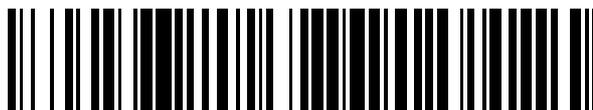


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 670 531**

51 Int. Cl.:

**C12N 15/10** (2006.01)

**C12N 15/82** (2006.01)

**C12N 15/85** (2006.01)

**C12N 15/90** (2006.01)

**A01K 67/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.05.2014 PCT/EP2014/061178**

87 Fecha y número de publicación internacional: **04.12.2014 WO14191518**

96 Fecha de presentación y número de la solicitud europea: **28.05.2014 E 14727479 (9)**

97 Fecha y número de publicación de la concesión europea: **28.03.2018 EP 3004349**

54 Título: **Un método para producir una escisión de ADN precisa utilizando la actividad nickasa de Cas9**

30 Prioridad:

**29.05.2013 DK 201370295**

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

**30.05.2018**

73 Titular/es:

**CELLECTIS S.A. (100.0%)  
8, rue de la Croix Jarry  
75013 Paris, FR**

72 Inventor/es:

**DUCHATEAU, PHILIPPE y  
BERTONATI, CLAUDIA**

74 Agente/Representante:

**ISERN JARA, Jorge**

ES 2 670 531 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

Un método para producir una escisión de ADN precisa utilizando la actividad nickasa de Cas9

## 5 Campo de la invención

La presente invención se refiere a un método de ingeniería genómica basado en el sistema CRISPR tipo II. En particular, la invención se refiere a un método para inducir de forma precisa una escisión de ácido nucleico en una secuencia genética de interés y prevenir la escisión "fuera de sitio" en un linfocito T. El método está basado en el uso de las arquitecturas de nickasa de Cas9 y de un único o múltiples ARNcr que albergan dos dianas distintas, lo que reduce el riesgo de producir escisión fuera de sitio. La presente invención también se refiere a linfocitos T y a aplicaciones terapéuticas relacionadas con el método descrito en este caso. También se divulgan en el presente documento polipéptidos, polinucleótidos, vectores y composiciones.

## 15 Antecedentes de la invención

Las nucleasas específicas de sitio son reactivos potentes para dirigirse a, y modificar de forma específica y eficaz una secuencia de ADN dentro de un genoma complejo. Existen numerosas aplicaciones de la ingeniería genómica mediante nucleasas específicas de sitio, que van de la investigación básica a las aplicaciones bioindustriales y la terapéutica humana. La reingeniería de una proteína de unión a ADN para este fin se ha limitado principalmente al diseño y producción de proteínas tales como las endonucleasas de asentamiento LADLIDADG (LHE, forma siglada del inglés *LADLIDADG homing endonucleases*) de origen natural, proteínas artificiales de dedos de zinc (ZFP, forma siglada de *zinc finger proteins*) y nucleasas efectoras tipo activadores de transcripción (nucleasas TALE, forma siglada de *transcription activator-like effectors nucleases*).

Recientemente, se ha desarrollado una nueva herramienta de ingeniería genómica basada en la nucleasa Cas9 guiada por ARN (Gasiunas, Barrangou *et al.*, 2012; Jinek, Chylinski *et al.*, 2012) del sistema inmunitario adaptativo CRISPR procariótico tipo II (grupos de repeticiones palindrómicas cortas en intervalos regulares). El sistema asociado a CRISPR (Cas, forma siglada de *CRISPR associated*) se descubrió primero en bacterias y actúa como defensa frente a ADN extraño, ya sea vírico o plasmídico. Hasta el momento se han identificado tres sistemas CRISPR bacterianos, denominados tipo I, II y III. El sistema tipo II es la base de la actual tecnología de ingeniería genómica disponible y, a menudo, se denomina simplemente CRISPR. Los locus de CRISPR/Cas tipo II están compuestos de un operón de genes que codifican en general las proteínas Cas9, Cas1, Cas2 y Csn2a, Csn2bor Cas4 (Chylinski, Le Rhun *et al.*, 2013), una agrupación CRISPR que consiste en una secuencia líder seguida de repeticiones idénticas intercaladas con espaciadores dirigidos al genoma exclusivos y una secuencia que codifica el ARNcrtra *trans*-activador.

El desarrollo y las aplicaciones de las tecnologías de edición del genoma de las TALEN (forma siglada de *transcription activator-like effector nucleases*: nucleasa efectoras tipo activador de la transcripción) y mediadas por CRISPR/Cas9 se resumen brevemente en Wei Chuanxian *et al.* (Journal of Genetics and Genomics, 40 (201), 281-289).

La inmunidad adaptativa mediada por CRISPR procede en tres fases distintas: adquisición de ADN extraño, biogénesis de ARN CRISPR (ARNcr) e interferencia de diana. (véase la revisión (Sorek, Lawrence *et al.* 2013)). En primer lugar, parece que la maquinaria CRISPR/Cas se dirige a una secuencia específica para la integración en el locus CRISPR. Las secuencias en el ADN extraño seleccionado para la integración se llaman espaciadores y estas secuencias a menudo están flanqueadas por un motivo de secuencia corto, denominado motivo adyacente al protoespaciador (PAM). La biogénesis de ARNcr en sistemas de tipo II es peculiar porque requiere un ARNcr transactivador (ARNtrac). El locus de CRISPR se transcribe inicialmente como un precursor largo de ARNcr (pre-ARNcr) a partir de una secuencia promotora en el líder. Cas9 actúa como un anclaje molecular que facilita el apareamiento de bases del ARNtrac con el pre-ARNcr para el posterior reconocimiento y escisión de las repeticiones del pre-ARNcr por la ARNasa III del hospedador (Deltcheva, Chylinski *et al.*, 2011). Después de los sucesos de procesamiento, el ARNcrtra permanece emparejado con el ARNcr y se une a la proteína Cas9. En este complejo ternario, la estructura doble de ARNtrac:ARNcr actúa como un ARN de guía que dirige la endonucleasa Cas9 al ADN diana relacionado (Jinek, Chylinski *et al.*, 2012). El reconocimiento de la diana por el complejo Cas9-ARNcrtra:ARNcr se inicia explorando la molécula de ADN invasora en cuanto a la homología entre la secuencia del protoespaciador en el ADN diana y la secuencia derivada del espaciador en el ARNcr. Además de la complementariedad del ADN protoespaciador-espaciador del ARNcr, el direccionamiento al ADN requiere la presencia de un motivo corto adyacente al protoespaciador (motivo adyacente al protoespaciador-PAM). Después del emparejamiento entre el ARN doble y la secuencia del protoespaciador, seguidamente Cas9 introduce una rotura bicatenaria roma 3 bases cadena arriba del motivo PAM (Garneau, Dupuis *et al.*, 2010).

La gran proteína Cas9 (> 1200 aminoácidos) contiene dos dominios predichos de nucleasa, a saber, un dominio de nucleasa HNH (de tipo McrA), que se emplaza en la parte media de la proteína, y un dominio de nucleasa de tipo RuvC separado (pliegue de ARNasa H) (Haft, Selengut *et al.*, 2005; Makarova, Grishin *et al.* 2006). Se ha encontrado que el dominio de nucleasa HNH y el dominio Ruv-C son esenciales para la actividad de escisión de la

5 doble cadena. Las mutaciones introducidas en estos dominios han conducido respectivamente a proteínas Cas9 que presentan actividad nickasa en lugar de actividad de escisión de la doble cadena. Una mutación (o mutaciones) de inactivación de los restos catalíticos distinta en los dominios similares a RuvC produce una nickasa capaz de cortar una cadena en la posición + 3 pb (opuesto al extremo 3') con respecto al emplazamiento del PAM. La mutación del resto catalítico del dominio HNH genera una nickasa capaz de cortar la otra cadena en la posición + 3 pb (opuesto al extremo 5') (Jinek, Chylinski *et al.*, 2012) (Figura 1).

10 Lei *et al.*, (Cell, Vol. 152, n.º 5, febrero de 2013) demostraron que una Cas9 catalíticamente "muerta" carente de actividad endonucleasa, cuando se coexpresa con un ARN guía, genera un complejo de reconocimiento de ADN que puede interferir específicamente con la elongación transcripcional. Unión de la ARN polimerasa o unión de factores de transcripción. Se informa que este sistema, denominado interferencia por CRISPR (CRISPRi), reprime de forma eficaz la expresión de genes diana en *Escherichia coli*.

15 También se ha demostrado que el sistema CRISPR-Cas funciona *in vivo* induciendo modificaciones genéticas dirigidas en embriones de pez cebra, con eficacias similares a las obtenidas utilizando nucleasas de dedos de zinc y nucleasas efectoras de tipo activador de la transcripción (Woong Y Hwang *et al.*, Nature Biotechnology, Vol. 31, n.º 3, enero de 2013, pág. 227-229) y proporcionando una actividad de endonucleasa guiada por ARN fuerte y específica en locus genómicos endógenos diana en *Saccharomyces cerevisiae* (Dicarlo *et al.*, Nucleic Acid Research, Vol. 41, n.º 7, marzo de 2013, pág. 4336-4343).

20 Jinek *et al.* (eLife, Vol. 2, enero de 2013, p.e00471) demostraron la capacidad del Cas9 programado por ARN para introducir roturas de doble cadena dirigidas en el ADN cromosómico humano, induciendo de este modo reacciones de edición del genoma específicas del sitio. Seung Woo Chung *et al.* (Nature Biotechnology, Vol. 31, n.º 3, marzo de 2013) emplearon el sistema CRISPR-Cas para la edición del genoma en células humanas, demostrando que los complejos de la proteína Cas9 y los ARN quiméricos artificiales escinden de forma eficaz dos sitios genómicos e inducen inserciones/deleciones con frecuencias de hasta el 33 %.

25 El sistema CRISPR de procarionota tipo II tiene la capacidad de reconocer cualquier posible secuencia diana de 12 a 20 nucleótidos, seguida de un motivo PAM específico en su extremo 3'. Sin embargo, la especificidad para el reconocimiento de la diana depende de solo 12 ácidos nucleicos (Jiang, Bikard *et al.*, 2013; Qi, Larson *et al.*, 2013), lo que es suficiente para garantizar genomas procarionotas con sitios de escisión poco comunes sobre una base estadística, pero que es crítico para genomas más grandes, como en las células eucariotas, donde se pueden encontrar varias veces secuencias de 12 ácidos nucleicos. Por lo tanto, existe una necesidad de desarrollar estrategias para mejorar la especificidad y reducir la posible actividad fuera de sitio utilizando el sistema CRISPR tipo II.

