

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 757 623**

51 Int. Cl.:

**A61K 48/00** (2006.01)

**C12N 15/62** (2006.01)

**C12N 15/63** (2006.01)

**C12N 15/90** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.07.2013 E 18200462 (2)**

97 Fecha y número de publicación de la concesión europea: **18.09.2019 EP 3494997**

54 Título: **Proteínas de unión a ADN inducibles y herramientas de perturbación genómica y aplicaciones de las mismas**

30 Prioridad:

**25.07.2012 US 201261675778 P**  
**01.11.2012 US 201261721283 P**  
**12.12.2012 US 201261736465 P**  
**15.03.2013 US 201361794458 P**  
**17.06.2013 US 201361835973 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.04.2020**

73 Titular/es:

**THE BROAD INSTITUTE, INC. (33.3%)**  
**7 Cambridge Center**  
**Cambridge, MA 02142, US;**  
**MASSACHUSETTS INSTITUTE OF TECHNOLOGY**  
**(33.3%) y**  
**PRESIDENT AND FELLOWS OF HARVARD**  
**COLLEGE (33.3%)**

72 Inventor/es:

**ZHANG, FENG;**  
**BRIGHAM, MARK;**  
**CONG, LE;**  
**KONERMANN, SILVANA y**  
**SANJANA, NEVILLE ESPI**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 757 623 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**

Proteínas de unión a ADN inducibles y herramientas de perturbación genómica y aplicaciones de las mismas

5 Solicitudes relacionadas

Esta solicitud reivindica prioridad a reivindica beneficio de las solicitudes de patente provisionales de Estados Unidos n.º 61/675778 presentada el 25 de julio de 2012, 61/721283 presentada el 1 de noviembre de 2012, 61/736465 presentada el 12 de diciembre de 2012, 61/794458 presentada el 15 de marzo de 2013 y 61/835973 presentada el 17 de junio de 2013 tituladas INDUCIBLE DNA BINDING PROTEINS AND GENOME PERTURBATION TOOLS AND APPLICATIONS THEREOF.

15 También se hace referencia a la solicitud provisional de Estados Unidos n.º 61/565171 presentada el 30 de noviembre de 2011 y las solicitudes de Estados Unidos n.º 13/554922 presentada el 30 de julio de 2012 y 13/604945 presentada el 6 de septiembre de 2012, tituladas NUCLEOTIDE-SPECIFIC RECOGNITION SEQUENCES FOR DESIGNER TAL EFFECTORS.

20 También se hace referencia a las solicitudes provisionales de Estados Unidos n.º 61/736527 presentada el 12 de diciembre de 2012; 61/748427 presentada el 2 de enero de 2013; 61/757972 presentada el 29 de enero de 2013, 61/768959 presentada el 25 de febrero de 2013 y 61/791409 presentada el 15 de marzo de 2013, tituladas SYSTEMS METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION.

25 También se hace referencia a las solicitudes provisionales de Estados Unidos n.º 61/758468 presentada el 30 de enero de 2013 y 61/769046 presentada el 15 marzo de 2013, tituladas ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION.

También se hace referencia a las solicitudes provisionales de Estados Unidos n.º 61/835931; 61/835936; 61/836080; 61/836101; 61/836123 y 61/836127 presentadas el 17 de junio de 2013.

30 También se hace referencia a la solicitud provisional de Estados Unidos n.º 61/842322 presentada el 2 de julio de 2013, titulada CRISPR-CAS SYSTEMS AND METHODS FOR ALTERING EXPRESSION OF GENE PRODUCTS y la solicitud provisional de Estados Unidos n.º 61/847537, presentada el 17 de julio de 2013, titulada DELIVERY, ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION AND APPLICATIONS.

35 Las solicitudes anteriores, y todos los documentos citados en los mismos o durante su proceso ("documentos citados en la solicitud") y todos los documentos citados o mencionados en los documentos citados en la solicitud, y todos los documentos citados o mencionados en este documento ("documentos citados en este documento") y todos los documentos citados o mencionados en los documentos citados de este documento, junto con cualquier instrucción del fabricante, descripción, especificaciones de producto y fichas técnicas para cualquier producto mencionado en este documento pueden emplearse en la práctica de la invención.

Campo de la invención

45 La presente invención se refiere en general a complejos de CRISPR-Cas, sistemas y usos relacionados.

Leyenda federal de financiación

50 Esta invención se realizó con el apoyo gubernamental según R01NS073124 y el Premio Pioneer 1DP1MH100706 otorgado por los National Institutes of Health. El gobierno tiene determinados derechos en la invención.

Antecedentes de la invención

55 Se necesitan tecnologías dirigidas al genoma precisas y eficaces para posibilitar la genomaniplulación inversa sistemática de variaciones genéticas causales permitiendo la perturbación selectiva de elementos genéticos individuales. Aunque las tecnologías de edición genómica tales como dedos de cinc de diseño (ZF) (M. H. Porteus, D. Baltimore, Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763 (2 de mayo de 2003); J. C. Miller *et al.*, An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25, 778 (julio de 2007); J. D. Sander *et al.*, Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* 8, 67 (enero de 2011) y A. J. Wood *et al.*, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (15 de julio de 2011)), efectores de tipo activador de la transcripción (TALE) (A. J. Wood *et al.*, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (15 de julio de 2011); M. Christian *et al.*, Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757 (octubre de 2010); F. Zhang *et al.*, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29, 149 (febrero de 2011); J. C. Miller *et al.*, A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29, 143 (febrero de 2011); D. Reyon *et al.*, FLASH assembly of TALENs for

high-throughput genome editing. *Nat Biotechnol* 30, 460 (mayo de 2012); J. Boch *et al.*, Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509 (11 de diciembre de 2009) y M. J. Moscou, A. J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501 (11 de diciembre de 2009)), y meganucleasas de asentamiento (B. L. Stoddard, Homing endonuclease structure and function. *Quarterly reviews of biophysics* 38, 49 (febrero de 2005)) han empezado a posibilitar modificaciones genómicas dirigidas, sigue habiendo una necesidad de nuevas tecnologías que se puedan cambiar de escala, se puedan producir y sean fáciles de diseñar. Aquí, los solicitantes informan del desarrollo de una nueva clase de herramientas de manipulación genómica de precisión basadas en la nucleasa Cas9 guiada por ARN (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (17 de agosto de 2012); G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109, E2579 (25 de septiembre de 2012) y J. E. Garneau *et al.*, The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67 (4 de noviembre de 2010)) de un sistema inmunitario adaptativo de CRISPR procarionota de tipo II (H. Deveau, J. E. Garneau, S. Moineau, CRISPR/Cas system and its role in phage-bacteria interactions. *Annual review of microbiology* 64, 475 (2010); P. Horvath, R. Barrangou, CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167 (8 de enero de 2010); K. S. Makarova *et al.*, Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9, 467 (junio de 2011) y D. Bhaya, M. Davison, R. Barrangou, CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45, 273 (2011)).

20 La mención o identificación de cualquier documento en esta solicitud no es una admisión de que dicho documento esté disponible como técnica anterior a la presente invención.

Sumario de la invención y de la presente divulgación

25 La presente invención se define por las reivindicaciones adjuntas. A continuación, se describen aspectos de la presente invención y/o presente divulgación. Los aspectos de la presente divulgación también pueden pertenecer a la invención como se define en las reivindicaciones adjuntas.

30 En particular, la presente invención proporciona un complejo de repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR)-Cas, que comprende:

- una Cas9, en la que la Cas9 tiene dos o más señales de localización nuclear;
- una secuencia de guía unida a una secuencia de acoplamiento tracr, y
- una secuencia tracr hibridable con la totalidad o una parte de la secuencia de acoplamiento tracr;

35 en el que la secuencia de guía se diseña para que tenga complementariedad con una secuencia diana en una célula eucariota y pueda dirigir la unión específica de secuencia del complejo de CRISPR a la secuencia diana en la célula eucariota.

40 La invención también proporciona un sistema de vector de repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR)-Cas, que comprende:

45 un primer elemento regulador unido de forma funcional a una o más secuencias que codifican (1) una secuencia de guía unida a una secuencia de acoplamiento tracr, y (2) una secuencia tracr hibridable con la totalidad o una parte de la secuencia de acoplamiento tracr;

un segundo elemento regulador unido de forma funcional a una secuencia que codifica una Cas9, en la que la Cas9 tiene dos o más señales de localización nuclear;

50 en el que los componentes (a) y (b) están ubicados en el mismo o en diferentes vectores del sistema y en el que la secuencia de guía se diseña para que tenga complementariedad con una secuencia diana en una célula eucariota.

La invención también proporciona el complejo de CRISPR-Cas de acuerdo con la invención o el sistema de vector de CRISPR-Cas de la invención para su uso en tratamiento.

55 Se divulga además el uso del complejo de CRISPR de acuerdo con la invención o el sistema de vector de CRISPR-Cas de acuerdo con la invención para manipulación genómica, en el que dicho uso no es un método de tratamiento del organismo animal o humano por tratamiento, y en el que dicho uso no es un proceso para modificar la identidad genética de la línea germinal de los seres humanos.

60 Aspectos adicionales de la divulgación proporcionan sistemas y métodos como se describe en este documento en los que el sistema de CRISPR puede comprender un sistema de vector que comprende: a) un primer elemento regulador unido de forma funcional ligado a un ARN de guía del sistema de CRISPR-Cas dirigido a un locus de interés, b) un segundo elemento inducible regulador unido de forma funcional a una proteína Cas, en el que los componentes (a) y  
65 (b) pueden estar ubicados en el mismo o en diferentes vectores del sistema, en el que el ARN de guía está dirigido a

ADN del locus de interés, en el que la proteína Cas y el ARN de guía no se producen de forma natural juntos. La proteína Cas es una enzima Cas9.

La invención también proporciona el vector que es un AAV o un lentivirus.

Un aspecto adicional de la divulgación es una manipulación genómica múltiple inducible usando sistema de CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas)/Cas9

La presente invención también se refiere a ácido nucleico que codifica los polipéptidos de la presente invención el ácido nucleico puede comprender un promotor, ventajosamente el promotor de sinapsina I humana (hSyn). En una realización particularmente ventajosa, el ácido nucleico puede empaquetarse en un vector de virus adenoasociado (AAV).

La invención también se refiere además a métodos de tratamiento o terapia que engloban los métodos y composiciones descritos en este documento.

Esta y otras realizaciones se divulgan y están englobadas por la siguiente descripción detallada.

Breve descripción de los dibujos

La siguiente descripción detallada, dada a modo de ejemplo, pero no destinada a limitar la invención, únicamente a las realizaciones específicas descritas, puede entenderse mejor junto con los dibujos adjuntos.

Las **figuras 1A-E** representan un locus de CRISPR de tipo II de *Streptococcus pyogenes* SF370 que puede reconstituirse en células de mamífero para facilitar DSB dirigidas de ADN. **(A)** La manipulación de SpCas9 y SpRNasa III con NLS posibilita la importación al núcleo de mamífero. **(B)** La expresión en mamífero de SpCas9 y SpRNasa III está dirigida por el promotor EF1a, mientras que la matriz de ARNtracr y preARNcr (DR-espaciador-DR) están dirigidos por el promotor U6. Se usa un protoespaciador (resaltado en azul) del locus *EMX1* humano con PAM como molde para el espaciador en la matriz de preARNcr. **(C)** Representación esquemática de emparejamiento de bases en el locus diana y el ARNcr dirigido a *EMX1*. La flecha roja indica el supuesto sitio de escisión. **(D)** Ensayo SURVEYOR para indels mediadas por SpCas9. **(E)** Un cromatograma ejemplar que muestra una microeliminación, así como secuencias representativas de alelos mutados identificados de 187 amplicones clonales. Guiones rojos, bases eliminadas; bases rojas, inserciones o mutaciones. Barra de escala=10  $\mu$ m.

Las **figuras 2A-C** representan que una SpCas9 puede programarse para que se dirija a múltiples locus genómicos en células de mamífero. **(A)** Esquema del locus *EMX1* humano que muestra la ubicación de cinco protoespaciadores, indicados por líneas azules con el correspondiente PAM en magenta. **(B)** Esquema del complejo de preARNcr:ARNtracr (parte superior) que muestra hibridación entre la región de repetición directa (gris) del preARNcr y ARNtracr. Esquema de un diseño de ARN quimérico (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816 (17 de agosto de 2012)) (parte inferior). La secuencia de ARNtracr se muestra en rojo y la secuencia espaciadora de 20 pb en azul. **(C)** Ensayo SURVEYOR que compara la eficacia de escisión mediada por Cas9 en cinco protoespaciadores en el locus *EMX1* humano. Cada protoespaciador se dirige usando el complejo procesado de preARNcr:ARNtracr (ARNcr) o ARN quimérico (ARNqui).

Las **figuras 3A-D** representan una evaluación de la especificidad de SpCas9 y la comparación de la eficacia con TALEN. **(A)** Se generaron ARNcr quiméricos dirigidos a *EMX1* con mutaciones puntuales individuales para evaluar los efectos de emparejamientos incorrectos de espaciador-protoespaciador. **(B)** Ensayo SURVEYOR que compara la eficacia de escisión de diferentes ARN quiméricos mutantes. **(C)** Esquema que muestra el diseño de TALEN dirigidas a *EMX1*. **(D)** Gel SURVEYOR que compara la eficacia de TALEN y SpCas9 (N=3).

