 se realizaron de manera análoga usando medio complementado con estreptomicina, con

la etapa adicional de inducir la expresión del transportador con IPTG (1 mM, 1 h) antes de sedimentar las células. Todo el medio después de la preparación de células electrocompetentes también se complementó con estreptomycin y IPTG (1 mM) para mantener la expresión del transportador.

5 Experimentos de replicación plasmídica *in vivo* con Cas9 (solo cultivo líquido)

Se transformaron células electrocompetentes YZ2 con diversos plásmidos guía pCas9 y se usaron clones individuales para inocular cultivos durante una noche. A continuación, las células se cultivaron, se prepararon y se sometieron a electroporación como se ha descrito anteriormente para YZ3, con las siguientes modificaciones: todo el medio se complementó adicionalmente con zeocina (para seleccionar pCas9) y glucosa al 0,2 %, las células electrocompetentes se almacenaron en DMSO al 10 % (v/v ddH₂O) a -80 °C hasta su uso, y los medios de crecimiento y recuperación se complementaron con dNaMTP (250 µM) y dTPT3TP (75 µM). Se añadieron concentraciones variables de IPTG (0-100 µM) al medio de cultivo (pero no al medio de recuperación) para inducir la expresión de Cas9. Los ARN_g correspondientes a la secuencia d(AXT) son las guías no diana para todas las secuencias excepto para la propia secuencia que contiene d(AXT), cuyas guías no diana corresponden a la secuencia d(GXT), y todos los experimentos con los ARN_g no diana se realizaron con la adición de IPTG (10 µM) al medio de cultivo. Para los experimentos de crecimiento y recambio, las células se cultivaron a una DO₆₀₀ de 3,5-4,0, a continuación se diluyeron 1:250 y se volvieron a cultivar a una DO₆₀₀ de 3,5-4,0, después de lo cual se aislaron los plásmidos.

20 Experimentos de replicación de plásmidos *in vivo* con Cas9 (siembra en placas y cultivo líquido)

Las células YZ4 electrocompetentes se cultivaron, se prepararon y se sometieron a electroporación como se ha descrito anteriormente para YZ2, con las siguientes modificaciones: el medio para el cultivo de células antes de la electroporación solo contenía cloranfenicol (es decir, sin zeocina), se usó zeocina para seleccionar pAIO (es decir, sin carbenicilina), y los medios de recuperación y crecimiento se complementaron con dNaMTP (150 µM) y dTPT3TP (37,5 µM). Después de la transformación con pAIO, las diluciones de la mezcla de recuperación se extendieron sobre medios sólidos que contenían cloranfenicol, zeocina, dNaMTP (150 µM), dTPT3TP (37,5 µM), glucosa al 0,2 % y diversas concentraciones de IPTG (0-50 µM). Después del crecimiento durante una noche (37 °C, -14 h), se usaron colonias individuales para inocular medio líquido de la misma composición que el medio sólido. Los experimentos realizados con pAIO2X se llevaron a cabo como se ha descrito anteriormente para YZ4 sin usar glucosa o células electrocompetentes congeladas. La segunda siembra en placas representada en la figura 4 se realizó extendiendo las células del cultivo líquido en medio sólido de la misma composición que el medio líquido y cultivo a 37 °C (~14 h). Se seleccionaron seis colonias al azar para continuar la propagación en el cultivo líquido.

35 Cálculo de duplicación celular

Las duplicaciones celulares para experimentos de crecimiento-dilución-recambio de cultivo líquido se calcularon mediante log₂ del factor de dilución (30.000 o 300.000) entre crecimientos, excepto para los crecimientos inoculados a partir de una colonia en placa, cuyas duplicaciones celulares se calcularon promediando, para cada clon individual, el tiempo desde la inoculación a la DO₆₀₀ diana (9,4 ± 1,1 h (1 D.E.) para la primera inoculación en placa, 10,2 ± 3 h para la segunda inoculación en placa) y dividiendo estos promedios por un tiempo de duplicación estimado de 40 min. Los tiempos de crecimiento variaron para cada clon porque las colonias se aislaron cuando apenas eran visibles a simple vista y, por lo tanto, no se intentó controlar la variabilidad en el número de células inoculadas en los cultivos líquidos. Cabe apreciar que las duplicaciones celulares notificadas eran solo una estimación de las duplicaciones en cultivo líquido, lo que subestimaba el número total de duplicaciones celulares, ya que no se intentó estimar el número de duplicaciones celulares que se produjeron durante cada uno de los crecimientos en medio sólido.

Ensayo de desplazamiento de biotina

La retención del uno o más UBP en plásmidos aislados se determinó y se validó de la siguiente manera: minipreparaciones de plásmidos o plásmidos con ensamblaje Golden Gate (0,5 l, 0,5-5 ng/µl) u oligonucleótidos que con tenían dNaM (0,5 fmol) se amplificaron por PCR con dNTP (400 µM), 1x SYBR Green, MgSO₄ (2,2 mM), cebadores (10 nM cada uno), d5SICSTP (65 µM), dMMO2^{Bio}TP (65 µM), ADN polimerasa OneTaq (0,018 U/µl) y ADN polimerasa DeepVent (0,007 U/µl) en tampón de reacción estándar OneTaq 1x (volumen final 15 µl), en las siguientes condiciones de termociclado: [20 x [95 °C 0:15 | x°C 0:15 | 68 °C 4:00]]; véase la Tabla 2 para obtener una lista de cebadores y sus correspondientes temperaturas de hibridación usadas en este ensayo. Tras la amplificación, se mezcló 1 µl de cada reacción con estreptavidina (2,5 µl, 2 µg/µl, Promega) y se incubó brevemente a 37 °C. Después de la incubación, las muestras se mezclaron con tampón de carga y se analizaron en un gel TBE de poliacrilamida al 6 % (acrilamida:bis-acrilamida, 29:1), a 120 V durante -30 min. A continuación, los geles se tiñeron con colorante 1x SYBR Gold (Thermo Fisher) y se tomaron imágenes usando un Molecular Imager Gel Doc XR+ (Bio-Rad) equipado con un filtro 520DF30 (Bio-Rad).

Cálculo de la retención de UBP

La retención de UBP se evaluó mediante análisis densitométrico de los geles (ImageJ o Image Studio Lite, LICOR) del ensayo de desplazamiento de biotina y el cálculo de un porcentaje de desplazamiento sin procesar, que equivale a la

intensidad de la banda desplazada por estreptavidina dividida por la suma de las intensidades de las bandas desplazadas y no desplazadas. Véase la figura 7 para geles representativos. Las retenciones de UBP notificadas son valores normalizados.

- 5 A menos que se indique lo contrario, para los experimentos que no involucran el cultivo en placas en medio sólido, la retención de UBP se normalizó dividiendo el porcentaje de desplazamiento sin procesar de cada muestra plasmídica propagada por el porcentaje de desplazamiento sin procesar del plásmido de entrada con ensamblaje Golden Gate. Se supuso que el contenido de UBP inicial de la población de plásmidos celulares era equivalente al contenido de UBP del plásmido de entrada, basándose en la inoculación directa de la transformación en cultivo líquido. Por lo tanto,
- 10 en estos experimentos, la retención de UBP normalizada era un valor relativo que relacionaba el contenido de UBP de la población de plásmidos propagados con el contenido de UBP de la población inicial, que no era del 100 % debido a la pérdida durante la PCR usada para generar la inserción para el ensamblaje de plásmidos de entrada (figura 7).

- 15 Para los experimentos que implican el cultivo en placas en medio sólido, la retención de UBP se normalizó dividiendo el porcentaje de desplazamiento sin procesar de cada muestra plasmídica propagada por el porcentaje de desplazamiento sin procesar del molde de oligonucleótido que contenía **dNaM** usado en el ensamblaje del plásmido de entrada. El cultivo en placas permitió el aislamiento clonal de plásmidos que contenían UBP a partir de plásmidos completamente naturales que surgieron durante la construcción plasmídica (algunos de los cuales pueden contener secuencias que no fueron reconocidas por el uno o más ARN_g empleados). Debido a que no hubo pérdida de UBP
- 20 mediada por PCR en el molde de oligonucleótidos, la normalización con respecto al molde de oligonucleótidos fue un mejor indicador de la retención absoluta de UBP que la normalización con respecto al plásmido de entrada. En las condiciones usadas en el ensayo de desplazamiento de biotina, la mayoría de los moldes de oligonucleótidos y los contextos de secuencia dieron >90 % de desplazamiento sin procesar, con un desplazamiento de <2 % para un molde completamente natural análogo (es decir, la incorporación errónea de UBP durante el ensayo de desplazamiento de biotina fue insignificante).
- 25

- El cultivo en placas permitió la diferenciación entre la pérdida de UBP que se produjo *in vivo* de la pérdida que se produjo *in vitro*, con la excepción de las muestras derivadas de clones que dieron un desplazamiento de <2 %, para las que no se pudo diferenciar entre si el UBP se perdió completamente *in vivo* o si la muestra provino de un transformante que originalmente recibió un plásmido completamente natural. Dichas muestras se excluyeron de los valores promedio informados cuando otras muestras de la misma transformación dieron cambios más altos.
- 30

Agotamiento por desplazamiento de biotina y análisis de mutación *in vivo*

- 35 Para determinar el espectro mutacional del UBP en muestras plasmídicas aisladas, se realizaron ensayos de desplazamiento de biotina como se ha descrito anteriormente. Las bandas no desplazadas, que correspondían a mutaciones naturales de las secuencias que contenían UBP, se escindieron y se extrajeron (37 °C, durante una noche) en una cantidad mínima de una solución acuosa de NaCl (200 mM) y EDTA (1 mM, pH 7), seguido de concentración y purificación por precipitación con etanol. Una muestra de extracto (1 µl) se amplificó por PCR en condiciones estándar (solo dNTP naturales), con ADN polimerasa OneTaq y los mismos cebadores usados para la PCR de desplazamiento de biotina, y los productos resultantes se secuenciaron mediante secuenciación de Sanger.
- 40

Caracterización funcional de un transportador *PtNTT2* mutante

- 45 La expresión del transportador de nucleósido trifosfato de *Phaeodactylum tricornutum* (*PtNTT2*) en *E. coli* permitió la importación de **dNaMTP** y **d5SICS**TP y la posterior replicación de **dNaM**-**d5SICS** UBP (figura 1A), pero su expresión también fue tóxica (figura 1B). En un SSO denominado en el presente documento como DM1, el transportador se expresó a partir de un promotor T7 en un plásmido multicopia (pCDF-lb) en C41(DE3) de *E. coli*, y su inducción se controló debido a la toxicidad asociada. En su célula de alga nativa, las secuencias señal del extremo N de *PtNTT2*
- 50 dirigen su localización subcelular y se eliminan mediante proteólisis. En algunos casos en el sistema de *E. coli*, la señal del extremo N se conservó y contribuyó a la toxicidad observada. La delección de los aminoácidos 1-65 y la expresión de la variante truncada en el extremo N resultante *PtNTT2*(66-575) en C41(DE3) de *E. coli* dio como resultado una menor toxicidad en relación con el *PtNTT2* de longitud completa, pero también redujo la captación de [α -³²P]-dATP (figura 5A y figura 5B), posiblemente debido a una expresión reducida. La expresión de *PtNTT2*(66-575)
- 55 en BL21(DE3) de *E. coli* dio como resultado niveles aumentados de captación de [α -³²P]-dATP con un pequeño aumento en la toxicidad en relación con un control de vector vacío (figura 5A y figura 5C), pero el nivel más alto de T7 RNAP en esta cepa era en sí mismo tóxico (figura 5A y figura 5C).

- Se exploró la expresión constitutiva de *PtNTT2*(66-575) a partir de un plásmido de bajo número de copias o un locus cromosómico, para eliminar la necesidad de producir niveles tóxicos de T7 RNAP, y para impartir al SSO una mayor autonomía, una expresión de transportador más homogénea y captación de trifosfato en una población de células y, en última instancia, mejorar la retención de UBP. Se exploró la expresión de *PtNTT2*(66-575) en BL21(DE3) de *E. coli* con los promotores de *E. coli* *P*_{lacI}, *P*_{bla} y *P*_{lac} de un plásmido pSC, y con *P*_{bla}, *P*_{lac}, *P*_{lacUV5}, *P*_{H207}, *P*_A, *P*_{tac} y *P*_{N25} del locus cromosómico *lacZYA* (véase la Tabla 2). También se exploró el uso de una variante de codones optimizados del transportador truncado (véase la Tabla 2). Aunque la captación de [α -³²P]-dATP se correlacionó negativamente con el tiempo de duplicación, cada cepa exhibió una mejor relación entre la captación y la aptitud en comparación con DM1
- 60
- 65

(figura 1B). La cepa YZ3, que expresó el PtNTT2(66-575) integrado cromosómicamente y con codones optimizados del promotor P_{lacUV5} , exhibió un fuerte crecimiento (<20 % de aumento en el tiempo de duplicación en relación con la cepa isogénica sin el transportador) y niveles razonables de captación de $[\alpha\text{-}^{32}\text{P}]\text{-dATP}$, y se seleccionó para una caracterización adicional.

Para determinar si el sistema transportador optimizado de YZ3 facilita una alta retención de UBP, se construyeron tres plásmidos que posicionan el UBP dentro de la secuencia TK1 de 75 nt (con un contexto de secuencia local de d(A-NaM-T)). Estos incluyen dos plásmidos derivados de pUC19 con alto número de copias, pUCX1 y pUCX2, así como un plásmido derivado de pBR322 con bajo número de copias, pBRX2 (figura 6). Además de examinar el efecto del número de copias en la retención de UBP, estos plásmidos posicionaron el UBP en las posiciones proximal (pUCX1) y distal (pUCX2 y pBRX2) en relación con el origen de la replicación. YZ3 y DM1 de *E. coli* se transformaron con pUCX1, pUCX2 o pBRX2 y se cultivaron directamente en medios de cultivo líquidos complementados con dNaMTP y d5SICS (e IPTG para DM1 para inducir el transportador), y se caracterizaron el crecimiento y la retención de UBP (a una DO_{600} de ~1) (véanse la sección Métodos y la figura 7A). Mientras que DM1 mostró niveles variables de retención y crecimiento reducido con los plásmidos de alto número de copias, YZ3 mostró niveles uniformemente altos de retención de UBP y fuerte crecimiento (figura 2a y figura 8A).

Para explorar el efecto del contexto de la secuencia local sobre la retención de UBP en YZ3, se construyeron dieciséis variantes de pUCX2 en las que el UBP estaba flanqueado por cada combinación posible de pares de bases naturales dentro de un fragmento de *gfp* (véase la Tabla 2). Bajo las mismas condiciones de crecimiento que anteriormente, se observó una amplia gama de retenciones de UBP, mostrando algunos contextos de secuencia una pérdida completa del UBP (figura 2B). Sin embargo, dado el desarrollo de DM1 con dNaM-d5SICS UBP, se determinó que la contracción anular y la derivatización de azufre de d5SICS, que producen dNaM-dTPT3 UBP (figura 1A), dieron como resultado una replicación *in vitro* más eficiente. Para explorar el uso *in vivo* de dNaM-dTPT3, los experimentos se repitieron con YZ3 y cada uno de los dieciséis plásmidos pUCX2 pero con crecimiento en medios complementados con dNaMTP y dTPT3TP. Las retenciones de UBP fueron claramente más altas con dNaM-dTPT3 que con dNaM-d5SICS (figura 2B).

Aunque dNaM-dTPT3 es un UBP más óptimo para el SSO que dNaM-d5SICS, su retención sigue siendo de moderada a mala en algunas secuencias. Además, varias secuencias que muestran una buena retención en YZ3 cultivadas en medios líquidos muestran una mala retención cuando el crecimiento incluye el cultivo en medio sólido (figura 8B). Para aumentar aún más la retención de UBP incluso con estas secuencias desafiantes y/o condiciones de crecimiento, se llevó a cabo la eliminación selectiva de plásmidos que pierden el UBP. En procariotas, el sistema de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR)/asociadas a CRISPR (Cas) proporciona inmunidad adaptativa contra virus y plásmidos. En los sistemas CRISPR-Cas de tipo II, tal como el de *Streptococcus pyogenes*, la endonucleasa Cas9 utiliza ARN codificados (o sus imitadores artificiales conocidos como ARN guía únicos (ARNgu)) para introducir roturas bicatenarias en el ADN complementario cadena arriba de un motivo adyacente de protoespaciador (PAM) de 5'-NGG-3' (figura 3A), que a continuación da como resultado la degradación del ADN por exonucleasas. *In vitro*, se encontró que la presencia de un UBP en el ADN diana generalmente reduce la escisión mediada por Cas9 en relación con las secuencias que son completamente complementarias al ARNgu proporcionado (figura 9). En algunos casos dentro de una célula, el Cas9 programado con uno o más ARNgu complementarios a las secuencias naturales que surgen de la pérdida de UBP forzaría la retención en una población de plásmidos, lo que se conoce como inmunidad a la pérdida de UBP. Para probar esto, se usó un plásmido p15A para construir pCas9, que expresa Cas9 a través de un promotor LacO inducible por IPTG, así como un ARNgu de 18 nt que es complementario a la secuencia TK1 que contiene la mutación dNaM-dTPT3 más común (dT-dA) a través del promotor ProK constitutivo (figura 6 y 10A). La cepa YZ2 (un precursor de YZ3 con un rendimiento de transportador ligeramente menos óptimo; figura 1A, figura 5D y figura 5F) que lleva el plásmido pCas9 se transformó con el plásmido pUCX2 correspondiente (es decir, la variante pUCX2 con el UBP incrustado dentro de la secuencia TK1 de tal forma que la pérdida de UBP produce una secuencia dirigida por el ARNgu codificado en pCas9), se cultivó a una DO_{600} de ~4, se diluyó 250 veces y se regeneró a la misma DO_{600} . La retención de UBP en experimentos de control con un ARNgu no diana se redujo al 17 % después del segundo crecimiento; en contraste, la retención de UBP en presencia del ARNgu correcto fue del 70 % (figura 10B). La secuenciación reveló que la mayoría de los plásmidos que carecían de UBP cuando se proporcionó el ARNgu correcto contenían una delección de un solo nucleótido en su lugar, lo que no se observó con el ARNgu no diana (figura 10C y figura 10D). Con un plásmido pCas9 que expresa dos ARNgu, uno dirigido a la mutación de sustitución más común y otro dirigido a la mutación de delección de un solo nucleótido (figura 6), y el mismo ensayo de crecimiento y recambio, la pérdida de UBP fue indetectable (figura 10B).

Para explorar más ampliamente la inmunidad mediada por Cas9 a la pérdida de UBP, se examinó la retención usando dieciséis variantes pUCX2 con secuencias que flanquean el UBP con cada combinación posible de pares de bases naturales, pero también varían su posición con respecto al PAM, y varían qué el nucleótido no natural esté presente en la cadena reconocida por los ARNgu (figura 11). También se construyó un conjunto correspondiente de dieciséis plásmidos pCas9 que expresan dos ARNgu, uno dirigido a una mutación de sustitución y otro dirigido a la mutación de delección de un solo nucleótido, para cada variante pUCX2. La cepa YZ2 que llevaba un plásmido pCas9 se transformó con su correspondiente variante pUCX2 y se cultivó en presencia de trifosfatos no naturales e IPTG (para inducir Cas9), y se evaluó la retención de UBP después de que las células alcanzaran una DO_{600} de ~1. Como control, los dieciséis plásmidos pUCX2 también se propagaron en YZ2 que llevaba un plásmido pCas9 con un ARNgu no diana. Para cuatro de las dieciséis secuencias exploradas, la pérdida de UBP ya era mínima sin inmunidad (ARNgu

no diana), pero era indetectable con la expresión del ARN_gu correcto (figura 3B). Las secuencias restantes mostraron una retención de moderada a nula sin inmunidad, y una retención significativamente mayor con esta, incluso en posiciones de hasta 15 nt del PAM

- 5 Para simplificar y agilizar aún más el SSO, la cepa YZ4 se construyó integrando un gen Cas9 inducible por IPTG en el locus *arsB* del cromosoma YZ3, lo que permite el uso de un solo plásmido que lleva un UBP y expresa los ARN_gu que fuerzan su retención. Dieciséis de estos plásmidos "todo en uno" (pAIO) se construyeron reemplazando el gen Cas9 en cada una de las variantes pCas9 con una secuencia UBP de la variante pUCX2 correspondiente (Datos ampliados en la figura 2). YZ4 e YZ3 (incluidos como un control sin Cas9 debido a la expresión con fugas de Cas9 en
- 10 YZ4) se transformaron con un solo plásmido pAIO y se cultivaron en medios de crecimiento sólidos complementados con trifosfatos no naturales y con o sin IPTG para inducir Cas9. Se usaron colonias individuales para inocular medios líquidos de la misma composición y se evaluó la retención de UBP después de que las células alcanzaran una DO₆₀₀ de ~1-2 (figura 3C). A pesar de los niveles variables de retención en ausencia de Cas9 (YZ3), con la inducción de la expresión de Cas9 en YZ4, la pérdida fue mínima o indetectable en 13 de las 16 secuencias. Si bien la retención con las tres secuencias problemáticas, d(C-**NaM**-C) d(C-**NaM**-A) y d(C-**NaM**-G), podría optimizarse, por ejemplo, a través de alteraciones en la expresión de Cas9 o ARN_gu, la pérdida indetectable del UBP con la mayoría de las secuencias después de un régimen que incluía crecimiento tanto en medios sólidos como líquidos, lo que no era posible con este SSO DM1 anterior, atestigua la vitalidad de YZ4.
- 15
- 20 Finalmente, se construyó un plásmido pAIO, pAIO2X, que contenía dos UBP: d**NaM** emparejado frente a d**TPT3** en la posición 451 de la cadena codificante del gen *gfp* y d**TPT3** emparejado frente a d**NaM** en la posición 35 de la cadena codificante del gen de ARNt *ser7*, así como codificación de los ARN_gu dirigidos a la mutación de sustitución más común esperada en cada secuencia (figura 6). YZ4 e YZ3 se transformaron con pAIO2X y se sometieron al régimen de crecimiento desafiante representado en la figura 4, que incluía un extenso crecimiento de alta densidad en medios de crecimiento líquidos y sólidos. Los plásmidos se recuperaron y se analizaron para determinar la retención de UBP
- 25 (figura 7B) cuando la DO₆₀₀ alcanzó 1-2 durante cada crecimiento líquido. En YZ3, que no expresa Cas9, o en ausencia de inducción de Cas9 (sin IPTG) en YZ4, la retención de UBP disminuyó constantemente con el crecimiento extendido (figura 4). Con la inducción de inmunidad (IPTG 20 o 40 µM) se observó una reducción marginal en la velocidad de crecimiento (menos del 14 % de aumento en el tiempo de duplicación: figura 12), y aproximadamente el 100 % de retención de UBP (pérdida no detectable) en ambos genes.
- 30

La Tabla 1 ilustra las secuencias descritas en el presente documento.

	Secuencia	SEQ ID NO:
PtNTT2 (longitud completa)	MRPYPTIALI SVFLSAATRI SATSSHQASA LPVKKGTHVP DSPKLSKLYI MAKTKSVSSSFDPPRGGSTV APTTPLATGG ALRKVRQAVF PIYGNQEVTK FLLIGSIKFF IILALTLTRD TKDTLIVTQC GAEIAFLKI YGVLPAAATF IALYSKMSNA MGKKMLFYST CIPFFTFGLFDVFIYPNAE RLHPSLEAVQ AILPGGAASG GMAVLAKIAT HWTSALFYVM AEIYSSVSVG LLFWQFANDV VNVDQAKRFY PLFAQMSGLAPVLAGQYVVR FASKAVNFEA SMHRLTAAVTFAGIMICIFY QLSSSYVERT ESAKPAADNE QSIKPKKKKP KMSMVESGKF LASSQYLRLI AMLVLGYGLS INFTEIMWKS LVKKQYPDPL DYQRFMGNFS SAVGLSTCIV IFFGVHVIRLLGWKVGALAT PGIMAILALP FFACILLGLD SPARLEIAVI FGTIQSLLSK TSKYALFDPT TQMAYIPLDD ESKVKGKAAI DVLGSRIGKS GGSIIQQGLV FVFGNIINAA PVVGVVYYSVLVAWMSAAGR LSGLFQAQTE MDKADKMEAK TNKEK	1

(continuación)

	Secuencia	SEQ ID NO:
<i>PtNTT2</i> (66-575)* ADN (codones optimizados) (*66-575 representa que el ADN que codifica los primeros 65 residuos de aminoácidos se ha eliminado en relación con el <i>PtNTT2</i> de longitud completa)	ATGGGTGGTAGCACCGTTGCACCGACCACACCGCTGGCAA CCGGTGGTGCACCTGCGTAAAGTTCGTGAGGCAGTTTTTCC GATTTATGGCAATCAAGAAGTGACCAAATTTCTGCTGATTG GCAGCATCAAATTCTTTATTATCCTGGCACTGACCCTGA CCCGTGATACCAAAGATACCCTGATTGTTACCCAGTGTGGT GCAGAAGCAATTGCATTTCTGAAAATCTATGGTGTCTG CCTGCAGCAACCGCATTATTGCACTGTATAGCAAAATGAG CAACGCAATGGGCAAAAAAATGCTGTTTTATAGCACCTG TATCCCGTTCTTTACCTTTTTTGGTCTGTTTCGATGTGTTCAATT TATCCGAATGCCGAACGTCTGCATCCGAGCCTGGAAG CAGTTCAGGCAATTCTGCCTGGTGGTGCCGCAAGCGGTGGT ATGGCAGTTCTGGCAAAAATTGCAACCCATTGGACCAGC GCACTGTTTTATGTTATGGCAGAAATCTATAGCAGCGTTAG CGTTGGTCTGCTGTTTTTGGCAGTTTGCAAATGATGTTGT TAATGTGGATCAGGCCAAACGTTTTTATCCGCTGTTTGCAC AGATGAGCGGTCTGGCACCGGTTCTGGCAGGTCAGTATG TTGTTCTGTTTTGCAAGCAAAGCCGTTAATTTTGAAGCAAGC ATGCATCGTCTGACCGCAGCAGTTACCTTTGCAGGTATT	2
	ATGATCTGCATCTTTTATCAGCTGAGCAGCTCATATGTTGA ACGTACCGAAAGCGCAAAACCGGCAGCAGATAATGAACA GAGCATTAACCGAAGAAAAAACCAGAAATGTGCGATG GTGGAAAGCGGTAAATTTCTGGCAAGCAGCCAGTATCTGC GTCTGATTGCAATGCTGGTTCTGGGTTATGGTCTGAGCATT AACTTTACCGAAATCATGTGGAAGCCTGGTGAAAAA CAGTATCCGGATCCGCTGGATTATCAGCGTTTTATGGGTAA TTTTAGCAGCGCAGTTGGTCTGAGTACCTGCATTGTTAT CTTTTTTGGCGTGCATGTTATTCGTCTGCTGGGTTGGAAAGT TGGTGCCCTGGCAACACCGGGTATTATGGCCATTCTGG CACTGCCGTTTTTTCATGTATTCTGCTGGGCCTGGATAGTC CGGCACGTCTGGAAATTGCAGTTATTTTTGGCACCATTT CAGAGCCTGCTGAGCAAAACCAGCAAATATGCACTGTTTG ATCCGACCACCCAGATGGCATATATCCCGCTGGATGATGA AAGCAAAGTTAAAGGCAAAGCAGCCATTGATGTTCTGGGT AGCCGTATTGGTAAATCAGGTGGTAGCCTGATTCAGCAGG GTCTGGTTTTTTGTTTTTGGCAATATTATCAATGCCGCACCGG TTGTTGGTGTGTGTATTATAGCGTTCTGGTTGCATGG ATGAGTGCAGCAGGTCTGCTGAGTGGTCTGTTTCAGGCACA GACCGAAATGGATAAAGCAGATAAAATGGAAGCCAAAAC CAACAAAGAAAAATGA	