35 Sumario de la invención

40 En este caso, los inventores han investigado distintas modificaciones en el sistema CRISPR tipo II para mejorar la especificidad y reducir la posible escisión fuera de sitio en un linfocito T. Inesperadamente, descubrieron que el uso de una versión mutada de Cas9 que tiene actividad nickasa, en lugar de actividad de escisión, puede utilizarse para producir la escisión en una diana de ADN dada y aumentar la especificidad al mismo tiempo. El método está basado en el uso simultáneo de la arquitectura de nickasa de Cas9 (dominio RuvC y/o dominio HNH) y ARN<sub>g</sub> (o los ARN<sub>g</sub>) (en inglés: sgRNA, *single guide RNA*, ADN guía único) que alberga dos secuencias complementarias distintas de dianas específicas, lo que reduce el riesgo de producir escisión fuera de sitio. Utilizando al menos un ARN guía que alberga dos secuencias complementarias distintas de dianas específicas o una combinación de al menos dos ARN guía, el requisito de especificidad pasa de 12 a 24 nucleótidos y, a su vez, la probabilidad de encontrar dos sitios de unión alternativos de Cas9 (distintos de los codificados en los dos ARN<sub>g</sub>) a una distancia el uno del otro eficaz para producir una división fuera de sitio es realmente baja. La invención se extiende a los mutantes de ARN<sub>cr</sub>, ARN<sub>tr</sub> y Cas diseñados para realizar este método y a los linfocitos T transfectados con el sistema CRISPR tipo II modificado resultante.

50 Breve descripción de las figuras

55 Figura 1: Representación esquemática de la rotura de doble cadena del ADN mediada por el sistema CRISPR/Cas tipo II. En el sistema CRISPR/Cas tipo II, Cas 9 está guiado por una estructura de dos ARN, denominada ARN guía (ARN<sub>g</sub>) formado por ARN<sub>cr</sub> y ARN<sub>tr</sub> para la diana de ácido nucleico bicatenario (ADN<sub>bc</sub>). El dominio RuvC de Cas9 induce un suceso de muesca (*nick*) (flecha) en una cadena en la posición +3 pb (opuesto al extremo 3') con respecto al emplazamiento del PAM y el dominio HNH de Cas9 induce un suceso de mella (flecha) en la otra cadena en la posición +3 pb (opuesto al extremo 5'). Para una mejor comprensión, la figura ilustra solo un aspecto de la ruptura de doble cadena mediada por el sistema CRISPR/Cas.

60 Figura 2: Representación esquemática del nuevo sistema CRISPR /Cas tipo II que usa una nickasa Cas9. A-B Se seleccionan dentro de una secuencia genética de interés dos dianas de ácido nucleico que comprenden cada una en una cadena un motivo PAM en los extremos 3'. Las dos dianas de ácido nucleico están espaciadas por

una distancia "d". Cas9 albergando un dominio RuvC o HNH no funcional (RuvC (-) (A-C) o (HNH (-) (B-D), respectivamente) es guiada por dos ARNg diseñados técnicamente, comprendiendo cada uno una secuencia complementaria a al menos 12 nucleótidos adyacentes al motivo PAM complementario de las primera y segunda dianas de ácido nucleico. AB. Cada uno de los motivos PAM de las dos dianas están presentes en distintas cadenas. La nickasa Cas9 induce una muesca (flecha) en las distintas cadenas que da como resultado una rotura de doble cadena dentro de la secuencia genética de interés. C-D. Cada uno de los motivos PAM de las dos dianas presentes en la misma cadena. La nickasa Cas9 induce una muesca en la misma cadena de la secuencia genética de interés, dando como resultado la delección de una secuencia de ácido nucleico monocatenaria entre los dos sucesos de muesca. La figura ilustra solo algunos aspectos del sistema CRISPR/Cas que utiliza la nickasa Cas9.

Figura 3: Representación esquemática del nuevo sistema CRISPR/Cas tipo II que utiliza dos nickasas Cas9 distintas. A-B Se seleccionan dentro de una secuencia genética de interés dos dianas de ácido nucleico que comprenden dos motivos PAM distintos (PAM1 and PAM2) en el extremo 3'. Las dos dianas de ácido nucleico están espaciadas por una distancia "d". Una primera Cas9 que alberga un RuvC no funcional es guiada por un ARNg diseñado técnicamente que comprende una secuencia complementaria a al menos 12 nucleótidos adyacentes al primer motivo PAM complementario. Una segunda Cas9 que alberga un dominio HNH no funcional es guiada por un ARNg diseñado técnicamente que comprende una secuencia complementaria a al menos 12 nucleótidos adyacentes al segundo motivo PAM complementario. A. Cada uno de los motivos PAM de las dos dianas están presentes en distintas cadenas. Las dos nickasas Cas9 inducen dos muescas (flechas) en la misma cadena de la secuencia genética de interés, dando como resultado la delección de una secuencia de ácido nucleico monocatenaria entre los dos sucesos de muesca. B. Cada uno de los motivos PAM de las dos dianas están presentes en la misma cadena. Las nickasas Cas9 inducen dos sucesos de muesca (flechas) en las distintas cadenas, dando como resultado una rotura de doble cadena dentro de la secuencia genética de interés. La figura ilustra solo algunos aspectos del sistema CRISPR/Cas que utiliza la nickasa Cas9.

#### Divulgación de la invención

A menos que se defina específicamente en el presente documento, todos los términos técnicos y científicos utilizados tienen el mismo significado que entiende habitualmente un experto en la materia en el campo de la terapia génica, la bioquímica, la genética, la biología molecular y la inmunología.

Todos los métodos y materiales similares o equivalentes a los descritos en el presente documento se pueden usar en la práctica o el ensayo de la presente invención, con los métodos y materiales adecuados que se describen en el presente documento.

La práctica de la presente invención empleará, salvo que se indique otra cosa, técnicas convencionales de biología celular, cultivo celular, biología molecular, biología transgénica, microbiología, ADN recombinante e inmunología, que están dentro de las habilidades en la técnica. Tales técnicas se explican por completo en la bibliografía. Véase, por ejemplo, *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, EE. UU.); *Molecular Cloning: A Laboratory Manual*, tercera edición, (Sambrook *et al*, 2001, Cold Spring Harbor, Nueva York: Cold Spring Harbor Laboratory Press); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al*. Patente de Estados Unidos N.º 4.683.195; *Nucleic Acid Hybridization* (B. D. Harries & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); la serie, *Methods In ENZYMOLOGY* (J. Abelson and M. Simon, eds.-in-chief, Academic Press, Inc., Nueva York), específicamente, vol. 154 y 155 (Wu *et al*. eds.) y vol. 185, "Gene Expression Technology" (D. Goeddel, ed.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, Londres, 1987); *Handbook Of Experimental Immunology*, Volúmenes I-IV (D. M. Weir y C. C. Blackwell, eds., 1986); y *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

#### Método para inducir con precisión una escisión de ácido nucleico en una secuencia genética

La presente invención se refiere así a un nuevo método basado en el sistema CRISPR/Cas para inducir con precisión una escisión en una diana de ácido nucleico bicatenario en un linfocito T. Este método procede de la herramienta del sistema inmunitario adaptativo CRISPR de ingeniería genómica que se ha desarrollado a base de la nucleasa Cas9 guiada por ARN (Gasiunas, Barrangou *et al.*, 2012; Jinek, Chylinski *et al*. 2012).

La presente invención se refiere a un método para inducir con precisión la escisión del ácido nucleico en una secuencia genética en una célula que comprende:

- (a) Seleccionar una primera y segunda dianas de ácido nucleico bicatenario en dicha secuencia genética, comprendiendo cada una de las dianas de ácido nucleico, en una cadena, un motivo PAM en una extremidad 3';
- (b) diseñar técnicamente dos ARNcr comprendiendo cada uno:

- una secuencia complementaria a una parte de la cadena opuesta de la diana de ácido nucleico que no comprende el motivo PAM, y
- una secuencia de extensión 3';

- 5 (c) proporcionar al menos un ARNcrtra que comprende una secuencia complementaria a una parte de las secuencias de extensión 3' de dichos ARNcr de b);
- (d) proporcionar al menos una nickasa cas9 que alberga un dominio de nucleasa de tipo RuvC no funcional o HNH no funcional y que reconoce específicamente dicho motivo (o motivos) PAM;
- 10 (e) introducir en la célula dichos ARNcr, dicho (o dichos) ARNcrtra y dicha nickasa Cas9; en donde la célula es un linfocito T,

de modo que cada complejo Cas9-ARNcrtra:ARNcr induce un suceso de muesca en las dianas de ácido nucleico bicatenario para escindir la secuencia genética entre dichas dianas de ácido nucleico. Dicha escisión puede ser el resultado de al menos un suceso de muesca en una cadena de ácido nucleico, preferentemente, dos sucesos de muesca en la misma cadena de ácido nucleico o más preferentemente dos sucesos de muesca en las cadenas de ácido nucleico opuestas. Cas9, también llamada Csn1 (COG3513 - SEQ ID NO: 1), es una proteína grande que participa tanto en la biogénesis de ARNcr como en la destrucción del ADN invasor. Cas9 se ha descrito en distintas especies bacterianas como *S. thermophilus* (Sapranaukas, Gasiunas *et al.*, 2011), *listeria innocua* (Gasiunas, Barrangou *et al.* 2012; Jinek, Chylinski *et al.* 2012) y *S. Pyogenes* (Deltcheva, Chylinski *et al.* 2011). La gran proteína Cas9 (> 1200 aminoácidos) contiene dos dominios predichos de nucleasa, a saber, un dominio de nucleasa HNH (de tipo McrA), que se emplaza en el medio de la proteína, y un dominio de nucleasa de tipo RuvC separado (pliegue de ARNasa H) (Haft, Selengut *et al.*, 2005; Makarova, Grishin *et al.* 2006).