Las **figuras 4A-G** representan aplicaciones de Cas9 para recombinación homóloga y manipulación genómica múltiple. **(A)** La mutación del dominio RuvC I convierte Cas9 en una enzima melladora (SpCas9n). **(B)** La coexpresión del ARN quimérico dirigido a *EMX1* con SpCas9 da lugar a indels, mientras que SpCas9n no (N=3). **(C)** Representación esquemática de la estrategia de recombinación. Se diseña un molde de reparación para insertar sitios de restricción en el locus *EMX1*. Los cebadores usados para amplificar la región modificada se muestran como flechas rojas. **(D)** Análisis en gel de polimorfismos de longitud de fragmentos de restricción. Las flechas indican los fragmentos generados por la digestión con *HindIII*. **(E)** Cromatograma ejemplar que muestra recombinación satisfactoria. **(F)** SpCas9 puede facilitar la modificación genómica múltiple usando una matriz de ARNcr que contiene dos espaciadores dirigidos a *EMX1* y *PVALB*. Esquema que muestra el diseño de la matriz de ARNcr (parte superior). Ambos espaciadores median la escisión eficaz de protoespaciador (parte inferior). **(G)** SpCas9 puede usarse para conseguir eliminación genómica precisa. Dos espaciadores dirigidos a *EMX1* (parte superior) mediaron una eliminación genómica de 118 pb (parte inferior).

La **figura 5** representa un esquema de la rotura bicatenaria de ADN mediada por CRISPR de tipo II. El locus de CRISPR de tipo II de *Streptococcus pyogenes* SF370 contiene un grupo de cuatro genes *Cas9*, *Cas1*, *Cas2*, y *Csn1*, así como dos elementos de ARN no codificantes, ARNtracr y una matriz característica de secuencias repetitivas (repeticiones directas) intercaladas por tramos cortos de secuencias no repetitivas (espaciadores, 30 pb cada uno) (15-18, 30, 31). Cada espaciador se obtiene típicamente de material genético exógeno (protoespaciador) y dirige la especificidad de la escisión de ácido nucleico mediada por CRISPR. En el ácido nucleico diana, cada protoespaciador está asociado con un motivo adyacente a protoespaciador (PAM) cuyo reconocimiento es específico para sistemas de CRISPR individuales (22, 23). El sistema de CRISPR de tipo II realiza rotura bicatenaria (DSB) de ADN diana en etapas secuenciales (M. Jinek *et al.*, Science 337, 816 (17 de agosto de 2012); Gasiunas, R. *et al.* Proc Natl Acad Sci USA 109, E2579 (25 de septiembre de 2012); J. E. Garneau *et al.*, Nature 468, 67 (4 de noviembre de 2010); R. Sapranuskas *et al.*, Nucleic Acids Res 39, 9275 (noviembre de 2011); A. H. Magadan *et al.* PLoS One 7, e40913 (2012)). En primer lugar, la matriz del preARNcr y el ARNtracr se transcriben del locus de CRISPR. En segundo lugar, el ARNtracr hibrida con las repeticiones directas del preARNcr y se asocia con Cas9 como una doble hélice, que media el procesamiento del preARNcr en ARNcr maduros que contienen secuencias espaciadoras individuales truncadas. En tercer lugar, la doble hélice de ARNcr:ARNtracr madura dirige Cas9 a la diana de ADN que consiste en el protoespaciador y el PAM necesario mediante formación de heterodúplex entre la región espaciadora del ARNcr y el ADN protoespaciador. Finalmente, Cas9 media la escisión del ADN diana en dirección 5' de PAM para crear una DSB dentro del protoespaciador.

Las **figuras 6A-C** representan una comparación de diferentes transcritos de ARNtracr para la dirección génica mediada por Cas9. **(A)** Esquema que muestra el diseño y secuencias de dos transcritos de ARNtracr ensayados (corto y largo). Cada transcrito está dirigido por un promotor U6. El sitio de inicio de la transcripción se marca como +1 y el terminador de la transcripción es como se indica. La línea azul indica la región cuya secuencia complementaria inversa se usa para generar sondas de transferencia de Northern para la detección de ARNtracr. **(B)** Ensayo SURVEYOR que compara la eficacia de la escisión mediada por hSpCas9 del locus *EMX1*. Se muestran dos replicas biológicas para cada transcrito de ARNtracr. **(C)** Análisis de transferencia Northern de ARN total extraído de células 293FT transfectadas con construcciones de expresión de U6 que portan ARNtracr largo o corto, así como SpCas9 y DR-*EMX1(1)*-DR. Los paneles de la izquierda y la derecha son de células 293FT transfectadas sin o con SpRNasa III respectivamente. U6 indica el control de carga transferido con una sonda dirigida a ARNnp U6 humano. La transfección de la construcción de expresión de ARNtracr corta dio lugar a niveles abundantes de la forma procesada de ARNtracr (~75 pb) (E. Deltcheva *et al.*, Nature 471, 602 (31 de marzo de 2011)). Se detectan cantidades muy bajas del ARNtracr largo en la transferencia de Northern. Como resultado de estos experimentos, los solicitantes eligieron usar ARNtracr corto para su aplicación en células de mamífero.

La **figura 7** representa un ensayo SURVEYOR para la detección de microinserciones y eliminaciones inducidas por rotura bicatenaria (D. Y. Guschin *et al.* Methods Mol Biol 649, 247 (2010)). Esquema del ensayo SURVEYOR usado para determinar la eficacia de escisión mediada por Cas9. En primer lugar, se usa PCR genómica (gPCR) para amplificar la región diana de Cas9 a partir de una población heterogénea de células modificadas y no modificadas, y los productos de gPCR se vuelven a hibridar lentamente para generar heterodúplex. Los heterodúplex rehibridados se escinden por nucleasas SURVEYOR, mientras que los homodúplex se dejan intactos. La eficacia de escisión mediada por Cas9 (% de indel) se calcula basándose en la fracción de ADN escindido.

La **figura 8A-B** representa un análisis de transferencia de Northern de procesamiento de ARNcr en células de mamífero. **(A)** Esquema que muestra el vector de expresión para un único espaciador flanqueado por dos repeticiones directas (DR-*EMX1(1)*-DR). El espaciador de 30 pb dirigido a protoespaciador 1 del locus *EMX1* humano (tabla 1) se muestra en azul y las repeticiones directas se muestran en gris. La línea naranja indica la región cuya secuencia complementaria inversa se usa para generar sondas de transferencia de Northern para la detección de ARNcr de *EMX1(1)*. **(B)** Análisis de transferencia de Northern de ARN total extraído de células 293FT transfectadas con construcciones de expresión de U6 que portan DR-*EMX1(1)*-DR. Los paneles de la izquierda y la derecha son de células 293FT transfectadas sin o con SpRNasa III respectivamente. DR-*EMX1(1)*-DR se procesó en ARNcr maduros únicamente en presencia de SpCas9 y ARNtracr corto, y no dependió de la presencia de SpRNasa III. El ARNcr maduro detectado del ARN total de 293FT transfectadas es ~33 pb y es más corto que el ARNcr maduro de 39-42 pb de *S. pyogenes* (E. Deltcheva *et al.*, Nature 471, 602 (31 de marzo de 2011)), lo que sugiere que el ARNcr maduro procesado en células 293FT humanas probablemente es diferente del ARNcr maduro bacteriano en *S. pyogenes*.

La **figura 9A-B** representa vectores de expresión bicistrónicos para la matriz de preARNcr o ARNcr quimérico con Cas9. **(A)** Esquema que muestra el diseño de un vector de expresión para la matriz de preARNcr. Pueden insertarse espaciadores entre dos sitios *BbsI* usando oligonucleótidos hibridados. El diseño de secuencia para los oligonucleótidos se muestra a continuación con los adaptadores de ligamiento apropiados indicados. **(B)** Esquema del vector de expresión para ARNcr quimérico. La secuencia de guía puede insertarse entre dos sitios *BbsI* usando oligonucleótidos hibridados. El vector ya contiene la repetición directa parcial (gris) y secuencias de ARNtracr parciales (rojo). WPRE, elemento regulador postranscripcional del virus de la hepatitis de marmota.

Las **figuras 10A-B** representan una selección de protoespaciadores en los locus *PVALB* humano y *Th* de ratón. Esquema de los locus *PVALB* humano **(A)** y *Th* de ratón **(B)** y la ubicación de los tres protoespaciadores dentro

del último exón de los genes *PVALB* y *Th*, respectivamente. Los protoespaciadores de 30 pb se indican por líneas negras y las secuencias de PAM adyacentes se indican por la barra magenta. Los protoespaciadores en las hebras con sentido y de antisentido se indican por encima y por debajo de las secuencias de ADN respectivamente.

5 Las **figuras 11A-C** representan las incidencias de secuencias PAM en el genoma humano. Histogramas de distancias entre el PAM del locus 1 de *Streptococcus pyogenes* SF370 adyacente (NGG) (**A**) y el PAM del locus 1 de *Streptococcus thermophilus* LMD9 (NNAGAAW) (**B**) en el genoma humano. (**C**) Distancias para cada PAM por cromosoma. Crom., cromosoma. Las supuestas dianas se identificaron usando tanto la hebra más como la hebra menos de las secuencias cromosómicas humanas. Dado que puede haber cromatina, metilación del ADN, estructura de ARN y otros factores que pueden limitar la actividad de escisión en algunas dianas protoespaciadoras, es importante apreciar que la capacidad de dirección real podría ser menor que el resultado de este análisis informático.

15 Las **figuras 12A-D** representan CRISPR de tipo II de *Streptococcus thermophilus* LMD-9 que también puede funcionar en células eucariotas. (**A**) Esquema del locus 2 de CRISPR de *Streptococcus thermophilus* LMD-9. (**B**) Diseño del sistema de expresión para el sistema de CRISPR de *S. thermophilus*. *hStCas9* de codones optimizados humana se expresa usando un promotor EF1a constitutivo. Las versiones maduras de ARNtracr y ARNcr se expresan usando el promotor U6 para asegurar un inicio preciso de la transcripción. Se muestran las secuencias para el ARNcr maduro y el ARNtracr. Una única base indicada por la letra minúscula "a" en la secuencia de ARNcr se usó para eliminar la secuencia de poli-U, que sirve como terminador transcripcional de ARN Pol III. (**C**) Esquema que muestra las dianas de secuencias protoespaciadoras y PAM correspondientes en el locus *EMX1* humano. Dos secuencias protoespaciadoras están resaltadas y sus secuencias PAM correspondientes que satisfacen el motivo NNAGAAW se indican por líneas magenta. Ambos protoespaciadores están dirigidos a la hebra de antisentido. (**D**) Ensayo SURVEYOR que muestra la escisión mediada por StCas9 en el locus diana. Los espaciadores 1 y 2 de guía de ARN indujeron un 14 % y un 6,4 %, respectivamente. El análisis estadístico de la actividad de escisión entre replicas biológicas en estos dos sitios protoespaciadores puede encontrarse en la tabla 1.

30 Las **figuras 13A-C** representan **producción de sobrenadante de AAV**. (**a**) Se usaron vectores lentivíricos y AAV que portan GFP para ensayar la eficacia de transducción. (**b**) Se transdujeron neuronas corticales embrionarias primarias con 300 y 250  $\mu$ l de sobrenadante derivado del mismo número de células 293FT transfectadas con AAV o lentivirus. Se recogieron imágenes representativas de la expresión de GFP a 7 d.p.i. Barras de escala=50  $\mu$ m. (**c**) El proceso representado se desarrolló para la producción de sobrenadante de AAV y posterior transducción de neuronas primarias. Las células 293FT se transfectaron con un vector AAV que porta el gen de interés, el vector de empaquetado de serotipo AAV1 (pAAV1) y el plásmido auxiliar (pDF6) usando PEI. Unas 48 h después, el sobrenadante se recogió y filtro a través de una membrana de PVDF de 0,45  $\mu$ m. Después se transdujeron neuronas primarias con sobrenadante y las alícuotas restantes se almacenaron a -80 °C. Se alcanzaron niveles estables de expresión de construcción de AAV después de 5-6 días. La producción de sobrenadante de AAV siguiendo este proceso puede usarse para la producción de hasta 96 construcciones víricas diferentes en formato de 96 pocillos.

40 Las **figuras 14A-G** representan que la proteína Cas9 de unión a ADN guiada por ARN puede usarse para dirigir dominios efectores de transcripción a locus genómicos específicos. (**a**) La nucleasa Cas9 guiada por ARN del sistema de CRISPR/Cas de *Streptococcus pyogenes* de tipo II puede convertirse en una proteína de unión a ADN guiada a ARN nucleolíticamente inactiva (Cas9\*\*) introduciendo dos sustituciones de alanina (D10A y H840A). Esquema que muestra que un ARN de guía sintético (ARNgs) puede dirigir la fusión de Cas9\*\*-efector a un locus específico en el genoma humano. El ARNgs contiene una secuencia de guía de 20 pb en el extremo 5' que especifica la secuencia diana. En el ADN genómico diana, el sitio diana de 20 pb tiene que estar seguido por un motivo PAM 5'-NGG. (**b, c**) Esquemas que muestran los sitios diana de ARNgs en los locus *KLF4* y *SOX2* humanos respectivamente. Cada sitio diana está indicado por la barra azul y la secuencia PAM correspondiente está indicada por la barra magenta. (**d, e**) Esquemas de las construcciones activadoras de la transcripción Cas9\*\*-VP64 y represoras de la transcripción SID4X-Cas9\*\*. (**f, g**) Activación de Cas9\*\*-VP64 y represión de SID4X-Cas9\*\*, respectivamente, mediadas por *KLF4* y *SOX2*. Todos los niveles de ARNm se midieron con respecto a las células de control transfectadas de forma simulada con GFP (media  $\pm$  e.t.m.;  $n=3$ ).