(continuación)

	Secuencia	SEQ ID NO:
ADN de <i>PtNTT2</i> (66-575) (sin codones optimizados)	ATGGGAGGCAGTACTGTTGCACCAACTACACCGTTGGCAA CCGGCGGTGCGCTCCGCAAAGTGCGACAAGCCGTCTTTCCC ATCTACGGAAACCAAGAAGTCACCAAATTTCTGCTCATCGG ATCCATTAAATTCTTTATAATCTTGGCACTCACGCTCACGC GTGATACCAAGGACACGTTGATTGTCACGCAATGTGGTGCC GAAGCGATTGCCTTTCTCAAAATATACGGGGTGCTACCCGC AGCGACCGCATTTATCGCGCTCTATTCCAAAATGTCCAACG CCATGGGCAAAAAAATGCTATTTTATTCCACTTGCATTCT TTCTTTACCTTTTTTCGGGCTGTTTGATGTTTTCATTTACCCG AACGCGGAGCGACTGCACCCTAGTTTGGAAGCCGTGCAGG CAATTCTCCCGGGCGGTGCCGCATCTGGCGGCATGGCGGTT CTGGCCAAGATTGCGACACACTGGACATCGGCCTTATTTTA CGTCATGGCGGAAATATATTCTTCCGTATCGGTGGGGCTAT TGTTTTGGCAGTTTGCGAACGACGTCGTCAACGTGGATCAG GCCAAGCGCTTTTATCCATTATTTGCTCAAATGAGTGGCCT CGCTCCAGTTTTAGCGGGCCAGTATGTGGTACGTTTTGCCA GCAAAGCGGTCAACTTTGAGGCATCCATGCATCGACTCACG GCGGCCGTAACATTTGCTGGTATTATGATTTGCATCTTTTAC CAACTCAGTTCGTCATATGTGGAGCGAACGGAATCAGCAA AGCCAGCGGCAGATAACGAGCAGTCTATCAAACCGAAAAA GAAGAAACCCAAAATGTCCATGGTTGAATCGGGGAAATTT CTCGCGTCAAGTCAGTACCTGCGTCTAATTGCCATGCTGGT GCTGGGATACGGCCTCAGTATTAACTTTACCGAAATCATGT GGAAAAGCTTGGTGAAGAAACAATATCCAGACCCGCTAGA TTATCAACGATTTATGGGTAACCTCTCGTCAGCGGTTGGTTT GAGCACATGCATTGTTATTTTCTTCGGTGTGCACGTGATCC GTTTGTGGGGTGGAAAGTCGGAGCGTTGGCTACACCTGGG ATCATGGCCATTCTAGCGTTACCCTTTTTTGCTTGCATTTTG TTGGGTTTGGATAGTCCAGCACGATTGGAGATCGCCGTAAT CTTTGGAACAATTCAGAGTTTGCTGAGCAAAACCTCCAAGT ATGCCCTTTTCGACCCTACCACACAAATGGCTTATATTCCTC TGGACGACGAATCAAAGGTCAAAGGAAAAGCGGCAATTGA TGTTTTGGGATCGCGGATTGGCAAGAGTGGAGGCTCACTGA TCCAGCAGGGCTTGGTCTTTGTTTTTGGAAATATCATTAAT GCCGCACCTGTAGTAGGGGTGTCTACTACAGTGTCTTGT TGCGTGGATGAGCGCAGCTGGCCGACTAAGTGGGCTTTTTTC AAGCACAAACAGAAATGGATAAGGCCGACAAAATGGAGG CAAAGACCAACAAAGAAAAGTAG	3
Proteína <i>PtNTT2</i> (66-575)	MGGSTVAPTTPLATGGALRKVRQAVFPIYGNQEVTKFLLIGSI KFFIILALTLTRDTKDTLIVTQCGAEIAFLKIYGVLPAAATAFIA LYSKMSNAMGKKMLFYSTCIPFFTFGLFDVFIYPNAERLHPS LEAVQAILPGGAASGGMVLAKEIATHWTSALFYVMAEYSSV SVGLLFWQFANDVVNVDQAKRFYPLFAQMSGLAPVLAGQYV VRFASKAVNFEASMHRLTAAVTFAGIMICIFYQLSSSYVERTE SAKPAADNEQSIKPKKKPKMSMVESGKFLASSQYLRLIAML VLGYGLSINFTEIMWKS LVKKQYPDPLDYQRFMGNFSSAVGL STCIVIFFGVHVIRLLGWKVGALATPGIMAILALPFFACILLGL DSPARLEIAVIFGTIQSLLSKTSKYALFDPTTQMAYIPLDDESK VKGKAAIDVLGSRIGKSGGSLIQQLVFVFGNIINAAPVVGVV YYSVLVAWMSAAGRLSGLFQAQTEMDKADKMEAKTNKEK	4

(continuación)

	Secuencia	SEQ ID NO:
<i>PtNTT2</i> (1-22,66-575)* ADN (* 1-22,66-575 representa que el ADN que codifica los residuos de aminoácidos 23-65 se ha eliminado en relación con el <i>PtNTT2</i> de longitud completa)	ATGAGACCATTTCGACGATTGCCTTGATTTCGGTTTTTCTT TCGGCGGCGACTCGCATTTTCGGCAGGAGGCAGTACTGTTGC ACCAACTACACCGTTGGCAACCGGCGGTGCGCTCCGCAA GTGCGACAAGCCGTCTTTCCCATCTACGGAAACCAAGAAGT CACCAAATTTCTGCTCATCGGATCCATTAAATTCTTTATAAT CTTGGCACTCACGCTCACGCTGATACCAAGGACACGTTGA TTGTCACGCAATGTGGTGCCGAAGCGATTGCCTTTCTCAA ATATACGGGGTGCTACCCGCAGCGACCGCATTTATCGCGCT CTATTCCAAAATGTCCAACGCCATGGGCAAAAAAATGCTAT TTTATTCCACTTGCATTCTTTCTTTACCTTTTTTCGGGCTGTT TGATGTTTTCATTTACCCGAACGCGGAGCGACTGCACCCTA GTTTGGAAGCCGTGCAGGCAATTCTCCCGGGCGGTGCCGCA TCTGGCGGCATGGCGGTTCTGGCCAAGATTGCGACACACTG GACATCGGCCTTATTTTACGTCATGGCGGAAATATATTCTT CCGTATCGGTGGGGCTATTGTTTTGGCAGTTTGCGAACGAC GTCGTCAACGTGGATCAGGCCAAGCGCTTTTATCCATTATT TGCTCAAATGAGTGGCCTCGCTCCAGTTTTAGCGGGCCAGT ATGTGGTACGGTTTGCCAGCAAAGCGGTCAACTTTGAGGCA TCCATGCATCGACTCACGGCGGCCGTAACATTTGCTGGTAT TATGATTTGCATCTTTTACCAACTCAGTTCGTCATATGTGGA GCGAACGGAATCAGCAAAGCCAGCGGCAGATAACGAGCA GTCTATCAAACCGAAAAAGAAACCCAAAATGTCCATG GTTGAATCGGGGAAATTTCTCGCGTCAAGTCAGTACCTGCG TCTAATTGCCATGCTGGTGCTGGGATACGGCCTCAGTATTA ACTTTACCGAAATCATGTGGAAGCTTGGTGAAGAAACA ATATCCAGACCCGCTAGATTATCAACGATTTATGGGTAAC TCTCGTCAGCGGTTGGTTTGAGCACATGCATTGTTATTTCT TCGGTGTGCACGTGATCCGTTTGTGGGGTGGAAGTCGGA GCGTTGGCTACACCTGGGATCATGGCCATTCTAGCGTTACC CTTTTTTGCTTGCAATTTGTTGGGTTTGGATAGTCCAGCACG ATTGGAGATCGCCGTAATCTTTGGAACAATTCAGAGTTTGC TGAGCAAAACCTCCAAGTATGCCCTTTTCGACCCTACCACA CAAATGGCTTATATTCCTCTGGACGACGAATCAAAGGTCAA AGGAAAAGCGGCAATTGATGTTTTGGGATCGCGGATTGGC AAGAGTGGAGGCTCACTGATCCAGCAGGGCTTGGTCTTTGT TTTTGGAAATATCATTAATGCCGCACCTGTAGTAGGGGTTG TCTACTACAGTGTCCTTGTGCGTGATGAGCGCAGCTGGC CGACTAAGTGGGCTTTTTCAAGCACAAACAGAAATGGATA AGGCCGACAAAATGGAGGCAAAGACCAACAAAGAAAAGT AG	5
<i>PtNTT2</i> (1-22, 66-575)	MRPFPTIALISVFLSAATRISAGGSTVAPTTPLATGGALRKVRQ AVFPIYGNQEVTKFLIGSIKFFILALTLTRDTKDLIVTQCGA EAIAFLKIYGVLPAAATAFIALYSKMSNAMGKKMLFYSTCIPFF	6

(continuación)

	Secuencia	SEQ ID NO:
Proteína	TFFGLFDVFIYPNAERLHPSLEAVQAILPGBAASGGMAVLAKI ATHWTSALFYVMAEYSSVSVGLLFWQFANDVVNVDAQRF YPLFAQMSGLAPVLAGQYVVRFASKAVNFEASMHRLTAAVT FAGIMICIFYQLSSSYVERTESAKPAADNEQSIKPKKKKPKMS MVESGKFLASSQYLRLIAMLVLGYGSLINFTEIMWKSLSVKKQ YPDPLDYQRFMGNFSSAVGLSTCIVIFFGVHVIRLLGWKVGAL ATPGIMAILALPFFACILLGLDSPARLEIAVIFGTIQSLLSKTSKY ALFDPTTQMAYIPLDDESKVKGKAIDVLGSRIGKSGGSLIQQ GLVFVFGNIINAAPVVGVVYYSVLVAWMSAAGRLSGLFQAQ TEMKDADKMEAKTNKEK	
PtNTT2 (23-575)* ADN (*23-575 representa que el ADN que codifica los residuos de aminoácidos 1-22 se ha eliminado en relación con el PtNTT2 de longitud completa)	ATGACTTCCTCTCATCAAGCAAGTGCACCTTCCTCTCAAAAA GGGAACGCATGTCCCGGACTCTCCGAAGTTGTCAAAGCTAT ATATCATGGCCAAAACCAAGAGTGTATCCTCGTCCTTCGAC CCCCCTCGGGGAGGCAGTACTGTTGCACCAACTACACCGTT GGCAACCGGCGGTGCGCTCCGCAAAGTGCGACAAGCCGTC TTTCCCATCTACGGAAACCAAGAAGTCACCAAATTTCTGCT CATCGGATCCATTAAATTCTTTATAATCTTGGCACTCACGCT CACGCGTGATACCAAGGACACGTTGATTGTACGCAATGTG GTGCCGAAGCGATTGCCTTTCTCAAAATATACGGGGTGCTA CCCGCAGCGACCGCATTTATCGCGCTCTATTCCAAAATGTC CAACGCCATGGGCAAAAAAATGCTATTTTATTCCACTTGCA TTCCTTTCTTTACCTTTTTTCGGGCTGTTTGATGTTTTATTTA CCCGAACGCGGAGCGACTGCACCCTAGTTTGGAAGCCGTG CAGGCAATTCTCCCGGGCGGTGCCGCATCTGGCGGCATGGC GGTTCTGGCCAAGATTGCGACACACTGGACATCGGCCTTAT TTTACGTCATGGCGGAAATATATTCTTCCGTATCGGTGGGG CTATTGTTTTTGGCAGTTTGCGAACGACGTCGTCAACGTGGA TCAGGCCAAGCGCTTTTATCCATTATTTGCTCAAATGAGTG GCCTCGCTCCAGTTTTAGCGGGCCAGTATGTGGTACGGTTT GCCAGCAAAGCGGTCAACTTTGAGGCATCCATGCATCGACT CACGGCGGCCGTAAACATTTGCTGGTATTATGATTTGCATCT TTTACCAACTCAGTTCGTATATGTGGAGCGAACGGAATCA GCAAAGCCAGCGGCAGATAACGAGCAGTCTATCAAACCGA AAAAGAAGAAACCCAAAATGTCCATGGTTGAATCGGGGAA ATTTCTCGCGTCAAGTCAGTACCTGCGTCTAATTGCCATGC TGGTGCTGGGATACGGCCTCAGTATTAACCTTTACCGAAATC ATGTGGAAGCTTGGTGAAGAAACAATATCCAGACCCGC TAGATTATCAACGATTTATGGGTAACTTCTCGTCAGCGGTT GGTTTGAGCACATGCATTGTTATTTTCTTCGGTGTGCACGTG ATCCGTTTGTGGGGTGGAAGTCGGAGCGTTGGCTACACC TGGGATCATGGCCATTCTAGCGTTACCCTTTTTTGCTTGCAT TTTGTGGGTTTGGATAGTCCAGCACGATTGGAGATCGCCG TAATCTTTGGAACAATTCAGAGTTTGCTGAGCAAAACCTCC AAGTATGCCCTTTTCGACCCTACCACACAAATGGCTTATAT TCCTCTGGACGACGAATCAAAGGTCAAAGGAAAAGCGGCA ATTGATGTTTTGGGATCGCGGATTGGCAAGAGTGGAGGCTC ACTGATCCAGCAGGGCTTGGTCTTTGTTTTTGGAAATATCA TTAATGCCGCACCTGTAGTAGGGGTTGTCTACTACAGTGTC CTTGTTGCGTGATGAGCGCAGCTGGCCGACTAAGTGGGCT TTTTCAAGCACAAACAGAAATGGATAAGGCCGACAAAATG GAGGCAAAGACCAACAAAGAAAAGTAG	7

(continuación)

	Secuencia	SEQ ID NO:
Proteína <i>Pt</i> NTT2 (23-575)	MTSSHQASALPLKKGTHVPDSPKLSKLYIMAKTKSVSSSFDP RGGSTVAPTTPLATGGALRKVRQAVFPIYGNQEVTKFLLIGSI KFFIILALTLTRDTKDTLIVTQCGAEIAFLKIYGVLPAAATAFIA LYSKMSNAMGKKMLFYSTCIPFFTFGLFDVFIYPNAERLHPS LEAVQAILPGGAASGGMAVLAKIATHWTSALFYVMAEIYSSV SVGLLFWQFANDVVNVDQAKRFYPLFAQMSGGLAPVLAGQYV	8
	VRFASKAVNFEASMHRLTAAVTFAGIMICIFYQLSSSYVERTE SAKPAADNEQSIKPKKKPKMSMVESGKFLASSQYLRLIAML VLGYGLSINFTEIMWKSLVKKQYPDPLDYQRFMGNFSSAVGL STCIVIFFGVHVIRLLGWKVGALATPGIMAILALPFFACILLGL DSPAILEIAVIFGTIQSLLSKTSKYALFDPTTQMAYIPLDDESK VKGKAAIDVLGSRIGKSGGSLIQQLVFVFGNIINAAPVVGVV YYSVLVAWMSAAGRLSGLFQAQTEMDKADKMEAKTNKEK	

(continuación)

	Secuencia	SEQ ID NO:
pBRX2P* (*N representa la posición del UBP)	AACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTG CTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCC TGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTG ATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT CAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTT TCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGG TGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAG CCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATG GCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTG ACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGC TGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTAC CGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCT CATCAGCGTGGTCGTGAAGCGATTACAGATGTCTGCCTGT TCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTT AATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGT TTTTCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGG ATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAG AGGATGCTCACGATACGGGTTACTGATGATGAACATGCCC GGTTACTGGAACGTTGTGAGGGTAAACAACCTGGCGGTATG GATGCGGCGGGACAGAGAAAAATCACTCAGGGTCAATGC CAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAG CCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGT GCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACAC GGAACCGAAGACCATTTCATGTTGTTGCTCAGGTGCGAG ACGTTTTGTCAGCAGCAGTCGCTTACGTTGCTCGCGTATC GGTGATTCAATTCTGCTAACCAGTAAGGCAACCCCGCCAG CCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCA CCCGTGGCCAGGACCCAACGCTGCCCCGAAATCTTGAAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTAAATGT CATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT TCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCT AAATACATTCAAATATGTATCCGCTCATGAGACAATAAC CCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTAT GAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTT GCGGCATTTTGCCTTCCTGTTTTTTGCTCACCCAGAAACGCTG GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACG AGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATC CTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGA GCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGT GTTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATAACAC TATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGA AAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG CAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACT TACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG CTTTTTTGCAACATGGGGGATCATGTAACCTCGCCTTGAT CGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGAC GAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGT TGCGCAAACCTATTAAGTGGCGAACTACTTACTCTAGCTTC CCGGCAACAATTAAGTGGCGAACTACTTACTCTAGCTTC CGGGCAACAATTAAGTGGCGAACTACTTACTCTAGCTTC GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGT	9

(continuación)

	Secuencia	SEQ ID NO:
	TTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGCTCTCGC GGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCC CGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTAT GGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTC ACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCAT ATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTA AAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACC AAAATCCCTTAACGTGAGTTTTTCGTTGGCTTTACACTTT ATGCTTCCGGCTCGTATGTTGTGTGGAANTGTGAGCGGATA ACAATTTACACAGGAAACAGCCACTGAGCGTCAGACCC CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTC TGC GCGTAATCTGCTGCTTGCAAACAAAAAACCACCGC TACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACT CTTTTTCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATA CCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCA CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGG ATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGA TACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGG TCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAA ACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCT GACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG AGCCTATGGAAA	

Tabla 2.

Cebador	Aplicación	Secuencia	SEQ ID NO:
Clonación de plásmidos transportadores e integración cromosómica			
YZ552	Clonación de <i>P</i> NTT2(66-575)	GGGAGGCAGTACTGTTGCAC	10
pCDF-1 b-fwd	Clonación de <i>P</i> NTT2(66-575)	GGTATATCTCCTTATTAAGTTAAACAAAAATTATTCT ACAGGGG	11
T7 seq	secuenciación de los plásmidos pCDF	TAATACGACTCACTATAGGG	12
T7 term seq	secuenciación de los plásmidos pCDF	GCTAGTTATTGCTCAGCGG	13
YZ580	clonación del transportador	TTACATTAAATTGCGTTGCGCTC	14
DM002	clonación del transportador	TTTTGGCGGATGGCATTTGAGAAGCACACGG	15
YZ576	clonación del transportador	ATTCTACCAATAAAAAACGCCCGG	16
YZ581	clonación del transportador	CCTGTAGAAATAATTTTGTAACTTTAATAAGGAG	17
DM052	secuenciación del transportador	CCCCGCGCTTGCGCCGATTC	18
DM053	secuenciación del transportador	GAAAGGCAATCAGCTGTTG	19
YZ50	secuenciación del transportador	CAGGCGAGGTCGTTAAATAG	20
YZ584	promotor lacI	GACACCATCGAATGGCGCAAAACCTTTCGCGGTATGG CATGATAGCGCCCGG	21
YZ585	promotor lacI	CCGGGCGCTATCATGCCATACCCGCGAAAGGTTTTCG CCATTCGATGGTGC	22
YZ582	promotor bla	ATTTTCTAAATACATTCAAAATATGTATCCGCTCATGA GACAATAACCCCTG	23
YZ583	promotor bla	CAGGGTTATTGTCTCATGAGCGGATACATATTGAAT GTATTTAGAAAAAT	24
YZ599	promotor lac	TTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTG GTATGTTGTGTGA	25

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
YZ600	promotor lac	TCCACACAACATACCAGCCGGAAGCATAAAGTGTAAG GCCTGGGGTGCCTAA	26
YZ595	promotor lacUV5	CTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTG GTATAATGTGTGGA	27
YZ596	promotor lacUV5	TCCACACATTATACCAGCCGGAAGCATAAAGTGTAAG GCCTGGGGTGCCTAG	28
YZ7	clonación de integración de transportador	GGGATAACGCGAGGAAAGAACATG	29
YZ12	clonación de integración de transportador	GCACTTTCGGGGAATGTGCG	30
YZ612	clonación de integración de transportador	AATTGCGGCTATATGGATGTTGGAAACCGTAAGAGAA ATAGACAGGCGGTCTGTGACGGAAGATCACTTCGCA G	31
YZ613	clonación de integración de transportador	TGCTCACATGTTCTTTCTTCTGCGTTATCCCCGCGTG AACCAGGC	32
YZ614	clonación de integración de transportador	ACCGCTGTCTATTCTCTTACGGTTCC	33
YZ615	clonación de integración de transportador	CGCGCTTAATGCGCGCTACAGGGGCGGTCGATTGGT GCCAGCGGCGCAG	34
YZ610	clonación de integración de transportador	GGTATATCTCTCTTATTAAAGTTAAACAAAAATTATTCT ACAGG	35
YZ616	<i>lacZYA</i> ::integración de transportadores	CAGCCACGTTTCTGCGAAAAAC	36
YZ617	<i>lacZYA</i> ::integración de transportadores	TACAGCGGTTCTTACTGGC	37
YZ618	PCR de colonias de integración de transportadores	GGTGGTGAATGTGAACACCAGTAACG	38
YZ619	PCR de colonias de integración de transportadores	CTGGGTGTTACTTCGGTCTG	39

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
YZ69	PCR de colonias de integración de transportadores	GGCCGTAATATCCAGCTGAAC	40
YZ587	PCR de colonias de integración de transportadores	ACTAGGGTGCAGTCGCTCCG	41
pCDF-1b-rev	clonación de integración	TTAACCTAGGCTGCTGCCACCG	42
YZ703	promotor tac	GCGCAACGCAATTAAATGTAATTCTGAAATGAGCTGTT	43
		GACAATTAAATCATCGGCTCG	
YZ704	promotor tac	AACAAAAATTATTTCTACAGGTCCACACATTATACGAG CCGATGATTAAATTGTCAAC	44
YZ707	promotor N25	GCGCAACGCAATTAAATGTAATCATAAAAAATTTATTT GCTTTCAGGAAAATTTTCTG	45
YZ708	promotor N25	AACAAAAATTATTTCTACAGGTGAATCTATTATACAGA AAAAATTTTCTGAAAGCAAATA	46
YZ709	promotor λ	GCGCAACGCAATTAAATGTAATTATCTCTGGCGGTGTT GACATAAATACCACTGGCG	47
YZ710	promotor λ	AACAAAAATTATTTCTACAGGTGTGCTCAGTATCACCG CCAGTGGTATTATGTCAAC	48
YZ711	Promotor H207	GCGCAACGCAATTAAATGTAATTTTAAAAAATTCATTT GCTAAACGCTTCAAATTCCTCG	49
YZ712	Promotor H207	AACAAAAATTATTTCTACAGGTGAAGTATATTATACGA GAATTGAAAGCGTTTACG	50
Clonación de plásmidos de destino pUCX y pBRX Golden Gate			
pUC19-lin-fwd	Eliminación del sitio TK1/línealización de pBRX1	TGGGGTGCCTAATGAGTGAGC	51

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
pUC19-lin-rev	Eliminación del sitio TK1/linealización de pBRX1	CTATGACCATGATTACGCCAAGCTTG	52
YZ51	Mutación del sitio Bsal <i>b/a</i>	TCTCGCGGTATCATTCGAGCACTG	53
YZ52	Mutación del sitio Bsal <i>b/a</i>	GCCACGCTCACCGGCTCC	54
YZ95	Linealización de pUCX2/pBRX2	AACGAAAACTCACGTTAAGGG	55
YZ96	Linealización de pUCX2/pBRX2	CCACTGAGCGTCAGACC	56
YZ93	Casete de relleno Bsal <i>zeo</i> ^R	GAGACCGTGTGACAATTAATCATCGGC	57
YZ94	Casete de relleno Bsal <i>zeo</i> ^R	GAGACCATTTCTCACCAATAAAAAACGCCCGG	58
Clonación de pCas9 y pAlO, integración cromosómica de Cas9			
JL126	Clonación de pCas9	CGGGGTACCATGAGACAAGAAGTACTCCATT	59
JL128	Clonación de pCas9	CTAGTCTAGATTACACCTTCTCTCTTCTTGGG	60
BL557	Eliminación del sitio pCas9 BsmBI	CTCCGGGGAACCCGCCGAAGCCACGCGGCTCAA	61
BL558	Eliminación del sitio pCas9 BsmBI	CTTCGGCGGTTTCCCGGAGTCGAACAGGAGGGCGCC AATGAGG	62
BL559	Clonación de pCas9-Multi	AGGAAGAAGACGTCTCACGCATCTTACTGCGCAGATA CGC	63
BL560	Clonación de pCas9-Multi	AAGATGCGTGAGACGTCTTCTTCTCGTCTCGGTCGAC AGTTTCATAGGTGATTGCTCAGG	64
YZ720	Integración de <i>arsB::CasQ</i>	GTCCCAATCGCAGCCAATCACATTG	65
YZ721	Integración de <i>arsB::Cas9</i>	GTCCGTACCATCGTATTGGTTATCTGGC	66
TG1	Secuenciación de Cas9	ATTAGAGGGCAGTGCCAGCTCGTTA	67
TG2	Secuenciación de Cas9	CTGCATTAGGTAGGCATCATGCGCA	68
TG3	Secuenciación de Cas9	CTGGGCTACCTGCAAGATTAGCGATG	69

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
TG4	Secuenciación de Cas9	TGAAGGACTGGGCAGAGGCCCCCTT	70
TG5	Secuenciación de Cas9	CGTAGGTGTCCTTTGCTCAGTTGAAGC	71
TG6	Secuenciación de Cas9	TAGCCATCTCATTACTAAAGATCTCCT	72
BL731	Clonación de pAIO-Multi (Δ Cas9, introducción de BsaI)	CGATAICGTTGGTCTCAACGACACAATTGTAAAGGTTAGATCT	73
BL732	Clonación de pAIO-Multi (Δ Cas9, introducción de BsaI)	CAACGATATCGGTCTCACACTGACTGGGCCCTTTCGTTT TATCT	74
BL450	Secuenciación de guías de pAIO	GCAATCACCTATGAACCTGTGCAC	75
Clonación de guías de Cas9			
la letra mayúscula indica la secuencia de reconocimiento BsmBI (6 nt), el saliente del sitio de restricción BsmBI (4 nt) o la secuencia diana variable (espaciador) (18-20 nt)			
BL691	1 ^{er} ARN _{gu} GG AGXG-T	aggaggagggaCGTCTCaTGGGccccgcgcattCACACAAATGTAGTGATCAgttttagagctagaataagc	76
BL627	1 ^{er} ARN _{gu} GG AGXA-T	aggaggagggaCGTCTCaTGGGccccgcgcattCCAGGATGGGTAC CACCgttttagagctagaataagc	77
BL707	1 ^{er} ARN _{gu} GG AGXC-T	aggaggagggaCGTCTCaTGGGccccgcgcattCACACAAATGTAGTCAgttttagagctagaataagc	78
BL642	1 ^{er} ARN _{gu} GG AGXT-G	aggaggagggaCGTCTCaTGGGccccgcgcattCCAGGATGGGCAC CAGCCgttttagagctagaataagc	79
BL659	1 ^{er} ARN _{gu} GG A AXG-T/CXT-T	aggaggagggaCGTCTCaTGGGccccgcgcattCCATGATGGGCACC ACCCgttttagagctagaataagc	80
BL623	1 ^{er} ARN _{gu} GG AAXA-T	aggaggagggaCGTCTCaTGGGccccgcgcattCACACAAATGTATAGATCAgttttagagctagaataagc	81