El motivo HNH es característico de muchas nucleasas que actúan sobre el ADN bicatenario, incluidas las colicinas, las enzimas de restricción y las endonucleasas de asentamiento. El dominio HNH (SMART ID: SM00507, nomenclatura SCOP: familia HNH) está asociado con una serie de proteínas de unión al ADN, que realizan una diversidad de funciones de unión y corte (Gorbalenya 1994; Shub, Goodrich-Blair *et al.* 1994). Varias de las proteínas son proteínas hipotéticas o supuestas con una función que no está bien definida. Las que tienen una función conocida están implicadas en una serie de procesos celulares que incluyen la toxicidad bacteriana, funciones de asentamiento en intrones de los grupos I y II e inteínas, recombinación, reordenamiento de ADN controlado por desarrollo, empaquetamiento de fagos y actividad de endonucleasa de restricción (Dalgaard, Klar *et al.*, 1997). Estas proteínas se encuentran en virus, arqueobacterias, eubacterias y eucariotas. Curiosamente, como con los motivos LAGLI-DADG y los GIY-YIG, el motivo HNH a menudo se asocia con dominios de endonucleasa de elementos autoprogantes como inteínas, intrones de grupo I y grupo II (Gorbalenya 1994; Dalgaard, Klar *et al.* 1997). El dominio HNH se puede caracterizar por la presencia de un resto Asp/His conservado flanqueado por restos His (amino terminales) e His/Asp/Glu (carboxilo terminales) conservados a cierta distancia. Un número considerable de estas proteínas también puede tener un motivo CX2C en cualquier lado del resto Asp/His central. Del punto de vista estructural, el motivo HNH aparece como una horquilla central de hojas plegadas  $\beta$  torsionadas, que están flanqueado en cada lado por una hélice  $\alpha$  (Kleanthous, Kuhlmann *et al.*, 1999). La otra ARNasaH de tipo RuvC (también llamada RuvC) de dominio catalítico de CRISPR, se encuentra en proteínas que muestran amplios espectros de funciones nucleolíticas, actuando tanto en ARN como en ADN (ARNasaH, RuvC, transposasas de ADN e integrasas retrovíricas, y el dominio PIWI de las proteínas Argonauta).

Recientemente, se ha demostrado que el dominio HNH es responsable de la producción de muescas de una cadena del ADN bicatenario diana y el dominio de pliegue de ARNasaH de tipo RuvC está implicado en la producción de muescas de la otra cadena (que comprende el motivo PAM) de la diana de ácido nucleico bicatenario (Jinek, Chylinski *et al.*, 2012). Sin embargo, en Cas9 de tipo silvestre, estos dos dominios dan como resultado una escisión roma del ADN invasor dentro de la misma secuencia diana (protoespaciador) en las inmediaciones de la PAM (Jinek, Chylinski *et al.*, 2012). En la presente invención, Cas 9 es una nickasa e induce un suceso de muesca dentro de distintas secuencias diana. Como ejemplo no limitante, Cas9 puede comprender una mutación (o mutaciones) en los restos catalíticos de los dominios HNH o de tipo RuvC, para inducir un suceso de muesca dentro de distintas secuencias diana. Como ejemplo no limitante, los restos catalíticos de la proteína Cas9 compacta son los correspondientes a los aminoácidos D10, D31, H840, H868, N882 y N891 de la SEQ ID NO: 1 o posiciones alineadas utilizando el método CLUSTALW en homólogos de miembros de la familia Cas. Cualquiera de estos restos puede reemplazarse por cualquier otro aminoácido, preferentemente por un resto de alanina. La mutación en los restos catalíticos significa la sustitución por otros aminoácidos, o la delección o adición de aminoácidos que inducen la inactivación de al menos un dominio catalítico de cas9. (*cf* (Sapranaukas, Gasiunas *et al.* 2011; Jinek, Chylinski *et al.* 2012). En una realización particular, Cas9 puede comprender una o varias de las mutaciones anteriores. En otra realización particular, Cas9 puede comprender solo uno de los dos dominios catalíticos RuvC y HNH. En la presente invención, se puede utilizar Cas9 de distintas especies, homólogos de Cas9, Cas9 diseñada técnicamente y variantes funcionales de la misma. La invención prevé el uso de tales variantes de Cas9 para realizar la escisión del ácido nucleico en una secuencia genética de interés. Dichas variantes de Cas9 tienen una secuencia de aminoácidos que comparte al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 % e incluso más preferentemente el 95 % de identidad con Cas9 de distintas especies, homólogos de Cas9, Cas9 diseñada técnicamente y variantes funcionales de la misma. Preferentemente, dichas variantes de Cas9 tienen

una secuencia de aminoácidos que comparte al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 % e incluso más preferentemente el 95 % de identidad con la SEQ ID NO: 1.

5 Las escisiones de ácido nucleico provocadas por nucleasas específicas de sitio se reparan habitualmente a través de mecanismos distintos de recombinación homóloga o por unión de extremos no homóloga (NHEJ, forma siglada de *non-homologous end joining*). Aunque la recombinación homóloga normalmente utiliza la cromátida hermana del ADN dañado como una secuencia exógena de ácido nucleico a partir de la cual se realiza la reparación perfecta de la lesión genética, la NHEJ es un proceso de reparación imperfecto que a menudo da como resultado cambios en la secuencia de ADN en el sitio de la escisión. Los mecanismos implican la reunión de lo que queda de los dos extremos de ADN a través de religamiento directo (Critchlow y Jackson 1998) o a través de la denominada unión de extremos mediada por microhomología (Ma, Kim et al., 2003). Además, la reparación a través de la unión de extremos no homóloga (NHEJ) a menudo da como resultado pequeñas inserciones o deleciones, y se puede utilizar para la creación de supresiones génicas específicas. Por lo tanto, un aspecto de la presente invención es inducir supresiones o introducir secuencias genéticas exógenas por recombinación homóloga en locus genéticos específicos.

20 Por secuencia genética de interés se entiende cualquier secuencia de ácido nucleico endógena, tal como, por ejemplo, un gen o una secuencia no codificante dentro de o adyacente a un gen, que es deseable modificar mediante escisión dirigida y/o recombinación homóloga dirigida. La secuencia de interés puede estar presente en un cromosoma, un episoma, un genoma de orgánulo, tal como el genoma mitocondrial o de cloroplasto, o material genético que puede existir independientemente del cuerpo principal del material genético, tal como un genoma vírico infeccioso, plásmidos, episomas, transposones, por ejemplo. Una secuencia de interés puede estar dentro de la secuencia codificante de un gen, dentro de una secuencia no codificante transcrita tal como, por ejemplo, secuencias líderes, secuencia de remolque o intrones, o dentro de una secuencia no transcrita, cadena arriba o cadena abajo de la secuencia codificante.

30 La primera y la segunda diana de ácido nucleico bicatenario están comprendidas dentro de la secuencia genética de interés, en la que se desea introducir una escisión y, por lo tanto, una modificación genética. Dicha modificación puede ser una deleción del material genético, inserción de nucleótidos en el material genético o una combinación de deleción e inserción de nucleótidos. Por "secuencia de ácido nucleico diana", "diana de ácido nucleico bicatenario" o "diana de ADN" significa un polinucleótido que puede ser procesado por el complejo Cas9-ARNcr: ARNcr de acuerdo con la presente invención. La secuencia diana de ácido nucleico bicatenario está definida por la secuencia 5' a 3' de una cadena de dicha diana. Estas expresiones se refieren a un emplazamiento de ADN específico dentro de la secuencia genética de interés. Las dos dianas pueden estar espaciadas entre sí de 1 a 500 nucleótidos, preferentemente entre 3 y 300 nucleótidos, más preferentemente entre 3 y 50 nucleótidos, también más preferentemente entre 1 y 20 nucleótidos.

40 Cualquier posible diana de ADN bicatenario seleccionada en la presente invención puede tener una secuencia específica en su extremo 3', llamada el motivo adyacente al protoespaciador o motivo asociado al protoespaciador (PAM). El PAM está presente en la cadena de la secuencia diana de ácido nucleico que no es complementaria al ARNcr. Preferentemente, el motivo adyacente al protoespaciador (PAM) puede corresponder a 2 a 5 nucleótidos, comenzando inmediatamente o cerca del protoespaciador en el extremo 3'. La secuencia y el emplazamiento del motivo PAM reconocido por la Cas9 específica varían entre los distintos sistemas. El motivo PAM puede ser, por ejemplo NNAGAA, NAG, NGG, NGGNG, AWG, CC, CCN, TCN, TTC, como ejemplos no limitantes (Shah, Erdmann *et al.* 2013). Distintos sistemas de tipo II tienen distintos requisitos de PAM. Por ejemplo, el sistema *S. pyogenes* requiere una secuencia NGG, donde N puede ser cualquier nucleótido. Los sistemas de *S. thermophilus* de tipo II requieren NGGNG (Horvath y Barrangou 2010) y NNAGAAW (Deveau, Barrangou *et al.* 2008), mientras que distintos sistemas de *S. mutant* toleran NGG o NAAR (van der Ploeg 2009). El PAM no se limita a la región adyacente al protoespaciador, sino que también puede ser parte del protoespaciador (Mojica, Diez-Villasenor *et al.*, 2009). En una realización particular, la proteína Cas9 se puede diseñar técnicamente para reconocer un motivo PAM no natural. En este caso, la secuencia diana seleccionada puede comprender un motivo PAM más pequeño o más grande con cualquier combinación de aminoácidos. Como ejemplo no limitante, los dos motivos PAM de las dos dianas de ácido nucleico pueden estar presentes en la misma cadena de ácido nucleico y, por lo tanto, la nickasa Cas9 que alberga un dominio de nucleasa RuvC o HNH no funcional induce dos sucesos de muesca en la misma cadena (Figuras 2C y D). En este caso, el ácido nucleico monocatenario resultante emplazado entre la primera y la segunda muesca puede deleccionarse. Esta deleción puede ser reparada por NHEJ o por mecanismos de recombinación homóloga. En otro aspecto de la invención, los dos motivos PAM de las dos dianas de ácido nucleico pueden estar presentes en las cadenas de ácido nucleico opuestas y, por lo tanto, la nickasa Cas9 que alberga un dominio de nucleasa RuvC o HNH no funcional induce dos sucesos de muesca en cada cadena de la secuencia génica de interés (Figuras 2A y B), lo que da como resultado una ruptura de doble cadena dentro de la secuencia genética de interés.