55 La **figura 15** representa los diseños de vector de Tet Cas9 para Cas9 inducible.

La **figura 16** representa un vector y la expresión de EGFP en células 293FT después de inducción con doxiciclina de Cas9 y EGFP.

60

#### Descripción detallada de la invención

65 La expresión "ácido nucleico" o "secuencia de ácido nucleico" se refiere a un oligonucleótido desoxirribonucleico o ribonucleico en forma monocatenaria o bicatenaria. La expresión engloba ácidos nucleicos, es decir, oligonucleótidos, que contienen análogos conocidos de nucleótidos naturales. La expresión también engloba estructuras de tipo ácido

nucleico con cadenas principales sintéticas, véase, por ejemplo, Eckstein, 1991; Baserga *et al.*, 1992; Milligan, 1993; documento WO 97/03211; documento WO 96/39154; Mata, 1997; Strauss-Soukup, 1997; y Samstag, 1996.

5 Como se usa en este documento, "recombinante" se refiere a un polinucleótido sintetizado o manipulado de otro modo *in vitro* (por ejemplo, "polinucleótido recombinante"), a métodos de uso de polinucleótidos recombinantes para producir productos génicos en células u otros sistemas biológicos, o a un polipéptido ("proteína recombinante") codificado por un polinucleótido recombinante. "Medio recombinante" engloba el ligamiento de ácidos nucleicos que tienen diversas regiones codificantes o dominios o secuencias promotoras de diferentes fuentes en un casete de expresión o vector para la expresión de, por ejemplo, expresión inducible o constitutiva de secuencias codificantes de polipéptido en los  
10 vectores de la invención.

15 El término "heterólogo", cuando se usa con referencia a un ácido nucleico, indica que el ácido nucleico está en una célula o un virus donde no se encuentra normalmente en la naturaleza; o comprende dos o más subsecuencias que no se encuentran en la misma relación entre sí que la encontrada normalmente en la naturaleza, o se manipula de forma recombinante para que su nivel de expresión o relación física con respecto a otros ácidos nucleicos u otras moléculas en una células, o estructura, no sea la encontrada normalmente en la naturaleza. Un término similar usado en este contexto es "exógeno". Por ejemplo, un ácido nucleico heterólogo típicamente se produce de forma recombinante, que tiene dos o más secuencias de genes no relacionados dispuestos de una manera no encontrada en la naturaleza; por ejemplo, un gen humano unido de forma funcional o una secuencia promotora insertada en un  
20 vector basado en adenovirus de la invención. Como un ejemplo, un ácido nucleico heterólogo de interés puede codificar un producto génico inmunógeno, en el que el adenovirus se administra terapéutica o profilácticamente como un vehículo o composición de fármaco-vacuna. Las secuencias heterólogas pueden comprender diversas combinaciones de promotores y secuencias, cuyos ejemplos se describen en detalle en este documento.

25 Un "ligando terapéutico" puede ser una sustancia que puede unirse a un receptor de una célula diana con efectos terapéuticos.

30 Un "efecto terapéutico" puede ser una consecuencia de un tratamiento médico de cualquier tipo, cuyos resultados se evalúan por un experto en el campo como deseables y beneficiosos. El "efecto terapéutico" puede ser un cambio de comportamiento o fisiológico que se produce como respuesta al tratamiento médico. El resultado puede ser esperado, inesperado o incluso una consecuencia no pretendida del tratamiento médico. Un "efecto terapéutico" puede incluir, por ejemplo, una reducción de los síntomas en un sujeto que padece infección por un patógeno.

35 Una "célula diana" puede ser una célula en que una alteración en su actividad puede inducir un resultado o respuesta deseada. Como se usa en este documento, una célula puede ser una célula *in vitro*. La célula puede ser una célula aislada que puede no tener capacidad de desarrollo en un organismo completo.

40 Un "ligando" puede ser cualquier sustancia que se una a y forme un complejo con una biomolécula para cumplir un fin biológico. Como se usa en este documento, "ligando" también puede referirse a un "antígeno" o "inmunógeno". Como se usa en este documento, "antígeno" e "inmunógeno" se usan indistintamente.

"Expresión" de un gen o ácido nucleico engloba no solamente expresión génica celular, sino también la transcripción y traducción de uno o más ácidos nucleicos en sistemas de clonación y en cualquier otro contexto.

45 Como se usa en este documento, un "vector" es una herramienta que permite o facilita la transferencia de una entidad de un entorno a otro. A modo de ejemplo, algunos vectores usados en técnicas de ADN recombinante permiten que entidades, tales como un segmento de ADN (tal como un segmento de ADN heterólogo, tal como un segmento de ADNc heterólogo), se transfiera a una célula diana. La presente invención comprende vectores recombinantes que pueden incluir vectores víricos, vectores bacterianos, vectores protozoicos, vectores de ADN o recombinantes de los  
50 mismos.

55 Con respecto al ADN exógeno para la expresión en un vector (por ejemplo, que codifica un epítipo de interés y/o un antígeno y/o un producto terapéutico) y documentos que proporcionan dicho ADN exógeno, así como con respecto a la expresión de factores de transcripción y/o traducción para potenciar la expresión de moléculas de ácido nucleico, y en cuanto a expresiones tales como "epítipo de interés", "producto terapéutico", respuesta inmunitaria", "respuesta inmunológica", "respuesta inmunitaria protectora", "composición inmunológica", "composición inmunógena" y "composición de vacuna", entre otros, se hace referencia a la patente de Estados Unidos n.º 5990091 expedida el 23 de noviembre de 1999 y los documentos WO 98/00166 y WO 99/60164, y los documentos citados en los mismos y los documentos de registro en el proceso de esa patente y esas solicitudes PCT. Por tanto, la patente de Estados Unidos n.º 5990091 y los documentos WO 98/00166 y WO 99/60164 y documentos citados en los mismos y documentos de registro en el proceso de esa patente y estas solicitudes PCT, y otros documentos citados en este documento, pueden consultarse en la práctica de esta invención; y todas las moléculas de ácido nucleico exógenas, promotores y vectores citados en los mismos pueden usarse en la práctica de esta invención. A este respecto, también puede hacerse mención de las patentes de Estados Unidos n.º 6706693; 6716823; 6348450; las solicitudes de patente de Estados  
60 Unidos n.º 10/424409; 10/052323; 10/116963; 10/346021; y el documento WO 99/08713, publicado el 25 de febrero de 1999, del documento PCT/US98/16739.

Aspectos de la invención comprenden los sistemas de CRISPR-Cas de la invención que se suministran a un organismo o una célula o un locus de interés mediante un sistema de suministro. Un medio de suministro es mediante un vector, en el que el vector es un vector vírico, tal como un vector lentivírico o baculovírico o preferiblemente adenovírico/vírico adenoasociado, pero se conocen otros medios de suministro (tales como sistema de levadura, microvesículas, pistolas génicas/medios de adhesión de vectores a nanopartículas de oro) y se proporcionan. En algunas realizaciones, puede suministrarse uno o más de los vectores víricos o plasmídicos mediante nanopartículas, exosomas, microvesículas o una pistola génica.

Como se usa en este documento, las expresiones "composición de fármaco" y "fármaco", "composición vacunal", "vacuna", "composición de vacuna", "composición terapéutica" y "composición terapéutica-inmunológica" cubren cualquier composición que induzca protección frente a un antígeno o patógeno. En algunas realizaciones, la protección puede deberse a una inhibición o prevención de infección por un patógeno. En otras realizaciones, la protección puede inducirse mediante una respuesta inmunitaria contra el uno o más antígenos de interés, o que protege de forma eficaz frente al antígeno; por ejemplo, después de la administración o inyección en el sujeto, provoca una respuesta inmunitaria protectora frente al antígeno o inmunógeno diana o proporciona protección eficaz frente al antígeno o inmunógeno expresado a partir de los vectores adenovíricos innovadores de la invención. La expresión "composición farmacéutica" significa cualquier composición que se suministra a un sujeto. En algunas realizaciones, la composición puede suministrarse para inhibir o prevenir la infección por un patógeno.

Una "cantidad terapéuticamente eficaz" es una cantidad o concentración del vector recombinante que codifica el gen de interés que, cuando se administra a un sujeto, produce una respuesta terapéutica o una respuesta inmunitaria contra el producto génico de interés.

La expresión "vector vírico", como se usa en este documento, incluye, aunque sin limitación, retrovirus, adenovirus, virus adenoasociados, alfavirus y virus del herpes simple.

Para los fines de la invención, las células pueden ser una célula eucariota, ventajosamente una célula de animal, más ventajosamente una célula de mamífero.

La presente invención también contempla una manipulación genómica múltiple usando sistema de CRISPR-Cas. La dilucidación funcional de las variantes o elementos genéticos causales requiere tecnologías de edición genómica precisas. El sistema inmunitario adaptativo de CRISPR procarionta de tipo II (repeticiones palindrómicas cortas intercalada de forma regular agrupadas) ha demostrado facilitar la escisión de ADN específica de sitio guiada por ARN. Los solicitantes manipularon dos sistemas de CRISPR de tipo II diferentes y demostraron que las nucleadas Cas9 pueden dirigirse mediante ARN cortos para inducir escisión precisa en locus genómicos endógenos en células humanas y de ratón. Cas9 también puede convertirse en una enzima melladora para facilitar la reparación dirigida por homología con actividad mutagénica mínima. Finalmente, pueden codificarse múltiples secuencias de guía en una única matriz de CRISPR para posibilitar la edición simultánea de varios sitios dentro del genoma de mamífero, lo que demuestra una capacidad de programación fácil y una aplicabilidad amplia de la tecnología de CRISPR.

En general "sistema de CRISPR" se refiere colectivamente a transcritos y otros elementos implicados en la expresión de o que dirigen la actividad de genes asociados a CRISPR ("Cas"), incluyendo secuencias que codifican un gen Cas, una secuencia tracr (trans-activadora de CRISPR) (por ejemplo, ARNtracr o un ARNtracr parcial activo), una secuencia de acoplamiento tracr (que engloba una "repetición directa" y una repetición directa parcial procesada de ARNtracr en el contexto de un sistema de CRISPR endógeno), una secuencia de guía (también denominada "espaciador" en el contexto de un sistema de CRISPR endógeno) u otras secuencias y transcritos de un locus de CRISPR. En algunas realizaciones, uno o más elementos de un sistema de CRISPR se obtienen de un sistema de CRISPR de tipo I, tipo II o tipo III. En algunas realizaciones, uno o más elementos de un sistema de CRISPR se obtienen de un organismo particular que comprende un sistema de CRISPR endógeno, tal como *Streptococcus pyogenes*. En general, un sistema de CRISPR se caracteriza por elementos que promueven la formación de un complejo de CRISPR en un sitio de una secuencia diana (también denominada protoespaciador en el contexto de un sistema de CRISPR endógeno). En el contexto de formación de un complejo de CRISPR, "secuencia diana" se refiere a una secuencia para la que se diseña una secuencia de guía para que tenga complementariedad, donde la hibridación entre una secuencia diana y una secuencia de guía promueve la formación de un complejo de CRISPR. Una secuencia diana puede comprender cualquier polinucleótido, tal como polinucleótidos de ADN o ARN. En algunas realizaciones, una secuencia diana está ubicada en el núcleo o citoplasma de una célula.

Típicamente, en el contexto de un sistema de CRISPR endógeno, la formación de un complejo de CRISPR (que comprende una secuencia de guía hibridada con una secuencia diana y en complejo con una o más proteínas Cas) provoca la escisión de una o ambas hebras en o cerca de (por ejemplo, a 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 o más pares de bases de) la secuencia diana. Sin el deseo de limitarse a teoría alguna, la totalidad o una parte de la secuencia tracr también puede formar parte de un complejo de CRISPR, tal como por hibridación a la totalidad o una parte de la secuencia de acoplamiento tracr que está unida de forma funcional a la secuencia de guía. En algunas realizaciones, uno o más vectores que dirigen la expresión de uno o más elementos de un sistema de CRISPR se introducen en una célula hospedadora de modo que la expresión de los elementos del sistema de CRISPR dirija la formación de un

complejo de CRISPR en uno o más sitios diana. Por ejemplo, una enzima Cas, una secuencia de guía unida a una secuencia de acoplamiento tracr y una secuencia tracr podrían estar cada una unidas de forma funcional a elementos reguladores diferentes en vectores diferentes. Como alternativa, dos o más de los elementos expresados a partir del mismo o diferentes elementos reguladores pueden combinarse en un único vector, con uno o más vectores adicionales que proporcionan cualquier componente del sistema de CRISPR no incluido en el primer vector. Los elementos del sistema de CRISPR que se combinan en un único vector pueden disponerse en cualquier orientación adecuada tal como un elemento ubicado 5' con respecto a ("cadena arriba" de) o 3' con respecto a ("cadena abajo" de) un segundo elemento. La secuencia codificante de un elemento puede estar ubicada en la misma hebra o en la hebra opuesta de la secuencia codificante de un segundo elemento, y orientada en la misma dirección o en la dirección opuesta. En algunas realizaciones, un único promotor dirige la expresión de un transcrito que codifica una enzima CRISPR y una o más de la secuencia guía, secuencia de acoplamiento tracr (opcionalmente unidad de forma funcional a la secuencia de guía) y una secuencia tracr incluida dentro de una o más secuencias intrónicas (por ejemplo, cada una en un intrón diferente, dos o más en al menos un intrón o todas en un único intrón). En algunas realizaciones la enzima CRISPR, la secuencia de guía, la secuencia de acoplamiento tracr y la secuencia tracr están unidas de forma funcional a y se expresan a partir del mismo promotor.