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
BL628	1 ^{er} ARN _{gu} GG AAXC-T	aggaggaaggacggtcttcaTGGGccccgcattCCAGGATGGGCAC CATCCgttttagagctagaataagc	82
BL567	1 ^{er} ARN _{gu} GG AAXT-A	aggaggaaggacggtcttcaTGGGccccgcattGTTGTGTGGAAATG TGAAGtttttagagctagaataagc	83
BL593	1 ^{er} ARN _{gu} GG ACXG-G	aggaggaaggacggtcttcaTGGGccccgcattTGTCACTACTCTTGA CCAGgttttagagctagaataagc	84
BL639	1 ^{er} ARN _{gu} GG ACXA-A	aggaggaaggacggtcttcaTGGGccccgcattTGTCACTACTCTTGA CCAAgttttagagctagaataagc	85
BL693	1 ^{er} ARN _{gu} GG ACXC-T	aggaggaaggacggtcttcaTGGGccccgcattCACACAATGTACTC ATCAgttttagagctagaataagc	86
BL660	1 ^{er} ARN _{gu} GG ACXT-A	aggaggaaggacggtcttcaTGGGccccgcattCCAAGATGGGCAC CACCCgttttagagctagaataagc	87
BL695	1 ^{er} ARN _{gu} GG ATXG-T	aggaggaaggacggtcttcaTGGGccccgcattCACACAATGTATTG ATCAgttttagagctagaataagc	88
BL629	1 ^{er} ARN _{gu} GG ATXA-A	aggaggaaggacggtcttcaTGGGccccgcattCACACAATGTATAA ATCAgttttagagctagaataagc	89
BL657	1 ^{er} ARN _{gu} GG ATXC-G	aggaggaaggacggtcttcaTGGGccccgcattCCAGGATGGGGAC CACCCgttttagagctagaataagc	90
BL620	1 ^{er} ARN _{gu} GG ATXT-T	aggaggaaggacggtcttcaTGGGccccgcattTTCACAATACTTTC TTTAgttttagagctagaataagc	91
BL701	2 ^o ARN _{gu} , Fwd GΔG	gcattTCACACAATGTAGGATCAGtttttagagctagaataagc	92
BL702	2 ^o ARN _{gu} , Rev GΔG	TGATCCTACATTGTGTGAaatcgggggcgcatcttact	93
BL617	2 ^o ARN _{gu} , Fwd GΔA	gcattACCAAGGATGGGACCACCCcgtttttagagctagaataagc	94

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
BL618	2° ARNgu, Rev GΔA	GGGTGGTCCCATCTGGTaatgcgggcgcatcttact	95
BL705	2° ARNgu, Fwd GΔC	gcattTCACACAAATGTAGCATCAGtttttagagctagaaatagc	96
BL706	2° ARNgu, Rev GΔC	TGATGCTACATTGTGTGAaatgcgggcgcatcttact	97
BL614	2° ARNgu, Fwd GΔT	gcattACCAAGGATGGGCACCAACCGtttttagagctagaaatagc	98
BL615	2° ARNgu, Rev GΔT	GGTGGTGCCCATCTGGTaatgcgggcgcatcttact	99
BL682	2° ARNgu, Fwd AΔG	gcattACCAAGATGGGCACCAACCGtttttagagctagaaatagc	100
BL683	2° ARNgu, Rev AΔG	GGGTGGTGCCCATCTGGTaatgcgggcgcatcttact	101
BL575	2° ARNgu, Fwd AΔA	gcattTCACACAAATGTAAGATCAGtttttagagctagaaatagc	102
BL576	2° ARNgu, Rev AΔA	TGATCTTACATTGTGTGAaatgcgggcgcatcttact	103
BL564	2° ARNgu, Fwd AΔT	gcattTGTGTGTGGAAATGTAGGtttttagagctagaaatagc	104
BL565	2° ARNgu, Rev AΔT	CTCACATTCCACACAAACAAatgcgggcgcatcttact	105
BL675	2° ARNgu, Fwd CΔG	gcattTTGTCACTACTCTGACCGgttttagagctagaaatagc	106
BL676	2° ARNgu, Rev CΔG	CGGTCAGAGTAGTGACAAaatgcgggcgcatcttact	107
BL673	2° ARNgu, Fwd CΔA	gcattTTGTCACTACTCTGACCAaatgctagagctagaaatagc	108
BL674	2° ARNgu, Rev CΔA	TGGTCAGAGTAGTGACAAaatgcgggcgcatcttact	109
BL703	2° ARNgu, Fwd CΔC	gcattTCACACAAATGTACCATCAGtttttagagctagaaatagc	110
BL704	2° ARNgu, Rev CΔC	TGATGGTACATTGTGTGAaatgcgggcgcatcttact	111
BL697	2° ARNgu, Fwd TΔG	gcattTCACACAAATGTATGATCAGtttttagagctagaaatagc	112
BL698	2° ARNgu, Rev TΔG	TGATCATACATTGTGTGAaatgcgggcgcatcttact	113
BL679	2° ARNgu, Fwd TΔA	gcattTCACACAAATGTATAATCAGtttttagagctagaaatagc	114
BL680	2° ARNgu, Rev TΔA	TGATTATACATTGTGTGAaatgcgggcgcatcttact	115
BL620	2° ARNgu, Fwd TΔT	gcattATTACACAAATACCTTTCAGtttttagagctagaaatagc	116

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
BL621	2º ARNgu, Rev TΔT	TAAAGAAGTATTGTGAATaatgcggggcgcatcttact	117
BL562	1ª construcción de ARNgu	agaaggaagaCGTCTCaCTGTcgaccacaaagcgtctctgtgagc	118
BL563	2ª construcción de ARNgu	aagaaggaCGTCTCaACAGtagtggcagcgctaactaag	119
BL566	Terminación de la construcción de ARNgu	aggagaggaCGTCTCtCGACcaaaaaagcgtctctgttgagcag	120
BL514	Clonación de guías de hEGFP natural	agtaagatgcgccccgcattGACCAGGATGGGCACCACCggttttaga gctaagaataag	121
BL515	Clonación de guías de hEGFP natural	ctatttctagctctaaaaacGGGTGGTGCCCATCCTGGTCaatgcggggc gcatcttact	122
BL464	Clonación de guías de hEGFP natural	GTTTAGAGCTAGAAAATAGCAAGTTAAAAATAAGGCTA GT	123
BL465	Clonación de guías de hEGFP natural	AATGCGGGGCGCATCTTACT	124
YZ310	ESer-69 guía fwd	cattGGCACCCGGTCTACTAAAC	125
YZ316	ESer-69 guía rev	aaacGTTTAGTAGACCCGGTGCC	126
YZ359	GFP151-69 guía fwd	cgcattCACACAATGTAAGTATCAgtttt	127
YZ360	GFP51-69 guía rev	ctctaaacTGATACTTACATTGTGTGaa	128
Moldes de ADN de escisión <i>in vitro</i> (IVC) y ARNgu			
BL487	Molde de ADN IVC	GTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACC AACCCGGTGAAACAGCTCCTCGCC	129
BL488	Molde de ADN IVC	GTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACC AGCCCGGTGAACAGCTCCTCGCC	130
BL489	Molde de ADN IVC	GTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACC ATCCCGGTGAACAGCTCCTCGCC	131

(continuación)			SEQ ID NO:
Cebador	Aplicación	Secuencia	
BL408	Molde de ADN IVC	GTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACC ACCCCGGTGAACAGCTCCTCGCC	132
BL415	Extensión del molde IVC	GCATCGCCCTCGCCCTCGCCGGACACGCTGAACCTTGT GGCCGTTTACGTCGCCGTCCAGC	133
BL416	Extensión del molde IVC	TGCAGTTTCAATTGATGCTCGATGAGTTATGTTGAGCA AGGGCGAGGAGCTGTTCAACCGG	134
BL484	ARNcr IVC	TTAATACGACTCACTATAGGGACCCAGGATGGGCACCA ACCGTTTATAGAGCTAGAAATAGCAAAGTTAAAAATAAGG CTAGT	135
BL485	ARNcr IVC	TTAATACGACTCACTATAGGGACCCAGGATGGGCACCA GCCGTTTATAGAGCTAGAAATAGCAAAGTTAAAAATAAGG CTAGT	136
BL486	ARNcr IVC	TTAATACGACTCACTATAGGGACCCAGGATGGGCACCA TCCGTTTATAGAGCTAGAAATAGCAAAGTTAAAAATAAGG CTAGT	137
BL318	ARNcr IVC	TTAATACGACTCACTATAGGGACCCAGGATGGGCACCA CCCGTTTATAGAGCTATGCTGTTTG	138
BL439	Conversión IVC de ARNcr en ARN _{Ng}	AAAAGCACCGACTCGGTGCCACTTTTTCAGTTGATA ACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCT AAAACGG	139
BL472	ARN _{Ng} de PCR	AAGAGGAAGAGGTTAATACGACTCACTATAGGGGAC	140
BL440	ARN _{Ng} de PCR	AAAAGCACCGACTCGGTGCC	141
BL473	ARN _{Ng} de PCR	ACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAA CGG	142

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
Cebadores para las inserciones Golden Gate que contienen UBP			
la letra en cursiva indica la secuencia de reconocimiento Bsal, el subrayado indica el sitio de restricción Bsal			
secuencia saliente			
los cebadores de inserción para pUCX2 también son compatibles con los plásmidos pBRX2 y pAIO GG			
# representa cebadores que también se usaron para el desplazamiento de biotina (los cebadores de desplazamiento de biotina dedicados correspondientes son idénticos en la secuencia de hibridación)			
YZ401 e YZ403 son muy sensibles a la temperatura de hibridación. La temperatura óptima de hibridación es de 54 °C			
BL528	Inserción de hEGFP para pUCX2 GG fwd	AAGAAAGGAGAA <u>GGTCTCT</u> <u>CT</u> AGTGGAGCAAGGGCGAGG AGCTGTTCAACCG	143
BL529	Inserción de hEGFP para pUCX2 GG rev	AAGAGAAGAGAGAGG <u>CTCTCA</u> <u>TC</u> ATCGTGTTTACGTGCGCGTC CAGCTC	144
YZ148	Inserción de GFP66 para pUCX2 GG fwd	ATGGGTCTCCAGTGGGGGCCAACACTTGTCACTAC	145
YZ149	Inserción de GFP66 para pUCX2 GG rev	ATGGG <u>TC</u> CTT <u>CGTT</u> CCGGATAACGGGAAAAGC	146
YZ150	Inserción de GFP151 para pUCX2 GG fwd	ATGGG <u>TC</u> <u>TC</u> CA <u>GTGGCTCGAGTACA</u> CACTTTAACTCACA C	147
YZ151	Inserción de GFP151 para pUCX2 GG rev	ATGGGTCCTTCGTTGATTCATCTTTTGTGTCTGC	148
YZ97	Inserción de TK1 para pUCX1 GG fwd	ATGGG <u>TC</u> <u>TC</u> TCATAGCTGTTTCTGTGTGAAAATTGTTA TCC	149
YZ98	Inserción de TK1 para pUCX1 GG rev	ATGGGTC <u>TC</u> ACCCCAAGGCTTTACACTTTATGCTTCCG	150
YZ99#	Inserción de TK1 para pUCX2 GG fwd	ATGGG <u>TC</u> <u>TC</u> CA <u>GTGGCTGTTTCT</u> GTGTGAAAATTGTTA TCC	151
YZ100#	Inserción de TK1 para pUCX2 GG rev	ATGGGTCCTTCGTTGGCTTTACACTTTATGCTTCCG	152
YZ118	Inserción de D8 para pUCX2 GG fwd	ATGGGTC <u>TC</u> CA <u>GTGGCACACAGGAA</u> ACAGCTATGAC	153

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
YZ119	Inserción de D8 para pUCX2 GG rev	ATGGGTCCTCTCGTTGGTTAAGCTTAACTTTAAGAAG GAG	154
YZ73#	Inserción de GFP151 para pAI02X GG fwd	ATGGGTCCTCACACA ~~AA~~ ACTCGAGTACAACTTTAACTCA CAC	155
YZ74#	Inserción de GFP151 para pAI02XGG rev	ATGGGTCCTCGATTCCATTCTTTTGTGTCTGC	156
YZ401#	Inserción de ESer para pAI02X GG fwd	ATTGGTCTCGGCCGAGCGGTTGAAGGCAC	157
YZ403#	Inserción de ESer para pAI02X GG rev	ATTGGTCTCTCTGGAACCCCTTTCGGGTCG	158
Cebadores de desplazamiento de biotina			
Temperatura de hibridación (°C) indicada entre paréntesis			
BL745	hEGFP fwd (48 x 3 ciclos, 54 x 20 ciclos)	GGCGAGGAGCTGTTACCCG	159
BL744	hEGFP rev (48 x 3 ciclos, 54 x 20 ciclos)	GTTTACGTCGCCGTCAGCTC	160
BL750	GFP66 fwd (50)	GGCCAACACTTGTCACTACT	161
BL751	GFP66 rev (50)	TCCGGATAACGGGAAAAAGC	162
YZ351	GFP151 fwd (50)	CTCGAGTACAACCTTTAACTCACAC	163
YZ352	GFP151 rev (50)	GATTCCATTCTTTTGTGTCTGC	164
BL748	TK1 fwd (50)	CTGTTTCCTGTGTGAAATTGTTATCC	165
BL749	TK1 rev (50)	GGCTTTACACTTTATGCTTCCG	166
BL774	D8 fwd (50)	CCCGGGTTATTACATGCGCTAGCACT	167
BL775	D8 rev (50)	GAAATTAAATACGACTCACTATAGGGTTAAGCTTAACT TTAAGAAGGAG	168
YZ17	pUCX2 fwd (60)	TGCAAGCAGCAGATTACGCGC	169

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
YZ18	pUCX2 rev (60)	GTAAGCTGTCAGACCAAGTTTACTC	170
Oligonucleótidos molde de UB			
* representa las secuencias usadas en los experimentos de Cas9			
X representa dNaM			
Nombre	Tipo de secuencia	Secuencia	
BL410*	hEGFP	GGCGAGGAGCTGTTTACCCGGGXTGGTGCCCATCCTGG TCGAGCTGGACGGCGACGTAAAC	171
BL413*	hEGFP	GTTTACGTCGCCGTCACAGCTCGACAGGATGGGXACC ACCCCGGTGAACAGCTCCTCGCC	172
BL411*	hEGFP	GTTTACGTCGCCGTCACAGCTCGACCAAGATGGGCACC ACCCCGGTGAACAGCTCCTCGCC	173
BL409*	hEGFP	GTTTACGTCGCCGTCACAGCTCGACAGGATGGGCACC AXCCCGGTGAACAGCTCCTCGCC	174
TK1*	TK1	CTGTTTCTCTGTGTGAAAATTGTTATCCGCTCACAXTTCC ACACAACATACGAGCCGGAAAGCATAAAGTGTAAGC C	175
DM510-16*	GFP151	CTCGAGTACAACCTTTAACTCACACAATGTAXAGATCA CGGCAGACAAACAAAGAAATGGAATC	176
DM510-13*	GFP66	CCGGATAACGGGAAAAGCAATTGAACACCCGXGGTCA GAGTAGTGACAAGTGTGGCCA	177
DM510-11*	GFP66	GGCCAAACACTTGTCACTACTCTGACCXAGGGTGTTC ATGCTTTTCCCGTTATCCGGA	178
BL412*	hEGFP	GGCGAGGAGCTGTTTACCCGGGTTGGTGCCCATCXTGG TCGAGCTGGACGGCGACGTAAAC	179

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
BL414*	hEGFP	GGCGAGGAGCTGTTACCCGGGTGGTXCCCATCCTGG TCGAGCTGGACGGCGACGTAAAC	180
DM510-20*	GFP151	GATTCACATCTTTTGTCTGCGGTGATTXATACATT GTGTGAGTTAAAGTTGTAATCGAGT	181
D8*	D8	CACACAGGAAACAGCTATGACCCGGGTTATTACATGC GCTAGCACTTGGAAATTCACAAATACXTCTTTAAGGAA ACCATAGTAAATCTCTTCTTAAAGTTAAGCTTAACCC TATAGTGAGTCGTATTAAATTTC	182
GFP151 -33	GFP151	CTCGAGTACAACTTTAACTCACACAATGTAAAXAATCA CGGCAGACAAACAAGAAATGGAAATC	183
GFP151 -35	GFP151	CTCGAGTACAACTTTAACTCACACAATGTAAAXCATCA CGGCAGACAAACAAGAAATGGAAATC	184
GFP151 -37	GFP151	CTCGAGTACAACTTTAACTCACACAATGTAAAXGATCA CGGCAGACAAACAAGAAATGGAAATC	185
GFP151 -39	GFP151	CTCGAGTACAACTTTAACTCACACAATGTCAXTGTCA CGGCAGACAAACAAGAAATGGAAATC	186
GFP151 -41	GFP151	CTCGAGTACAACTTTAACTCACACAATGTACXAATCA CGGCAGACAAACAAGAAATGGAAATC	187
GFP151 -43*	GFP151	CTCGAGTACAACTTTAACTCACACAATGTACXCATCA CGGCAGACAAACAAGAAATGGAAATC	188
DM510 -19	GFP151	CTCGAGTACAACTTTAACTCACACAATGTACXGATCA CGGCAGACAAACAAGAAATGGAAATC	189
GFP151 -47	GFP151	CTCGAGTACAACTTTAACTCACACAATGTACXTATCA CGGCAGACAAACAAGAAATGGAAATC	190

(continuación)

Cebador	Aplicación	Secuencia	SEQ ID NO:
GFP151 -49	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTAGXAATCA CGGCAGACAAAACAAAAGAAATGGAATC	191
GFP151 -51*	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTAGXCATCA CGGCAGACAAAACAAAAGAAATGGAATC	192
GFP151 -53*	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTAGXGATCA CGGCAGACAAAACAAAAGAAATGGAATC	193
GFP151 -55	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTAGXTATCA CGGCAGACAAAACAAAAGAAATGGAATC	194
GFP151 -57	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTATXAATCA CGGCAGACAAAACAAAAGAAATGGAATC	195
GFP151 -59	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTATXCATCA CGGCAGACAAAACAAAAGAAATGGAATC	196
GFP151 -61*	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTATXGATCA CGGCAGACAAAACAAAAGAAATGGAATC	197
GFP151 -63	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTATXTATCAC GGCAGACAAAACAAAAGAAATGGAATC	198
GFP151 -69*	GFP151	CTCGAGTACAACTTTTAACTCACACAATGTAAAGXATCA CGGCAGACAAAACAAAAGAAATGGAATC	199
ESer-69*	ESer	CTCTGGAACCCCTTTCGGGTCGCCGGTTTAGXAGACCG GTGCC'TTCAACCGCTCGGC	200

La Tabla 3 ilustra las secuencias del péptido señal descritas en el presente documento.

ID de péptido de señal	Secuencias	SEQ ID NO:
pelB_SP1	MKYLLPTAEAGLLLLAAQPAIA	225
malE_SP2	MKIKTGARILALSELTMMFSASALA	226
phoA_SP3	M KQSTIALALLPLLFTPVTKA	227
treA_SP4	MKSPAPSRPQKMALIPACIFLCFAALSVQA	228
pcoE_SP5	M KKILVSFVAIMAVASSAMA	229
Quitosanasa (Csn) - SP6	MKISMQKADFWKKAAILLVFTMFFTLMMSETVFAAGL NK	230
OmpA_SP7	MKKTAAIAVALAGFATVAQASAGLNKD	231
DsbAss	Proteína: MKKIWLALAGLVLAFSASA	232
	ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTT AGCGTTTAGCGCATCGGCG ADN:	233
PelBss	Proteína: MKYLLPTAAAGLLLLAAQPAMA	234
	ATGAAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCT GCTCCTCGCTGCCAGCCGGCGATGGCG ADN:	235
PhoAss	Proteína: MKQSTIALALLPLLFTPVTKA	227
	ATGAAACAAAGCACTATTGCACTGGCACTCTTACCGT TACTGTTTACCCCTGTGACAAAAGCG ADN:	236
NTss	Proteína: MKTHIVSSVTTLGLLSILMNPV ANA	237
	ATGAAAACACATATAGTCAGCTCAGTAACAACAACAC TATTGCTAGGTTCCATATAATGAATCCTGTCGCTAAT GCC ADN:	238
NSP1	Proteína: MKYLLPWLALAGLVLAFSASA	239
	ATGAAATACCTGCTGCCGTGGCTGGCGCTGGCTGGTT TAGTTTTAGCGTTTAGCGCATCGGCG ADN:	240
NSP2	Proteína: MKKITAAAGLLLLAAFSASA	241
	ATGAAAAAGATTACCGCTGCTGCTGGTCTGCTGCTCCT CGCTGCGTTTAGCGCATCGGCG ADN:	242
NSP3	Proteína: MKKIWLALAGLVLAQPAMA	243
	ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTT AGCCAGCCGGCGATGGCG ADN:	244
NSP3a	Proteína: MKKILVLGALALWAQPAMA	245
	ATGAAAAAGATTTTAGTTTTAGGTGCTCTGGCGCTGTG GGCCAGCCGGCGATGGCG ADN:	246
NSP3b	Proteína: MKKIWLALVLLAGAQPAMA	247
	ATGAAAAAGATTTGGCTGGCGTTAGTTTTACTGGCTG GTGCCAGCCGGCGATGGCG ADN:	248
NSP3c	Proteína: MKKILAGWLALVLAQPAMA	249
	ATGAAAAAGATTCTGGCTGGTTGGCTGGCGTTAGTTTT AGCCAGCCGGCGATGGCG ADN:	250
NSP3d	Proteína: MKKILVLLAGWLAAQPAMA	251
	ATGAAAAAGATTTTAGTTTTACTGGCTGGTTGGCTGGC GGCCAGCCGGCGATGGCG ADN:	252
NSP4	Proteína: MKKITAAAGLLLLAAQPAMA	253
	ATGAAAAAGATTACCGCTGCTGCTGGTCTGCTGCTCCT CGCTGCCAGCCGGCGATGGCG ADN:	254
NSP4a	Proteína: MKKILLLLGTAAAAAQPAMA	255
	ATGAAAAAGATTCTGCTGCTCCTCGGTACCGCTGCTG CTGCTGCCAGCCGGCGATGGCG ADN:	256
NSP4b	Proteína: MKKILLLLLLLLLAQPAMA	257
	ATGAAAAAGATTCTGCTGCTCCTCCTGCTGCTCCTCCT GCTCGCCAGCCGGCGATGGCG ADN:	258

(continuación)

ID de péptido de señal	Secuencias	SEQ ID NO:
NSP4c	Proteína: MKKIAAAAAAAAAAQPAMA	259
	ATGAAAAAGATTGCTGCTGCTGCTGCGGCGGCGGCGG CTGCGGCCAGCCGCGCATGGCG ADN:	260
NSP5	Proteína: MKYLLPWLALAGLVLAQPAMA	261
	ATGAAATACCTGCTGCCGTGGCTGGCGCTGGCTGGTT TAGTTTTAGCCCAGCCGCGCATGGCG ADN:	262
NSP6	Proteína: MKYLLPTAAAGLLLLAASFASA	263
	ATGAAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCT GCTCCTCGCTGCGTTTAGCGCATCGGCG ADN:	264

Si bien se han mostrado y descrito realizaciones preferidas de la presente divulgación en el presente documento, será obvio para los expertos en la materia que tales realizaciones se proporcionan solamente a modo de ejemplo. Numerosas variaciones, cambios y sustituciones se les ocurrirán ahora a los expertos en la materia sin apartarse de la divulgación. Debe entenderse que se pueden emplear diversas alternativas a los casos de la divulgación descritos en el presente documento para llevar a la práctica la divulgación. Se pretende que las siguientes reivindicaciones definan el alcance de la divulgación y que los métodos y estructuras dentro del alcance de estas reivindicaciones y sus equivalentes estén cubiertos de ese modo.