65 En una realización particular, el método de la presente invención utiliza dos nickasas Cas9, cada una capaz de reconocer distintos motivos PAM dentro de las dos dianas de ácido nucleico. Como ejemplo no limitante, la primer Cas9 es capaz de reconocer el motivo PAM NGG y la segunda Cas9 es capaz de reconocer el motivo PAM NNAGAAW.

En particular, la presente invención se refiere a un método que comprende una o varias de las siguientes etapas:

- (a) seleccionar una primera y segunda secuencias diana de ácido nucleico bicatenario que comprenden cada una un motivo PAM en una cadena en su extremo 3', en donde dichos motivos PAM son distintos;
- (b) diseñar técnicamente dos ARNcr que comprenden cada uno una secuencia complementaria a una parte de la otra cadena de las primera y segunda dianas de ácido nucleico bicatenario, y que tienen una secuencia de extensión 3';
- (c) proporcionar al menos un ARNcrtra que comprende una secuencia complementaria con una parte de las secuencias de extensión 3' de dichos ARNcr;
- (d) proporcionar una primera nucleasa Cas9 que reconoce específicamente el motivo PAM de la primera diana y que alberga un dominio de nucleasa RuvC o HNH no funcional;
- (e) proporcionar una segunda Cas9 que reconoce específicamente el motivo PAM de la segunda diana y que alberga un dominio de nucleasa RuvC o HNH no funcional;
- (f) introducir en la célula dichos ARNcr, dichos ARNcrtra (o los ARNcrtra), dichas nucleasas Cas9, de forma que cada complejo Cas9-ARNcrtra:ARNcr induzca un suceso de muesca en la diana de ácido nucleico bicatenario,

en donde la célula es un linfocito T.

A modo de ejemplos no limitantes, pueden introducirse en la célula la Cas9 de *S. pyogenes* que carece de un dominio catalítico RuvC o HNH funcional y Cas9 de *S. thermophilus* que carece del dominio catalítico RuvC o HNH funcional, para reconocer específicamente el motivo PAM NGG en la primera secuencia de ácido nucleico diana y el motivo PAM NNAGAAW en la segunda secuencia de ácido nucleico diana, respectivamente. En una realización en particular, los dos motivos PAM distintos de las dos dianas de ácido nucleico pueden estar presentes en la misma cadena de ácido nucleico y, por lo tanto, las nickasas Cas9 que albergan un dominio de nucleasa RuvC o HNH no funcional inducen dos sucesos de muesca en la misma cadena. En este caso, el ácido nucleico monocatenario resultante emplazado entre la primera y la segunda muesca puede delecionarse. En otra realización, los dos motivos PAM distintos de las dos dianas de ácido nucleico pueden estar presentes en las cadenas de ácido nucleico opuestas y, por lo tanto, las nickasas Cas9 que albergan un dominio de nucleasa de tipo RuvC o HNH no funcional inducen dos sucesos de muesca en cada cadena de la secuencia génica de interés, lo que da como resultado una rotura de doble cadena dentro de la secuencia genética de interés.

En otra realización particular, la primera nickasa Cas9 alberga un dominio de nucleasa de tipo RuvC no funcional y la segunda nickasa Cas9 alberga un dominio de nucleasa HNH no funcional. Los distintos motivos PAM de las dos dianas de ácido nucleico pueden estar en la misma cadena, por lo tanto, las dos nickasas Cas9 inducen un suceso de muesca en cada cadena (Figura 3A), dando como resultado una rotura de doble cadena dentro de la secuencia genética de interés. Los dos motivos PAM también pueden estar en cadenas opuestas y, por lo tanto, las dos nickasas Cas9 inducen un suceso de muesca en la misma hebra de la secuencia genética de interés (Figura 3B). En este caso, el ácido nucleico monocatenario resultante emplazado entre la primera y la segunda muesca puede delecionarse. Esta delección puede ser reparada por NHEJ o por mecanismos de recombinación homóloga.

El método de la presente invención comprende diseñar técnicamente dos ARNcr con distintas regiones complementarias a cada diana de ácido nucleico. En el sistema CRISPR tipo II natural, las secuencias de direccionamiento del ARN de direccionamiento de CRISPR (ARNcr) se transcriben a partir de las secuencias de ADN conocidas como protoespaciadores. Los protoespaciadores se agrupan en el genoma bacteriano en un grupo llamado agrupación CRISPR. Los protoespaciadores son secuencias cortas de ADN extraño conocidas separadas por una repetición palindrómica corta y mantenidas como un registro frente a futuros encuentros. Para crear el ARNcr, la agrupación CRISPR se transcribe y el RNA se procesa para separar las secuencias de reconocimiento individuales entre las repeticiones. El locus CRISPR que contiene los espaciadores se transcribe en un preARNcr largo. El procesamiento del transcrito de la agrupación CRISPR (preARNcr) en los ARNcr individuales depende de la presencia de un ARNcr transactivador (ARNcrtra) que tiene una secuencia complementaria a la repetición palindrómica. El ARNcrtra híbrida con las regiones de repetición que separan los espaciadores del preARNcr, iniciando la escisión del ARNcr por la ARNasa III endógena, seguido por un segundo suceso de escisión por Cas9 dentro de cada espaciador, produciendo los ARNcr maduros que permanecen asociados con el ARNcrtra y Cas9, y forman el complejo Cas9-ARNcrtra:ARNcr. El ARNcr modificado con ARNcrtra es capaz de dirigirse a una secuencia de ácido nucleico seleccionada, obviando la necesidad de ARNasa III y el procesamiento de ARNcr en general (Jinek, Chylinski *et al.*, 2012).

En la presente invención, se diseñan técnicamente dos ARNcr para que comprendan distintas secuencias complementarias a una parte de una cadena de las dos dianas de ácido nucleico, de manera que sean capaces de dirigir, preferentemente inducir, un suceso de muesca en cada una de las dianas de ácido nucleico. En una realización en particular, las dos dianas de ácido nucleico están espaciadas entre sí de 1 a 300 pb, preferentemente de 3 a 250 pb, preferentemente de 3 a 200 pb, más preferentemente de 3 a 150 pb, 3 a 100 pb, 3 a 50 pb, 3 a 25 pb, 3 a 10 pb.

La secuencia de ARNcr es complementaria a una cadena de la diana de ácido nucleico, esta cadena no comprende el motivo PAM en el extremo 3' (Figura 1). En una realización particular, cada ARNcr comprende una secuencia de 5

a 50 nucleótidos, preferentemente de 8 a 20 nucleótidos, más preferentemente de 12 a 20 nucleótidos, que es complementaria a la secuencia de ácido nucleico diana. En una realización más particular, el ARNcr es una secuencia de al menos 30 nucleótidos que comprende al menos 10 nucleótidos, preferentemente 12 nucleótidos, complementarios a la secuencia de ácido nucleico diana. En particular, cada ARNcr puede comprender una secuencia complementaria seguida de 4-10 nucleótidos en el extremo 5' para mejorar la eficacia del direccionamiento (Cong, Ran *et al.* 2013; Mali, Yang *et al.* 2013; Qi, Larson *et al.* 2013). En una realización preferente, la secuencia complementaria del ARNcr está seguida en el extremo 3' de una secuencia de ácido nucleico llamada secuencia de repetición o secuencia de extensión 3'.

El ARNcr de acuerdo con la presente invención también puede modificarse para aumentar la estabilidad de la estructura secundaria y/o su afinidad de unión por Cas9. En una realización particular, el ARNcr puede comprender un fosfato 2',3' cíclico. El extremo fosfato 2',3' cíclico parece estar implicado en muchos procesos celulares, es decir, corte y empalme de ARNt, escisión endonucleolítica por varias ribonucleasas, en la autoescisión por ARN ribozima y en respuesta a diversos estreses celulares, incluida la acumulación de proteína no plegada en el retículo endoplasmático y el estrés oxidativo (Schutz, Hesselberth *et al.*, 2010). Los inventores han especulado que el fosfato 2',3' cíclico potencia la estabilidad del ARNcr o su afinidad/especificidad para Cas9. Por lo tanto, la presente invención se refiere al ARNcr modificado que comprende un fosfato 2',3' cíclico y a los métodos para ingeniería genómica basados en el sistema CRISPR/cas (Jinek, Chylinski *et al.*, 2012; Cong, Ran *et al.* 2013; Mali, Yang *et al.* 2013) que comprenden utilizar ARNcr modificado.

En una realización particular, el ARNcr se puede diseñar técnicamente para reconocer al menos las dos secuencias de ácido nucleico diana de forma simultánea. En este caso, el mismo ARNcr comprende al menos dos secuencias complementarias a una porción de las dos secuencias de ácido nucleico diana. En una realización preferente, dichas secuencias complementarias están espaciadas por una secuencia de repetición.

El ARN CRISPR transactivador de acuerdo con la presente invención se caracteriza por una secuencia de antirrepetición que tiene la capacidad del emparejamiento de bases con al menos una parte de la secuencia de extensión 3' del ARNcr, para formar un ARNcrtra:ARNcr, también llamado ARN guía (ARNg). El ARNcrtra comprende una secuencia complementaria a una región del ARNcr.

Un ARN guía único (ARNgu) sintético que comprende una fusión de ARNcr y ARNcrtra que forma una horquilla que imita el complejo ARNcrtra-ARNcr (Cong, Ran *et al.* 2013; Mali, Yang *et al.*, 2013) se puede utilizar para dirigir la escisión del ácido nucleico diana mediada por endonucleasa Cas9. Se ha demostrado que este sistema funciona en una diversidad de células eucariotas, incluyendo las humanas, de pez cebra y levaduras. El ARNgu puede comprender dos secuencias distintas complementarias a una porción de las dos secuencias de ácido nucleico diana, preferentemente espaciadas por una secuencia de repetición.

Los métodos de la invención implican introducir ARNcr, el ARNcrtra, ARNgu y Cas9 en un linfocito T. El ARNcr, el ARNcrtra, el ARNgu o la Cas9 se pueden sintetizar *in situ* en la célula como resultado de la introducción de un polinucleótido que codifica ARN o polipéptidos en la célula. Como alternativa, el ARNcr, el ARNcrtra, el ARNgu, el ARN de Cas9 o los polipéptidos de Cas9, podrían producirse fuera de la célula e introducirse después en la misma. Los métodos de introducción de una construcción de polinucleótido en animales son conocidos en la técnica e incluyen, como ejemplos no limitantes, métodos de transformación estable en donde la construcción de polinucleótido está integrada en el genoma de la célula, métodos de transformación transitoria en donde la construcción de polinucleótido no está integrada en el genoma de la célula y métodos mediados por virus. Dichos polinucleótidos pueden introducirse en una célula mediante, por ejemplo, vectores víricos recombinantes (por ejemplo, retrovirus, adenovirus), liposomas y similares. Por ejemplo, los métodos de transformación transitoria incluyen, por ejemplo, microinyección, electroporación o bombardeo de partículas. Dichos polinucleótidos pueden incluirse en vectores, más particularmente plásmidos o virus, en vista de expresarlos en células T.