En algunas realizaciones, un vector comprende uno o ms sitios de inserción, tal como una secuencia de reconocimiento de endonucleasa de restricción (también denominada "sitio de clonación"). En algunas realizaciones, uno o más sitios de inserción (por ejemplo, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más sitios de inserción) están ubicados en dirección 5' y/o dirección 3' de uno o más elementos de secuencia de uno o más vectores. En algunas realizaciones, un vector comprende un sitio de inserción en dirección 5' de una secuencia de acoplamiento tracr y opcionalmente en dirección 3' de un elemento regulador unido de forma funcional a la secuencia de acoplamiento tracr, de modo que después de la inserción de una secuencia de guía en el sitio de inserción y tras la expresión de la secuencia de guía dirige la unión específica de secuencia de un complejo de CRISPR a una secuencia diana en una célula eucariota. En algunas realizaciones, un vector comprende dos o más sitios de inserción, estando ubicado cada sitio de inserción entre dos secuencias de acoplamiento tracr para permitir la inserción de una secuencia de guía en cada sitio. En dicha disposición, las dos o más secuencia de guía pueden comprender dos o más copias de una única secuencia de guía, dos o más secuencias de guía diferentes o combinaciones de estas. Cuando se usan múltiples secuencias de guía diferentes, puede usarse una única construcción de expresión para dirigir la actividad de CRISPR a múltiples secuencias diana correspondientes diferentes dentro de una célula. Por ejemplo, un único vector puede comprender aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 o más secuencias de guía. En algunas realizaciones, pueden proporcionarse aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más de dichos vectores que contienen secuencia de guía, y opcionalmente suministrarse a una célula.

En algunas realizaciones, un vector comprende un elemento regulador unido de forma funcional a una secuencia codificante de enzima que codifica una enzima CRISPR, tal como una proteína Cas. Ejemplos no limitantes de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homólogos de las mismas o versiones modificadas de las mismas. En algunas realizaciones, la enzima CRISPR no modificada tiene actividad de escisión de ADN, tal como Cas9. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o ambas hebras en la ubicación de una secuencia diana, tal como dentro de la secuencia diana y/o dentro del complemento de la secuencia diana. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o ambas hebras en aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 o más pares de bases desde el primer o último nucleótido de una secuencia diana. En algunas realizaciones, un vector codifica una enzima CRISPR que se muta con respecto a una enzima de tipo silvestre correspondiente de modo que la enzima CRISPR mutada carezca de la capacidad de escindir una o ambas hebras de un polinucleótido diana que contiene una secuencia diana. Por ejemplo, una sustitución de aspartato a alanina (D10A) en el dominio catalítico RuvC I de Cas9 de *S. pyogenes* convierte Cas9 de una nucleasa que escinde ambas hebras en una nickasa (escinde una única hebra). Otros ejemplos de mutaciones que hacen que Cas9 sea una nickasa incluyen, sin limitación H840A, N854A y N863A. Como ejemplo adicional, pueden mutarse dos o más dominios catalíticos de Cas9 (RuvC I, RuvC II y RuvC III) para producir una Cas9 mutada que carece sustancialmente de toda la actividad de escisión de ADN. En algunas realizaciones, una mutación D10A se combina con una o más mutaciones H840A, N854A o N863A para producir una enzima Cas9 que carece sustancialmente de toda actividad de escisión de ADN. En algunas realizaciones, una enzima CRISPR se considera que carece sustancialmente de toda la actividad de escisión de ADN cuando la actividad de escisión de ADN de la enzima mutada es menor de aproximadamente un 25 %, 10 %, 5 %, 1 %, 0,1 %, 0,01 % o inferior con respecto a su forma no mutada.

En algunas realizaciones, la secuencia codificante de enzima que codifica una enzima CRISPR tiene los codones optimizados para su expresión en células particulares, tales como células eucariotas. Las células eucariotas pueden ser aquellas de o derivadas de un organismo particular, tal como un mamífero incluyendo, aunque sin limitación, ser humano, ratón, rata, conejo, perro o primate no humano. En general, la optimización de codones se refiere a un proceso de modificación de una secuencia de ácido nucleico para la expresión potenciada en las células hospedadoras de interés reemplazando al menos un codón (por ejemplo, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 o más codones) de la secuencia nativa con codones que se usan más frecuentemente o mucho más

frecuentemente en los genes de esa célula hospedadora manteniendo al mismo tiempo la secuencia de aminoácidos nativa. Diversas especies muestran una desviación particular por determinados codones de un aminoácido particular. La desviación de codones (diferencias en el uso de codones entre organismos) a menudo se correlaciona con la eficacia de traducción del ARN mensajero (ARNm), que, a su vez, se cree que depende de, entre otras cosas, las propiedades de los codones que se están traduciendo y la disponibilidad de moléculas particulares de ARN transferente (ARNt). La predominancia de ARNt seleccionados en una célula en general es un reflejo de los codones usados mucho más frecuentemente en la síntesis de péptidos. Por consiguiente, los genes pueden adaptarse para expresión génica óptima en un organismo dado basándose en la optimización de codones. Hay tablas de uso de codones fácilmente disponibles, por ejemplo, en la "base de datos de Uso de Codones" disponible en [www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/) (visitada el 9 de julio de 2002) y estas tablas pueden adaptarse de varias maneras. Véase Nakamura, Y., *et al.* "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000). También están disponibles algoritmos informáticos para la optimización de codones de una secuencia particular para su expresión en una célula hospedadora particular, tal como Gene Forge (Aptagen; Jacobus, PA) que también está disponible. En algunas realizaciones, uno o más codones (por ejemplo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 o más, o todos los codones) en una secuencia que codifica una enzima CRISPR corresponden al codón más frecuentemente usado para un aminoácido particular.

En algunas realizaciones, un vector codifica una enzima CRISPR que comprende una o más secuencias de localización nuclear (NLS), tal como aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS. En algunas realizaciones, la enzima CRISPR comprende aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS en o cerca del extremo amino, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS en o cerca del extremo carboxi o una combinación de estos (por ejemplo, una o más NLS en el extremo amino y una o más NLS en el extremo carboxi). Cuando hay más de una NLS presente, cada una puede seleccionarse independientemente de las otras, de modo que una única NLS puede estar presente en más de una copia y/o en combinación con una o más NLS distintas presentes en una o más copias. En algunas realizaciones, una NLS se considera cerca del extremo N o C cuando el amino ácido más cercano a la NLS está en aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 o más aminoácidos a lo largo de la cadena polipeptídica desde el extremo N o C. Ejemplos no limitantes de NLS incluyen una secuencia de NLS derivado de: la NLS del antígeno T grande del virus SV40, que tiene la secuencia de aminoácidos PKKKRKV; la NLS de nucleoplasmina (por ejemplo, la NLS bipartita de nucleoplasmina con la secuencia KRPAATKKAGQAKKKK); la NLS de c-myc que tiene la secuencia de aminoácidos PAAKRVKLD o RQRRNELKRSP; la NLS hRNP A1 M9 que tiene la secuencia NQSSNFGPMKGGNFGGRSSGPGYGGGGQYFAKPRNQGGY; la secuencia RMRIZFKNKGKDTAELRRRVEVSVELRKAKKDEQILKRRNV del dominio IBB de importina alfa; las secuencias VSRKRPRP y PPKKARED de la proteína T de mioma; la secuencia POPKKKPL de p53 humana; la secuencia SALIKKKKKMAP de c-abl IV de ratón; las secuencias DRLRR y PKQKKRK de NS1 del virus de la gripe; la secuencia RKLKKKIKKL del antígeno delta del virus de la hepatitis; la secuencia REKKKFLKRR de la proteína Mx1 de ratón; la secuencia KRKGDEVDGVDEVAKKKSKK de la poli(ADP ribosa) polimerasa humana; y la secuencia RKCLQAGMNLARKTKK de los receptores de hormonas esteroideas (humanos) glucocorticoesteroides.

En general la una o más NLS son de suficiente fuerza para dirigir la acumulación de la enzima CRISPR en una cantidad detectable en el núcleo de una célula eucariota. En general, la fuerza de la actividad de localización nuclear puede derivar del número de NLS en la enzima CRISPR, la una o más NLS particulares usadas o una combinación de estos factores. La detección de acumulación en el núcleo puede realizarse por cualquier técnica adecuada. Por ejemplo, un marcador detectable puede fusionarse a la enzima CRISPR, de modo que la ubicación dentro de una célula pueda visualizarse, tal como en combinación con un medio para detectar la ubicación del núcleo (por ejemplo, un tinte específico para el núcleo tal como DAPI). Los núcleos celulares también pueden aislarse de las células, cuyos contenidos entonces pueden analizarse por cualquier proceso adecuado para detectar proteínas. Tal como inmunohistoquímica, transferencia Western o ensayo de actividad enzimática. La acumulación en el núcleo también puede determinarse indirectamente, tal como mediante un ensayo para el efecto de formación del complejo de CRISPR (por ejemplo, ensayo para la escisión de ADN o mutación en la secuencia diana, o ensayo para actividad de expresión génica alterada afectada por la formación del complejo de CRISPR y/o una actividad de enzima CRISPR), en comparación con un control no expuesto a la enzima o complejo de CRISPR o expuesto a una enzima CRISPR que carece de la una o más NLS.

En otra realización de la presente invención, la invención se refiere a un CRISPR inducible que puede comprender una Cas9 inducible.

El sistema de CRISPR puede estar codificado dentro de un sistema de vector que puede comprender uno o más vectores que pueden comprender I. un primer elemento regulador unido de forma funcional a una secuencia polinucleotídica de ARN quimérica (ARNqui) del sistema de CRISPR/Cas, en el que la secuencia polinucleotídica puede comprender (a) una secuencia de guía que puede hibridar con una secuencia diana en una célula eucariota, (b) una secuencia de acoplamiento tracr y (c) una secuencia tracr, y II. un segundo elemento regulador unido de forma funcional a una secuencia que codifica enzima que codifica una enzima CRISPR que puede comprender al menos una o más secuencias de localización nuclear, en el que (a), (b) y (c) están dispuestos en una orientación 5' a 3', en el que los componentes I y II están ubicados en el mismo vector o vectores diferentes del sistema, en el que cuando se transcribe, la secuencia de acoplamiento tracr hibrida con la secuencia tracr y la secuencia de guía dirige la unión

específica de secuencia de un complejo de CRISPR a la secuencia diana, y en el que el complejo de CRISPR puede comprender la enzima CRISPR en complejo con (1) la secuencia de guía que hibrida con la secuencia diana y (2) la secuencia de acoplamiento tracr que hibrida con la secuencia tracr, en el que la secuencia que codifica enzima que codifica la enzima CRISPR codifica además un dominio funcional heterólogo.

5 En una realización ventajosa, la Cas9 inducible puede prepararse en un lentivirus. Por ejemplo, la figura 61 representa diseños de vector de Tet Cas9 y la figura 62 representa un vector y expresión de EGFP en células 293FT. En particular, se contempla un sistema de tetraciclina inducible para un CRISPR inducible. El vector puede diseñarse como se describe en Markusic *et al.*, Nucleic Acids Research, 2005, Vol. 33, n.º 6 e63. El sistema regulador de la transcripción dependiente de tetraciclina se basa en el operador de resistencia a tetraciclina Tn10 de *Escherichia coli* que consiste en la proteína represora de tetraciclina (TetR) y un sitio de unión a ADN específico, la secuencia operadora de tetraciclina (TetO). En ausencia de tetraciclina, TetR dimeriza y se une a TetO. Puede unirse tetraciclina o doxiciclina (un derivado de tetraciclina) e inducir un cambio conformacional en TetR que da lugar a su disociación de TetO. En una realización ventajosa, el vector puede ser un único vector lentivírico Tet-On con expresión autorregulada de rtTA para la expresión regulada del complejo de CRISPR. La tetraciclina o doxiciclina pueden contemplarse para activar el complejo de CRISPR inducible.

20 En otra realización, se contempla un sistema de cambio génico de cumate para un CRISPR inducible. Un sistema similar se describe en Mullick *et al.*, BMC Biotechnology 2006, 6:43 doi:10.1186/1472-6750-6-43. El sistema cumate inducible implica mecanismos reguladores de operones bacterianos (cmt y cym) para regular la expresión génica en células de mamífero usando tres estrategias diferentes. En la configuración represora, la regulación está mediada por la unión del represor (CymR) al sitio operador (CuO), colocado en dirección 3' de un promotor constitutivo fuerte. La adición de cumate, una molécula pequeña, alivia la represión. En la configuración transactivadora, una proteína transactivadora quimérica (cTA), formada por la fusión de CymR con el dominio de activación de VP16, puede activar la transcripción cuando se une a múltiples copias de CuO, colocado en dirección 5' del promotor mínimo de CMV. La adición de cumate anula la unión a ADN y, por lo tanto, la transactivación por cTA. La invención también contempla un activador de cumate inverso (rcTA), que activa la transcripción en presencia en lugar de en ausencia de cumate. CymR puede usarse como represor que bloquea de forma reversible la expresión de un promotor fuerte, tal como CMV. Determinados aspectos del sistema represor/operador de cumate se describen adicionalmente en la patente de Estados Unidos n.º 7745592.