LISTADO DE SECUENCIAS

<110> THE SCRIPPS RESEARCH INSTITUTE

<120> TRANSPORTADOR DE NUCLEÓSIDO TRIFOSFATO NOVEDOSO Y USOS DEL MISMO

<130> 46085-707.601

<140>

<141>

<150> 62/354.650

<151> 24/06/2016

<160> 264

<170> PatentIn versión 3.5

<210> 1

<211> 575

<212> PRT

<213> Phaeodactylum tricornutum

<400> 1

ES 2 929 047 T3

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

Met Gly Lys Lys Met Leu Phe Tyr Ser Thr Cys Ile Pro Phe Phe Thr
 165 170 175
 Phe Phe Gly Leu Phe Asp Val Phe Ile Tyr Pro Asn Ala Glu Arg Leu
 180 185 190
 His Pro Ser Leu Glu Ala Val Gln Ala Ile Leu Pro Gly Gly Ala Ala
 195 200 205
 Ser Gly Gly Met Ala Val Leu Ala Lys Ile Ala Thr His Trp Thr Ser
 210 215 220
 Ala Leu Phe Tyr Val Met Ala Glu Ile Tyr Ser Ser Val Ser Val Gly
 225 230 235 240
 Leu Leu Phe Trp Gln Phe Ala Asn Asp Val Val Asn Val Asp Gln Ala
 245 250 255
 Lys Arg Phe Tyr Pro Leu Phe Ala Gln Met Ser Gly Leu Ala Pro Val
 260 265 270
 Leu Ala Gly Gln Tyr Val Val Arg Phe Ala Ser Lys Ala Val Asn Phe
 275 280 285
 Glu Ala Ser Met His Arg Leu Thr Ala Ala Val Thr Phe Ala Gly Ile
 290 295 300
 Met Ile Cys Ile Phe Tyr Gln Leu Ser Ser Ser Tyr Val Glu Arg Thr
 305 310 315 320
 Glu Ser Ala Lys Pro Ala Ala Asp Asn Glu Gln Ser Ile Lys Pro Lys
 325 330 335
 Lys Lys Lys Pro Lys Met Ser Met Val Glu Ser Gly Lys Phe Leu Ala
 340 345 350
 Ser Ser Gln Tyr Leu Arg Leu Ile Ala Met Leu Val Leu Gly Tyr Gly
 355 360 365
 Leu Ser Ile Asn Phe Thr Glu Ile Met Trp Lys Ser Leu Val Lys Lys
 370 375 380
 Gln Tyr Pro Asp Pro Leu Asp Tyr Gln Arg Phe Met Gly Asn Phe Ser
 385 390 395 400
 Ser Ala Val Gly Leu Ser Thr Cys Ile Val Ile Phe Phe Gly Val His
 405 410 415

Val Ile Arg Leu Leu Gly Trp Lys Val Gly Ala Leu Ala Thr Pro Gly
420 425 430

Ile Met Ala Ile Leu Ala Leu Pro Phe Phe Ala Cys Ile Leu Leu Gly
435 440 445

Leu Asp Ser Pro Ala Arg Leu Glu Ile Ala Val Ile Phe Gly Thr Ile
450 455 460

Gln Ser Leu Leu Ser Lys Thr Ser Lys Tyr Ala Leu Phe Asp Pro Thr
465 470 475 480

Thr Gln Met Ala Tyr Ile Pro Leu Asp Asp Glu Ser Lys Val Lys Gly
485 490 495

Lys Ala Ala Ile Asp Val Leu Gly Ser Arg Ile Gly Lys Ser Gly Gly
500 505 510

Ser Leu Ile Gln Gln Gly Leu Val Phe Val Phe Gly Asn Ile Ile Asn
515 520 525

Ala Ala Pro Val Val Gly Val Val Tyr Tyr Ser Val Leu Val Ala Trp
530 535 540

Met Ser Ala Ala Gly Arg Leu Ser Gly Leu Phe Gln Ala Gln Thr Glu
545 550 555 560

Met Asp Lys Ala Asp Lys Met Glu Ala Lys Thr Asn Lys Glu Lys
565 570 575

<210> 2

<211> 1536

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Polinucleótido sintético

<400> 2

```

atgggtggta gcaccgttgc accgaccaca ccgctggcaa ccggtgggtgc actgcgtaaa      60
gttcgtcagg cagtttttcc gatttatggc aatcaagaag tgaccaaatt tctgctgatt      120
ggcagcatca aattctttat tatcctggca ctgaccctga cccgtgatac caaagatacc      180
ctgattgtta cccagtgtgg tgcagaagca attgcatttc tgaaaatcta tgggtgttctg      240
cctgcagcaa ccgcatttat tgcactgtat agcaaaatga gcaacgcaat gggcaaaaaa      300
atgctgtttt atagcacctg tatcccgctt tttacctttt ttggtctgtt cgatgtgttc      360

```

atztatccga atgccgaacg tctgcatccg agcctggaag cagttcaggc aattctgcct	420
ggtggtgccg caagcgggtg tatggcagtt ctggcaaaaa ttgcaacca ttggaccagc	480
gcactgtttt atgttatggc agaaatctat agcagcgtta gcgttggtct gctgttttgg	540
cagtttgcaa atgatgttgt taatgtggat caggccaaac gtttttatcc gctgtttgca	600
cagatgagcg gtctggcacc ggttctggca ggtcagtatg ttgttcgttt tgcaagcaaa	660
gccgttaatt ttgaagcaag catgcatcgt ctgaccgcag cagttacctt tgcaagcatt	720
atgatctgca tcttttatca gctgagcagc tcatatgttg aacgtaccga aagcgcaaaa	780
ccggcagcag ataatgaaca gagcattaaa ccgaagaaaa aaaaaccgaa aatgtcgatg	840
gtggaaagcg gtaaatttct ggcaagcagc cagtatctgc gtctgattgc aatgctgggt	900
ctgggttatg gtctgagcat taactttacc gaaatcatgt ggaaaagcct ggtgaaaaaa	960
cagtatccgg atccgctgga ttatcagcgt tttatgggta attttagcag cgcaagttgt	1020
ctgagtacct gcattgttat cttttttggc gtgcatgtta ttcgtctgct gggttgaaaa	1080
gttggtgcc tggcaacacc ggttattatg gccattctgg cactgccgtt ttttgcattg	1140
attctgctgg gcctggatag tccggcacgt ctggaaattg cagttatttt tggcaccatt	1200
cagagcctgc tgagcaaaac cagcaaatat gcaactgttg atccgaccac ccagatggca	1260
tatatccgc tggtgatga aagcaaagtt aaaggcaaag cagccattga tgttctgggt	1320
agccgtattg gtaaatacag tggtagcctg attcagcagg gtctggtttt tgtttttggc	1380
aatattatca atgccgcacc ggttgttggg gttgtgtatt atagcgttct ggttgcattg	1440
atgagtgcag caggtcgtct gagggtctg tttcaggcac agaccgaaat ggataaagca	1500
gataaaatgg aagccaaaac caacaaagaa aaatga	1536

<210> 3
 <211> 1536
 <212> ADN
 <213> Phaeodactylum tricornutum
 <400> 3

5

atgggaggca gtactgttgc accaactaca ccgttggcaa ccggcgggtgc gctccgcaaa	60
gtgcgacaag ccgtctttcc catctacgga aaccaagaag tcaccaaatt tctgctcatc	120
ggatccatta aattctttat aatcttggca ctcacgctca cgcgtgatac caaggacacg	180
ttgattgtca cgcaatgtgg tgccgaagcg attgcctttc tcaaaatata cggggtgcta	240
cccgcagcga ccgcatttat cgcgctctat tccaaaatgt ccaacgccat gggcaaaaaa	300
atgctatttt attccacttg cattcctttc tttacctttt tcgggctggt tgatgttttc	360
atttaccga acgcggagcg actgcaccct agtttggaag ccgtgcaggc aattctcccg	420
ggcggtgccg catctggcgg catggcgggt ctggccaaga ttgcgacaca ctggacatcg	480
gccttatttt acgtcatggc ggaaatatat tcttccgtat cggtggggct attgttttgg	540
cagtttgca acgacgtcgt caacgtggat caggccaagc gcttttatcc attatttgct	600
caaatgagtg gcctcgctcc agttttagcg ggccagtatg tggtagcggt tgccagcaaa	660
gcggtcaact ttgaggcata catgcatacga ctcacggcgg ccgtaacatt tgctggtatt	720
atgatttgca tcttttacca actcagttcg tcatatgtgg agcgaacgga atcagcaaaag	780
ccagcggcag ataacgagca gtctatcaaa ccgaaaaaga agaaacccaa aatgtccatg	840
gttgaatcgg ggaaatttct cgcgtcaagt cagtacctgc gtctaattgc catgctggtg	900
ctgggatacg gcctcagtat taactttacc gaaatcatgt ggaaaagctt ggtgaagaaa	960
caatatccag acccgctaga ttatcaacga tttatgggta acttctcgtc agcgggttgg	1020
ttgagcacat gcattgttat tttcttcggg gtgcacgtga tccgtttggt ggggtggaaa	1080
gtcggagcgt tggctacacc tgggatcatg gccattctag cgttaccctt ttttgcttgc	1140
attttggttg gtttgatag tccagcacga ttggagatcg ccgtaatctt tggaacaatt	1200
cagagtttgc tgagcaaaac ctccaagtat gcccttttcg accctaccac acaaattggct	1260
tatattcctc tggacgacga atcaaaggtc aaaggaaaag cggcaattga tgttttgga	1320
tcgcggattg gcaagagtgg aggtcactg atccagcagg gcttgggtct tgttttgga	1380
aatatcatta atgccgcacc tgtagtaggg gttgtctact acagtgtcct tgttgcggtg	1440
atgagcgcag ctggccgact aagtgggctt tttcaagcac aaacagaaat ggataaggcc	1500
gacaaaatgg aggcaaagac caacaaagaa aagtag	1536

<210> 4
 <211> 511
 <212> PRT
 <213> Phaeodactylum tricornutum

<400> 4

ES 2 929 047 T3

Met Gly Gly Ser Thr Val Ala Pro Thr Thr Pro Leu Ala Thr Gly Gly
1 5 10 15

Ala Leu Arg Lys Val Arg Gln Ala Val Phe Pro Ile Tyr Gly Asn Gln
20 25 30

Glu Val Thr Lys Phe Leu Leu Ile Gly Ser Ile Lys Phe Phe Ile Ile
35 40 45

Leu Ala Leu Thr Leu Thr Arg Asp Thr Lys Asp Thr Leu Ile Val Thr
50 55 60

Gln Cys Gly Ala Glu Ala Ile Ala Phe Leu Lys Ile Tyr Gly Val Leu
65 70 75 80

Pro Ala Ala Thr Ala Phe Ile Ala Leu Tyr Ser Lys Met Ser Asn Ala
 85 90 95

Met Gly Lys Lys Met Leu Phe Tyr Ser Thr Cys Ile Pro Phe Phe Thr
 100 105 110

Phe Phe Gly Leu Phe Asp Val Phe Ile Tyr Pro Asn Ala Glu Arg Leu
 115 120 125

His Pro Ser Leu Glu Ala Val Gln Ala Ile Leu Pro Gly Gly Ala Ala
 130 135 140

Ser Gly Gly Met Ala Val Leu Ala Lys Ile Ala Thr His Trp Thr Ser
 145 150 155 160

Ala Leu Phe Tyr Val Met Ala Glu Ile Tyr Ser Ser Val Ser Val Gly
 165 170 175

Leu Leu Phe Trp Gln Phe Ala Asn Asp Val Val Asn Val Asp Gln Ala
 180 185 190

Lys Arg Phe Tyr Pro Leu Phe Ala Gln Met Ser Gly Leu Ala Pro Val
 195 200 205

Leu Ala Gly Gln Tyr Val Val Arg Phe Ala Ser Lys Ala Val Asn Phe
 210 215 220

Glu Ala Ser Met His Arg Leu Thr Ala Ala Val Thr Phe Ala Gly Ile
 225 230 235 240

Met Ile Cys Ile Phe Tyr Gln Leu Ser Ser Ser Tyr Val Glu Arg Thr
 245 250 255

Glu Ser Ala Lys Pro Ala Ala Asp Asn Glu Gln Ser Ile Lys Pro Lys
 260 265 270

Lys Lys Lys Pro Lys Met Ser Met Val Glu Ser Gly Lys Phe Leu Ala
 275 280 285

Ser Ser Gln Tyr Leu Arg Leu Ile Ala Met Leu Val Leu Gly Tyr Gly
 290 295 300

Leu Ser Ile Asn Phe Thr Glu Ile Met Trp Lys Ser Leu Val Lys Lys
 305 310 315 320

Gln Tyr Pro Asp Pro Leu Asp Tyr Gln Arg Phe Met Gly Asn Phe Ser
 325 330 335

Ser Ala Val Gly Leu Ser Thr Cys Ile Val Ile Phe Phe Gly Val His
340 345 350

Val Ile Arg Leu Leu Gly Trp Lys Val Gly Ala Leu Ala Thr Pro Gly
355 360 365

Ile Met Ala Ile Leu Ala Leu Pro Phe Phe Ala Cys Ile Leu Leu Gly
370 375 380

Leu Asp Ser Pro Ala Arg Leu Glu Ile Ala Val Ile Phe Gly Thr Ile
385 390 395 400

Gln Ser Leu Leu Ser Lys Thr Ser Lys Tyr Ala Leu Phe Asp Pro Thr
405 410 415

Thr Gln Met Ala Tyr Ile Pro Leu Asp Asp Glu Ser Lys Val Lys Gly
420 425 430

Lys Ala Ala Ile Asp Val Leu Gly Ser Arg Ile Gly Lys Ser Gly Gly
435 440 445

Ser Leu Ile Gln Gln Gly Leu Val Phe Val Phe Gly Asn Ile Ile Asn
450 455 460

Ala Ala Pro Val Val Gly Val Val Tyr Tyr Ser Val Leu Val Ala Trp
465 470 475 480

Met Ser Ala Ala Gly Arg Leu Ser Gly Leu Phe Gln Ala Gln Thr Glu
485 490 495

Met Asp Lys Ala Asp Lys Met Glu Ala Lys Thr Asn Lys Glu Lys
500 505 510

<210> 5

<211> 1599

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: Polinucleótido sintético

<400> 5

atgagaccat ttccgacgat tgccttgatt tcggtttttc tttcggcggc gactcgcatt	60
tcggcaggag gcagtactgt tgcaccaact acaccgttgg caaccggcgg tgcgctccgc	120
aaagtgcgac aagccgtctt tcccatctac ggaaaccaag aagtcaccaa atttctgctc	180
atcggatcca ttaaattctt tataatcttg gcactcacgc tcacgcgtga taccaaggac	240

acgttgattg tcacgcaatg tgggtgccgaa gcgattgcct ttctcaaaat atacggggtg 300
ctacccgcag cgaccgcatt tatcgcgctc tattccaaaa tgtccaacgc catgggcaaa 360
aaaatgctat ttattccac ttgcattcct ttctttacct ttttcgggct gtttgatgtt 420
ttcatttacc cgaacgcgga gcgactgcac cctagtttgg aagccgtgca ggcaattctc 480
ccgggcggtg ccgcatctgg cggcatggcg gttctggcca agattgacgac aactggaca 540
tcggccttat ttacgtcat ggcgaaaata tattcttccg tatcggtggg gctattgttt 600
tggcagtttg cgaacgacgt cgtcaacgtg gatcaggcca agcgctttta tccattattt 660
gctcaaatga gtggcctcgc tccagtttta gcgggccagt atgtggtacg gtttgccagc 720
aaagcgttca actttgaggc atccatgcat cgactcacgg cggccgtaac atttgctggt 780
attatgattt gcatctttta ccaactcagt tcgtcatatg tggagcgaac ggaatcagca 840
aagccagcgg cagataacga gcagtctatc aaaccgaaaa agaagaaacc caaatgtcc 900
atggttgaat cggggaaatt tctcgcgta agtcagtacc tgcgtctaata tgccatgctg 960
gtgctgggat acggcctcag tattaacttt accgaaatca tgtggaaaag cttggtgaag 1020
aaacaatatc cagacccgct agattatcaa cgatttatgg gtaacttctc gtcagcgggt 1080
ggtttgagca catgcattgt tattttcttc ggtgtgcacg tgatccgttt gttggggtgg 1140
aaagtgcggag cgttggctac acctgggatc atggccattc tagcgttacc cttttttgct 1200
tgcattttgt tgggtttgga tagtccagca cgattggaga tcgccgtaat ctttgaaca 1260
attcagagtt tgctgagcaa aacctccaag tatgcccttt tcgaccctac cacacaaatg 1320
gcttatattc ctctggacga cgaatcaaag gtcaaaggaa aagcggcaat tgatgttttg 1380
ggatcgcgga ttggcaagag tggaggctca ctgatccagc agggcttggg ctttgttttt 1440
ggaaatatca ttaatgccgc acctgtagta ggggttgtct actacagtgt cttgtgtgag 1500
tggatgagcg cagctggccg actaagtggg ctttttcaag cacaacaga aatggataag 1560
gccgacaaaa tggaggcaaa gaccaacaaa gaaaagtag 1599

<210> 6

<211> 532

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: Polipéptido sintético

<400> 6

Met Arg Pro Phe Pro Thr Ile Ala Leu Ile Ser Val Phe Leu Ser Ala
1 5 10 15

Ala Thr Arg Ile Ser Ala Gly Gly Ser Thr Val Ala Pro Thr Thr Pro

ES 2 929 047 T3

20					25					30					
Leu	Ala	Thr	Gly	Gly	Ala	Leu	Arg	Lys	Val	Arg	Gln	Ala	Val	Phe	Pro
		35					40					45			
Ile	Tyr	Gly	Asn	Gln	Glu	Val	Thr	Lys	Phe	Leu	Leu	Ile	Gly	Ser	Ile
	50					55					60				
Lys	Phe	Phe	Ile	Ile	Leu	Ala	Leu	Thr	Leu	Thr	Arg	Asp	Thr	Lys	Asp
65					70					75					80
Thr	Leu	Ile	Val	Thr	Gln	Cys	Gly	Ala	Glu	Ala	Ile	Ala	Phe	Leu	Lys
				85					90					95	
Ile	Tyr	Gly	Val	Leu	Pro	Ala	Ala	Thr	Ala	Phe	Ile	Ala	Leu	Tyr	Ser
			100					105					110		
Lys	Met	Ser	Asn	Ala	Met	Gly	Lys	Lys	Met	Leu	Phe	Tyr	Ser	Thr	Cys
		115					120					125			
Ile	Pro	Phe	Phe	Thr	Phe	Phe	Gly	Leu	Phe	Asp	Val	Phe	Ile	Tyr	Pro
	130					135					140				
Asn	Ala	Glu	Arg	Leu	His	Pro	Ser	Leu	Glu	Ala	Val	Gln	Ala	Ile	Leu
145					150					155					160
Pro	Gly	Gly	Ala	Ala	Ser	Gly	Gly	Met	Ala	Val	Leu	Ala	Lys	Ile	Ala
				165					170					175	
Thr	His	Trp	Thr	Ser	Ala	Leu	Phe	Tyr	Val	Met	Ala	Glu	Ile	Tyr	Ser
			180					185					190		
Ser	Val	Ser	Val	Gly	Leu	Leu	Phe	Trp	Gln	Phe	Ala	Asn	Asp	Val	Val
		195					200					205			
Asn	Val	Asp	Gln	Ala	Lys	Arg	Phe	Tyr	Pro	Leu	Phe	Ala	Gln	Met	Ser
	210					215					220				
Gly	Leu	Ala	Pro	Val	Leu	Ala	Gly	Gln	Tyr	Val	Val	Arg	Phe	Ala	Ser
225					230					235					240
Lys	Ala	Val	Asn	Phe	Glu	Ala	Ser	Met	His	Arg	Leu	Thr	Ala	Ala	Val
				245					250					255	
Thr	Phe	Ala	Gly	Ile	Met	Ile	Cys	Ile	Phe	Tyr	Gln	Leu	Ser	Ser	Ser
			260					265					270		

ES 2 929 047 T3

Tyr Val Glu Arg Thr Glu Ser Ala Lys Pro Ala Ala Asp Asn Glu Gln
 275 280 285
 Ser Ile Lys Pro Lys Lys Lys Lys Pro Lys Met Ser Met Val Glu Ser
 290 295 300
 Gly Lys Phe Leu Ala Ser Ser Gln Tyr Leu Arg Leu Ile Ala Met Leu
 305 310 315 320
 Val Leu Gly Tyr Gly Leu Ser Ile Asn Phe Thr Glu Ile Met Trp Lys
 325 330 335
 Ser Leu Val Lys Lys Gln Tyr Pro Asp Pro Leu Asp Tyr Gln Arg Phe
 340 345 350
 Met Gly Asn Phe Ser Ser Ala Val Gly Leu Ser Thr Cys Ile Val Ile
 355 360 365
 Phe Phe Gly Val His Val Ile Arg Leu Leu Gly Trp Lys Val Gly Ala
 370 375 380
 Leu Ala Thr Pro Gly Ile Met Ala Ile Leu Ala Leu Pro Phe Phe Ala
 385 390 395 400
 Cys Ile Leu Leu Gly Leu Asp Ser Pro Ala Arg Leu Glu Ile Ala Val
 405 410 415
 Ile Phe Gly Thr Ile Gln Ser Leu Leu Ser Lys Thr Ser Lys Tyr Ala
 420 425 430
 Leu Phe Asp Pro Thr Thr Gln Met Ala Tyr Ile Pro Leu Asp Asp Glu
 435 440 445
 Ser Lys Val Lys Gly Lys Ala Ala Ile Asp Val Leu Gly Ser Arg Ile
 450 455 460
 Gly Lys Ser Gly Gly Ser Leu Ile Gln Gln Gly Leu Val Phe Val Phe
 465 470 475 480
 Gly Asn Ile Ile Asn Ala Ala Pro Val Val Gly Val Val Tyr Tyr Ser
 485 490 495
 Val Leu Val Ala Trp Met Ser Ala Ala Gly Arg Leu Ser Gly Leu Phe
 500 505 510
 Gln Ala Gln Thr Glu Met Asp Lys Ala Asp Lys Met Glu Ala Lys Thr
 515 520 525

Asn Lys Glu Lys
530

5 <210> 7
<211> 1665
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Polinucleótido sintético
<400> 7

atgacttcct	ctcatcaagc	aagtgcactt	cctctcaaaa	agggaacgca	tgtcccggac	60
tctccgaagt	tgtcaaagct	atatatcatg	gccaaaacca	agagtgtatc	ctcgtccttc	120
gacccccctc	ggggaggcag	tactgttgca	ccaactacac	cgttggaac	cggcggcg	180
ctccgcaaag	tgcgacaagc	cgtctttccc	atctacggaa	accaagaagt	caccaaattt	240
ctgctcatcg	gatccattaa	attctttata	atcttggcac	tcacgctcac	gcgtgatacc	300
aaggacacgt	tgattgtcac	gcaatgtggt	gccgaagcga	ttgcctttct	caaaatatac	360
ggggtgctac	ccgcagcgac	cgcattttatc	gcgctctatt	ccaaaatgtc	caacgccatg	420
ggcaaaaaaa	tgctatttta	ttccacttgc	attcctttct	ttaccttttt	cgggctgttt	480
gatgttttca	tttaccgaa	cgcggagcga	ctgcacccta	gtttggaagc	cgtgcaggca	540
attctcccgg	gcggtgccgc	atctggcggc	atggcggttc	tggccaagat	tgcgacacac	600
tggacatcgg	ccttatttta	cgtcatggcg	gaaatatatt	cttccgtatc	ggtggggcta	660
ttgttttggc	agtttgcgaa	cgacgtcgtc	aacgtggatc	aggccaagcg	cttttatcca	720
ttatttgctc	aaatgagtgg	cctcgctcca	gttttagcgg	gccagtatgt	ggtacggttt	780
gccagcaaag	cgggtcaactt	tgaggcatcc	atgcatcgac	tcacggcggc	cgtaacattt	840
gctggtatta	tgatttgcat	cttttaccaa	ctcagttcgt	catatgtgga	gcgaacggaa	900
tcagcaaagc	cagcggcaga	taacgagcag	tctatcaaac	cgaaaaagaa	gaaacccaaa	960
atgtccatgg	ttgaatcggg	gaaattttctc	gcgtcaagtc	agtacctgcg	tctaattgcc	1020
atgctggtgc	tgggatacgg	cctcagtatt	aactttaccg	aatcatgtg	gaaaagcttg	1080
gtgaagaaac	aatatccaga	cccgctagat	tatcaacgat	ttatgggtaa	cttctcgtca	1140
gcggttggtt	tgagcacatg	cattgtttatt	ttcttcgggtg	tgcacgtgat	ccgtttgttg	1200
gggtggaaag	tccgagcggt	ggctacacct	gggatcatgg	ccattctagc	gttacccttt	1260
tttgcttgca	ttttgttggg	tttggatagt	ccagcacgat	tggagatcgc	cgtaatcttt	1320
ggaacaattc	agagtttgct	gagcaaaacc	tccaagtatg	cccttttcga	ccctaccaca	1380
caaatggctt	atattcctct	ggacgacgaa	tcaaagggtca	aaggaaaagc	ggcaattgat	1440

gttttgggat cgcggattgg caagagtgga ggctcactga tccagcaggg cttgggtcttt 1500
 gtttttggaa atatcattaa tgccgcacct gtagtagggg ttgtctacta cagtgtcctt 1560
 gttgcgtgga tgagcgcagc tggccgacta agtgggcttt ttcaagcaca aacagaaatg 1620
 gataaggccg acaaaatgga ggcaaagacc aacaaagaaa agtag 1665

<210> 8
 <211> 554
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Polipéptido sintético

<400> 8

ES 2 929 047 T3

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

ES 2 929 047 T3

Ala Val Gln Ala Ile Leu Pro Gly Gly Ala Ala Ser Gly Gly Met Ala
180 185 190

Val Leu Ala Lys Ile Ala Thr His Trp Thr Ser Ala Leu Phe Tyr Val
195 200 205

Met Ala Glu Ile Tyr Ser Ser Val Ser Val Gly Leu Leu Phe Trp Gln
210 215 220

Phe Ala Asn Asp Val Val Asn Val Asp Gln Ala Lys Arg Phe Tyr Pro
225 230 235 240

Leu Phe Ala Gln Met Ser Gly Leu Ala Pro Val Leu Ala Gly Gln Tyr
245 250 255

Val Val Arg Phe Ala Ser Lys Ala Val Asn Phe Glu Ala Ser Met His
260 265 270

Arg Leu Thr Ala Ala Val Thr Phe Ala Gly Ile Met Ile Cys Ile Phe
275 280 285

Tyr Gln Leu Ser Ser Ser Tyr Val Glu Arg Thr Glu Ser Ala Lys Pro
290 295 300

Ala Ala Asp Asn Glu Gln Ser Ile Lys Pro Lys Lys Lys Lys Pro Lys
305 310 315 320

Met Ser Met Val Glu Ser Gly Lys Phe Leu Ala Ser Ser Gln Tyr Leu
325 330 335

Arg Leu Ile Ala Met Leu Val Leu Gly Tyr Gly Leu Ser Ile Asn Phe
340 345 350

Thr Glu Ile Met Trp Lys Ser Leu Val Lys Lys Gln Tyr Pro Asp Pro
355 360 365

Leu Asp Tyr Gln Arg Phe Met Gly Asn Phe Ser Ser Ala Val Gly Leu
370 375 380

Ser Thr Cys Ile Val Ile Phe Phe Gly Val His Val Ile Arg Leu Leu
385 390 395 400

Gly Trp Lys Val Gly Ala Leu Ala Thr Pro Gly Ile Met Ala Ile Leu
405 410 415

Ala Leu Pro Phe Phe Ala Cys Ile Leu Leu Gly Leu Asp Ser Pro Ala
420 425 430

Arg Leu Glu Ile Ala Val Ile Phe Gly Thr Ile Gln Ser Leu Leu Ser
435 440 445

Lys Thr Ser Lys Tyr Ala Leu Phe Asp Pro Thr Thr Gln Met Ala Tyr
450 455 460

Ile Pro Leu Asp Asp Glu Ser Lys Val Lys Gly Lys Ala Ala Ile Asp
465 470 475 480

Val Leu Gly Ser Arg Ile Gly Lys Ser Gly Gly Ser Leu Ile Gln Gln
485 490 495

Gly Leu Val Phe Val Phe Gly Asn Ile Ile Asn Ala Ala Pro Val Val
500 505 510

Gly Val Val Tyr Tyr Ser Val Leu Val Ala Trp Met Ser Ala Ala Gly
515 520 525

Arg Leu Ser Gly Leu Phe Gln Ala Gln Thr Glu Met Asp Lys Ala Asp
530 535 540

Lys Met Glu Ala Lys Thr Asn Lys Glu Lys
545 550

<210> 9

<211> 3013

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Polinucleótido sintético

<220>

<221> base modificada

<222> (2349)..(2349)

<223> Cualquier nucleótido no natural

<400> 9

aacgccagca acgcggcctt tttacgggtc ctggcctttt gctggccttt tgctcacatg	60
ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct	120
gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa	180
gagcgcctga tgcggtattt tctccttaag catctgtgcg gtatttcaca ccgcatatgg	240
tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat	300
cgctacgtga ctgggtcatg gctgcgcccc gacacccgcc aacacccgct gacgcgccct	360
gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct	420

gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct	480
catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcattccgcg tccagctcgt	540
tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcggggccatg ttaagggcgg	600
ttttttcctg tttggtcact gatgcctccg tgtaaggggg atttctgttc atgggggtaa	660
tgataccgat gaaacgagag aggatgctca cgatacgggt tactgatgat gaacatgccc	720
ggttactgga acgttgtgag ggtaaacaaac tggcgggatg gatgcggcgg gaccagagaa	780
aaatcactca gggccaatgc cagcgcttcg ttaatacaga tgtaggtgtt ccacagggtta	840
gccagcagca tcctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg	900
tttccagact ttacgaaaca cggaaaccga agaccattca tgttggtgct caggtcgcag	960
acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac	1020
cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgcgca	1080
cccgtggcca ggaccaacg ctgcccgaat ttcttgaaga cgaaagggcc tcgtgatacg	1140
cctattttta taggttaatg tcatgataat aatgggtttct tagacgtcag gtggcacttt	1200
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta	1260
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat	1320
gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt	1380
ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg	1440
agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga	1500
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccg	1560
tgttgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt	1620
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg	1680
cagtgtgcc ataaccatga gtgataaac tgccggccaac ttacttctga caacgatcgg	1740
aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttgta	1800
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc	1860
tgcagcaatg gcaacaacgt tgcgcaaaact attaaactggc gaactactta ctctagcttc	1920
ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc	1980
ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtggctctcg	2040
cggatatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac	2100
gacggggagt caggcaacta tggatgaacg aaatagacag atcgttgaga taggtgcctc	2160
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt	2220
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac	2280
caaaatccct taacgtgagt tttcgttggc tttaactttt atgcttccgg ctcgatatgtt	2340

gtgtggaant gtgagcggat aacaatttca cacaggaaac agccactgag cgtcagaccc 2400
 cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt 2460
 gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac 2520
 tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg tccttctagt 2580
 gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct 2640
 gctaatacctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga 2700
 ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac 2760
 acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg 2820
 agaaagcgcc acgcttcccg aaggagaaa ggcggacagg tatccggtaa gcggcagggg 2880
 cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc 2940
 tgtcggggtt cgccacctct gacttgagcg tcgatttttg tgatgctcgt cagggggggcg 3000
 gagcctatgg aaa 3013