También se divulgan en el presente documento polinucleótidos, en particular el ADN o ARN que codifica los polipéptidos y proteínas descritas anteriormente. Estos polinucleótidos pueden incluirse en vectores, más particularmente plásmidos o virus, en vista de expresarlos en células T.

La presente invención contempla la modificación de la secuencia del polinucleótido Cas9 de forma que el uso de codones se optimice para el organismo en el que se está introduciendo. Por lo tanto, por ejemplo, la secuencia del polinucleótido Cas9 obtenida de *pyogenes* o *S. Thermophilus* con codones optimizados para su uso en humanos, se expone en (Cong, Ran *et al.* 2013; Mali, Yang *et al.* 2013).

En realizaciones particulares, los polinucleótidos Cas9 de acuerdo con la presente invención pueden comprender al menos un motivo de localización subcelular. Un motivo de localización subcelular se refiere a una secuencia que facilita el transporte o confinamiento de una proteína en un emplazamiento subcelular definido que incluye al menos uno de núcleo, citoplasma, membrana plasmática, retículo endoplasmático, aparato de Golgi, endosomas, peroxisomas y mitocondrias. Se conocen bien en la técnica los motivos de localización subcelular. Un motivo de localización subcelular requiere una orientación específica, por ejemplo, N- y/o C-terminal para la proteína. Como un ejemplo no limitante, la señal de localización nuclear (NLS, forma siglada de *nuclear localization signal*) del antígeno

T grande del virus de simio 40 puede orientarse en el extremo N y/o C. La NLS es una secuencia de aminoácidos que actúa para dirigir la proteína al núcleo de la célula a través del Complejo del Poro Nuclear y para dirigir una proteína recién sintetizada al núcleo a través de su reconocimiento por receptores citosólicos de transporte nuclear. Normalmente, una NLS consiste en una o más secuencias cortas de aminoácidos cargados positivamente, tales como lisinas o argininas.

La presente invención también se refiere a un método para modificar una secuencia genética de interés, que comprende adicionalmente la etapa de expresar un dominio catalítico adicional en una célula hospedadora. En una realización más preferente, la presente invención se refiere a un método para aumentar la mutagénesis, en donde dicho dominio catalítico adicional es una enzima para el procesamiento final de ADN. Los ejemplos no limitantes de enzimas para el procesamiento final de ADN incluyen exonucleasas 5-3', exonucleasas 3-5', exonucleasas alcalinas 5-3', endonucleasas flap 5', helicasas, fosfatasa, hidrolasas y polimerasas de ADN independientes de molde. Los ejemplos no limitantes de tal dominio catalítico comprenden un dominio de proteína o derivado catalíticamente activo del dominio de proteína seleccionado del grupo que consiste en Exolh (EXO1\_HUMAN), Exol de levadura (EXO1\_YEAST), Exol de *E. coli*, TREX2 de ser humano, TREX1 de ratón, TREX1 de ser humano, TREX1 de bovino, TREX1 de rata, TdT (desoxinucleotidil transferasa terminal) DNA2 Humana, DNA2 de levadura (DNA2\_YEAST\_YEAST). En una realización preferente, dicho dominio catalítico adicional tiene una actividad exonucleasa 3'-5' y en una realización más preferente, dicho dominio catalítico adicional tiene actividad de exonucleasa de TREX, más preferentemente actividad de TREX2. En otra realización preferente, dicho dominio catalítico está codificado por un polipéptido TREX de cadena única.

Se sabe que las roturas endonucleolíticas estimulan la tasa de recombinación homóloga. Por lo tanto, en otra realización preferente, la presente invención se refiere a un método para inducir el direccionamiento del gen homólogo en la secuencia genética de interés, que comprende adicionalmente proporcionar a la célula un ácido nucleico exógeno que comprende al menos una secuencia homóloga a una porción de la secuencia genética de interés, de forma que la recombinación homóloga se produce entre la secuencia genética de interés y el ácido nucleico exógeno.

En realizaciones particulares, dicho ácido nucleico exógeno comprende unas primera y segunda porciones que son homólogas a la región 5' y 3' de la secuencia genética de interés, respectivamente. Dicho ácido nucleico exógeno también comprende en estas realizaciones una tercera porción posicionada entre la primera y la segunda porción, que no comprende una homología con las regiones 5' y 3' de la secuencia genética de interés. Después de la escisión de la secuencia genética de interés, se estimula un suceso de recombinación homóloga entre la secuencia del ácido nucleico diana y el ácido nucleico exógeno. Preferentemente, se utilizan secuencias homólogas de al menos 50 pb, preferentemente de más de 100 pb y más preferentemente de más de 200 pb, dentro de dicho ácido nucleico exógeno. Por lo tanto, el ácido nucleico exógeno es preferentemente de 200 pb a 6000 pb, más preferentemente de 1000 pb a 2000 pb. De hecho, las homologías de ácido nucleico compartidas se emplazan en regiones flanqueantes cadena arriba y cadena abajo de la escisión inducida y la secuencia de ácidos nucleicos que debe introducirse debe estar emplazada entre los dos brazos.

Dependiendo del emplazamiento de la secuencia genética de interés en donde se ha producido el suceso de ruptura, tal ácido nucleico exógeno se puede utilizar para suprimir un gen, por ejemplo, cuando el ácido nucleico exógeno se emplaza dentro de la fase de lectura abierta de dicho gen, o para introducir secuencias nuevas o genes de interés. Las inserciones de secuencias utilizando tal ácido nucleico exógeno se pueden utilizar para modificar un gen existente que se tiene como objetivo, mediante corrección o reemplazo de dicho gen (intercambio de alelos como un ejemplo no limitante), o para regular al alza o a la baja la expresión del gen que se tiene como objetivo (intercambio de promotor como ejemplo no limitante), dicha corrección o reemplazo del gen dirigido.

#### Células modificadas y kits

En el método de acuerdo con la presente invención se utilizan linfocitos T. Los linfocitos T pueden ser líneas celulares para cultivos *in vitro* o células primarias de origen animal.

Por "célula primaria" o "células primarias" se entiende células tomadas directamente del tejido vivo (es decir, material de biopsia) y establecidas para el cultivo *in vitro*, que han experimentado muy pocas duplicaciones de su población y son, por lo tanto, más representativas de los principales componentes funcionales y las características de los tejidos de los que se obtienen, en comparación con líneas celulares continuas tumorígenas o inmortalizadas de forma artificial. Por lo tanto, estas células representan un modelo más valioso para el estado *in vivo* al que se refieren.

En el marco de la presente invención, "linfocitos T" se refiere a linfocitos T o a una línea de linfocitos T animales obtenidos de los organismos enumerados a continuación y establecidos para el cultivo *in vitro*.

Más preferentemente, la célula animal es del género *Homo*, *Rattus*, *Mus*, *Sus*, *Bos*, *Danio*, *Canis*, *Felis*, *Equus*, *Salmo*, *Oncorhynchus*, *Gallus*, *Meleagris*, más preferentemente, la célula animal es de las especies *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Sus scrofa*, *Bos taurus*, *Danio rerio*, *Canis lupus*, *Felis catus*, *Equus caballus*, *Salmo salar*, *Oncorhynchus mykiss*, *Gallus gallus*, *Meleagris gallopavo*.

En la presente invención, el linfocito T es preferentemente una línea de células o células de mamífero para cultivos *in vitro*, o células primarias tomadas directamente de tejido vivo y establecidas para el cultivo *in vitro*. Como ejemplos no limitantes, las líneas celulares se pueden seleccionar del grupo que consiste en células Jurkat y células Molt 4.

5 Estas líneas celulares se pueden modificar mediante el método de la presente invención, para proporcionar modelos de líneas celulares para producir, expresar, cuantificar, detectar, estudiar un gen o una proteína de interés; estos modelos también se pueden utilizar para explorar moléculas biológicamente activas de interés en investigación y producción, y en diversos campos tales como el de los productos químicos, los biocombustibles, los productos terapéuticos y la agronomía, como ejemplos no limitantes. Un aspecto particular de la presente invención se refiere a  
 10 una célula aislada como se describe anteriormente, obtenida mediante el método de acuerdo con la invención. Normalmente, dicha célula aislada comprende las nickasas Cas9, el ARNcr (o los ARNcr) y el ARNcrtra o el ARNgu. La célula aislada resultante comprende una secuencia genética modificada de interés en la que se ha producido una escisión. La célula modificada resultante se puede utilizar como una línea celular para una diversidad de aplicaciones, tales como la terapia celular mediante el uso de linfocitos T. Los métodos de la invención son útiles para diseñar técnicamente genomas y reprogramar células. También se divulga en el presente documento un kit  
 15 para la transformación de células que comprende uno o varios de los componentes del sistema CRISPR tipo II modificado de acuerdo con la invención, como se describe anteriormente. Este kit comprende de forma más particular:

- 20 - dos ARNcr que comprenden una secuencia complementaria a una cadena de una primera y segunda secuencias diana de ácido nucleico bicatenarias que comprenden un motivo PAM en la otra cadena y que tienen una secuencia de extensión 3';
- 25 - al menos un ARNcrtra que comprende una secuencia complementaria a las secuencias de extensión 3' de dichos ARNcr;
- al menos una nucleasa cas9 que alberga un dominio de nucleasa de tipo RuvC o HNH no funcional, o un polinucleótido que codifica el mismo.