35 Se contemplan otros sistemas inducibles tales como, aunque sin limitación, regulación por metales pesados [Mayo KE *et al.*, Cell 1982, 29:99-108; Searle PF *et al.*, Mol Cell Biol 1985, 5:1480-1489 y Brinster RL *et al.*, Nature (Londres) 1982, 296:39-42], hormonas esteroideas [Hynes NE *et al.*, Proc Natl Acad Sci USA 1981, 78:2038-2042; Klock G *et al.*, Nature (Londres) 1987, 329:734-736 y Lee F *et al.*, Nature (Londres) 1981, 294:228-232.], choque térmico [Nouer L: Heat Shock Response. Boca Raton, FL: CRC; 1991] y otros reactivos que se han desarrollado [Mullick A, Massie B: Transcription, translation and the control of gene expression. En Encyclopedia of Cell Technology Editada por: Speir RE. Wiley; 2000:1140-1164 y Fussenegger M. Biotechnol Prog 2001, 17:1-51]. Sin embargo, hay limitaciones con estos promotores de mamífero inducibles tales como "filtración" del estado "inactivo" y efectos pleiotrópicos de los inductores (choque térmico, metales pesados, glucocorticoesteroides, etc.). El uso de hormonas de insectos (ecdisona) se ha propuesto en un intento por reducir la interferencia con procesos celulares en células de mamífero [No D *et al.*, Proc Natl Acad Sci USA 1996, 93:3346-3351]. Otro sistema elegante usa rapamicina como inductor [Rivera VM *et al.*, Nat Med 1996, 2:1028-1032], pero la función de rapamicina como inmunosupresor fue una limitación importante a su uso *in vivo* y, por lo tanto, fue necesario encontrar un compuesto biológicamente inerte [Saez E *et al.*, Proc Natl Acad Sci USA 2000, 97:14512-14517] para el control de la expresión génica.

50 La presente invención también se refiere a ácido nucleico que codifica los polipéptidos de la presente invención. El ácido nucleico puede comprender un promotor, ventajosamente el promotor de sinapsina I humana (hSyn). En una realización particularmente ventajosa, el ácido nucleico puede empaquetarse en un vector de virus adenoasociado (AAV).

También se contemplan por la presente invención, vectores recombinantes y adenovirus recombinantes que pueden comprender partículas subvéricas de más de un serotipo de adenovirus. Por ejemplo, se sabe que los vectores adenovíricos pueden presentar un tropismo alterado por tejidos o tipos celulares específicos (Havenga, M.J.E. *et al.*, 2002) y, por lo tanto, la mezcla y acoplamiento de diferentes cápsidas adenovíricas, es decir, proteínas fibra o pentona de diversos serotipos adenovíricos puede ser ventajoso. La modificación de las cápsidas adenovíricas, incluyendo fibra y pentona puede producir un vector adenovírico con un tropismo que es diferente del adenovirus no modificado. Los vectores adenovíricos que se modifican y optimizan en su capacidad de infectar células diana pueden permitir una reducción significativa en la dosis terapéutica o profiláctica, provocando toxicidad local y diseminada reducida.

60 Los sistemas de suministro génico de vector vírico se usan habitualmente en transferencia génica y aplicaciones de genoterapia. Diferentes sistemas de vector vírico tienen sus propias ventajas y desventajas únicas. Los vectores víricos que pueden usarse para expresar el ligando derivado de patógeno de la presente invención incluyen, aunque sin limitación, vectores adenovíricos, vectores víricos adenoasociados, vectores alfavíricos, vectores víricos del herpes simples y vectores retrovíricos, descritos en mayor detalle a continuación.

Las características generales adicionales de los adenovirus son tales que la biología del adenovirus se caracteriza en detalle; el adenovirus no se asocia con patología humana grave; el adenovirus es extremadamente eficaz en introducir su ADN en la célula hospedadora; el adenovirus puede infectar una amplia diversidad de células y tiene un amplio rango de hospedador; el adenovirus puede producirse en grandes cantidades con relativa facilidad; y el adenovirus puede hacerse defectuoso en la replicación y/o no replicante por eliminaciones en la región temprana 1 ("E1") del genoma vírico.

El adenovirus es un virus de ADN sin envuelta. El genoma del adenovirus es una molécula de ADN bicatenaria lineal de aproximadamente 36 000 pares de bases ("pb") con una proteína terminal de 55 kDa unida covalentemente al extremo 5' de cada hebra. El ADN del adenovirus contiene repeticiones terminales invertidas ("ITR") idénticas de aproximadamente 100 pb, dependiendo la longitud exacta del serotipo. Los orígenes de replicación víricos están ubicados dentro de las ITR exactamente en los extremos del genoma. La síntesis del ADN se produce en dos fases. En primer lugar, la replicación transcurre por desplazamiento de hebra, generando una molécula de doble hélice hija y una hebra desplazada precursora. La hebra desplazada es monocatenaria y puede formar un intermedio "de franja", que permite el inicio de la replicación y la generación de una molécula de doble hélice hija. Como alternativa, la replicación puede transcurrir desde ambos extremos del genoma simultáneamente, obviando la necesidad de formar la estructura de franja.

Durante el ciclo de infección productivo, los genes víricos se expresan en dos fases: la fase temprana, que es el periodo hasta la replicación del ADN vírico y la fase tardía que coincide con el inicio de la replicación del ADN vírico. Durante la fase temprana, únicamente los productos génicos tempranos, codificados por las regiones E1, E2, E3 y E4, se expresan, que pueden realizar varias funciones que preparan a la célula para la síntesis de proteínas estructurales víricas (Berk, A.J., 1986). Durante la fase tardía, los productos génicos víricos tardíos se expresan además de los productos génicos tempranos y el ADN de la célula hospedadora y la síntesis de proteínas se inactivan. Por consiguiente, la célula queda dedicada a la producción de ADN vírico y de proteínas estructurales víricas (Tooze, J., 1981).

La región E1 del adenovirus es la primera región del adenovirus expresada después de la infección de la célula diana. Esta región consiste en dos unidades transcripcionales, los genes E1A y E1B, que son los dos necesarios para la transformación oncogénica de cultivos de roedor primarios (embrionarios). Las funciones principales de los productos génicos E1A son inducir la entrada de las células quiescentes en el ciclo celular y reanudar la síntesis de ADN celular, y activar transcripcionalmente el gen E1B y las otras regiones tempranas (E2, E3 y E4) del genoma vírico. La transfección de células primarias con el gen E1A en solitario puede inducir proliferación ilimitada (inmortalización), pero no provoca transformación completa. Sin embargo, la expresión de E1A, en la mayoría de los casos, provoca la inducción de muerte celular programada (apoptosis) y se obtiene inmortalización únicamente de forma ocasional (Jochemsen *et al.*, 1987). La coexpresión del gen E1B es necesaria para evitar la inducción de apoptosis y para que se produzca transformación morfológica completa. En líneas celulares inmortales establecidas, un alto nivel de expresión de E1A puede causar transformación completa en ausencia de E1B (Roberts, B.E. *et al.*, 1985).

Las proteínas codificadas E1B ayudan a E1A a redirigir las funciones celulares para permitir la replicación vírica. Las proteínas E1B de 55 kD y E4 de 33 kD, que forman un complejo que está esencialmente ubicado en el núcleo, funcionan inhibiendo la síntesis de las proteínas hospedadoras y facilitando la expresión de genes víricos. Su influencia principal es establecer el transporte selectivo de ARNm víricos desde el núcleo hasta el citoplasma, simultáneamente con la aparición de la fase tardía de la infección. La proteína E1B de 21 kD es importante para el control temporal correcto del ciclo de infección productivo, evitando de ese modo la muerte prematura de la célula hospedadora antes de que se haya completado el ciclo vital del virus. Los virus mutantes que no pueden expresar el producto génico E1B de 21 kDa muestran un ciclo de infección acortado que está acompañado por degradación excesiva de ADN cromosómico de la célula hospedadora (fenotipo deg) y en un efecto citopático potenciado (fenotipo cit; Telling *et al.*, 1994). Los fenotipos deg y cit se suprimen cuando además el gen E1A está mutado, lo que indica que estos fenotipos son una función de E1A (White, E. *et al.*, 1988). Además, la proteína E1B de 21 kDa ralentiza la velocidad a la que E1A activa los otros genes víricos. Aún no se conocen los mecanismos por los que E1B de 21 kD inactiva estas funciones dependientes de E1A.

En contraste con los retrovirus, por ejemplo, los adenovirus no se integran de forma eficaz en el genoma de la célula hospedadora, pueden infectar células que no están en división y pueden transferir de forma eficaz genes recombinantes *in vivo* (Brody *et al.*, 1994). Estas características hacen que los adenovirus sean candidatos atractivos para transferencia génica *in vivo* de, por ejemplo, un antígeno o inmunógeno de interés en células, tejidos o sujetos que lo necesitan.

Los vectores adenovíricos que contienen múltiples eliminaciones son preferidos tanto para aumentar la capacidad de transporte del vector como para reducir la probabilidad de recombinación para generar adenovirus competentes en la replicación (RCA). Cuando el adenovirus contiene múltiples eliminaciones, no es necesario que cada una de las eliminaciones, si se presentan en solitario, provoquen un adenovirus deficiente en la replicación y/o no replicante. Siempre que una de las eliminaciones haga que el adenovirus sea deficiente en la replicación o no replicante, pueden incluirse las eliminaciones adicionales para otros fines, por ejemplo, para aumentar la capacidad de transporte del genoma adenovírico para secuencias de nucleótidos heterólogas. Preferiblemente, más de una de las eliminaciones

evita la expresión de una proteína funcional y hace que el adenovirus sea defectuoso en la replicación y/o no replicante y/o atenuado. Más preferiblemente, todas las eliminaciones son eliminaciones que harían que el adenovirus fuera deficiente en la replicación y/o no replicante y/o atenuado. Sin embargo, la invención también engloba adenovirus y vectores adenovíricos que son competentes en la replicación y/o de tipo silvestre, es decir, comprende todos los genes adenovíricos necesarios para la infección y replicación en un sujeto.

Las realizaciones de la invención que emplean recombinantes adenovíricos pueden incluir E1 defectuosa o eliminada, o E3 defectuosa o eliminada, o E4 defectuosa o eliminada o vectores adenovíricos que comprenden eliminaciones de E1 y E3, o E1 y E4, o E3 y E4, o E1, E3, y E4 eliminada, o el vector adenovírico "gutless" en que todos los genes víricos están eliminados. Los vectores adenovíricos pueden comprender mutaciones en los genes E1, E3 o E4, o eliminaciones en estos o todos los genes adenovíricos. La mutación de E1 aumenta el margen de seguridad del vector porque los mutantes de adenovirus de E1 defectuosa se dice que son defectuosos en la replicación y/o no replicantes en células no permisivas y están, en última instancia, muy atenuados. La mutación de E3 potencia la inmunogenia del antígeno alterando el mecanismo por el que el adenovirus regula por disminución las moléculas MHC de clase I. La mutación E4 reduce la inmunogenia del vector adenovírico suprimiendo la expresión del gen tardío, por tanto, puede permitir la utilización en revacunación repetida del mismo vector. La presente invención comprende vectores adenovíricos de cualquier serotipo o serogrupo que tienen eliminación o mutación en E1, o E3, o E4, o E1 y E3, o E1 y E4. La eliminación o mutación de estos genes adenovíricos provoca actividad alterada o pérdida sustancialmente completa de la actividad de estas proteínas.

El vector adenovírico "cobarde" es otro tipo de vector en la familia de vectores adenovíricos. Su replicación requiere un virus auxiliar y una línea celular 293 humana especial que expresa tanto E1a como Cre, una condición que no existe en su entorno natural; el vector está desprovisto de todos los genes víricos, por tanto, el vector como vehículo de vacuna no es inmunógeno y puede inocularse múltiples veces para revacunación. El vector adenovírico "cobarde" también contiene un espacio de 36 kb para acomodar el antígeno o el uno o más inmunógenos de interés, permitiendo de este modo el cosuministro de un gran número de antígeno o inmunógenos en las células.

El virus adenoasociado (AAV) es un parvovirus de ADN bicatenario que es endógeno para la población humana. Aunque tiene capacidad de infección productiva en células de una diversidad de especies, el AAV es un dependovirus, que requiere funciones auxiliares de adenovirus o herpesvirus para su propia replicación. En ausencia de funciones auxiliares de estos herpesvirus, AAV infectará células, perderá la cubierta en el núcleo e integrará su genoma en el cromosoma hospedador, pero no se replicará o producirá nuevas partículas víricas.

El genoma de AAV se ha clonado en plásmidos bacterianos y está bien caracterizado. El genoma vírico consiste en 4682 bases que incluyen dos repeticiones terminales de 145 bases cada una. Estas repeticiones terminales sirven como orígenes de replicación de ADN para el virus. Algunos investigadores han propuesto que tienen funciones potenciadoras. El resto del genoma se divide en dos dominios funcionales. La parte izquierda del genoma codifica las funciones rep que regulan la replicación del ADN vírico y la expresión génica vital. El lado derecho del genoma vital contiene los genes cap que codifican las proteínas estructurales de la cápsida VP1, VP2 y VP3. Las proteínas codificadas por el gen rep y también el gen cap funcionan en trans durante la replicación productiva de AAV.