- 5 <210> 10
 <211> 20
 <212> ADN
 <213> Secuencia artificial
- 10 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 10
 gggaggcagt actgttgac 20
- 15 <210> 11
 <211> 45
 <212> ADN
 <213> Secuencia artificial
- 20 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 11
 ggtatatctc ctattaaag ttaaacaaaa ttattctac agggg 45
- 25 <210> 12
 <211> 20
 <212> ADN
 <213> Secuencia artificial
- 30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 12
 taatacgact cactataggg 20
- 35 <210> 13
 <211> 19
 <212> ADN
 <213> Secuencia artificial
- 40 <220>

<223> Descripción de la secuencia artificial: Cebador sintético

<400> 13
gctagttatt gctcagcgg 19

5

<210> 14
<211> 22
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 14
ttacattaat tgcgttgccg tc 22

15

<210> 15
<211> 31
<212> ADN
<213> Secuencia artificial

20

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

25

<400> 15
tttggcgga tggcattga gaagcacag g 31

30

<210> 16
<211> 25
<212> ADN
<213> Secuencia artificial

35

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 16
atttcacca ataaaaaacg cccgg 25

40

<210> 17
<211> 36
<212> ADN
<213> Secuencia artificial

45

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 17
cctgtagaaa taatttgg tttaacttaaat aaggag 36

50

<210> 18
<211> 20
<212> ADN
<213> Secuencia artificial

55

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 18
ccccgcgcgt tggccgattc 20

60

<210> 19
<211> 19

<212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 19
 gaagggcaat cagctgttg 19
 10 <210> 20
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 20
 cagggcaggg tcgttaaata g 21
 20 <210> 21
 <211> 52
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 21
 30 gacaccatcg aatggcgcaa aaccttcgc ggtatggcat gatagcgccc gg 52
 <210> 22
 <211> 52
 <212> ADN
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40 <400> 22
 ccgggcgcta tcatgccata ccgcgaaagg tttgcgcca ttcgatggtg tc 52
 <210> 23
 <211> 51
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 50 <400> 23
 attttctaa atacattcaa atatgtatcc gctcatgaga caataaccct g 51
 <210> 24
 <211> 51
 55 <212> ADN
 <213> Secuencia artificial
 <220>
 60 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 24
 cagggttatt gtctcatgag cggatacata ttgaatgta ttagaaaaa t 51

5 <210> 25
 <211> 52
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 10 <400> 25
 ttaggcaccc caggctttac actttatgct tccggctggt atgttgtgtg ga 52

 <210> 26
 <211> 52
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 20 <400> 26
 tccacacaac ataccagccg gaagcataaa gtgtaaagcc tgggggcct aa 52

 <210> 27
 <211> 52
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 30 <400> 27
 ctaggcacc caggctttac actttatgct tccggctggt ataattgtgtg ga 52

 <210> 28
 <211> 52
 35 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 28
 tccacacatt ataccagccg gaagcataaa gtgtaaagcc tgggggcct ag 52

 45 <210> 29
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 29
 ggggataacg caggaaagaa catg 24

 55 <210> 30
 <211> 22
 <212> ADN
 60 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 30
gcacttttcg gggaaatgtg cg 22

5 <210> 31
<211> 75
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 31

aattgcggcc tatatggatg ttggaaccgt aagagaaata gacaggcggc cctgtgacgg
aagatcactt cgcag

15 <210> 32
<211> 46
<212> ADN
20 <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

25 <400> 32
tgctcacatg ttcttctcg cggtatcccc gcgtggtgaa ccaggc 46

<210> 33
<211> 28
30 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

35 <400> 33
accgcctgtc tatttctctt acggttcc 28

<210> 34
<211> 49
40 <212> ADN
<213> Secuencia artificial

<220>
45 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 34
cgcgcttaat gcgccgtac agggcgcgtc gattggtgcc agcgcgag 49

50 <210> 35
<211> 43
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 35

ggtatatctc cttattaaag ttaacaaaa ttattctac agg 43

5 <210> 36
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 36
 cagccacgtt tctgcgaaaa c 21

15 <210> 37
 <211> 20
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 37
 tacagcgggt ccttactggc 20

25 <210> 38
 <211> 26
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 38
 gggtggtgaa tgtgaaacca gtaacg 26

35 <210> 39
 <211> 21
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 39
 45 ctgggtgttt acttcggtct g 21

50 <210> 40
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

55 <400> 40
 ggccgtaata tccagtgaa c 21

60 <210> 41
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>

<223> Descripción de la secuencia artificial: Cebador sintético

<400> 41
actagggtgc agtcgctccg 20

5

<210> 42
<211> 22
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 42
ttaacctagg ctgctgccac cg 22

15

<210> 43
<211> 57
<212> ADN
<213> Secuencia artificial

20

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

25

<400> 43
gcgcaacgca attaatgtaa ttctgaaatg agctgttgac aattaatcat cggctcg 57

30

<210> 44
<211> 56
<212> ADN
<213> Secuencia artificial

35

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 44
aacaaaatta ttctacagg tccacacatt atacgagccg atgattaatt gtcaac 56

40

<210> 45
<211> 58
<212> ADN
<213> Secuencia artificial

45

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 45
gcgcaacgca attaatgtaa tcataaaaaa ttatttgct tcaggaaaa ttttctg 58

50

<210> 46
<211> 59
<212> ADN
<213> Secuencia artificial

55

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 46
aacaaaatta ttctacagg tgaatctatt atacagaaaa atttctctga aagcaaata 59

60

<210> 47
<211> 56

<212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 47
 gcgcaacgca attaatgtaa ttatctctgg cgggtgtgac ataaatacca ctggcg 56
 10 <210> 48
 <211> 57
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 48
 aacaaaatta ttctacagg tgtgctcagt atcaccgcca gtggtatta tgtcaac 57
 20 <210> 49
 <211> 58
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 49
 30 gcgcaacgca attaatgtaa ttttaaaaaa ttcatttgct aaacggtca aattctcg 58
 <210> 50
 <211> 55
 <212> ADN
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40 <400> 50
 aacaaaatta ttctacagg tgaagtatat tatacgagaa ttgaagcgt ttagc 55
 <210> 51
 <211> 21
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 50 <400> 51
 tgggtgcct aatgagtgag c 21
 55 <210> 52
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 <220>
 60 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 52

ctatgaccat gattacgcca agcttg 26

5 <210> 53
 <211> 24
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 53
 tctcgcggta tcattgcagc actg 24

15 <210> 54
 <211> 18
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 54
 gccacgctca ccggctcc 18

25 <210> 55
 <211> 21
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 55
 aacgaaaact cacgtaagg g 21

35 <210> 56
 <211> 17
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 56
 ccactgagcg tcagacc 17

45 <210> 57
 <211> 30
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

55 <400> 57
 gagacccgctc gttgacaatt aatcatcggc 30

60 <210> 58
 <211> 31
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 58
 gagaccattc tcaccaataa aaaacgccc g 31
 5

 <210> 59
 <211> 30
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 15 <400> 59
 cggggtacca tggacaagaa gtactccatt 30

 <210> 60
 <211> 34
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 25 <400> 60
 ctagtctaga ttacacctc ctctctctt tggg 34

 <210> 61
 30 <211> 33
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 61
 ctccggggaa accgccgaag ccacgcggct caa 33
 40 <210> 62
 <211> 44
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 62
 cttcgccggt ttcccggag tcgaacagga gggcgccaat gagg 44
 50 <210> 63
 <211> 40
 <212> ADN
 <213> Secuencia artificial
 55 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 63
 60 aggaagaaga cgtctcacgc atcttactgc gcagatacgc 40

 <210> 64
 <211> 60

<212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 64
 aagatgcgtg agacgtcttc ttctcgtct cggtcgacag tcataggtg attgctcagg 60
 10 <210> 65
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 65
 gtcccaaadc gcagccaadc acattg 26
 20 <210> 66
 <211> 28
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 66
 30 gtctgacca tcgtattggt tatctggc 28
 <210> 67
 <211> 26
 <212> ADN
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40 <400> 67
 atttagaggg cagtgccagc tcgtta 26
 <210> 68
 <211> 26
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 50 <400> 68
 ctgcattcag gtaggcata tgcga 26
 <210> 69
 55 <211> 26
 <212> ADN
 <213> Secuencia artificial
 <220>
 60 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 69
 ctgggctacc tgcaagatta gcatg 26

5 <210> 70
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 10 <400> 70
 tgaaggactg ggcagaggcc ccctt 25

 15 <210> 71
 <211> 26
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 20 <400> 71
 cgtaggtgtc ttgctcagt tgaagc 26

 25 <210> 72
 <211> 27
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 30 <400> 72
 tagccatctc attactaaag atctcct 27

 35 <210> 73
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 73
 cgatatcgtt ggtctcaacg acacaattgt aaaggtaga tct 43

 45 <210> 74
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 74
 caacgatatc ggtctcacac tgactgggcc ttctgttta tct 43

 55 <210> 75
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 60 <210> 76
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

	<400> 75		
	gcaatcacct atgaactgtc gac 23		
5	<210> 76 <211> 69 <212> ADN <213> Secuencia artificial		
10	<220> <223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 76		
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tagtgatcag ttttagagct	60	
	agaaatagc	69	
15	<210> 77 <211> 69 <212> ADN <213> Secuencia artificial		
20	<220> <223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 77		
25	aggaggaagg acgtctcatg cgccccgcat tccaggatgg gtaccacccg ttttagagct	60	
	agaaatagc	69	
30	<210> 78 <211> 69 <212> ADN <213> Secuencia artificial		
35	<220> <223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 78		
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tagtcatcag ttttagagct	60	
	agaaatagc	69	
40	<210> 79 <211> 69 <212> ADN <213> Secuencia artificial		
45	<220> <223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 79		
	aggaggaagg acgtctcatg cgccccgcat tccaggatgg gcaccagccg ttttagagct	60	
50	agaaatagc	69	
55	<210> 80 <211> 69 <212> ADN <213> Secuencia artificial		

	<220>		
	<223> Descripción de la secuencia artificial: Cebador sintético		
5	<400> 80		
	aggaggaagg acgtctcatg cgccccgcat tccatgatgg gcaccacccg ttttagagct	60	
	agaaatagc	69	
10	<210> 81		
	<211> 69		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 81		
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tatagatcag ttttagagct	60	
	agaaatagc	69	
20	<210> 82		
	<211> 69		
	<212> ADN		
	<213> Secuencia artificial		
25	<220>		
	<223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 82		
30	aggaggaagg acgtctcatg cgccccgcat tccaggatgg gcaccatccg ttttagagct	60	
	agaaatagc	69	
35	<210> 83		
	<211> 69		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 83		
	aggaggaagg acgtctcatg cgccccgcat tgttgtgtgg aaatgtgagg ttttagagct	60	
	agaaatagc	69	
45	<210> 84		
	<211> 69		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> Descripción de la secuencia artificial: Cebador sintético		
	<400> 84		

	aggaggaagg acgtctcatg cgccccgcat ttgtcactac tctgaccagg ttttagagct	60
	agaaatagc	69
5	<210> 85 <211> 69 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 85	
	aggaggaagg acgtctcatg cgccccgcat ttgtcactac tctgaccaag ttttagagct	60
	agaaatagc	69
15	<210> 86 <211> 69 <212> ADN <213> Secuencia artificial	
20	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 86	
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tactcatcag ttttagagct	60
25	agaaatagc	69
30	<210> 87 <211> 69 <212> ADN <213> Secuencia artificial	
35	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 87	
	aggaggaagg acgtctcatg cgccccgcat tccaagatgg gcaccaccg ttttagagct	60
	agaaatagc	69
40	<210> 88 <211> 69 <212> ADN <213> Secuencia artificial	
45	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 88	
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tattgatcag ttttagagct	60
50	agaaatagc	69
	<210> 89 <211> 69 <212> ADN	

	<213> Secuencia artificial	
	<220>	
5	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 89	
	aggaggaagg acgtctcatg cgccccgcat tcacacaatg tataaatcag ttttagagct	60
	agaaatagc	69
10	<210> 90	
	<211> 69	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 90	
	aggaggaagg acgtctcatg cgccccgcat tccaggatgg ggaccacccg ttttagagct	60
20	agaaatagc	69
	<210> 91	
	<211> 69	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
30	<400> 91	
	aggaggaagg acgtctcatg cgccccgcat tttcacaata ctttcttttag ttttagagct	60
	agaaatagc	69
35	<210> 92	
	<211> 43	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 92	
	gcatttcaca caatgtagga tcagttttag agctagaaat agc 43	
45	<210> 93	
	<211> 38	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 93	
	tgatcctaca ttgtgtgaaa tgcggggcgc atcttact 38	
55	<210> 94	
	<211> 43	

<212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 94
 gcattaccag gatgggacca cccgttttag agctagaaat agc 43
 10 <210> 95
 <211> 38
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 95
 gggtgtccc atcctggtaa tgcggggcgc atcttact 38
 20 <210> 96
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 96
 30 gcatttcaca caatgtagca tcagttttag agctagaaat agc 43
 <210> 97
 <211> 38
 <212> ADN
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40 <400> 97
 tgatgctaca ttgtgtgaaa tgcggggcgc atcttact 38
 <210> 98
 <211> 43
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 50 <400> 98
 gcattaccag gatgggcacc accgttttag agctagaaat agc 43
 <210> 99
 <211> 38
 55 <212> ADN
 <213> Secuencia artificial
 <220>
 60 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 99
 gggtgtccc atcctggtaa tgcggggcgc atcttact 38

5 <210> 100
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 10 <400> 100
 gcattaccag atgggcacca cccgttttag agctagaaat agc 43

 15 <210> 101
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 20 <400> 101
 gggtggtgcc catctggtaa tgcggggcgc atcttact 38

 25 <210> 102
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 30 <400> 102
 gcattcaca caatgtaaga tcagttttag agctagaaat agc 43

 35 <210> 103
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 103
 tgatcttaca ttgtgtgaaa tgcggggcgc atcttact 38

 45 <210> 104
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 104
 gcatttgttg tgtggaatgt gaggttttag agctagaaat agc 43

 55 <210> 105
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 105
ctcacattcc acacaacaaa tgcggggcgc atcttact 38

5 <210> 106
 <211> 43
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 106
 gcattttgtc actactctga ccggttttag agctagaaat agc 43

15 <210> 107
 <211> 38
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 107
 cggtcagagt agtgacaaaa tgcggggcgc atcttact 38

25 <210> 108
 <211> 43
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 108
 gcattttgtc actactctga ccagtttttag agctagaaat agc 43

35 <210> 109
 <211> 38
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

45 <400> 109
 tggtcagagt agtgacaaaa tgcggggcgc atcttact 38

 <210> 110
 <211> 43
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

55 <400> 110
 gcatttcaca caatgtacca tcagtttttag agctagaaat agc 43

60 <210> 111
 <211> 38
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

5 <400> 111
tgatggtaca ttgtgtgaaa tgcggggcgc atcttact 38

10 <210> 112
 <211> 43
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 112
gcatttcaca caatgtatga tcagtttag agctagaaat agc 43

20 <210> 113
 <211> 38
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 113
tgatcataca ttgtgtgaaa tgcggggcgc atcttact 38

30 <210> 114
 <211> 43
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 114
gcatttcaca caatgtataa tcagtttag agctagaaat agc 43

40 <210> 115
 <211> 38
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

50 <400> 115
tgattataca ttgtgtgaaa tgcggggcgc atcttact 38

55 <210> 116
 <211> 43
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 116
gcattattca caatacttct ttagtttag agctagaaat agc 43

 <210> 117

<211> 38
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 117
 taaagaagta ttgtgaataa tgcggggcgc atcttact 38

10 <210> 118
 <211> 47
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 118
 agaaggaaga cgtctcactg tcgacaaaa aagcctgctc gttgagc 47

20 <210> 119
 <211> 40
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

30 <400> 119
 aagaaggacg tctcaacagt agtggcagcg gctaactaag 40

 <210> 120
 <211> 44
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

40 <400> 120
 aggagaggac gtctctcgac caaaaaagcc tgctcgttga gcag 44

45 <210> 121
 <211> 59
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 121
 agtaagatgc gccccgcatt gaccaggatg ggcaccaccc gtttagagc tagaaatag 59

55 <210> 122
 <211> 59
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 122
ctatttctag ctctaaaacg ggtggtgcc atcctggtca atcgggggcg catcttact 59

5 <210> 123
<211> 39
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 123
gttttagagc tagaaatagc aagtaaaat aaggctagt 39

15 <210> 124
<211> 20
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 124
aatgcggggc gcatcttact 20

25 <210> 125
<211> 22
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 125
cattggcacc ggtctactaa ac 22

35 <210> 126
<211> 22
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

45 <400> 126
aaacgttttag tagaccggtg cc 22

50 <210> 127
<211> 29
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 127
cgcatcaca caatgtaagt atcagtttt 29

60 <210> 128
<211> 29
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 128
5 ctctaaaact gatactaca ttgtgtgaa 29

<210> 129
<211> 60
<212> ADN
10 <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

15 <400> 129
gtttacgtcg ccgtccagct cgaccaggat gggcaccaac ccggtgaaca gtcctcgcc 60

<210> 130
<211> 60
20 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

25 <400> 130
gtttacgtcg ccgtccagct cgaccaggat gggcaccagc ccggtgaaca gtcctcgcc 60

<210> 131
<211> 60
30 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

35 <400> 131
gtttacgtcg ccgtccagct cgaccaggat gggcaccatc ccggtgaaca gtcctcgcc 60

<210> 132
<211> 60
40 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Cebador sintético

45 <400> 132
gtttacgtcg ccgtccagct cgaccaggat gggcaccacc ccggtgaaca gtcctcgcc 60

50 <210> 133
<211> 60
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 133
60 gcatcgccct cgccctcgcc ggacacgctg aactgtggc cgttacgtc gccgtccagc 60

<210> 134
<211> 60

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 134	
	tcaggttca ttgatgctc gatgagttat ggtgagcaag ggcgaggagc tgttcaccgg	60
10	<210> 135	
	<211> 79	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 135	
	ttaatacgcac tcactatagg gaccaggatg ggcaccaacc gttttagagc tagaaatagc	60
20	aagttaaaat aaggctagt	79
	<210> 136	
	<211> 79	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
30	<400> 136	
	ttaatacgcac tcactatagg gaccaggatg ggcaccagcc gttttagagc tagaaatagc	60
	aagttaaaat aaggctagt	79
	<210> 137	
35	<211> 79	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
40	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 137	
	ttaatacgcac tcactatagg gaccaggatg ggcaccatcc gttttagagc tagaaatagc	60
	aagttaaaat aaggctagt	79
45	<210> 138	
	<211> 62	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> Descripción de la secuencia artificial: Cebador sintético	
	<400> 138	
55		

	ttaatacgcac tcactatagg gaccaggatg ggcaccaccc gtttagagc tatgctgttt	60
	tg	62
5	<210> 139 <211> 82 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 139	
	aaaagcaccg actcgggtgcc actttttcaa gttgataacg gactagcctt attttaactt	60
	gctattttcta gctctaaaac gg	82
15	<210> 140 <211> 35 <212> ADN <213> Secuencia artificial	
20	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 140 aagaggaaga ggtaatacg actcactata gggac 35	
25	<210> 141 <211> 20 <212> ADN <213> Secuencia artificial	
30	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 141 aaaagcaccg actcgggtgcc 20	
35	<210> 142 <211> 41 <212> ADN <213> Secuencia artificial	
40	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 142 actagcctta tttaacttg ctatttctag ctctaaaacg g 41	
45	<210> 143 <211> 48 <212> ADN <213> Secuencia artificial	
50	<220> <223> Descripción de la secuencia artificial: Cebador sintético <400> 143 aagaaggaga aggtctctag tggagcaagg gcgaggagct gttcaccg 48	
55		

<210> 144
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 144
 aagagaagag aggtctcatc gtgtttacgt cgccgtccag ctc 43
 10
 <210> 145
 <211> 34
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 20
 <400> 145
 atgggtctcc agtggggcca acactgtca ctac 34
 <210> 146
 <211> 34
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 30
 <400> 146
 atgggtctct tcgtttccgg ataacgggaa aagc 34
 <210> 147
 <211> 39
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40
 <400> 147
 atgggtctcc agtggctcga gtacaacttt aactcacac 39
 45
 <210> 148
 <211> 39
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 148
 atgggtctct tcgttgatgc cattcttttg ttgtctgc 39
 55
 <210> 149
 <211> 41
 <212> ADN
 <213> Secuencia artificial
 60
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 149
atgggtctct catagctgtt tcctgtgtga aattgtatc c 41

5 <210> 150
 <211> 37
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 150
atgggtctca cccaggtt tacacttat gcttccg 37

15 <210> 151
 <211> 41
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 151
atgggtctcc agtggctgtt tcctgtgtga aattgtatc c 41

25 <210> 152
 <211> 37
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 152
atgggtctct tcgttggtt tacacttat gcttccg 37

35 <210> 153
 <211> 36
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

45 <400> 153
atgggtctcc agtggcacac aggaaacagc tatgac 36

50 <210> 154
 <211> 41
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 154
atgggtctct tcgttggtt aagcttaact ttaagaagga g 41

60 <210> 155
 <211> 40
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 155
 5 atgggtctca cacaaactcg agtacaactt taactcacac 40

 <210> 156
 <211> 33
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 15 <400> 156
 atgggtctcg attccattct ttgtttgtc tgc 33

 <210> 157
 <211> 29
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 25 <400> 157
 attggtctcg gccgagcggg tgaaggcac 29

 <210> 158
 30 <211> 29
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 158
 attggtctct ctggaaccct ttcgggtcg 29

 40 <210> 159
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 159
 50 gccgaggagc tggtcaccg 19

 <210> 160
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 55

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 160
 60 gtttacgtcg ccgtccagct c 21

 <210> 161
 <211> 20

<212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 161
 ggccaacact tgcactact 20
 10 <210> 162
 <211> 19
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 162
 tccggataac gggaaaagc 19
 20 <210> 163
 <211> 24
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 163
 30 ctcgagtaca actttaactc acac 24
 <210> 164
 <211> 24
 <212> ADN
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 40 <400> 164
 gattccattc tttgttgt ctgc 24
 <210> 165
 <211> 26
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 50 <400> 165
 ctgtttcctg tgtgaaattg ttatcc 26
 <210> 166
 <211> 22
 55 <212> ADN
 <213> Secuencia artificial
 <220>
 60 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 166
 ggctttacac ttatgcttc cg 22

5 <210> 167
 <211> 26
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 10 <400> 167
 cccgggttat tacatgcgct agcact 26

 15 <210> 168
 <211> 48
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 20 <400> 168
 gaaattaata cgactcacta tagggtaag ctaactta agaaggag 48

 25 <210> 169
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 30 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 169
 tgcaagcagc agattacgcg c 21

 35 <210> 170
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 170
 gtaactgtca gaccaagttt actc 24

 45 <210> 171
 <211> 60
 <212> ADN
 <213> Secuencia artificial

 50 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 55 <220>
 <221> base_modificada
 <222> (22)..(22)
 <223> dNaM

 <400> 171
 60 ggcgaggagc tgttcacgg gntggtgcc atcctggtcg agctggacgg cgacgtaaac 60

 <210> 172
 <211> 60
 <212> ADN

	<213> Secuencia artificial	
	<220>	
5	<223> Descripción de la secuencia artificial: Oligonucleótido sintético	
	<220>	
	<221> base_modificada	
	<222> (34)..(34)	
10	<223> dNaM	
	<400> 172	
	gtttacgtcg ccgtccagct cgaccaggat gggnaccacc ccggtgaaca gctcctcgcc 60	
15	<210> 173	
	<211> 60	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético	
	<220>	
	<221> base_modificada	
25	<222> (27)..(27)	
	<223> dNaM	
	<400> 173	
	gtttacgtcg ccgtccagct cgaccangat gggcaccacc ccggtgaaca gctcctcgcc 60	
30	<210> 174	
	<211> 60	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético	
	<220>	
40	<221> base_modificada	
	<222> (39)..(39)	
	<223> dNaM	
	<400> 174	
45	gtttacgtcg ccgtccagct cgaccaggat gggcaccanc ccggtgaaca gctcctcgcc 60	
	<210> 175	
	<211> 75	
	<212> ADN	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético	
55	<220>	
	<221> base_modificada	
	<222> (34)..(34)	
	<223> dNaM	
60	<400> 175	
	ctgttttcctg tgtgaaattg ttatccgctc acanttcac acaacatacg agccggaagc 60	
	ataaagtgtg aagcc 75	

5 <210> 176
 <211> 63
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

10 <220>
 <221> base_modificada
 <222> (31)..(31)
 <223> dNaM

15 <400> 176
 ctcgagtaca actttaactc acacaatgta nagatcacgg cagacaaaca aaagaatgga 60
 atc 63

20 <210> 177
 <211> 58
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

30 <220>
 <221> base_modificada
 <222> (31)..(31)
 <223> dNaM

<400> 177
 ccggataacg ggaaaagcat tgaacaccgc nggtcagagt agtgacaagt gttggcca 58

35 <210> 178
 <211> 58
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

45 <220>
 <221> base_modificada
 <222> (27)..(27)
 <223> dNaM

<400> 178
 ggccaacact tgtcactact ctgaccnagg gtgttcaatg ctttcccgt tatccgga 58

50 <210> 179
 <211> 60
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

60 <220>
 <221> base_modificada
 <222> (34)..(34)
 <223> dNaM

<400> 179
ggcgaggagc tgttcaccgg ggtggtgccc atcntggtcg agctggacgg cgacgtaaac 60

5 <210> 180
<211> 60
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético

15 <220>
<221> base_modificada
<222> (27)..(27)
<223> dNaM

<400> 180
ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg agctggacgg cgacgtaaac 60

20 <210> 181
<211> 64
<212> ADN
<213> Secuencia artificial

25 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético

30 <220>
<221> base_modificada
<222> (32)..(32)
<223> dNaM

<400> 181
gattccattc tttgtttgt ctgccgtgat tnatacattg tgtgagtaa agttgtactc 60

gagt 64

35 <210> 182
<211> 134
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Descripción de la secuencia artificial: Polinucleótido sintético

45 <220>
<221> base_modificada
<222> (63)..(63)
<223> dNaM

50 <400> 182

cacacaggaa acagctatga cccgggttat tacatgctgct agcacttgga attcacaata 60

ctntcttttaa ggaaccata gtaaactctcc ttcttaaagt taagcttaac cctatagtga 120

gtcgtattaa ttcc 134

55 <210> 183
<211> 63
<212> ADN
<213> Secuencia artificial

	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
5	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		
10	<400> 183		
	ctcgagtaca actttaactc acacaatgta anaatcacgg cagacaaaca aaagaatgga	60	
	atc	63	
15	<210> 184		
	<211> 63		
	<212> ADN		
	<213> Secuencia artificial		
20	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
25	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		
	<400> 184		
	ctcgagtaca actttaactc acacaatgta ancatcacgg cagacaaaca aaagaatgga	60	
	atc	63	
30	<210> 185		
	<211> 63		
	<212> ADN		
	<213> Secuencia artificial		
35	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
40	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		
	<400> 185		
	ctcgagtaca actttaactc acacaatgta angatcacgg cagacaaaca aaagaatgga	60	
45	atc	63	
50	<210> 186		
	<211> 63		
	<212> ADN		
	<213> Secuencia artificial		
55	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		