30 En otra realización, el kit comprende:

- Dos ARNcr que comprenden una secuencia complementaria a una cadena de una primera y segunda secuencias diana de ácido nucleico bicatenarias que comprenden distintos motivos PAM en la otra cadena y que tienen una secuencia de extensión 3';
- 35 - al menos un ARNcrtra que comprende una secuencia complementaria a las secuencias de extensión 3' de dichos ARNcr;
- una primera nucleasa Cas9 que reconoce específicamente el motivo PAM de la primera diana de ácido nucleico y que alberga un tipo RuvC no funcional, o un polinucleótido que codifica la misma.
- 40 - una segunda nucleasa Cas9 que reconoce específicamente el motivo PAM de la segunda diana de ácido nucleico y que alberga un dominio de nucleasa HNH no funcional, o un polinucleótido que codifica la misma.

#### 45 Aplicaciones terapéuticas

El método divulgado en el presente documento puede tener una diversidad de aplicaciones. En una realización, el método puede utilizarse para aplicaciones clínicas o terapéuticas. El método puede utilizarse para reparar o corregir genes que provocan enfermedades. El método puede utilizarse para corregir mutaciones de la zona de unión del empalme, deleciones, inserciones y similares en otros genes o secuencias cromosómicas que desempeñan un papel  
 50 en una enfermedad o patología particular.

Dichos métodos también se pueden utilizar para modificar genéticamente linfocitos T primarios, con vistas a inyectar tales células en un paciente, para tratar una enfermedad o una infección. Dichos programas de terapia celular se desarrollan más particularmente para tratar el cáncer, una infección vírica tal como la provocada por el CMV o el VIH, o enfermedades autoinmunitarias.

#### Definiciones

60 En la descripción anterior se utilizan de forma extensa varios términos. Las siguientes definiciones se proporcionan para facilitar la comprensión de las presentes realizaciones.

Como se utiliza en el presente documento, "un", "una" o "uno" pueden significar uno o más de uno.

Los restos de aminoácido en una secuencia de polipéptido se designan en el presente documento de acuerdo con el código de una letra, en el cual, por ejemplo, Q significa Gln o resto de glutamina, R significa Arg o resto de arginina, y D significa Asp o resto de ácido aspártico.

- 5 Sustitución de aminoácidos significa el reemplazo de un resto de aminoácido por otro, por ejemplo, el reemplazo de un resto de arginina por un resto de glutamina en una secuencia peptídica es una sustitución de aminoácido.

Los nucleótidos se designan de la siguiente manera: el código de una letra se utiliza para designar la base de un nucleósido: a es adenina, t es timina, c es citosina y g es guanina. Para los nucleótidos degenerados, r representa g o a (nucleótidos de purina), k representa g o t, s representa g o c, w representa a o t, m representa a o c, y representa t o c (nucleótidos de pirimidina), d representa g, a o t, v representa g, a o c, b representa g, t o c, h representa a, t o c, y n representa g, a, t o c.

15 Como se utiliza en el presente documento, "ácido nucleico" o "polinucleótido" se refiere a nucleótidos y/o polinucleótidos, tal como el ácido desoxirribonucleico (ADN) o el ácido ribonucleico (ARN), oligonucleótidos, fragmentos generados por la reacción en cadena de la polimerasa (PCR) y fragmentos generados por cualquiera de ligamiento, corte, acción de endonucleasa y acción exonucleasa. Las moléculas de ácido nucleico pueden componerse de monómeros que son nucleótidos de origen natural (tal como ADN y ARN) o análogos de nucleótidos de origen natural (por ejemplo, formas enantioméricas de nucleótidos de origen natural), o una combinación de ambos. Los nucleótidos modificados pueden tener modificaciones en fracciones de azúcar y/o en fracciones de base pirimidínica o purínica. Las modificaciones de azúcares incluyen, por ejemplo, el reemplazo de uno o más grupos hidroxilo por halógenos, grupos alquilo, las aminas y los grupos azido, o azúcares, se pueden funcionalizar como éteres o ésteres. Además, toda la fracción de azúcar puede reemplazarse por estructuras estéricamente y electrónicamente similares, tales como aza-azúcares y análogos carbocíclicos de azúcar. Los ejemplos de modificaciones en una fracción de base incluyen purinas y pirimidinas alquiladas, purinas o pirimidinas aciladas, u otros sustitutos heterocíclicos bien conocidos. Los monómeros de ácido nucleico se pueden unir mediante enlaces fosfodiéster o análogos de tales enlaces. Los ácidos nucleicos pueden ser monocatenarios o bicatenarios.

30 Por "secuencia complementaria" se entiende la parte de la secuencia de un polinucleótido (por ejemplo, parte de ARNcr o ARNcrtra) que puede hibridar con otra parte de polinucleótidos (por ejemplo, la secuencia de ácido nucleico diana o el ARNcr, respectivamente), en condiciones convencionales de baja rigurosidad. Tales condiciones pueden ser, por ejemplo, a temperatura ambiente durante 2 horas utilizando un tampón que contiene formamida al 25 %, SSC 4x, tampón NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> 50 mM; pH 7,0, sol. Denhardt 5x, EDTA 1 mM, DNA 1 mg/ml + 20 a 200 ng/ml de la sonda a analizar (aprox. 20 - 200 ng/ml). Esto también puede predecirse mediante el cálculo convencional de hibridación utilizando el número de bases complementarias dentro de la secuencia y el contenido en G-C a temperatura ambiente, como se proporciona en la bibliografía. De preferencia, las secuencias son complementarias entre sí, de acuerdo con la complementariedad entre dos cadenas de ácido nucleico basadas en el emparejamiento entre las cadenas de Watson-Crick, es decir, el emparejamiento de bases inherente entre los nucleótidos adenina y timina (A-T) y los nucleótidos guanina y citosina (G-C). El emparejamiento preciso de bases que se equipara con el emparejamiento de bases de Watson-Crick incluye el emparejamiento de bases entre nucleósidos convencionales y modificados, y el emparejamiento de bases entre nucleósidos modificados, donde los nucleósidos modificados son capaces de sustituir los nucleósidos convencionales apropiados de acuerdo con el apareamiento de Watson-Crick. La secuencia complementaria del oligonucleótido monocatenario puede ser de cualquier longitud que sustente la hibridación específica y estable entre los dos oligonucleótidos monocatenarios en las condiciones de reacción. La secuencia complementaria generalmente permite un solapamiento parcial de doble cadena entre los dos oligonucleótidos hibridados de más de 3 pb, preferentemente más de 5 pb, preferentemente más de 10 pb. La secuencia complementaria se selecciona de forma ventajosa para que no sea homóloga a ninguna secuencia en el genoma, para evitar la recombinación fuera de la diana o la recombinación que no implique la secuencia de ácido nucleico exógena completa (es decir, solo un oligonucleótido).

50 Por "secuencia de ácido nucleico homóloga" se entiende una secuencia de ácido nucleico con suficiente identidad con otra como para conducir a la recombinación homóloga entre secuencias, de forma más particular, que tiene al menos el 80 % de identidad, preferentemente al menos el 90 % de identidad y más preferentemente al menos el 95 %, e incluso más preferentemente el 98 % de identidad. "Identidad" se refiere a la identidad de secuencia entre dos moléculas de ácido nucleico o polipéptidos. La identidad se puede determinar comparando una posición en cada secuencia que puede alinearse con fines de comparación. Cuando una posición en la secuencia comparada está ocupada por la misma base, entonces las moléculas son idénticas en esa posición. Un grado de similitud o identidad entre secuencias de ácido nucleico o de aminoácidos es una función del número de nucleótidos idénticos o que coinciden en las posiciones compartidas por las secuencias de ácido nucleico. Para calcular la identidad entre dos secuencias se pueden utilizar diversos algoritmos y/o programas de alineamiento, incluyendo FASTA o BLAST, que están disponibles como parte del paquete de análisis de secuencias GCG (Universidad de Wisconsin, Madison, Wis.), y pueden utilizarse con, por ejemplo, parámetros predeterminados.

65 Los términos "vector" o "vectores" hacen referencia a una molécula de ácido nucleico que tiene la capacidad de transportar otro ácido nucleico al que se lo ha unido. Un "vector" en la presente invención incluye, pero sin limitación, un vector vírico, un plásmido, un vector de ARN o una molécula de ADN o ARN lineal o circular que puede consistir

en ácidos nucleicos cromosómicos y no cromosómicos, semisintéticos o sintéticos. Los vectores preferentes son los que tienen la capacidad de replicación autónoma (vector episómico) y/o de expresión de los ácidos nucleicos a los que están unidos (vectores de expresión). Los expertos en la materia conocen grandes cantidades de vectores adecuados disponibles en el mercado. Los vectores víricos incluyen retrovirus, adenovirus, parvovirus (por ejemplo, virus adenoasociados), coronavirus, virus de ARN de cadena negativa, tales como ortomixovirus (por ejemplo, virus de la gripe), rhabdovirus (por ejemplo, virus de la rabia y de la estomatitis vesicular), paramixovirus (por ejemplo, sarampión y Sendai), virus de ARN de cadena positiva tales como picornavirus y alfavirus, y virus de ADN bicatenario, que incluyen adenovirus, herpesvirus (por ejemplo, virus Herpes Simple tipos 1 y 2, Virus de Epstein-Barr, citomegalovirus) y poxvirus (por ejemplo, vaccinia, virus de la viruela aviar y de la viruela del canario). Otros virus incluyen el virus de Norwalk, los togavirus, los flavivirus, los reovirus, los papovavirus, los hepadnavirus y el virus de la hepatitis, por ejemplo. Los ejemplos de retrovirus incluyen: de la leucosis-sarcoma aviar, de mamífero tipo C, virus de tipo B, virus tipo D, grupo HTLV-BLV, lentivirus, espumavirus (Coffin, J. M., *Retroviridae: The viruses and their replication*, en *Fundamental Virology*, tercera edición, B. N. Fields, *et al.*, Eds., Lippincott-Raven Publishers, Filadelfia, 1996).

La presente invención se ilustra por referencia a determinados ejemplos específicos.