El AAV se considera un candidato ideal para su uso como vector transductor, y se ha usado de esta manera. Dichos vectores transductores de AAV comprenden suficientes funciones de acción en cis para replicarse en presencia de funciones auxiliares de adenovirus o herpesvirus proporcionadas en trans. Se ha construido AAV recombinante (rAAV) en varios laboratorios y se ha usado para transportar genes exógenos a células de una diversidad de linajes. En estos vectores, los genes cap y/o rep de AAV se eliminan del genoma vírico y se reemplazan con un segmento de ADN de elección. Los vectores actuales pueden acomodar hasta 4300 bases de ADN insertado.

Para producir rAAV, los plásmidos que contienen la construcción vital deseada se transfectan en células infectadas con adenovirus. Además, se cotransfecta un segundo plásmido auxiliar en estas células para proporcionar los genes rep y cap de AAV que son obligatorios para la replicación y empaquetado de la construcción vírica recombinante. En estas condiciones, las proteínas rep y cap de AAV actúan en trans para estimular la replicación y empaquetado de la construcción de rAAV. Tres días después de la transfección, el rAAV se recoge de las células junto con el adenovirus. El adenovirus contaminante entonces se inactiva por tratamiento térmico.

El virus del herpes simple 1 (VHS-1) es un virus de ADN bicatenario con envuelta con un genoma de 153 kb que codifica más de 80 genes. Su amplio rango de hospedador se debe a la unión de las glucoproteínas de la envuelta vírica a las moléculas de sulfato de heparina extracelulares encontradas en membranas celulares (WuDunn & Spear, 1989). La internalización del virus entonces requiere la glucoproteína gD de la envuelta y el receptor del factor de crecimiento de fibroblastos (Kaner, 1990). El VHS puede infectar células de forma lítica o puede establecer latencia. Los vectores de VHS se han usado para infectar una amplia diversidad de tipos celulares (Lowenstein, 1994; Huard, 1995; Miyanochara, 1992; Liu, 1996; Goya, 1998).

Hay dos tipos de vectores de VHS, llamados los vectores de VHS recombinante y los vectores de amplicón. Los vectores de VHS recombinantes se generan por inserción de unidades de transcripción directamente en el genoma

de VHS, a través de eventos de recombinación homóloga. Los vectores de amplicón se basan en plásmidos que alberga la unidad de transcripción de elección, un origen de replicación y una señal de empaquetado.

5 Los vectores de VHS tienen las ventajas obvias de una gran capacidad de inserción de genes exógenos, la capacidad de establecer latencia en neuronas, un amplio rango de hospedador y la capacidad de conferir expresión transgénica al SNC durante hasta 18 meses (Carpenter & Stevens, 1996).

10 Los retrovirus son virus de ARN monocatenario con envuelta, que se han usado ampliamente en protocolos de transferencia génica. Los retrovirus tienen un genoma diploide de aproximadamente 7-10 kb, compuesto de cuatro regiones génicas llamadas *gag*, *pro*, *pol* y *env*. Estas regiones génicas codifican proteínas estructurales de la cápsida, proteasa vírica, integrasa y retrotranscriptasa vírica, y glucoproteínas de la envuelta, respectivamente. El genoma también tiene una señal empaquetado y secuencias de acción en cis, llamadas repeticiones terminales largas (LTR), en cada extremo, que tienen una función en el control transcripcional y la integración.

15 Los vectores retrovéricos más habitualmente usados se basan en el virus de la leucemia murina de Moloney (Mo-MLV) y tienen tropismos celulares variables, dependiente del dominio superficial de unión a receptor de la glucoproteína de la envuelta.

20 A los vectores retrovéricos recombinantes se les elimina todos los genes retrovéricos, que se reemplazan con genes marcadores o terapéuticos, o ambos. Para propagar retrovirus recombinantes, es necesario proporcionar los genes víricos, *gag*, *pol* y *env* en *trans*.

25 Los lentivirus son retrovirus complejos que tienen la capacidad de infectar y expresar sus genes tanto en células mitóticas como en células posmitóticas. El lentivirus más comúnmente conocido es el virus de la inmunodeficiencia humana (VIH), que usa las glucoproteínas de la envuelta de otros virus para abordar una amplia gama de tipos celulares.

30 Los alfavirus, incluyendo el virus Sindbis prototipo (SIN), el virus del bosque Semliki (VBS) y el virus de la encefalitis equina venezolana (EEV), constituyen un grupo de virus con envuelta que contienen genomas de ARN de hebra más, dentro de cápsidas icosaédricas.

35 Los vectores víricos de la presente invención son útiles para el suministro de ácidos nucleicos que expresan antígenos o inmunógenos a células tanto *in vitro* como *in vivo*. En particular, los vectores de la invención pueden emplearse ventajosamente para suministrar o transferir ácidos nucleicos a células, más preferiblemente células de mamífero. Los ácidos nucleicos de interés incluyen ácidos nucleicos que codifican péptidos y proteínas, preferiblemente péptidos o proteínas terapéuticos (por ejemplo, para usos médicos o veterinarios) o inmunógenos (por ejemplo, para vacunas).

40 Preferiblemente, los codones que codifican el antígeno o inmunógeno de interés son codones "optimizados", es decir, los codones son aquellos que aparecen frecuentemente en, por ejemplo, genes muy expresados en la especie del sujeto, en lugar de los codones que se usan frecuentemente por, por ejemplo, un virus de la gripe. Dicho uso de codones proporciona expresión eficaz del antígeno o inmunógeno en células animales. En otras realizaciones, por ejemplo, cuando el antígeno o inmunógeno de interés se expresa en bacterias, levadura u otro sistema de expresión, el patrón de uso de codones se altera para que represente la desviación de codones para genes muy expresados en el organismo en que se está expresando el antígeno o inmunógeno. Los patrones de uso de codones son conocidos en la bibliografía para genes muy expresados de muchas especies (por ejemplo, Nakamura *et al.*, 1996; Wang *et al.*, 45 1998; McEwan *et al.* 1998).

50 Como alternativa adicional, los vectores víricos pueden usarse para infectar una célula en cultivo para que exprese un producto génico deseado, por ejemplo, para producir una proteína o péptido de interés. Preferiblemente, la proteína o péptido se secreta en el medio y puede purificarse del mismo usando técnicas habituales conocidas en la técnica. Se conocen en la técnica secuencias de péptido señal que dirigen la secreción extracelular de proteínas y secuencias de nucleótidos que codifican las mismas pueden unirse de forma funcional a la secuencia de nucleótidos que codifica el péptido o proteína de interés por técnicas habituales conocidas en la técnica. Como alternativa, las células pueden lisarse y la proteína recombinante expresada puede purificarse del lisado celular. Preferiblemente, la célula es una 55 célula animal, más preferiblemente una célula de mamífero. También se prefieren células que sean competentes para la transducción por vectores víricos particulares de interés. Dichas células incluyen células PER.C6, células 911 y células HEK293.

60 Un medio de cultivo para cultivar células hospedadoras incluye un medio habitualmente usado para cultivo tisular, tal como M199-earle base, MEM de Eagle (E-MEM), MEM de Dulbecco (DMEM), SC-UCM102, UP-SFM (GIBCO BRL), EX-CELL302 (Nichirei), EX-CELL293-S (Nichirei), TFBM-01 (Nichirei), ASF104, entre otros. Los medios de cultivo adecuados para tipos celulares específicos pueden encontrarse en la American Type Culture Collection (ATCC) o la European Collection of Cell Cultures (ECACC). Los medios de cultivo pueden complementarse con aminoácidos tales como L-glutamina, sales, agentes antifúngicos o antibacterianos tales como Fungizone®, penicilina-estreptomocina, 65 suero animal y similares. El medio de cultivo celular opcionalmente puede ser sin suero.

Las composiciones terapéuticas o de diagnóstico de la invención se administran a un individuo en cantidades suficientes para tratar o diagnosticar trastornos. La cantidad eficaz puede variar de acuerdo con una diversidad de factores tales como el estado del individuo, el peso, el sexo y la edad. Otros factores incluyen el modo de administración.

Las composiciones farmacéuticas pueden proporcionarse al individuo mediante una diversidad de vías tales como subcutánea, tópica, oral o intramuscular.

Los compuestos identificados de acuerdo con los métodos divulgados en este documento pueden usarse en solitario a dosificaciones apropiadas. Como alternativa, puede ser deseable la coadministración o administración secuencial de otros agentes.

La presente invención también tiene el objetivo de proporcionar formulaciones farmacéuticas tópicas, orales, sistémicas y parenterales adecuadas para su uso en los métodos novedosos de tratamiento de la presente invención. Las composiciones que contienen compuestos identificados de acuerdo con esta invención como ingrediente activo pueden administrarse en una amplia diversidad de formas galénicas terapéuticas en vehículos convencionales para su administración. Por ejemplo, los compuestos pueden administrarse en formas galénicas orales tales como comprimidos, cápsulas (incluyendo cada una, formulaciones de liberación programada y liberación mantenida), píldoras, polvos, gránulos, elixires, tinturas, soluciones, suspensiones, jarabes y emulsiones, o por inyección. Asimismo, también pueden administrarse en forma intravenosa (tanto en bolo como en infusión), intraperitoneal, subcutánea, tópica con o sin oclusión o intramuscular, todas usando formas bien conocidas por los expertos en las técnicas farmacéuticas.

Ventajosamente, los compuestos de la presente invención pueden administrarse en una única dosis diaria, o la dosificación diaria total puede administrarse en dosis divididas de dos, tres o cuatro veces al día. Además, los compuestos para la presente invención pueden administrarse en forma intranasal mediante uso tópico de vehículos intranasales adecuados, o mediante vías transdérmicas, usando aquellas formas de parches cutáneos transdérmicos bien conocidas por los expertos en la materia. Para administrarse en forma de un sistema de suministro transdérmico, la administración de la dosificación será, por supuesto, continua en lugar de intermitente en toda la pauta posológica.

Para tratamiento de combinación con más de un agente activo, donde los agentes activos están en formulaciones diferentes de dosificación, los agentes activos pueden administrarse simultáneamente, o cada uno puede administrarse en momentos escalonados por separado.

La pauta posológica que utiliza los compuestos de la presente invención se selecciona de acuerdo con una diversidad de factores que incluyen el tipo, especie, edad, peso, sexo y estado médico del paciente; la gravedad de la afección a tratar; la vía de administración; la función renal, hepática y cardiovascular del paciente; y el compuesto particular del mismo empleado. Un médico habitual puede determinar fácilmente y prescribir la cantidad eficaz del fármaco necesaria para prevenir, contrarrestar o detener el progreso de la afección, la precisión óptima a la hora de conseguir concentraciones de fármaco dentro del intervalo que produzca eficacia sin toxicidad requiere un régimen basado en la cinética de la disponibilidad del fármaco en los sitios diana, esto implica una consideración de la distribución, equilibrio y eliminación de un fármaco.

Aunque la presente invención y sus ventajas se han descrito en detalle, debe entenderse que pueden hacerse diversos cambios, sustituciones y alteraciones en este documento sin alejarse del alcance de la invención que se define en las reivindicaciones adjuntas.

La presente invención se ilustrará adicionalmente en los siguientes ejemplos que se dan con fines de ilustración únicamente y no pretenden limitar la invención de ninguna manera.

## Ejemplos

### Ejemplo 1

#### *Ingeniería genómica múltiple usando sistemas de CRISPR/Cas*

La dilucidación funcional de las variantes y elementos genéticos causales requiere tecnologías de edición genómica precisas. El sistema inmunitario adaptativo de CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas) procariota de tipo II ha demostrado facilitar la escisión de ADN específica de sitio guiada por ARN. Los solicitantes manipularon dos sistemas de CRISPR de tipo II diferentes y demuestran que las nucleasas Cas9 pueden dirigirse por ARN cortos para inducir la escisión precisa en locus genómicos endógenos en células humanas y de ratón. Cas9 también puede convertirse en una enzima melladora para facilitar la reparación dirigida por homología con actividad mutagénica mínima. Finalmente, pueden codificarse múltiples secuencias de guía en una única matriz de CRISPR para posibilitar la edición simultánea de varios sitios dentro del genoma de mamífero, lo que demuestra una fácil capacidad de programación y amplia aplicabilidad de la tecnología de CRISPR.

Los sistemas inmunitarios adaptativos de CRISPR procariontes pueden reconstituirse y manipularse para mediar edición genómica múltiple en células de mamífero.