	<400> 186	
	ctcgagtaca actttaactc acacaatgtc antgtcacgg cagacaaaca aaagaatgga	60
	atc	63
5	<210> 187 <211> 63 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
15	<220> <221> base_modificada <222> (32)..(32) <223> dNaM	
	<400> 187	
	ctcgagtaca actttaactc acacaatgta cnaatcacgg cagacaaaca aaagaatgga	60
20	atc	63
	<210> 188 <211> 63 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
30	<220> <221> base_modificada <222> (32)..(32) <223> dNaM	
35	<400> 188	
	ctcgagtaca actttaactc acacaatgta cncatcacgg cagacaaaca aaagaatgga	60
	atc	63
40	<210> 189 <211> 63 <212> ADN <213> Secuencia artificial	
45	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
50	<220> <221> base_modificada <222> (32)..(32) <223> dNaM	
	<400> 189	
	ctcgagtaca actttaactc acacaatgta cngatcacgg cagacaaaca aaagaatgga	60
55	atc	63

<210> 190
 <211> 63
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 <220>
 10 <221> base_modificada
 <222> (32)..(32)
 <223> dNaM
 <400> 190
 15
 ctcgagtaca actttaactc acacaatgta cntatcacgg cagacaaaca aaagaatgga 60
 atc 63
 <210> 191
 <211> 63
 20 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 25
 <220>
 <221> base_modificada
 <222> (32)..(32)
 <223> dNaM
 30
 <400> 191
 ctcgagtaca actttaactc acacaatgta gnaatcacgg cagacaaaca aaagaatgga 60
 atc 63
 35 <210> 192
 <211> 63
 <212> ADN
 <213> Secuencia artificial
 40 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 <220>
 <221> base_modificada
 45 <222> (32)..(32)
 <223> dNaM
 <400> 192
 ctcgagtaca actttaactc acacaatgta gncatcacgg cagacaaaca aaagaatgga 60
 50 atc 63
 <210> 193
 <211> 63
 <212> ADN
 55 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		
5	<400> 193		
	ctcgagtaca actttaactc acacaatgta gngatcacgg cagacaaaca aaagaatgga	60	
	atc	63	
10	<210> 194		
	<211> 63		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
	<220>		
	<221> base_modificada		
20	<222> (32)..(32)		
	<223> dNaM		
	<400> 194		
	ctcgagtaca actttaactc acacaatgta gntatcacgg cagacaaaca aaagaatgga	60	
25	atc	63	
	<210> 195		
	<211> 63		
	<212> ADN		
30	<213> Secuencia artificial		
	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
35	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
	<223> dNaM		
40	<400> 195		
	ctcgagtaca actttaactc acacaatgta tnaatcacgg cagacaaaca aaagaatgga	60	
	atc	63	
	<210> 196		
45	<211> 63		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
50	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
	<220>		
	<221> base_modificada		
	<222> (32)..(32)		
55	<223> dNaM		
	<400> 196		

	ctcgagtaca actttaactc acacaatgta tncatcacgg cagacaaaca aaagaatgga	60
	atc	63
5	<210> 197 <211> 63 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
15	<220> <221> base_modificada <222> (32)..(32) <223> dNaM	
	<400> 197	
	ctcgagtaca actttaactc acacaatgta tngatcacgg cagacaaaca aaagaatgga	60
	atc	63
20	<210> 198 <211> 63 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
30	<220> <221> base_modificada <222> (32)..(32) <223> dNaM	
	<400> 198	
	ctcgagtaca actttaactc acacaatgta tntatcacgg cagacaaaca aaagaatgga	60
35	atc	63
40	<210> 199 <211> 63 <212> ADN <213> Secuencia artificial	
45	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético	
50	<220> <221> base_modificada <222> (33)..(33) <223> dNaM	
	<400> 199	
	ctcgagtaca actttaactc acacaatgta agnatcacgg cagacaaaca aaagaatgga	60
	atc	63
55	<210> 200 <211> 56 <212> ADN	

<213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 5
 <220>
 <221> base_modificada
 <222> (31)..(31)
 <223> dNaM
 10
 <400> 200
 ctctggaacc cttcggtgc gccggttag nagaccggtg cctcaaccg ctcggc 56
 <210> 201
 15 <211> 23
 <212> ADN
 <213> Secuencia artificial
 <220>
 20 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 <220>
 <221> base_modificada
 <222> (18)..(18)
 25 <223> dC, dT, dA, dG, dTPT3 o dNaM
 <220>
 <223> Véase la memoria descriptiva según se presentó para descripción detallada de sustituciones y realizaciones preferidas
 30
 <400> 201
 ctggtcctac ccgtggtngg tcc 23
 <210> 202
 35 <211> 20
 <212> ADN
 <213> Secuencia artificial
 <220>
 40 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 <220>
 <223> Descripción de molécula de ADN/ARN combinada: Oligonucleótido sintético
 45
 <220>
 <221> base_modificada
 <222> (18)..(18)
 <223> g, a, u o c
 50
 <220>
 <223> Véase la memoria descriptiva según se presentó para descripción detallada de sustituciones y realizaciones preferidas
 <400> 202
 55 gaccaggatg ggcaccancc 20
 <210> 203
 <211> 20
 <212> ARN
 60 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético
 65
 <400> 203

gaccaggaug ggcaccaccc 20

5 <210> 204
 <211> 18
 <212> ARN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 <400> 204
 guugugugga aaugugag 18

15 <210> 205
 <211> 18
 <212> ARN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 <400> 205
 uguugugugg aaugugag 18

25 <210> 206
 <211> 22
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

35 <220>
 <223> Descripción de molécula de ADN/ARN combinada: Oligonucleótido sintético

 <220>
 <221> base_modificada
 <222> (16)..(16)
 <223> dTPT3

40 <400> 206
 guatgttg tggaantgtg ag 22

45 <210> 207
 <211> 13
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

55 <220>
 <221> CDS
 <222> (1)..(12)

 <400> 207

 tca cat ttc cac a
 Ser His Phe His
 1

60 <210> 208
 <211> 4
 <212> PRT

<213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 5 <400> 208

 Ser His Phe His
 1

 10 <210> 209
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 <220>
 <221> base_modificada
 20 <222> (14)..(14)
 <223> dNaM

 <400> 209
 tcacacaatg tagngatcac gg 22
 25
 <210> 210
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 <220>
 <221> base_modificada
 35 <222> (12)..(12)
 <223> dNaM

 <400> 210
 40 accaggatgg gnaccacccc gg 22

 <210> 211
 <211> 22
 <212> ADN
 45 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 50 <220>
 <221> base_modificada
 <222> (14)..(14)
 <223> dNaM

 55 <400> 211
 tcacacaatg tagncatcac gg 22

 <210> 212
 <211> 22
 60 <212> ADN
 <213> Secuencia artificial

 <220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<220>
 <221> base_modificada
 5 <222> (17)..(17)
 <223> dTPT3

<400> 212
 accaggatgg gcaccanccc gg 22

10 <210> 213
 <211> 22
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<220>
 20 <221> base_modificada
 <222> (5)..(5)
 <223> dNaM

<400> 213
 25 accangatgg gcaccacccc gg 22

<210> 214
 <211> 22
 <212> ADN
 30 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

35 <220>
 <221> base_modificada
 <222> (13)..(13)
 <223> dNaM

40 <400> 214
 tcacacaatg tanagatcac gg 22

<210> 215
 <211> 22
 45 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

50 <220>
 <221> base_modificada
 <222> (17)..(17)
 <223> dNaM

55 <400> 215
 accaggatgg gcaccanccc gg 22

60 <210> 216
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<220>
 <221> base_modificada
 5 <222> (13)..(13)
 <223> dTPT3

<400> 216
 tgttgtgtgg aantgtgagc gg 22

10 <210> 217
 <211> 22
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<220>
 20 <221> base_modificada
 <222> (18)..(18)
 <223> dTPT3

<400> 217
 25 ttgtcactac tctgaccngc gg 22

<210> 218
 <211> 22
 <212> ADN
 30 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

35 <220>
 <221> base_modificada
 <222> (18)..(18)
 <223> dNaM

40 <400> 218
 ttgtcactac tctgaccnag gg 22

<210> 219
 <211> 22
 45 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

50 <220>
 <221> base_modificada
 <222> (14)..(14)
 <223> dNaM

55 <400> 219
 tcacacaatg tacncatcac gg 22

60 <210> 220
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 5 <220>
 <221> base_modificada
 <222> (5)..(5)
 <223> dTPT3

 <400> 220
 10 accangatgg gcaccacccc gg 22

 <210> 221
 <211> 22
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 20 <220>
 <221> base_modificada
 <222> (14)..(14)
 <223> dNaM

 25 <400> 221
 tcacacaatg tatngatcac gg 22

 <210> 222
 <211> 22
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 35 <220>
 <221> base_modificada
 <222> (14)..(14)
 <223> dTPT3

 40 <400> 222
 tcacacaatg tatnaatcac gg 22

 <210> 223
 45 <211> 22
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 50 <220>
 <221> base_modificada
 <222> (12)..(12)
 55 <223> dTPT3

 <400> 223
 accaggatgg gnaccacccc gg 22

 60 <210> 224
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

5 <220>
 <221> base_modificada
 <222> (13)..(13)
 <223> dNaM

10 <400> 224
 attcacaata ctnctttaa gg 22

15 <210> 225
 <211> 22
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 225

Met	Lys	Tyr	Leu	Leu	Pro	Thr	Ala	Glu	Ala	Gly	Leu	Leu	Leu	Leu	Ala
1				5					10					15	

Ala	Gln	Pro	Ala	Ile	Ala
				20	

25 <210> 226
 <211> 26
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 226

Met	Lys	Ile	Lys	Thr	Gly	Ala	Arg	Ile	Leu	Ala	Leu	Ser	Glu	Leu	Thr
1				5					10					15	

Thr	Met	Met	Phe	Ser	Ala	Ser	Ala	Leu	Ala
			20					25	

35 <210> 227
 <211> 21
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

45 <400> 227

Met	Lys	Gln	Ser	Thr	Ile	Ala	Leu	Ala	Leu	Leu	Pro	Leu	Leu	Phe	Thr
1				5					10					15	

Pro	Val	Thr	Lys	Ala
			20	

5 <210> 228
 <211> 30
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Polipéptido sintético

 10 <400> 228

 Met Lys Ser Pro Ala Pro Ser Arg Pro Gln Lys Met Ala Leu Ile Pro
 1 5 10 15

 Ala Cys Ile Phe Leu Cys Phe Ala Ala Leu Ser Val Gln Ala
 20 25 30

 15 <210> 229
 <211> 20
 <212> PRT
 <213> Secuencia artificial

 <220>
 20 <223> Descripción de la secuencia artificial: Péptido sintético

 <400> 229

 Met Lys Lys Ile Leu Val Ser Phe Val Ala Ile Met Ala Val Ala Ser
 1 5 10 15

 Ser Ala Met Ala
 20
 25
 <210> 230
 <211> 40
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de la secuencia artificial: Polipéptido sintético

 <400> 230
 35
 Met Lys Ile Ser Met Gln Lys Ala Asp Phe Trp Lys Lys Ala Ala Ile
 1 5 10 15

 Ser Leu Leu Val Phe Thr Met Phe Phe Thr Leu Met Met Ser Glu Thr
 20 25 30

 Val Phe Ala Ala Gly Leu Asn Lys
 35 40

 40 <210> 231
 <211> 28
 <212> PRT
 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 231

5

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
1 5 10 15

Thr Val Ala Gln Ala Ser Ala Gly Leu Asn Lys Asp
20 25

<210> 232

<211> 19

10

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

15

<400> 232

Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
1 5 10 15

Ala Ser Ala

20

<210> 233

<211> 57

<212> ADN

<213> Secuencia artificial

25

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 233

30

atgaaaaaga ttggctggc gctggctggt ttgttttag cgttagcgc atcggcg 57

<210> 234

<211> 22

<212> PRT

35

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

40

<400> 234

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met Ala
20

45

<210> 235

<211> 66

<212> ADN

<213> Secuencia artificial

	<220>		
	<223> Descripción de la secuencia artificial: Oligonucleótido sintético		
5	<400> 235		
	atgaaatacc tgctgccgac cgctgctgct ggtctgctgc tcctcgctgc ccagccggcg	60	
	atggcg	66	
10	<210> 236 <211> 63 <212> ADN <213> Secuencia artificial		
15	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético		
	<400> 236		
	atgaaacaaa gcactattgc actggcactc ttaccgttac tgtttaccac tgtgacaaaa	60	
20	gcg	63	
	<210> 237 <211> 26 <212> PRT <213> Secuencia artificial		
25	<220> <223> Descripción de la secuencia artificial: Péptido sintético		
	<400> 237		
30	Met Lys Thr His Ile Val Ser Ser Val Thr Thr Thr Leu Leu Leu Gly		
	1 5 10 15		
	Ser Ile Leu Met Asn Pro Val Ala Asn Ala		
	20 25		
35	<210> 238 <211> 78 <212> ADN <213> Secuencia artificial		
40	<220> <223> Descripción de la secuencia artificial: Oligonucleótido sintético		
	<400> 238		
	atgaaaacac atatagtcag ctacagtaaca acaacactat tgctaggttc catattaatg	60	
	aatcctgtcg ctaatgcc	78	
45	<210> 239 <211> 21 <212> PRT <213> Secuencia artificial		
50	<220>		

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 239

Met Lys Tyr Leu Leu Pro Trp Leu Ala Leu Ala Gly Leu Val Leu Ala
1 5 10 15

Phe Ser Ala Ser Ala
20

5

<210> 240

<211> 63

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

15 <400> 240

atgaaatacc tgctgccgtg gctggcgctg gctgggttag ttttagcgtt tagcgcatcg 60

gcg 63

<210> 241

20 <211> 20

<212> PRT

<213> Secuencia artificial

<220>

25 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 241

Met Lys Lys Ile Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Phe
1 5 10 15

Ser Ala Ser Ala
20

30

<210> 242

<211> 60

<212> ADN

35 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 242

40 atgaaaaaga ttaccgctgc tgctggtctg ctgctcctcg ctgcgttag cgcacggcg 60

<210> 243

<211> 19

<212> PRT

45 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

50 <400> 243

Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Gln Pro
1 5 10 15

Ala Met Ala

5 <210> 244
<211> 57
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 244
atgaaaaaga ttggctggc gctggctggt ttagtttag ccagccggc gatggcg 57

15 <210> 245
<211> 19
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: Péptido sintético

<400> 245

Met Lys Lys Ile Leu Val Leu Gly Ala Leu Ala Leu Trp Ala Gln Pro
1 5 10 15

Ala Met Ala

25 <210> 246
<211> 57
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 246
35 atgaaaaaga ttttagttt aggtgctctg gcgctgtggg ccagccggc gatggcg 57

40 <210> 247
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Péptido sintético

45 <400> 247

Met Lys Lys Ile Trp Leu Ala Leu Val Leu Leu Ala Gly Ala Gln Pro
1 5 10 15

Ala Met Ala

<210> 248

<211> 57
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 248
 atgaaaaaga ttggctggc gtagtttta ctggctggtg cccagccggc gatggcg 57

10 <210> 249
 <211> 19
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

20 <400> 249

Met	Lys	Lys	Ile	Leu	Ala	Gly	Trp	Leu	Ala	Leu	Val	Leu	Ala	Gln	Pro
1				5					10					15	

Ala Met Ala

25 <210> 250
 <211> 57
 <212> **ADN**
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 250
 atgaaaaaga ttctggctgg ttggctggcg ttagtttttag cccagccggc gatggcg 57

35 <210> 251
 <211> 19
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 251

Met	Lys	Lys	Ile	Leu	Val	Leu	Leu	Ala	Gly	Trp	Leu	Ala	Ala	Gln	Pro
1				5					10					15	

Ala Met Ala

45 <210> 252
 <211> 57
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 252

atgaaaaaga ttttagtttt actggctggt tggctggcgg cccagccggc gatggcg 57

<210> 253

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 253

Met Lys Lys Ile Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln
1 5 10 15

Pro Ala Met Ala
20

<210> 254

<211> 60

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 254

atgaaaaaga ttaccgtgc tgctggtctg ctgctcctcg ctgccagcc ggcgatggcg 60

<210> 255

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 255

Met Lys Lys Ile Leu Leu Leu Leu Gly Thr Ala Ala Ala Ala Ala Gln
1 5 10 15

Pro Ala Met Ala
20

<210> 256

<211> 60

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 256

atgaaaaaga ttctgtgct cctcggtacc gctgctgctg ctgccagcc ggcgatggcg 60

<210> 257

<211> 20

<212> PRT

<213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

5 <400> 257

Met Lys Lys Ile Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Ala Gln
 1 5 10 15

Pro Ala Met Ala
 20

10 <210> 258
 <211> 60
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 258
 atgaaaaaga ttctgtctgt cctctgtctg ctctctctgc tcgccagcc ggcgatggcg 60

20 <210> 259
 <211> 20
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 259

Met Lys Lys Ile Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gln
 1 5 10 15

Pro Ala Met Ala
 20

30 <210> 260
 <211> 60
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

40 <400> 260
 atgaaaaaga ttgtgtctgc tgctgcggcg gcggcggtg cggcccagcc ggcgatggcg 60

45 <210> 261
 <211> 21
 <212> PRT
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 261

ES 2 929 047 T3

Met Lys Tyr Leu Leu Pro Trp Leu Ala Leu Ala Gly Leu Val Leu Ala
1 5 10 15

Gln Pro Ala Met Ala
20

<210> 262

<211> 63

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 262

atgaaatacc tgctgccgtg gctggcgctg gctggtttag ttttagccca gccggcgatg 60

gcg 63

<210> 263

<211> 22

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 263

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Phe Ser Ala Ser Ala
20

<210> 264

<211> 66

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

<400> 264

atgaaatacc tgctgccgac cgctgctgct ggtctgctgc tctcgctgc gtttagcgca 60

tcggcg 66

REIVINDICACIONES

- 5 1. Un transportador de nucleósido trifosfato modificado de *Phaeodactylum tricornutum* (PtNTT2) que comprende una
delección, en donde el transportador de nucleósido trifosfato modificado tiene al menos un 95 % de identidad de
secuencia con el SEQ ID NO: 4.
2. El transportador de nucleósido trifosfato modificado de la reivindicación 1, en donde el transportador de nucleósido
trifosfato modificado comprende además un péptido señal.
- 10 3. El transportador de nucleósido trifosfato modificado de la reivindicación 1, en donde el transportador de nucleósido
trifosfato modificado se obtiene a partir de una célula genomanipulada.
4. El transportador de nucleósido trifosfato modificado de la reivindicación 3, en donde la célula genomanipulada
comprende una célula procariota.
- 15 5. El transportador de nucleósido trifosfato modificado de la reivindicación 3, en donde la célula genomanipulada es
E. coli.
- 20 6. El transportador de nucleósido trifosfato modificado de una cualquiera de las reivindicaciones 1-5, en donde el
transportador de nucleósido trifosfato modificado tiene al menos un 97 % de identidad de secuencia con el SEQ ID
NO: 4.
7. El transportador de nucleósido trifosfato modificado de una cualquiera de las reivindicaciones 1-5, en donde la
secuencia del transportador de nucleósido trifosfato modificado consiste en el SEQ ID NO: 4.
- 25 8. Una célula genomanipulada que comprende:
una primera molécula de ácido nucleico que codifica un transportador de nucleósido trifosfato modificado de
Phaeodactylum tricornutum (PtNTT2), en donde el transportador de nucleósido trifosfato modificado comprende una
delección y tiene al menos un 95 % de identidad de secuencia con el SEQ ID NO: 4.
- 30 9. La célula genomanipulada de la reivindicación 8, en donde la célula genomanipulada comprende una célula
procariota.
10. La célula genomanipulada de la reivindicación 8, en donde la célula genomanipulada es *E. coli*.
- 35 11. La célula genomanipulada de acuerdo con una cualquiera de las reivindicaciones 8-10, en donde el transportador
de nucleósido trifosfato modificado tiene al menos un 97 % de identidad de secuencia con el SEQ ID NO: 4.
12. La célula genomanipulada de acuerdo con una cualquiera de las reivindicaciones 8-10, en donde la secuencia del
transportador de nucleósido trifosfato modificado consiste en el SEQ ID NO: 4.
- 40 13. La célula genomanipulada de acuerdo con una cualquiera de las reivindicaciones 8-12, en donde el ácido nucleico
del transportador de nucleósido trifosfato modificado está incorporado en la secuencia genómica de la célula
genomanipulada.
- 45 14. La célula genomanipulada de acuerdo con una cualquiera de las reivindicaciones 8-12, en donde la célula
genomanipulada comprende un plásmido que comprende dicho primer ácido nucleico que codifica el transportador de
nucleósido trifosfato modificado.

Fig. 1A

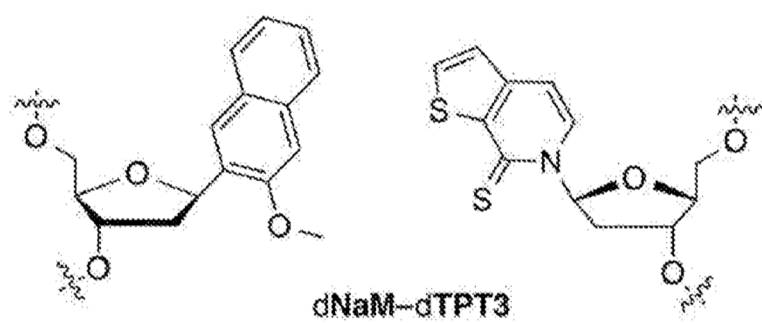
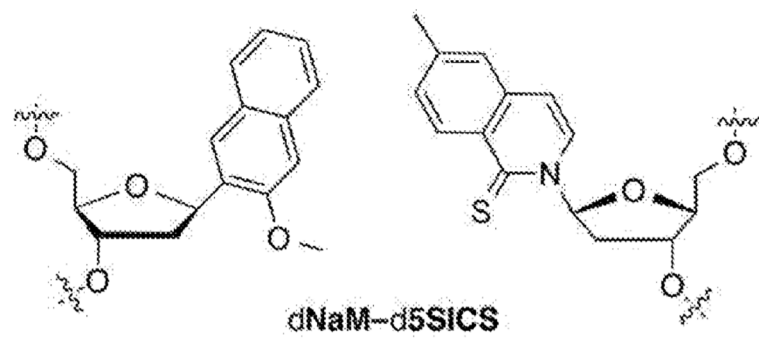
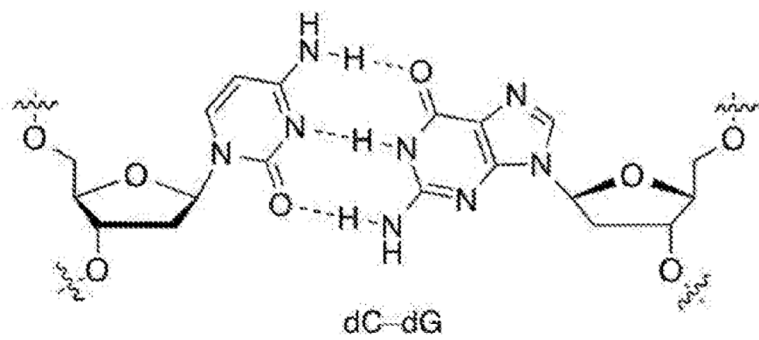