#### REFERENCIAS

- Chylinski, K., A. Le Rhun, *et al.* (2013). "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems." *RNA Biol* 10(5).
- Cong, L., F. A. Ran, *et al.* (2013). "Multiplex genome engineering using CRISPR/Cas systems." *Science* 339(6121): 819-23.
- Critchlow, S. E. y S. P. Jackson (1998). "DNA end-joining: from yeast to man." *Trends Biochem Sci* 23(10): 394-8.
- Dalgaard, J. Z., A. J. Klar, *et al.* (1997). "Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family." *Nucleic Acids Res* 25(22): 4626-38.
- Deltcheva, E., K. Chylinski, *et al.* (2011). "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." *Nature* 471(7340): 602-7.
- Deveau, H., R. Barrangou, *et al.* (2008). "Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*." *J Bacteriol* 190(4): 1390-400.
- Garneau, J. E., M. E. Dupuis, *et al.* (2010). "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." *Nature* 468(7320): 67-71.
- Gasiunas, G., R. Barrangou, *et al.* (2012). "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria." *Proc Natl Acad Sci U S A* 109(39): E2579-86.
- Gorbalenya, A. E. (1994). "Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family." *Protein Sci* 3(7): 1117-20.
- Haft, D. H., J. Selengut, *et al.* (2005). "A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes." *PLoS Comput Biol* 1(6): e60.
- Horvath, P. y R. Barrangou (2010). "CRISPR/Cas, the immune system of bacteria and archaea." *Science* 327(5962): 167-70.
- Jiang, W., D. Bikard, *et al.* (2013). "RNA-guided editing of bacterial genomes using CRISPR-Cas systems." *Nat Biotechnol* 31(3): 233-9.
- Jinek, M., K. Chylinski, *et al.* (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." *Science* 337(6096): 816-21.
- Kleanthous, C., U. C. Kuhlmann, *et al.* (1999). "Structural and mechanistic basis of immunity toward endonuclease colicins." *Nat Struct Biol* 6(3): 243-52.
- Ma, J. L., E. M. Kim, *et al.* (2003). "Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences." *Mol Cell Biol* 23(23): 8820-8.

- Makarova, K. S., N. V. Grishin, *et al.* (2006). "A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action." *Biol Direct* 1: 7.
- 5 Mali, P., L. Yang, *et al.* (2013). "RNA-guided human genome engineering via Cas9." *Science* 339(6121): 823-6.
- Mojica, F. J., C. Diez-Villasenor, *et al.* (2009). "Short motif sequences determine the targets of the prokaryotic CRISPR defence system." *Microbiology* 155(Pt 3): 733-40.
- 10 Qi, L. S., M. H. Larson, *et al.* (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell* 152(5): 1173-83.
- Sapranauskas, R., G. Gasiunas, *et al.* (2011). "The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*." *Nucleic Acids Res* 39(21): 9275-82.
- 15 Schutz, K., J. R. Hesselberth, *et al.* (2010). "Capture and sequence analysis of RNAs with terminal 2',3'-cyclic phosphates." *Rna* 16(3): 621-31.
- Shah, S. A., S. Erdmann, *et al.* (2013). "Protospacer recognition motifs: Mixed identities and functional diversity." *RNA Biol* 10(5).
- 20 Shub, D. A., H. Goodrich-Blair, *et al.* (1994). "Amino acid sequence motif of group I intron endonucleases is conserved in open reading frames of group II introns." *Trends Biochem Sci* 19(10): 402-4.
- Sorek, R., C. M. Lawrence, *et al.* (2013). "CRISPR-mediated Adaptive Immune Systems in Bacteria and Archaea." *Annu Rev Biochem*.
- 25 van der Ploeg, J. R. (2009). "Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M 102-like bacteriophages." *Microbiology* 155(Pt 6): 1966-76.
- 30

LISTADO DE SECUENCIAS

- <110> Collectis
- 35 <120> UN MÉTODO PARA PRODUCIR UNA ESCISIÓN DE ADN PRECISA UTILIZANDO LA ACTIVIDAD NICKASA DE CAS9
- <130> P81303480PCT00
- 40 <150> PA201370295  
<151> 29-05-2013
- <160> 1
- 45 <170> PatentIn version 3.5
- <210> 1  
<211> 1368  
<212> PRT
- 50 <213> *Streptococcus pyogenes* de serotipo M1
- <220>  
<223> Cas9
- 55 <400> 1

ES 2 670 531 T3

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val  
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe  
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile  
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu  
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
 100 105 110

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

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
 130 135 140

ES 2 670 531 T3

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His  
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro  
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr  
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala  
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn  
 210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn  
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe  
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp  
 260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp  
 275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp  
 290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser  
 305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys  
 325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe  
 340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser  
 355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp  
 370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg  
 385 390 395 400

ES 2 670 531 T3

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu  
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe  
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile  
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp  
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu  
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr  
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser  
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys  
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln  
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr  
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp  
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly  
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp  
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr  
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala  
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr  
645 650 655

ES 2 670 531 T3

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp  
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe  
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe  
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu  
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly  
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly  
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln  
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile  
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro  
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu  
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg  
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys  
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg  
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys  
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys  
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp

ES 2 670 531 T3

900					905					910					
Lys	Ala	Gly	Phe	Ile	Lys	Arg	Gln	Leu	Val	Glu	Thr	Arg	Gln	Ile	Thr
		915					920					925			
Lys	His	Val	Ala	Gln	Ile	Leu	Asp	Ser	Arg	Met	Asn	Thr	Lys	Tyr	Asp
	930					935					940				
Glu	Asn	Asp	Lys	Leu	Ile	Arg	Glu	Val	Lys	Val	Ile	Thr	Leu	Lys	Ser
945					950					955					960
Lys	Leu	Val	Ser	Asp	Phe	Arg	Lys	Asp	Phe	Gln	Phe	Tyr	Lys	Val	Arg
				965					970					975	
Glu	Ile	Asn	Asn	Tyr	His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn	Ala	Val
			980					985					990		
Val	Gly	Thr	Ala	Leu	Ile	Lys	Lys	Tyr	Pro	Lys	Leu	Glu	Ser	Glu	Phe
		995					1000					1005			
Val	Tyr	Gly	Asp	Tyr	Lys	Val	Tyr	Asp	Val	Arg	Lys	Met	Ile	Ala	Lys
	1010					1015					1020				
Ser	Glu	Gln	Glu	Ile	Gly	Lys	Ala	Thr	Ala	Lys	Tyr	Phe	Phe	Tyr	Ser
1025					1030					1035					1040
Asn	Ile	Met	Asn	Phe	Phe	Lys	Thr	Glu	Ile	Thr	Leu	Ala	Asn	Gly	Glu
				1045					1050					1055	
Ile	Arg	Lys	Arg	Pro	Leu	Ile	Glu	Thr	Asn	Gly	Glu	Thr	Gly	Glu	Ile
			1060				1065						1070		
Val	Trp	Asp	Lys	Gly	Arg	Asp	Phe	Ala	Thr	Val	Arg	Lys	Val	Leu	Ser
		1075					1080				1085				
Met	Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr	Glu	Val	Gln	Thr	Gly	Gly
	1090					1095					1100				
Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys	Arg	Asn	Ser	Asp	Lys	Leu	Ile
1105					1110					1115					1120
Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro	Lys	Lys	Tyr	Gly	Gly	Phe	Asp	Ser
				1125					1130					1135	
Pro	Thr	Val	Ala	Tyr	Ser	Val	Leu	Val	Val	Ala	Lys	Val	Glu	Lys	Gly
			1140				1145						1150		

ES 2 670 531 T3

Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile  
 1155 1160 1165

Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala  
 1170 1175 1180

Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys  
 1185 1190 1195 1200

Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser  
 1205 1210 1215

Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr  
 1220 1225 1230

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser  
 1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His  
 1250 1255 1260

Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val  
 1265 1270 1275 1280

Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys  
 1285 1290 1295

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu  
 1300 1305 1310

Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp  
 1315 1320 1325

Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp  
 1330 1335 1340

Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile  
 1345 1350 1355 1360

Asp Leu Ser Gln Leu Gly Gly Asp  
 1365

**REIVINDICACIONES**

1. Un método *in vitro* para inducir con precisión una escisión de ácido nucleico en una secuencia genética en una célula, que comprende:

- 5 (a) Seleccionar una primera y una segunda diana de ácido nucleico bicatenario en dicha secuencia genética, comprendiendo cada diana de ácido nucleico, en una cadena, un motivo adyacente al protoespaciador (PAM) en una extremidad 3';
- 10 (b) diseñar técnicamente dos ARN de direccionamiento de CRISPR (los ARNcr), comprendiendo cada uno:
  - una secuencia complementaria a una parte de la cadena opuesta de la diana de ácido nucleico que no comprende el motivo PAM, y
  - una secuencia de extensión 3';
- 15 (c) proporcionar al menos un ARN de direccionamiento de CRISPR transactivador (ARNcrtra) que comprende una secuencia complementaria a una parte de las secuencias de extensión 3' de dichos ARNcr de b);
- (d) proporcionar al menos una nickasa cas9 que es una nucleasa cas9 que alberga un dominio de nucleasa de tipo RuvC no funcional o HNH no funcional y que reconoce dicho motivo (o motivos) PAM;
- 20 (e) introducir en la célula dichos ARNcr, dicho (o dichos) ARNcrtra y dicha nickasa Cas9; en donde la célula es un linfocito T,

de forma que cada complejo Cas9-ARNcrtra:ARNcr induzca un suceso de muesca en las dianas de ácido nucleico bicatenario para escindir la secuencia genética entre dichas primera y segunda dianas de ácido nucleico.

- 25 2. El método de la reivindicación 1, en donde los dos motivos PAM están presentes en cadenas de ácido nucleico opuestas.
- 30 3. El método de la reivindicación 1, en donde los dos motivos PAM están presentes en la misma cadena de ácido nucleico.
- 4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde la primera y la segunda dianas de ácido nucleico bicatenario comprenden distintos motivos PAM reconocidos específicamente por dos nickasas Cas9 distintas.
- 35 5. El método de la reivindicación 4, en donde dicho método implica una primera nickasa Cas9 que alberga un tipo RuvC no funcional y una segunda nickasa Cas9 que alberga un dominio de nucleasa HNH no funcional.
- 40 6. El método de acuerdo con las reivindicaciones 1 a 5, en donde al menos una nickasa Cas9 comprende al menos una mutación en el dominio RuvC.
- 7. El método de acuerdo con las reivindicaciones 1 a 5, en donde al menos una nickasa Cas 9 comprende al menos una mutación en el dominio HNH.
- 45 8. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en donde cada ARNcr comprende una secuencia complementaria de 12 a 20 nucleótidos.
- 50 9. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 8, que comprende en la etapa b) diseñar técnicamente un ARNcr que comprende dos secuencias complementarias a una parte de cada una de las secuencias de ácido nucleico diana.
- 10. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 9, en donde el ARNcr y el ARNcrtra se fusionan para formar un ARN guía único.
- 55 11. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 10, en donde la primera y la segunda secuencias diana de ácido nucleico están separadas entre sí por una región espaciadora de 1 a 300 pb, preferentemente de 3 a 250 pb.
- 60 12. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 11, que comprende adicionalmente introducir una secuencia de ácido nucleico exógena que comprende al menos una secuencia homóloga a al menos una porción de la secuencia genética, de forma que se produce recombinación homóloga entre dicha secuencia exógena y la secuencia genética.
- 65 13. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 12, en donde dicho linfocito T es un linfocito T primario.