5 Son necesarias tecnologías de dirección genómica precisas y eficaces para posibilitar la manipulación inversa sistemática de variaciones genéticas causales permitiendo la perturbación selectiva de elementos genéticos individuales. Aunque las tecnologías de edición genómica tales como dedos de cinc de diseño (ZF) (M. H. Porteus, D. Baltimore, Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763 (May 2, 2003); J. C. Miller *et al.*, An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25, 778 (julio de 2007); J. D. Sander *et al.*, Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* 8, 67 (enero de 2011) y A. J. Wood *et al.*, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (15 de julio de 2011)), los efectores de tipo activador de la transcripción (TALE) (A. J. Wood *et al.*, Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (15 de julio de 2011); M. Christian *et al.*, Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757 (octubre de 2010); F. Zhang *et al.*, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29, 149 (febrero de 2011); J. C. Miller *et al.*, A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29, 143 (febrero de 2011); D. Reyon *et al.*, FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 30, 460 (mayo de 2012); J. Boch *et al.*, Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509 (11 de diciembre de 2009) y M. J. Moscou, A. J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501 (11 de diciembre de 2009)) y meganucleasas de asentamiento (B. L. Stoddard, Homing endonuclease structure and function. *Quarterly reviews of biophysics* 38, 49 (febrero de 2005)) han empezado a posibilitar modificaciones genómicas dirigidas, sigue existiendo una necesidad de nuevas tecnologías que se puedan cambiar de escala, se puedan producir y sean fáciles de diseñar. Aquí, los solicitantes informan del desarrollo de una nueva clase de herramientas de manipulación genómica de precisión basadas en la nucleasa Cas9 guiada por ARN (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (17 de agosto de 2012); G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109, E2579 (25 de septiembre de 2012) y J. E. Garneau *et al.*, The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67 (4 de noviembre de 2010)) a partir del sistema inmunitario adaptativo de CRISPR procarionte de tipo II (H. Deveau, J. E. Garneau, S. Moineau, CRISPR/Cas system and its role in phage-bacteria interactions. *Annual review of microbiology* 64, 475 (2010); P. Horvath, R. Barrangou, CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167 (8 de enero de 2010); K. S. Makarova *et al.*, Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9, 467 (enero de 2011) y D. Bhaya, M. Davison, R. Barrangou, CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45, 273 (2011)).

El locus de CRISPR de tipo II de *Streptococcus pyogenes* SF370 consiste en cuatro genes, incluyendo la nucleasa Cas9, así como dos ARN no codificantes: ARNtracr y una matriz de preARNcr que contiene secuencias de guía de nucleasa (espaciadores) intercaladas por repeticiones directas idénticas (DR) (figura 5) (E. Deltcheva *et al.*, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602 (31 de marzo de 2011)). Los solicitantes buscan aprovechar este sistema de nucleasa programable por ARN procarionte para introducir roturas bicatenarias (DSB) dirigidas en cromosomas de mamífero mediante expresión heteróloga de los componentes clave. Se ha demostrado previamente que la expresión de ARNtracr, preARNcr, RNasa III de factor hospedador y nucleasa Cas9 son necesarios y suficientes para la escisión de ADN *in vitro* (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 (17 de agosto de 2012) y G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109, E2579 (25 de septiembre de 2012)) y en células procariontes (R. Sapranauskas *et al.*, The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39, 9275 (noviembre de 2011) y A. H. Magadan, M. E. Dupuis, M. Villion, S. Moineau, Cleavage of phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas system. *PLoS One* 7, e40913 (2012)). Los solicitantes optimizaron los codones de Cas9 de *S. pyogenes* (SpCas9) y RNase III (SpRNase III) y adhirieron señales de localización nuclear (NLS) para asegurar la compartimentación nuclear en células de mamífero. La expresión de estas construcciones en células 293FT humanas reveló que se requieren dos NLS para dirigir SpCas9 al núcleo (figura 1A). Para reconstituir los componentes de ARN no codificantes de CRISPR, los solicitantes expresaron un ARNtracr de 89 nucleótidos (nt) (figura 6) bajo el promotor de la ARN polimerasa III U6 (figura 1B). Asimismo, los solicitantes usaron el promotor U6 para dirigir la expresión de un matriz de preARNcr que comprende un único espaciador de guía flanqueado por DR (figura 1B). Los solicitantes diseñaron un espaciador inicial para abordar un sitio de 30 pares de bases (pb) (protoespaciador) en el locus *EMXH1* humano que precede a un NGG, el motivo adyacente a protoespaciador (PAM) necesario (figura 1C y figura 5) (Deveau *et al.*, Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190, 1390 (febrero de 2008) y F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733 (marzo de 2009)).

Para ensayar si la expresión heteróloga del sistema de CRISPR (SpCas9, SpRNasa III, ARNtracr y preARNcr) puede conseguir la escisión dirigida de cromosomas de mamífero, los solicitantes transfectaron células 293FT con diferentes combinaciones de componentes de CRISPR. Como las DSB en ADN de mamífero se reparan parcialmente por la ruta

de unión de extremos no homólogos (NHEJ) que forma indel, los solicitantes usaron el ensayo SURVEYOR (figura 7) para detectar la escisión de la diana endógena (figura 1D y figura 6B). La cotransfección de los cuatro componentes de CRISPR necesarios provocó la escisión eficaz del protoespaciador (figura 1D y figura 6B), que posteriormente se verifica por secuenciación de Sanger (figura 1E). De forma interesante, la SpRNasa III no fue necesaria para la escisión del protoespaciador (figura 1D) y el ARNtracr de 89 nt se procesa en su ausencia (figura 6C). Asimismo, la maduración del preARNcr no requiere RNasa III (figura 1D y figura 8), lo que sugiere que puede haber RNasas de mamífero endógenas que ayuden en la maduración del preARNcr (M. Jinek, J. A. Doudna, A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457**, 405 (22 de enero de 2009); C. D. Malone, G. J. Hannon, Small RNAs as guardians of the genome. *Cell* **136**, 656 (20 de febrero de 2009) y G. Meister, T. Tuschl, Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343 (16 de septiembre de 2004)). Eliminar cualquier ARN restante o componente de Cas9 anulaba la actividad de escisión del genoma del sistema de CRISPR (figura 1D). Estos resultados definen un sistema de tres componentes mínimo para la modificación genómica mediada por CRISPR eficaz en células de mamífero.

A continuación, los solicitantes exploraron la capacidad de generalizar la escisión mediada por CRISPR en células eucariotas por dirección a protoespaciadores adicionales dentro del locus *EMX1* (figura 2A). Para mejorar el cosuministro, los solicitantes diseñaron un vector de expresión para dirigir tanto el preARNcr como SpCas9 (figura 9). En paralelo, los solicitantes adaptaron un diseño híbrido quimérico de ARNcr-ARNtracr (figura 2B, parte superior) recientemente validado *in vitro* (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816 (17 de agosto de 2012)), donde se fusiona un ARNcr maduro a un ARNtracr parcial mediante un tallo-bucle sintético para imitar la doble hélice natural de ARNcr:ARNtracr (figura 2B, parte inferior). Los solicitantes observaron escisión de todas las dianas protoespaciadoras cuando SpCas9 se coexpresa con preARNcr (DR-espaciador-DR) y ARNtracr. Sin embargo, no todos los diseños de ARN quimérico podían facilitar la escisión de sus dianas genómicas (figura 2C, tabla 1). Los solicitantes entonces ensayaron la dirección a locus genómicos adicionales tanto en células humanas como de ratón diseñando preARNcr y ARN quiméricos dirigidos a los locus *PVALB* humano y *Th* de ratón (figura 10). Los solicitantes consiguieron modificación eficaz en las tres dianas *Th* de ratón y una *PVALB* usando el diseño de ARNcr:ARNtracr, demostrando de este modo la amplia aplicabilidad del sistema de CRISPR en la modificación de locus diferentes entre múltiples organismos (tabla 1). Para las mismas dianas protoespaciadoras, las eficacias de escisión de los ARN quiméricos fueron incluso inferiores que las de las dobles hélices de ARNcr:ARNtracr o indetectables. Esto puede deberse a diferencias en la expresión y estabilidad de los ARN, la degradación por la maquinaria de iARN endógena o estructuras secundarias que dan lugar a carga de Cas9 o reconocimiento de diana ineficaz.

La edición genómica eficaz requiere que las nucleasas se dirijan a locus genómicos específicos tanto con precisión como con eficacia elevadas. Para investigar la especificidad de la escisión mediada por CRISPR, los solicitantes analizaron emparejamientos incorrectos de un solo nucleótido entre el espaciador y su diana protoespaciadora de mamífero (figura 3A). Los solicitantes observaron que el emparejamiento incorrecto de una sola base hasta 12 pb 5' del PAM anulaba completamente la escisión genómica por SpCas9, mientras que espaciadores con mutaciones mucho más 5' retenían la actividad contra la diana protoespaciadora (figura 3B). Esto es coherente con los estudios bacterianos e *in vitro* previos de la especificidad de Cas9 (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816 (17 de agosto de 2012) y R. Sapranaukas *et al.*, The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* **39**, 9275 (noviembre de 2011)). Además, CRISPR puede mediar la escisión genómica tan eficazmente como una pareja de nucleasas TALE (TALEN) dirigidas al mismo protoespaciador *EMX1* (figura 3, C y D).

La modificación dirigida de genomas evita de forma ideal mutaciones que surgen del mecanismo de NHEJ propenso a errores. La SpCas9 de tipo silvestre puede mediar DSB específicas de sitio, que pueden repararse mediante NHEJ o reparación dirigida por homología (HDR). Los solicitantes diseñaron una sustitución de aspartato a alanina (D10A) en el dominio RuvC I de SpCas9 para convertir la nucleasa en una ADN nickasa (SpCas9n, figura 4A) (M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816 (17 de agosto de 2012); G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579 (25 de septiembre de 2012) y R. Sapranaukas *et al.*, The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* **39**, 9275 (Nov, 2011)), porque el ADN genómico mellado típicamente se repara impecablemente o a través de HDR de alta fidelidad. SURVEYOR (figura 4B) y la secuenciación de 327 amplicones no detectaron ningún indel inducido por SpCas9n. Sin embargo, hay que apreciar que el ADN mellado puede procesarse en casos infrecuentes mediante un intermedio de DSB y provocar un evento de NHEJ (M. T. Certo *et al.*, Tracking genome engineering outcome at individual DNA breakpoints. *Nat Methods* **8**, 671 (agosto de 2011)). Los solicitantes entonces ensayaron HDR mediada por Cas9 en el mismo locus *EMX1* con un molde de reparación por homología para introducir un par de sitios de restricción cerca del protoespaciador (figura 4C). SpCas9 y SpCas9n catalizaron la integración del molde de reparación en el locus *EMX1* a niveles similares (figura 4D), que los solicitantes verificaron adicionalmente mediante secuenciación de Sanger (figura 4E). Estos resultados demuestran la utilidad de CRISPR para facilitar las inserciones genómicas dirigidas. Dada la especificidad de diana de 14 pb (12 pb de la secuencia fundamental y 2 pb de PAM) (figura 3B) de la SpCas9 de tipo silvestre, el uso de una nickasa puede reducir mutaciones inespecíficas.

Finalmente, la arquitectura natural de los locus de CRISPR con espaciadores ordenados (figura 5) sugiere la posibilidad de manipulación genómica múltiple. Usando una matriz de CRISPR individual que codifica un par de espaciadores dirigidos a *EMX1* y *PVALB*, los solicitantes detectaron escisión eficaz en ambos locus (figura 4F). Los solicitantes ensayaron además la eliminación dirigida de regiones genómicas más grandes mediante DSB simultáneas usando espaciadores contra dos dianas dentro de *EMX1* espaciado por 119 pb, y observaron una eficacia de eliminación de un 1,6 % (3 de los 182 amplicones; figura 4G), demostrando de este modo que el sistema de CRISPR puede mediar la edición múltiple dentro de un único genoma.

La capacidad de usar ARN para programar la escisión de ADN específica de secuencia define una nueva clase de herramientas de manipulación genómica. Aquí, los solicitantes han demostrado que el sistema de CRISPR de *S. pyogenes* puede reconstituirse de forma heteróloga en células de mamífero para facilitar la edición genómica eficaz; un estudio adjunto ha confirmado independientemente la dirección genómica mediada por CRISPR de alta eficacia en varias líneas celulares humanas (Mali *et al.*). Sin embargo, varios aspectos del sistema de CRISPR pueden mejorarse adicionalmente para aumentar su eficacia y versatilidad. El requisito de una NGG PAM restringe el espacio diana de CRISPR de *S. pyogenes* a cada 8 pb de promedio en el genoma humano (figura 11), sin tener en cuenta las restricciones potenciales aportadas por la estructura secundaria de ARNcr o la accesibilidad genómica debido a los estados de metilación de la cromatina y el ADN. Algunas de estas restricciones pueden superarse explotando la familia de enzimas de Cas9 y sus diferentes requisitos de PAM (H. Deveau *et al.*, Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* **190**, 1390 (febrero de 2008) y F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733 (marzo de 2009)) a través de la diversidad microbiana (K. S. Makarova *et al.*, Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467 (junio de 2011)). De hecho, otros locus de CRISPR probablemente son trasplantables a células de mamífero; por ejemplo, el CRISPR1 de *Streptococcus thermophilus* LMD-9 también puede mediar la escisión genómica de mamífero (figura 12). Finalmente, la capacidad de realizar edición genómica múltiple en células de mamífero posibilita aplicaciones potentes en ciencia básica, biotecnología y medicina (P. A. Carr, G. M. Church, Genome engineering. *Nat Biotechnol* **27**, 1151 (diciembre de 2009)).

## Ejemplo 2

### *Manipulación genómica múltiple usando sistema de CRISPR/Cas: Material complementario*

**Cultivo celular y transfección.** La línea celular 293FT de riñón embrionario humano (HEK) (Life Technologies) se mantuvo en medio Eagle modificado por Dulbecco (DMEM) complementado con suero bovino fetal al 10 % (HyClone), GlutaMAX 2 mM (Life Technologies), 100 U/ml de penicilina y 100 µg/ml de estreptomycin a 37 °C con incubación en CO<sub>2</sub> al 5 %. La línea celular de ratón neuro2A (N2A) (ATCC) se mantuvo con DMEM complementado con suero bovino fetal al 5 % (HyClone), GlutaMAX 2 mM (Life Technologies), 100 U/ml de penicilina y 100 µg/ml de estreptomycin a 37 °C con CO<sub>2</sub> al 5 %.