Fig. 1B

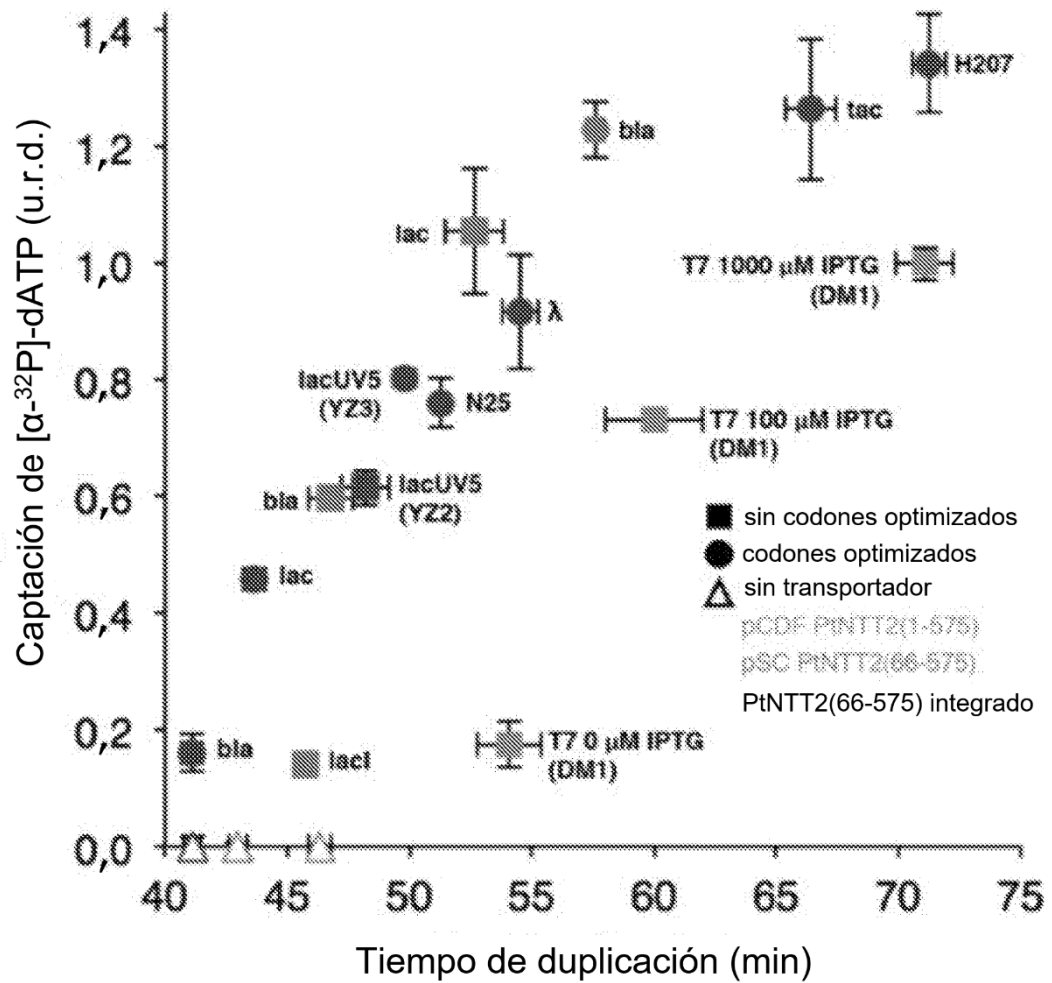


Fig. 2A

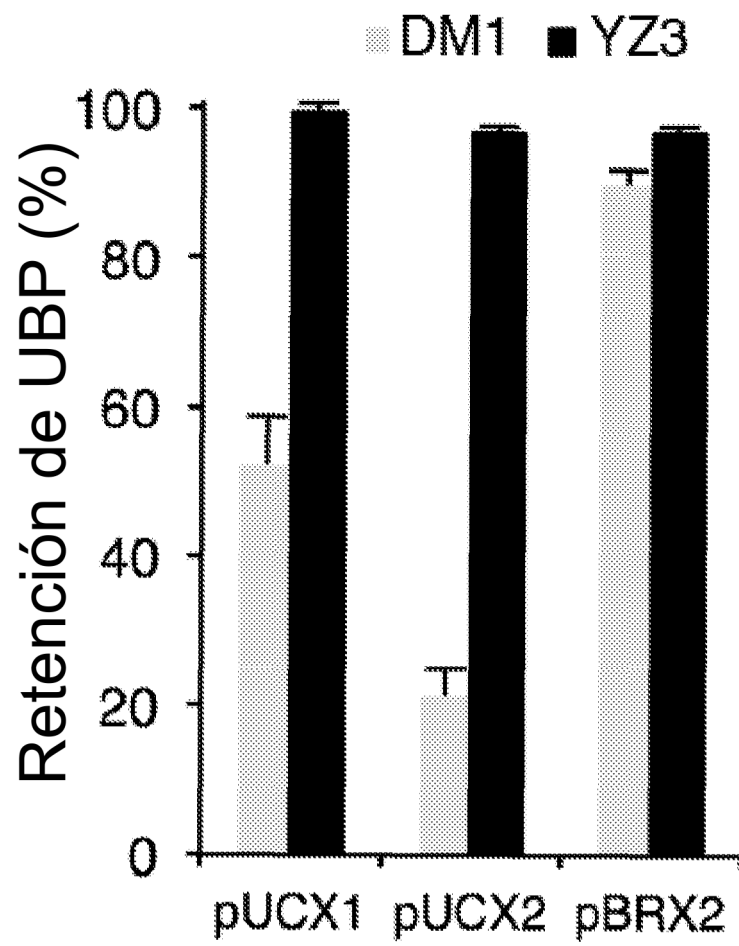


Fig. 2B

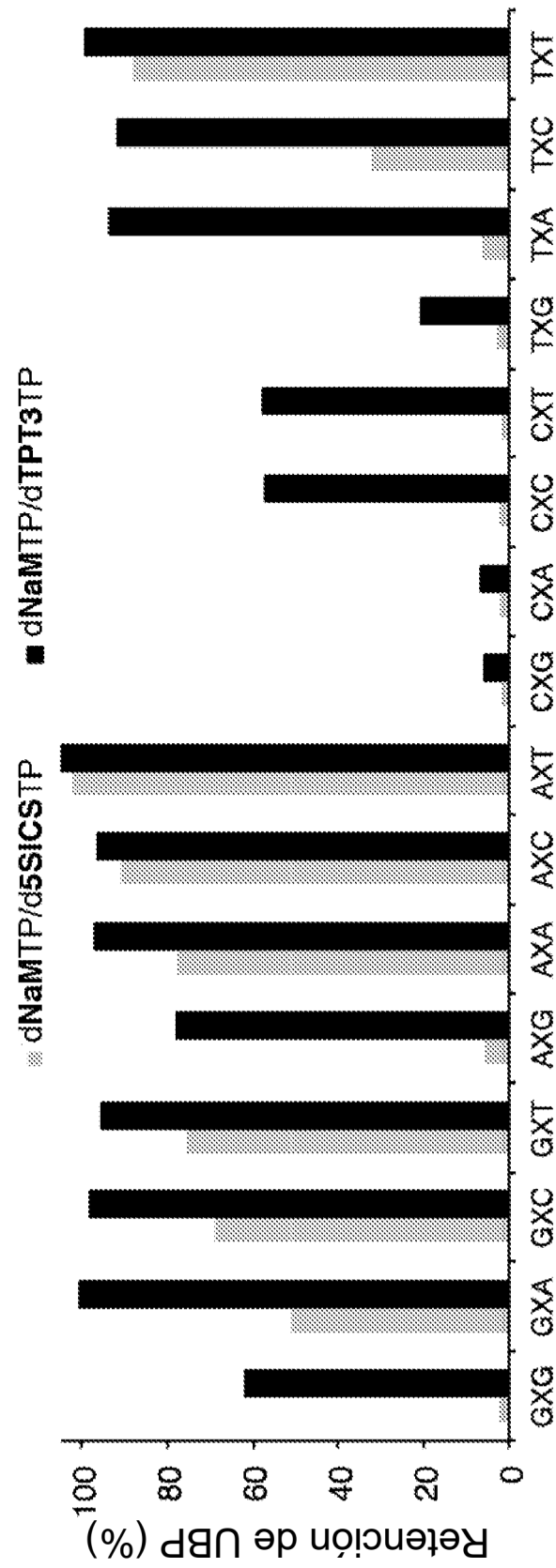
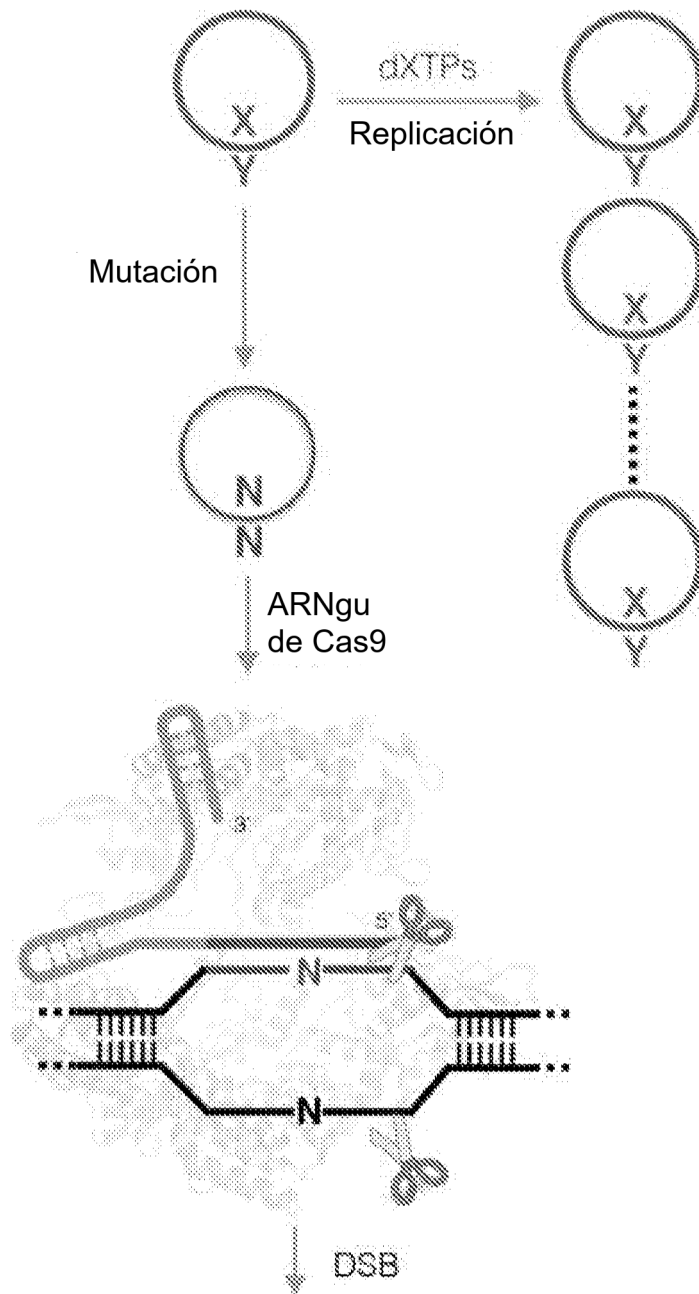