14. Un linfocito T aislado que comprende:

- dos ARNcr que comprenden secuencias complementarias a una primera y una segunda secuencias diana de ácido nucleico bicatenario y que tienen una secuencia de extensión 3';
- al menos un ARNcrtra que comprende una secuencia complementaria a las secuencias de extensión 3' de dichos ARNcr;
- al menos una nickasa cas9 que es una nucleasa cas9 que alberga un dominio de tipo RuvC no funcional o un dominio de nucleasa HNH no funcional, o un polinucleótido que codifica la misma.

5

10 15. El linfocito T de acuerdo con la reivindicación 14, para su uso en el tratamiento del cáncer, de una infección vírica o de enfermedades autoinmunitarias.

16. El linfocito T para su uso de acuerdo con la reivindicación 15, en donde el linfocito T es un linfocito T primario.

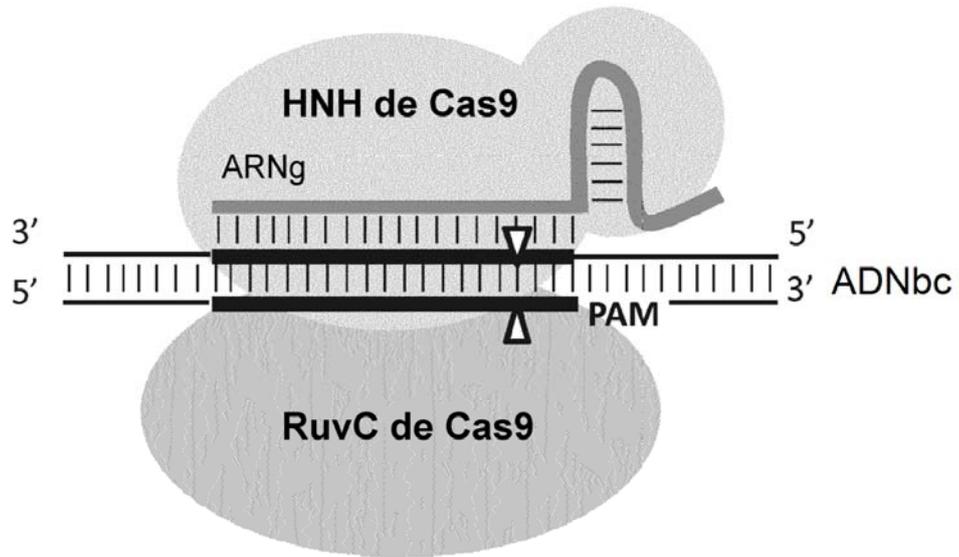


Figura 1

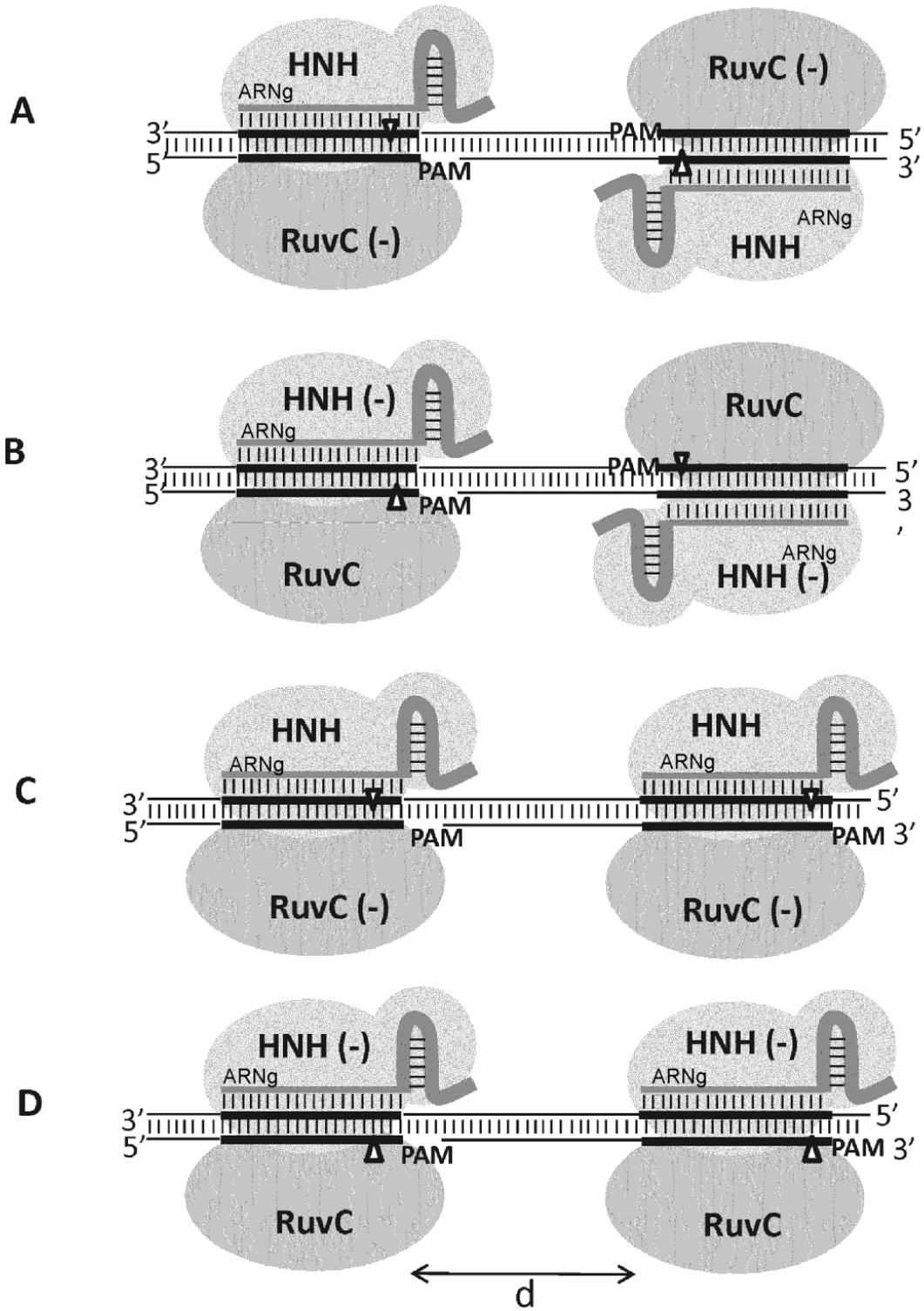


Figura 2

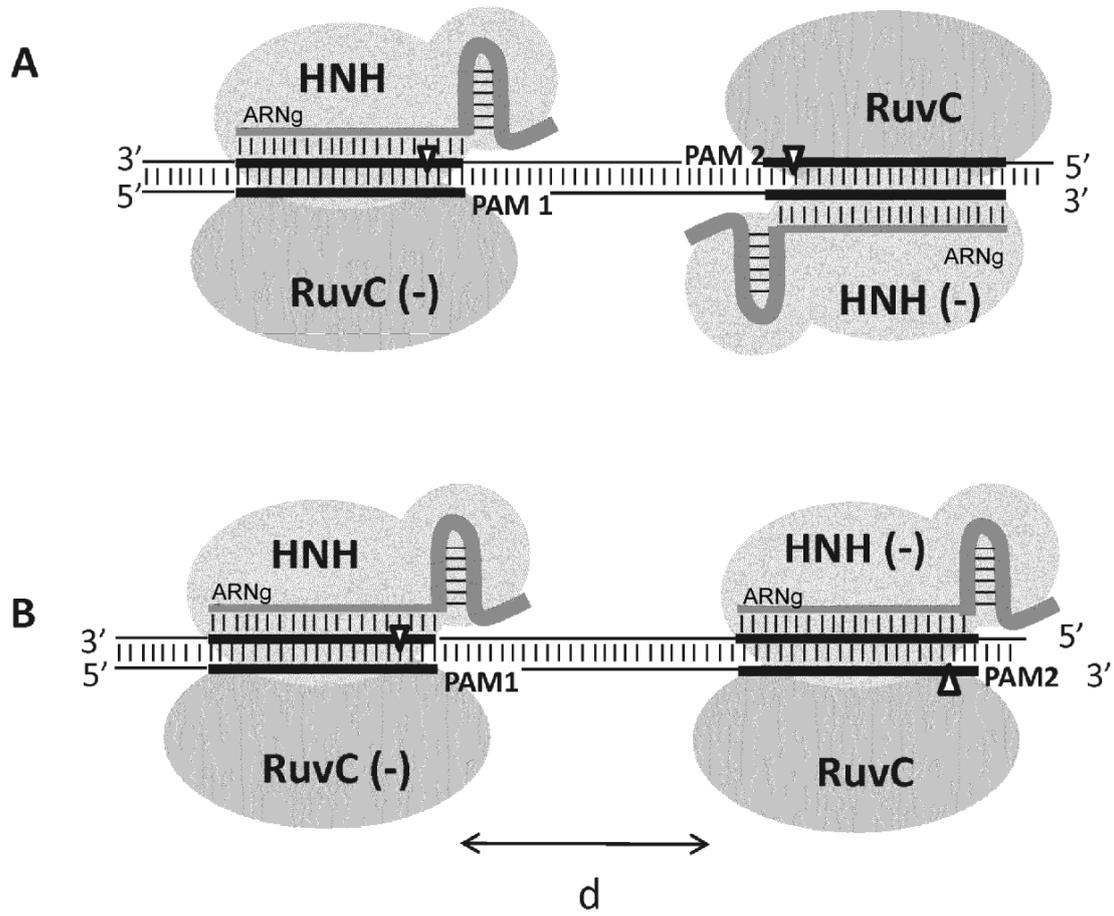


Figura 3