Fig. 3A



Degradación plasmídica

Fig. 3B

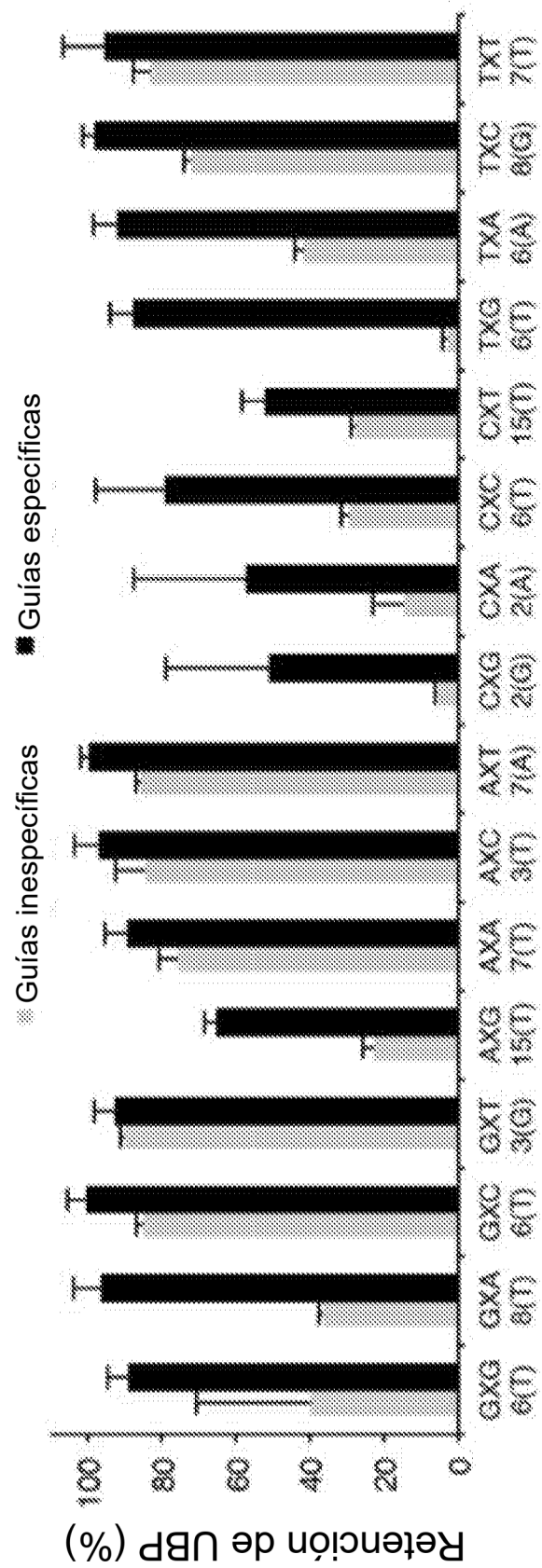


Fig. 3C

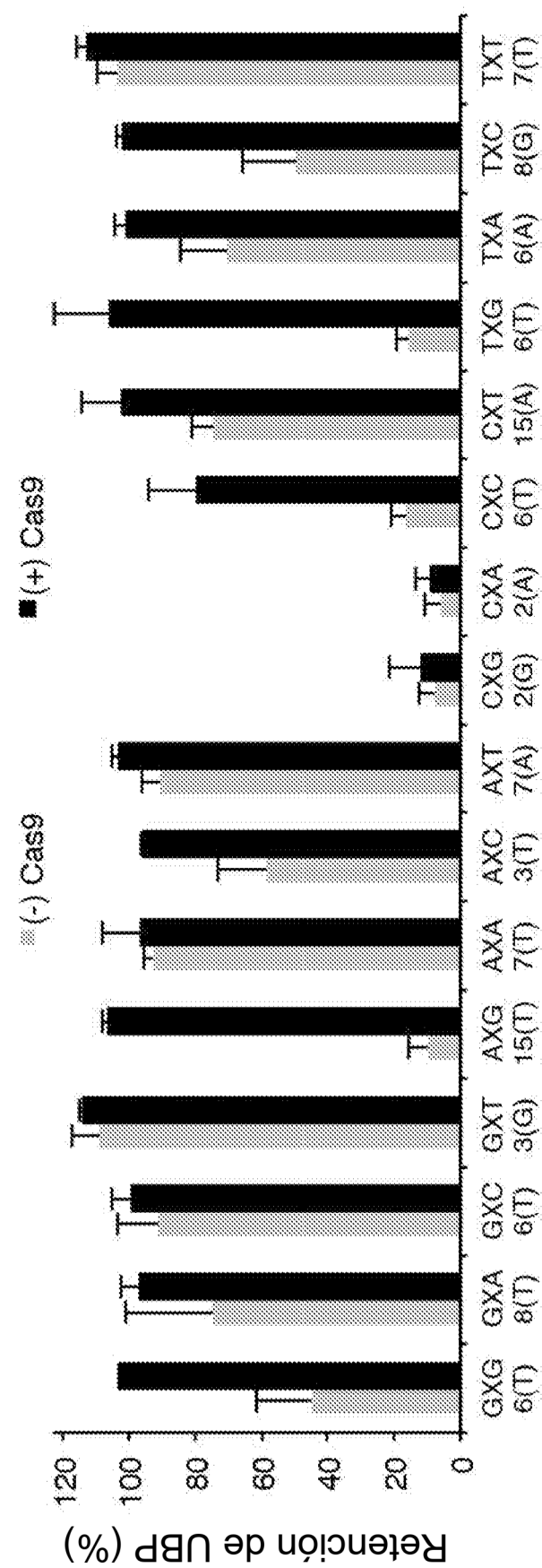


Fig. 4

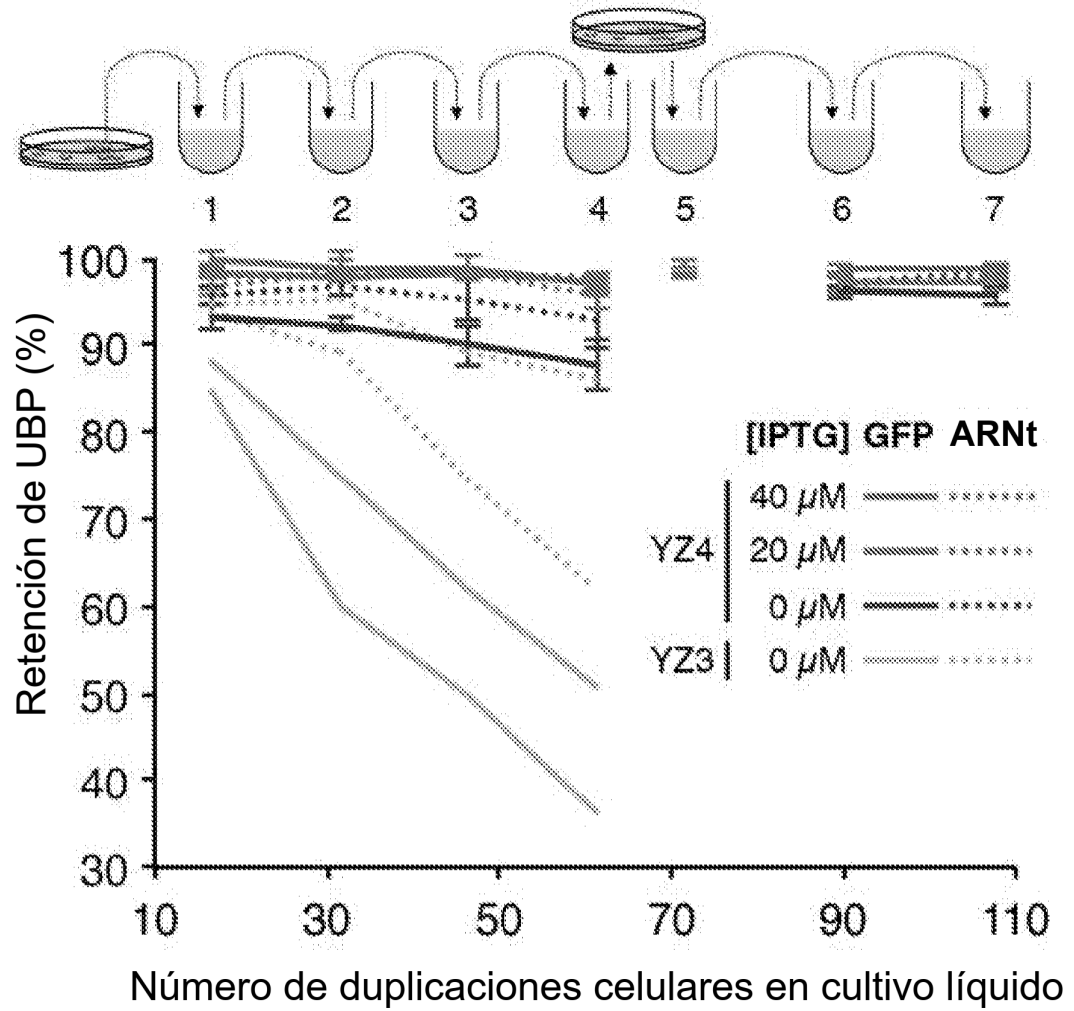


Fig. 5A



Fig. 5B

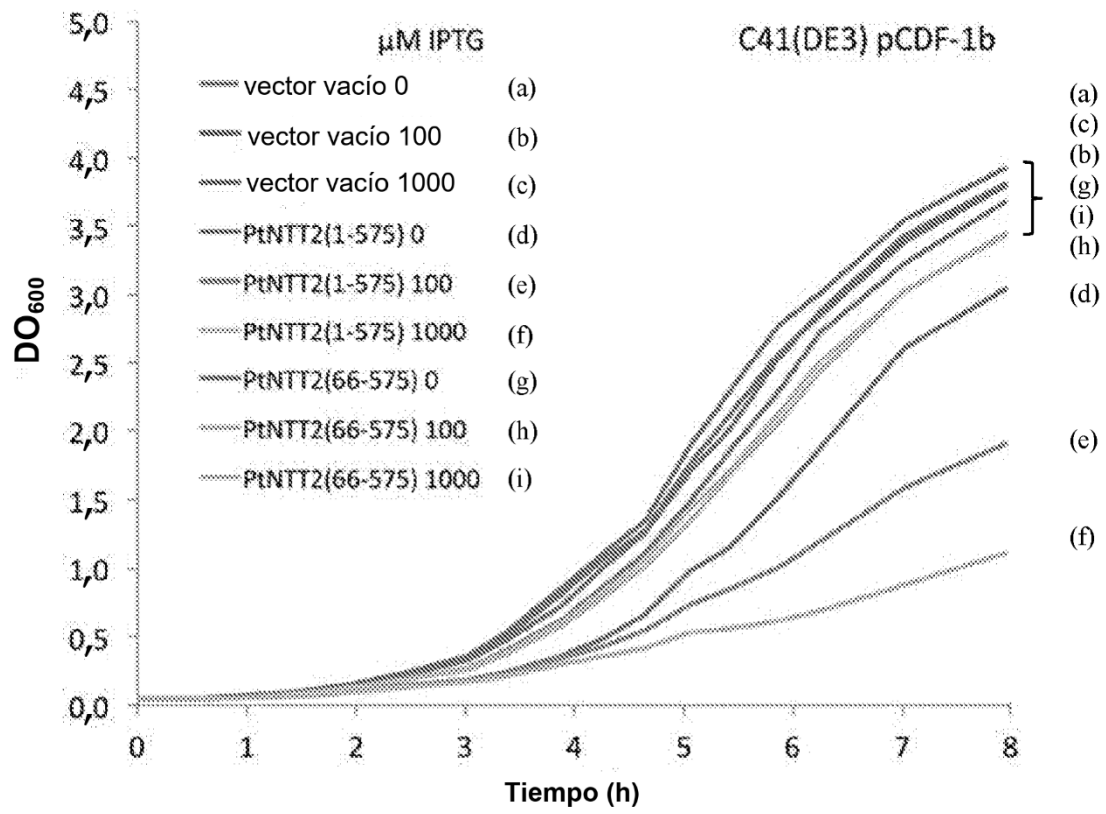


Fig. 5C

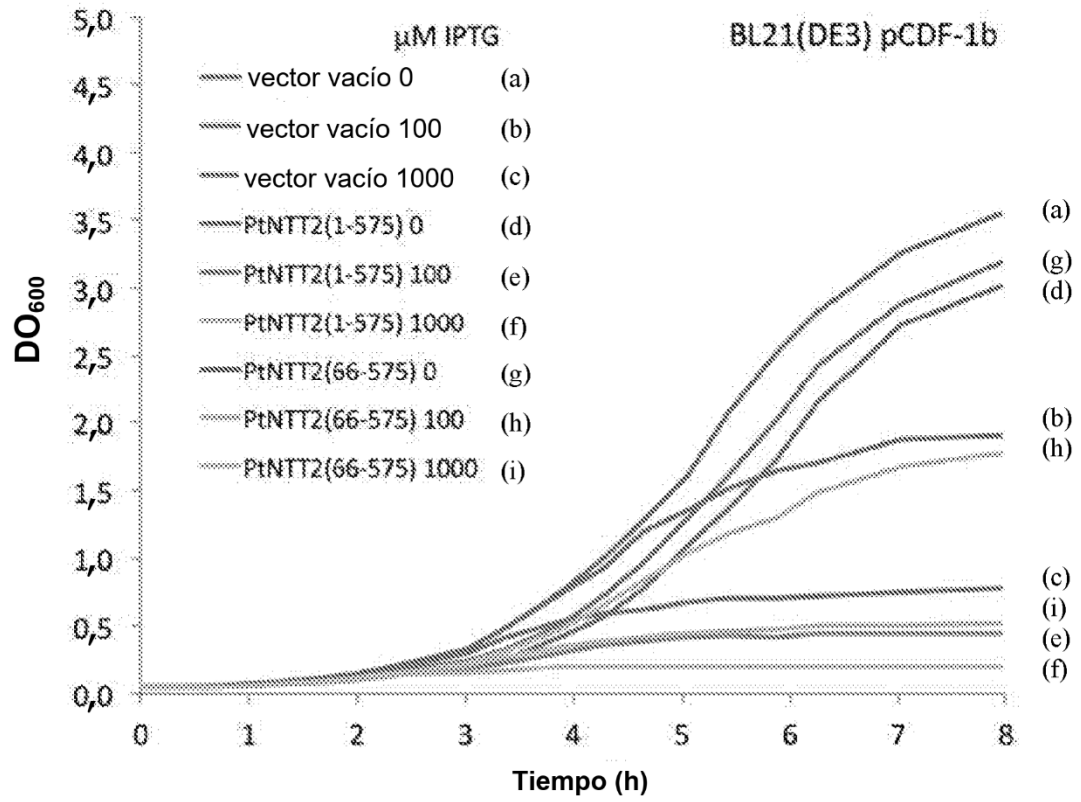


Fig. 5D

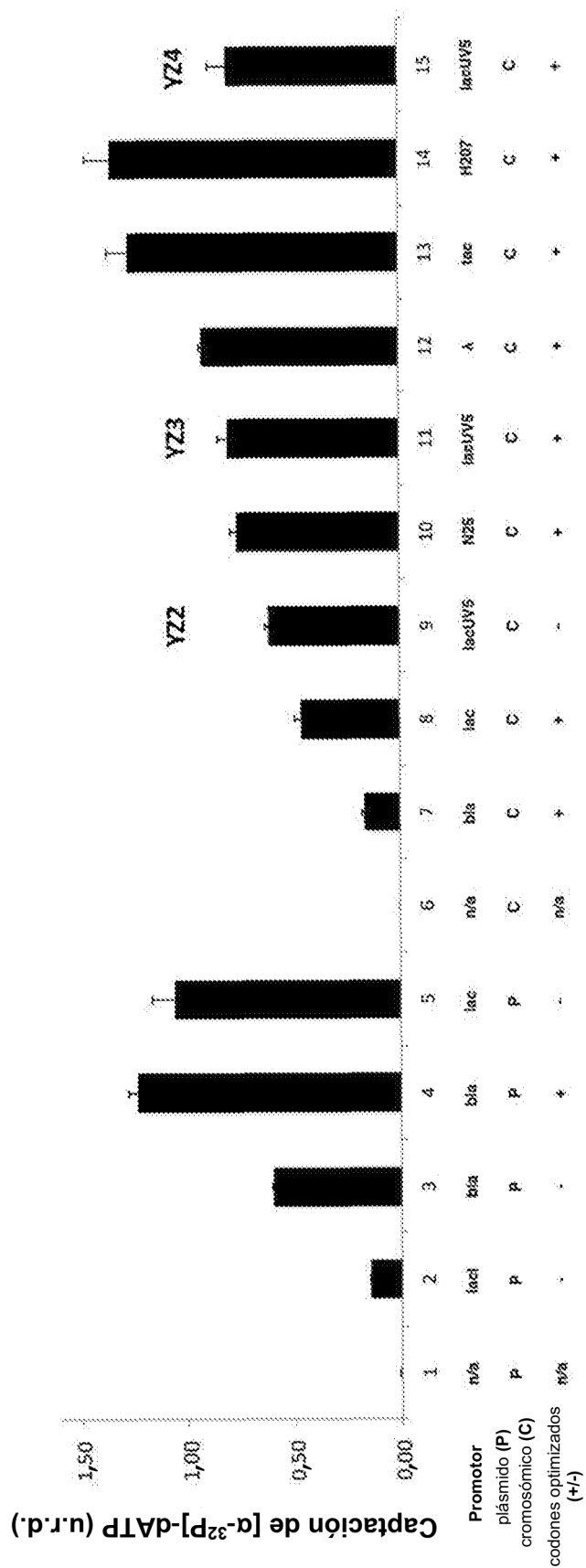


Fig. 5E

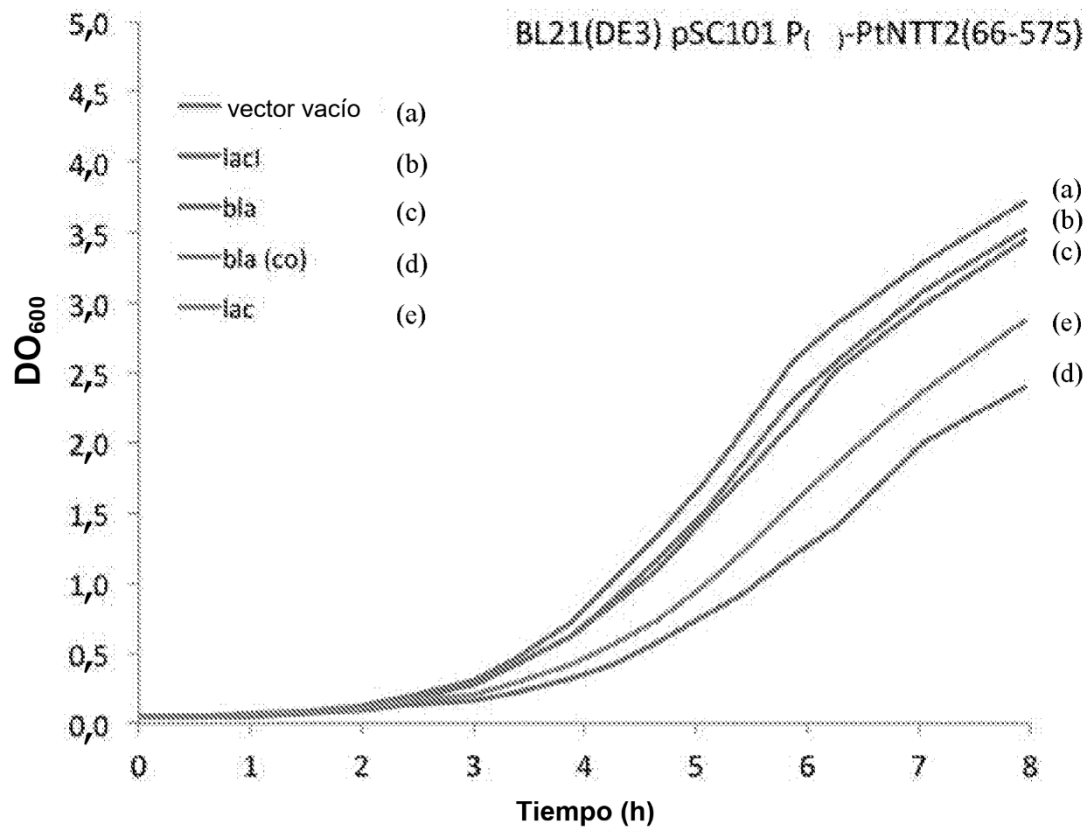


Fig. 5F

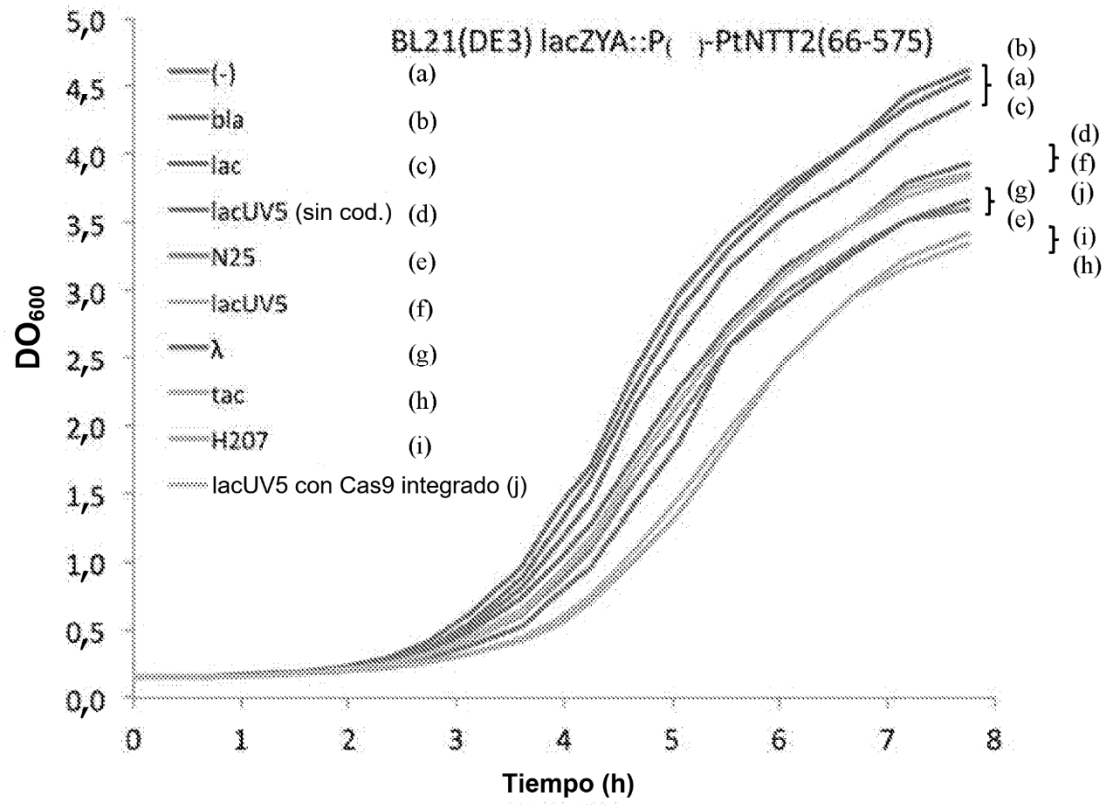


Fig. 6

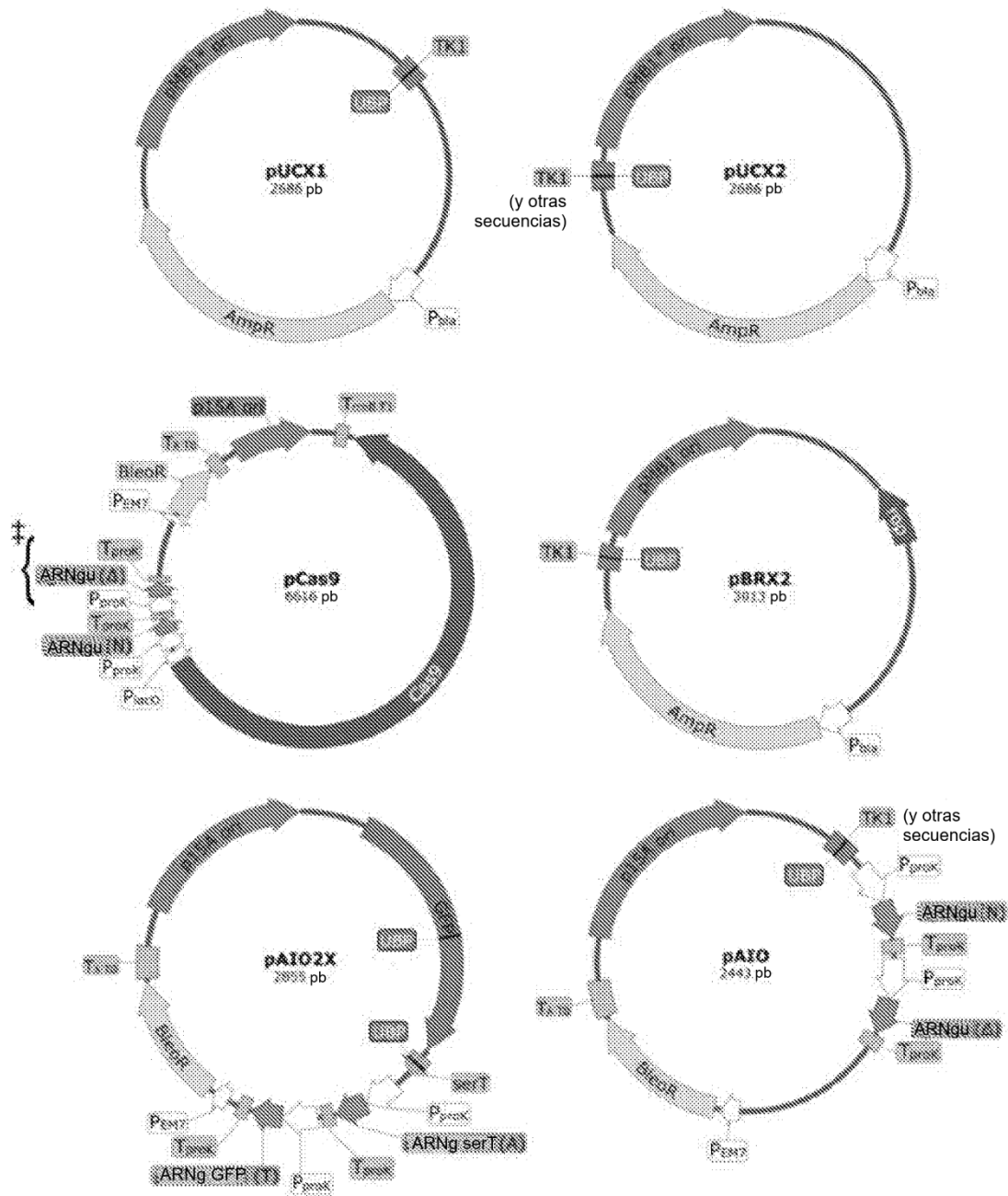


Fig. 7A

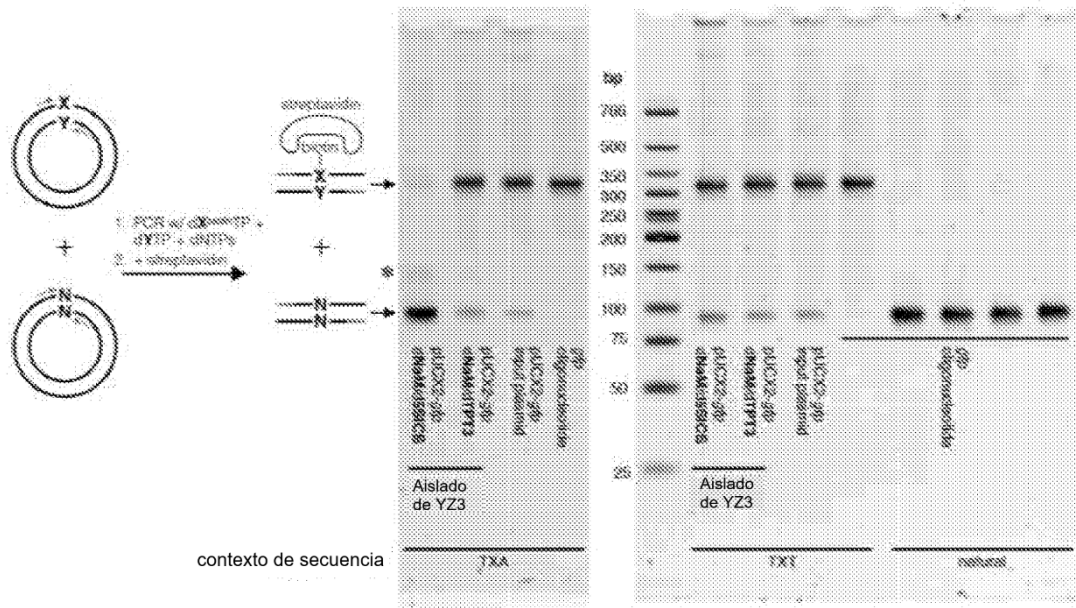


Fig. 7B

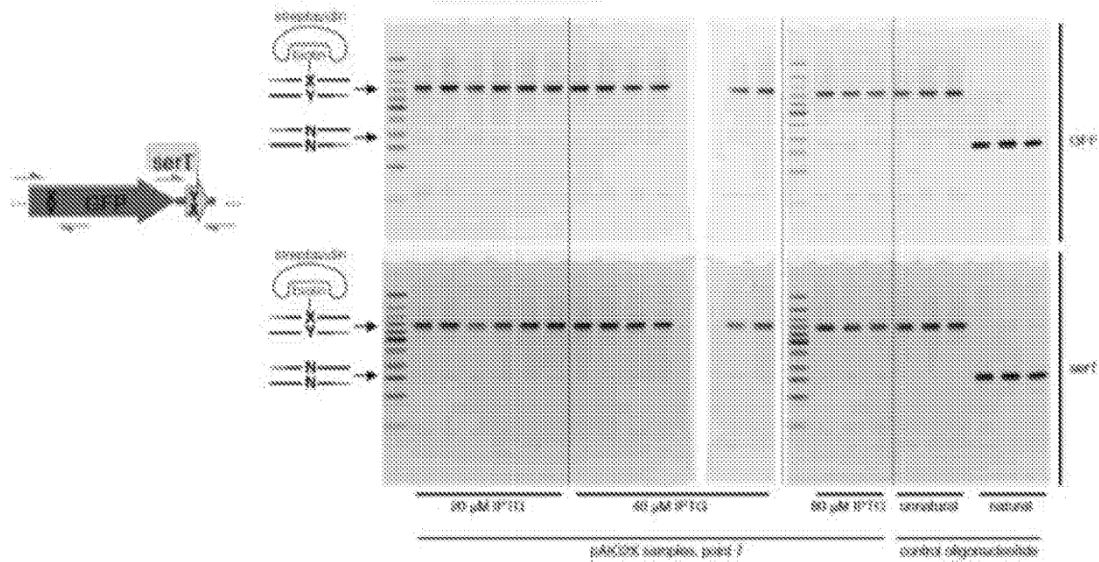


Fig. 8A

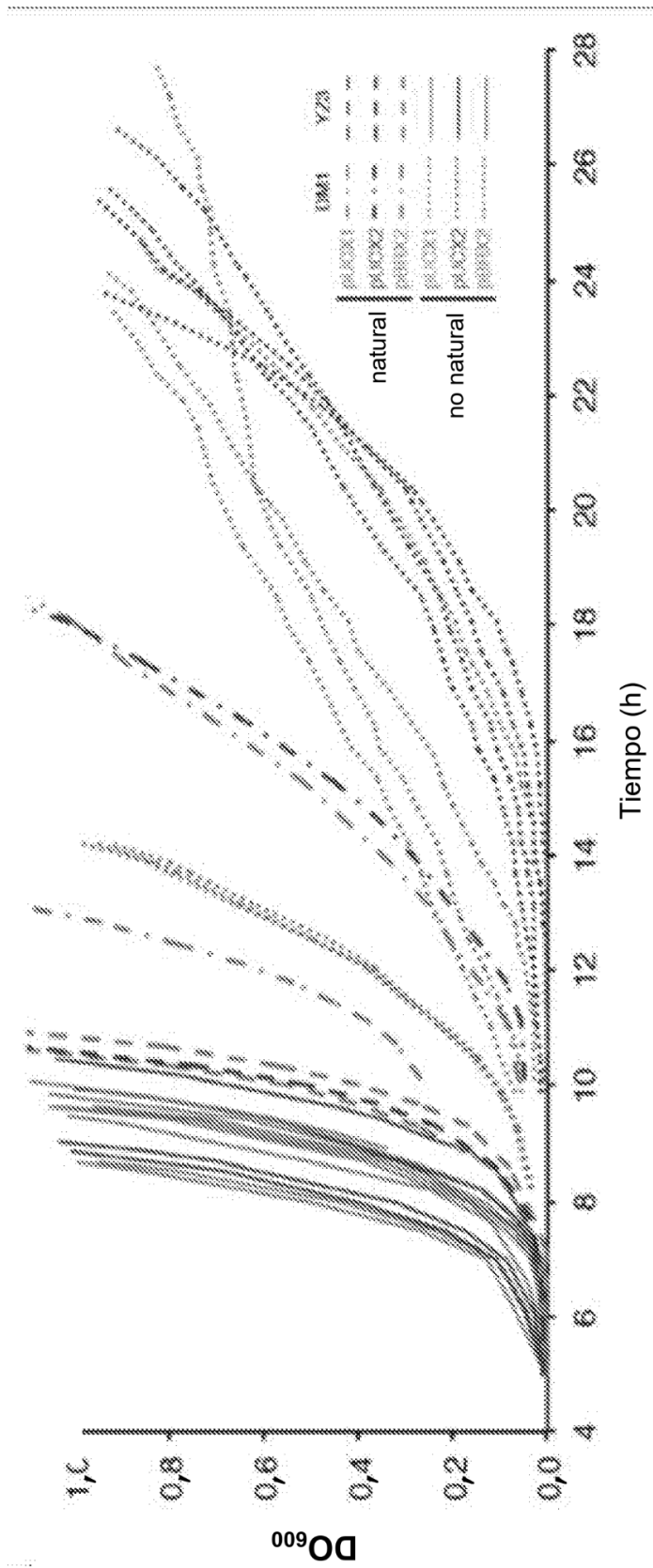
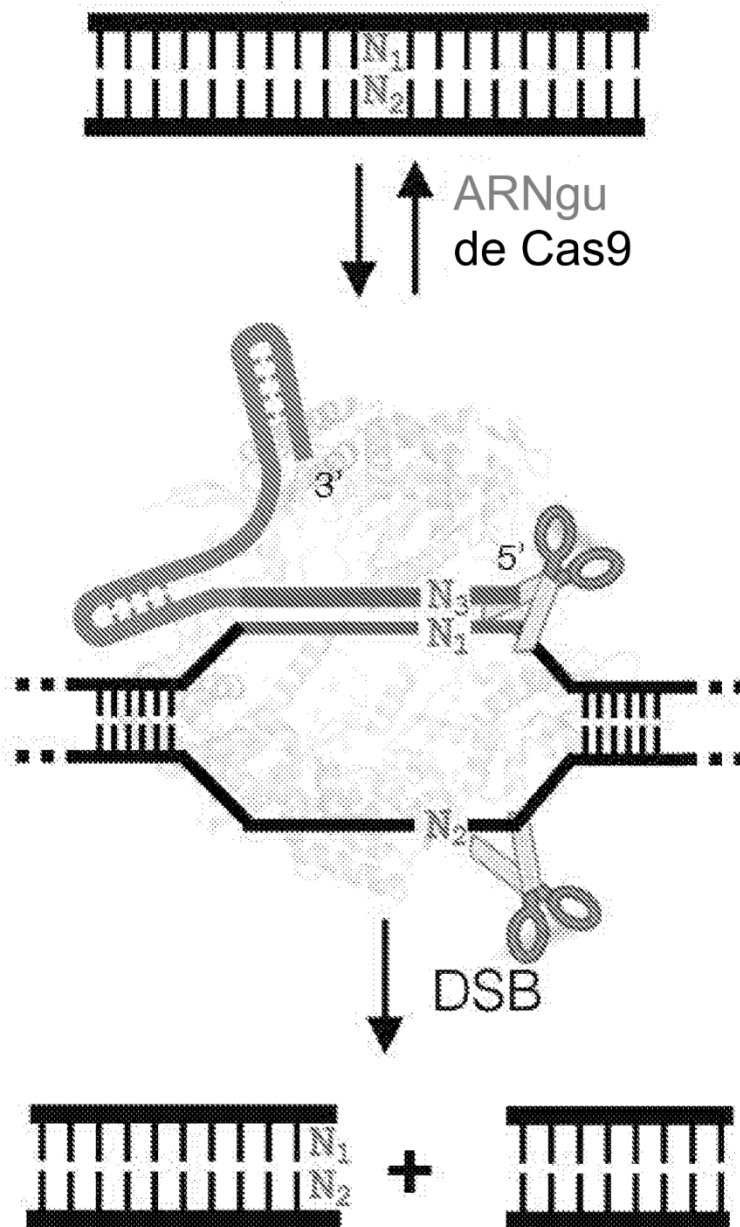


Fig. 8B

pUCX2 UBP Contexto	Retenciones de UBP (%)					
	Crecimiento líquido solo	Colonias (crecimiento sólido y líquido)				
		1	2	3	4	5
GXA	91	35	33	31	29	1
GXC	84	89	81	94	82	1
GXT	83	88	97	1	50	97
AXG	67	6	1	17	8	
AXA	91	47	47	26	57	43
AXC	83	0	92	92		
AXT	91	80	70	71	94	1
TXA	80	1	0	1	1	22
TXT	82	62	74	85	87	82

Fig. 9A

ADN de PCR de 110 pb



Análisis y cuantificación PAGE

Fig. 9B

		N_1					
		dC	dT	dA	dG	dTPT3	dNaM
N_3	G	20 ± 7% (100%)	1 ± 1% (5%)	1 ± 1% (5%)	5 ± 4% (25%)	3 ± 3% (15%)	1 ± 1% (5%)
	A	11 ± 2% (22%)	51 ± 6% (100%)	38 ± 9% (75%)	11 ± 2% (22%)	21 ± 3% (41%)	51 ± 6% (100%)
	U	8 ± 2% (50%)	1 ± 1% (6%)	16 ± 3% (100%)	16 ± 4% (100%)	3 ± 1% (8%)	8 ± 5% (50%)
	C	3 ± 4% (6%)	31 ± 7% (65%)	25 ± 8% (52%)	19 ± 7% (40%)	15 ± 5% (31%)	48 ± 23% (100%)

Diana de ADN: CTGGTCCTACCCGTGGT N_1 GGTCC
 Molde de ARN_g: GACCAGGATGGGCACCA N_3 CC

Fig. 10A

ARN guía	Secuencia
hEGFP	GACCAGGAUGGGCACCACCC
TK1-A	GUUGUGUGGAAUGUGAG
TK1-A/ Δ	1) GUUGUGUGGAAUGUGAG 2) UGUUGUGUGGAAUGUGAG
Diana de ADN:	GUATGTTGTGTGGAATGTGAG

Fig. 10B

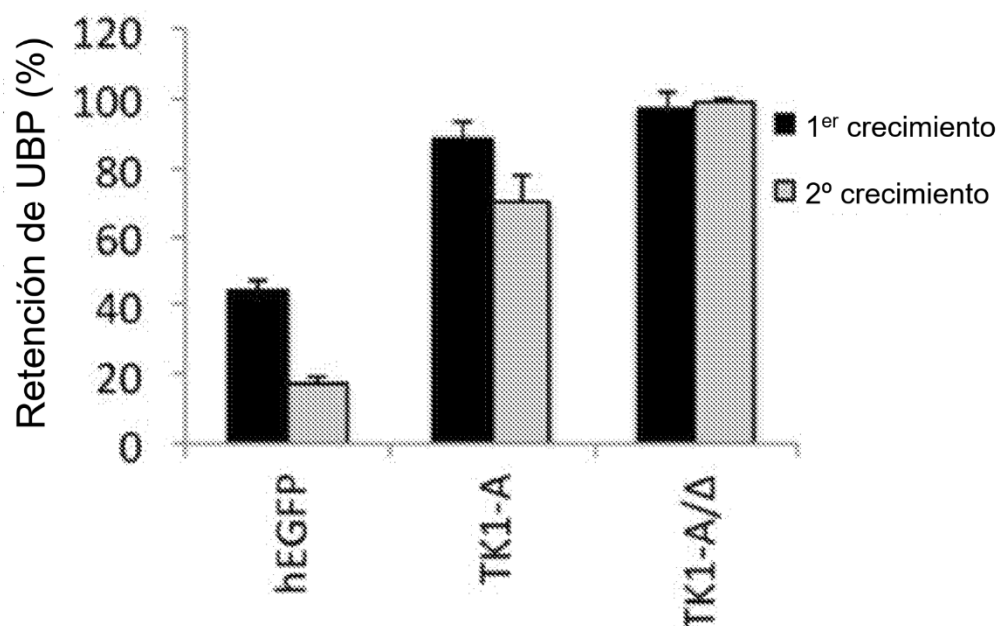


Fig. 10C

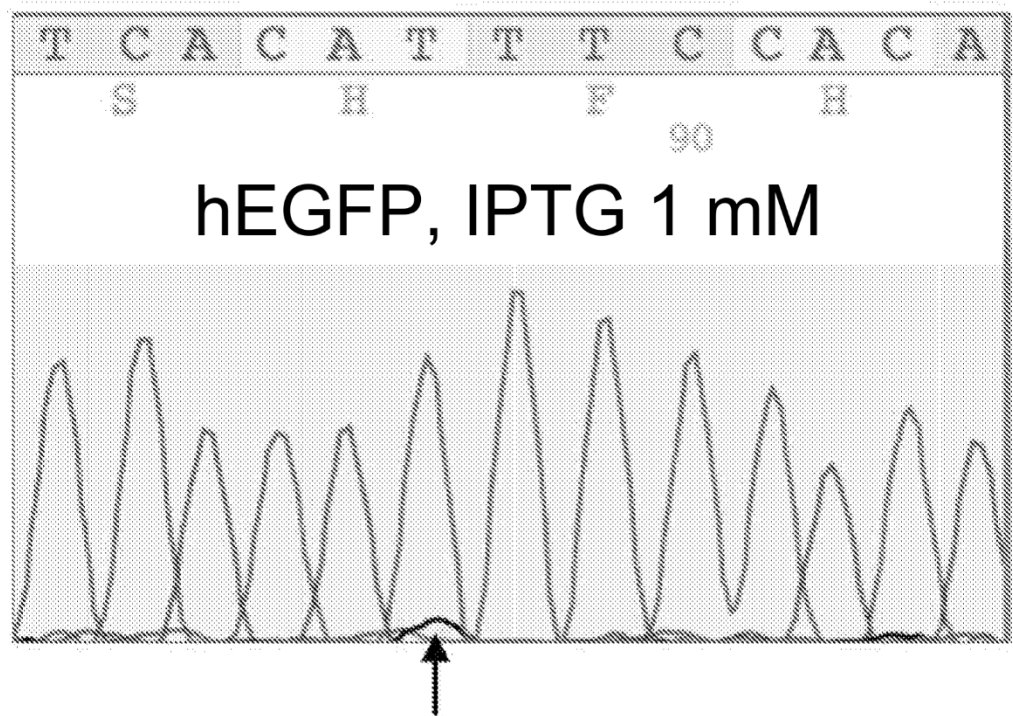
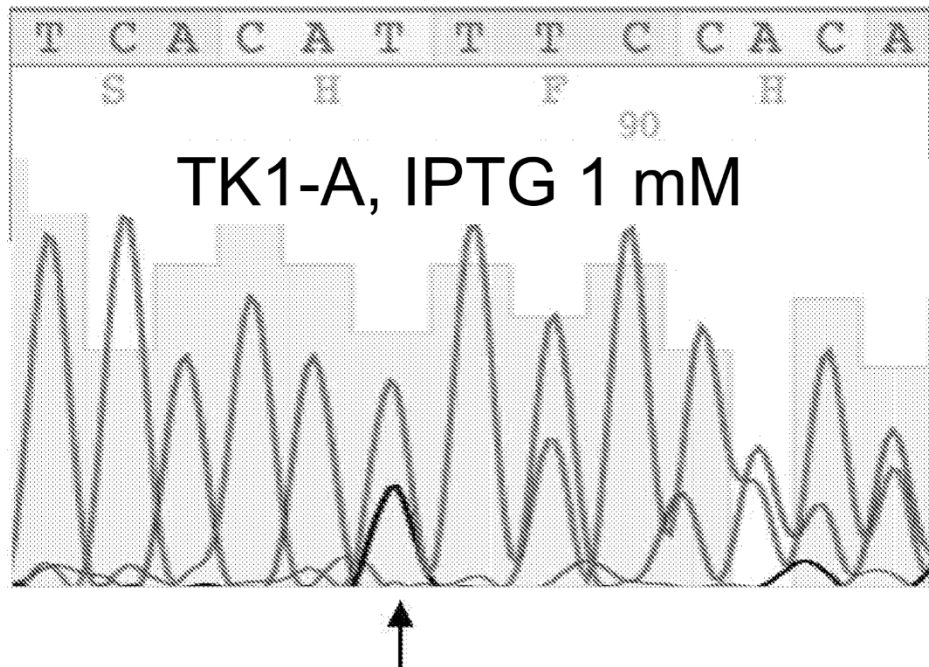


Fig. 10D



Mutación principal: dNaM $\rightarrow$ Δ , dG

Fig. 11

pAIO UBP Contexto	Secuencia	Guías (N/A)	Retenciones de UBPs (%) en YZ3 (-Cas9)								Retenciones de UBPs (%) en YZ4 (+Cas9)						
			Colonias								[IPTG] μM	Colonias					
			1	2	3	4	5	6	7	8		1	2	3	4	5	6
GXG	TCACACAATGTAGXGATCACCG	T/Δ	59	23	40	57					50	103	103	103			
GXA	ACCAGGATGGGXACCACCCCGG	T/Δ	36	90	90	83					10	94	102	103	95	88	98
GXC	TCACACAATGTAGXCATCACCG	T/Δ	98	97	102	88	65	96	92		10	95	97	106			
GXT	ACCAGGATGGGCACCAATGCCCG	G/Δ	111	97	115	113					25	115	113	114			
AXG	ACCAXGATGGGCACCAACCCCGG	T/Δ	3	13	13						50	107	106	105	109	105	
AXA	TCACACAATGTAXAGATCACCG	T/Δ	96	91	91						50	104	103	101	98	76	
AXC	ACCAGGATGGGCACCAACCCCGG	T/Δ	43	49	72	70					0	95	96	96			
AXT	TGTTGTGTGAAATGTGAGCGG	A/Δ	82	92	89	93	85	84	100		50	103	103	105	105	102	99
CXG	TTGTCACTACTCTGAGCCYCCG	G/Δ	2	2	6	13	7	14	4	12	50	8	16	23	1		
CXA	TTGTCACTACTCTGACCAAGG	A/Δ	1	5	10	0	9	11			50	8	3	16	10	7	
CXC	TCACACAATGTACXCATCACCG	T/Δ	16	21	12						50	66	95	77			
CXT	ACCAGGATGGGCACCAACCCCGG	A/Δ	84	73	72	69					50	106	112	88			
TXG	TCACACAATGTATXGATCACCG	T/Δ	13	26	12	17					50	117	86	114			
TXA	TCACACAATGTATXAATCACCG	A/Δ	54	79	78						50	101	104	97			
TXC	ACCAGGATGGGYACCACCCCGG	G/Δ	43	38	68						10	101	104	100			
TXT	ATTCACAATACTTCTTTAAGG	T/Δ	104	111	103	96					0	115	114	109			

Fig. 12