Las células 293FT o N2A se sembraron en placas de 24 pocillos (Corning) un día antes de la transfección a una densidad de 200 000 células por pocillo. Las células se transfectaron usando Lipofectamine 2000 (Life Technologies) siguiendo el protocolo recomendado por el fabricante. Para cada pocillo de una placa de 24 pocillos, se usó un total de 800 ng de plásmidos.

**Ensayo surveyor y análisis de secuenciación para modificación genómica.** Las células 293FT o N2A se transfectaron con ADN plasmídico como se describe anteriormente. Las células se incubaron a 37 °C durante 72 horas después de la transfección antes de la extracción del ADN genómico. El ADN genómico se extrajo usando el kit de extracción de ADN QuickExtract (Epicentre) siguiendo el protocolo del fabricante. En resumen, las células se resuspendieron en solución QuickExtract y se incubaron a 65 °C durante 15 minutos y 98 °C durante 10 minutos.

La región genómica que rodea el sitio diana de CRISPR para cada gen se amplificó por PCR y los productos se purificaron usando la columna de centrifugación QiaQuick Spin (Qiagen) siguiendo el protocolo del fabricante. Se mezcló un total de 400 ng de los productos de PCR purificados con 2 µl de tampón de PCR de polimerasa Taq 10X (Enzymatics) y agua ultrapura hasta un volumen final de 20 µl, y se sometieron a un proceso de rehibridación para posibilitar la formación de heterodúplex: 95 °C durante 10 min, 95 °C a 85 °C en rampa a - 2 °C/s, 85 °C a 25 °C a - 0,25 C/s y mantenimiento a 25 °C durante 1 minuto. Después de la rehibridación, los productos se trataron con nucleasa SURVEYOR y potenciador S SURVEYOR (Transgenomics) siguiendo el protocolo recomendado por el fabricante, y se analizaron en 4-20 geles de poliacrilamida Novex TBE (Life Technologies). Los geles se tiñeron con tinte de ADN SYBR Gold (Life Technologies) durante 30 minutos y se tomaron imágenes con un sistema de imágenes de gel Gel Doc (Biorad). La cuantificación se basó en las intensidades de banda relativas.

**Ensayo de polimorfismos de longitud de fragmentos de restricción para la detección de recombinación homóloga.** Se transfectaron células HEK 293FT y N2A con ADN plasmídico y se incubaron a 37 °C durante 72 horas antes de la extracción del ADN genómico como se describe anteriormente. La región genómica diana se amplificó por PCR usando cebadores fuera de los brazos de homología del molde de recombinación homóloga (HR). Los productos de PCR se separaron en un gel de agarosa al 1 % y se extrajeron con el kit de extracción del gel MinElute (Qiagen). Los productos

purificados se digirieron con HindIII (Fermentas) y se analizaron en un gel de poli(acrilamida) Novex TBE al 6 % (Life Technologies).

5 *Extracción y purificación de ARN.* Las células HEK 293FT se mantuvieron y transfectaron como se indica previamente. Las células se recogieron por tratamiento con tripsina seguido de lavado en solución salina tamponada con fosfato (PBS). Se extrajo el ARN celular total con reactivo TRI (Sigma) siguiendo el protocolo del fabricante. El ARN total extraído se cuantificó usando Nanodrop (Thermo Scientific) y se normalizó a la misma concentración.

10 *Análisis de transferencia de Northern de la expresión de ARNcr y ARNtracr en células de mamífero.* Los ARN se mezclaron con volúmenes iguales de tampón de carga 2X (Ambion), se calentaron hasta 95 °C durante 5 min, se enfriaron en hielo durante 1 min y después se cargaron en geles de poli(acrilamida) desnaturizantes al 8 % (SequaGel, National Diagnostics) después de preprocesar el gel durante al menos 30 minutos. Las muestras se sometieron a electroforesis durante 1,5 horas a un límite de 40 W. Después de ello, el ARN se transfirió a membrana Hybond N+ (GE Healthcare) a 300 mA en un aparato de transferencia semiseca (Bio-rad) a temperatura ambiente durante 1,5 horas. El ARN se reticuló con la membrana usando el botón de autorreticulación en el reticulador por UV de Stratagene Stratalinker (Stratagene). La membrana se prehibridó en tampón de hibridación ULTRAhyb-Oligo (Ambion) durante 30 minutos con rotación a 42 °C y después se añadieron las sondas y se hibridaron durante una noche. Las sondas se ordenaron a partir de IDT y se marcaron con [gamma 32P] ATP (Perkin Elmer) con polinucleótido cinasa T4 (New England Biolabs). La membrana se lavó una vez con SSC 2x precalentado (42 °C), SDS al 0,5 % durante 20 1 min seguido de dos lavados de 30 minutos a 42 °C. La membrana se expuso a un filtro de fósforo durante una hora o durante una noche a temperatura ambiente y después se exploró con phosphorimager (Typhoon).

25 Tabla 1. Secuencias protoespaciadoras y eficacias de modificación de dianas genómicas de mamífero. Las dianas protoespaciadoras se diseñaron basándose en locus de CRISPR de tipo II de *Streptococcus pyogenes* y CRISPR1 de *Streptococcus thermophilus* con sus PAM necesarios frente a tres genes diferentes en genomas humano y de ratón. Las células se transfectaron con Cas9 y preARNcr/ARNtracr o ARN quimérico. Las células se analizaron 72 horas después de la transfección. El porcentaje de indel se calcula basándose en los resultados del ensayo SURVEYOR a partir de las líneas celulares indicadas, N=3 para todas las dianas protoespaciadoras, los errores son E.T.M. N.D., no detectable usando el ensayo SURVEYOR; N.T., no ensayado en ese estudio.

30

Cas9	especie de la diana	gen	ID del protoespaciador	secuencia del protoespaciador (5' a 3')	PAM	hebra	línea celular ensayada	% deindel (preARNcr + ARNtracr)	% deindel (ARN quimérico)
CRISPR de tipo II de <i>S. pyogenes</i> SF370	<i>Homo sapiens</i>	EMX1	1	GGAAGGGCCCTGAGATCCGAGCAGAAGAAGAA	GGG	+	293FT	20 ± 1,8	6,7 ± 0,62
		EMX1	2	CATTGGAGGTGACATCGATGTCCTCCCCAT	TGG	-	293FT	2,1 ± 0,31	N.D.
		EMX1	3	GGACATCGATGTCACCTCCAATGACTAGGG	TGG	+	293FT	14 ± 1,1	N.D.
		EMX1	4	CATCGATGCTCCCCATTGGCCTGCTTCG	TGG	-	293FT	11 ± 1,7	N.D.
		EMX1	5	TTCGTGGCAATGCGCCACCCTGGTTGATGTA	TGG	-	293FT	4,3 ± 0,46	2,1 ± 0,51
		EMX1	6	TCGTGGCAATGCGCCACCCTGGTTGATGTA	GGG	-	293FT	4,0 ± 0,66	0,41 ± 0,25
		EMX1	7	TCCAGCTTCTGCCGTTTGTACTTTGTCCCTC	CGG	-	293FT	1,5 ± 0,12	N.D.
		EMX1	8	GGAGGGAGGGCCACAGATGAGAACTCAGG	AGG	-	293FT	7,8 ± 0,83	2,3 ± 1,2
<i>Homo sapiens</i>	PVALB	9	AGGGCCCGAGATTGGGTGTTCCAGGGCAGAG	AGG	+	293FT	21 ± 2,6	6,5 ± 0,32	
	PVALB	10	ATGCAGGGAGGTGGCGAGAGGGGCCGAGAT	TGG	+	293FT	N.D.	N.D.	
	PVALB	11	GGTGGCGAGAGGGGCCGAGATTGGGTGTTCC	AGG	+	293FT	N.D.	N.D.	
	Th	12	CAAGCACGTAGTGCCATTAGCTAAATGTCAT	AGG	-	Neuro2A	27 ± 4,3	4,1 ± 2,2	
<i>Mus musculus</i>	Th	13	AATGCATAGGTACCACCCACAGGTGCCAG	GGG	-	Neuro2A	4,8 ± 1,2	N.D.	
	Th	14	ACACACATGGGAAAGCCCTCTGGGCCAGGAA	AGG	+	Neuro2A	11,3 ± 1,3	N.D.	
	EMX1	15	GGAGGAGGTAGTATACAGAACACACAGAGAA	TAGAAI	-	293FT	14 ± 0,88	N.T.	
CRISPR1 de <i>S. thermophilus</i> LMD-9	<i>Homo sapiens</i>	EMX1	16	AGAATGTAGAGGAGTCACAGAAGCTCAGCA	CTAGAAA	-	293FT	7,8 ± 0,77	N.T.

Tabla 2. Secuencias para cebadores y sondas usados para ensayo el SURVEYOR, ensayo de RFLP, secuenciación genómica y transferencia de Northern.

Nombre del cebador	Ensayo	Diana genómica	Secuencia del cebador
Sp-EMX1-F	ensayo SURVEYOR, secuenciación	EMX1	AAAACCACCCTTCTCTCTGGC
Sp-EMX1-R	ensayo SURVEYOR, secuenciación	EMX1	GGAGATTGGAGACACGGAGAG
Sp-PVALB-F	ensayo SURVEYOR, secuenciación	PVALB	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	ensayo SURVEYOR, secuenciación	PVALB	GGCAGCAAACCTCCTTGTCT
Sp-Th-F	ensayo SURVEYOR, secuenciación	Th	GTGCTTTGCAGAGGCCTACC
Sp-Th-R	ensayo SURVEYOR, secuenciación	Th	CCTGGAGCGCATGCAGTAGT
St-EMX1-F	ensayo SURVEYOR, secuenciación	EMX1	ACCTTCTGTGTTTCCACCATTC
St-EMX1-R	ensayo SURVEYOR, secuenciación	EMX1	TTGGGGAGTGCACAGACTTC
Sp-EMX1-RFLP-F	RFLP, secuenciación	EMX1	GGCTCCCTGGGTTCAAAGTA
Sp-EMX1-RFLP-R	RFLP, secuenciación	EMX1	AGAGGGGTCTGGATGTCGTAA
Pb_EMX1_sp1	sonda de transferencia de Northern	No aplicable	TAGCTCTAAAACCTTCTTCTCTGCTCGGAC
Pb_tracrRNA	sonda de transferencia de Northern	No aplicable	CTAGCCTTATTTAACTTGCTATGCTGTTT

5 Secuencias complementarias

> U6-ARNtracr corto (*Streptococcus pyogenes* SF370)

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT  
 GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
 TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTT  
 AAAATGGACTATCATATGCTTACCCTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT  
 ATATCTTGTGGAAAGGACGAAACACCGGAACATTCAAACAGCATAGCAAGTTAA  
 AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

10

> U6-ARNtracr largo (*Streptococcus pyogenes* SF370)

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT  
 GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
 TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTT  
 AAAATGGACTATCATATGCTTACCCTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT  
 ATATCTTGTGGAAAGGACGAAACACCGGTAGTATTAAGTATTGTTTTATGGCTGATA  
 AATTTCTTTGAATTTCTCCTTGATTATTTGTTATAAAAAGTTATAAAATAATCTTGTTG  
 GAACCATTCAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA  
 AAAAGTGGCACCGAGTCGGTGCTTTTTTTT

15

> U6-DR estructura de BbsI-DR (*Streptococcus pyogenes* SF370)

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT  
 GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
 TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTT  
 AAAATGGACTATCATATGCTTACCCTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT  
 ATATCTTGTGGAAAGGACGAAACACCGGGTTTTAGAGCTATGCTGTTTTGAATGGTC  
 CCAAACGGGTCTTCGAGAAGACGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAA  
 AC

20

> U6-ARN quimérico-estructura de BbsI (*Streptococcus pyogenes* SF370)

ES 2 757 623 T3

GAGGGCCTATTTCCCATGATTCCCTTCATATTTGCATATACGATACAAGGCT  
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAA  
TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTT  
AAAATGGACTATCATATGCTTACCCTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT  
ATATCTTGTGGAAAGGACGAAACACCCGGGTCTTCGAGAAGACCTGTTTTAGAGCTA  
GAAATAGCAAGTTAAAATAAGGCTAGTCCG

> 3xFLAG-NLS-SpCas9-NLS

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
GGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTC  
TGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATTCAAGG  
TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTG  
TTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAA  
GATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG  
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA  
AGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGG  
CCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGC  
ACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTC  
CGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAA  
GCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAA  
CGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGAC  
GGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCCGGC  
AACCTGATTGCCCTGAGCCTGGGCCTGACCCCAACTTCAAGAGCAACTTCGACCTG  
GCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAA  
CCTGCTGGCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCT  
GTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGG  
CCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACC  
CTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC  
GACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG  
AGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTG  
CTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGG  
CAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG  
AAGATTTTACCCATTCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACC  
TTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGG  
ATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTCGAGGAAGTGGTGGAA  
CAAGGGCGCTTCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC  
TGCCCAACGAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT  
ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAAGCCCGCCTTC

5

ES 2 757 623 T3

CTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAA  
AGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT  
CCGTGGAAATCTCCGGCGTGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC  
GATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGA  
CATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA  
GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAAGTGATGAAGCAGCTGA  
AGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC  
CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGC  
CAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGA  
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGG  
CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
GCGGGATATGTACGTGGACCAGGAAGTGGACATCAACCGGCTGTCCGACTACGATG  
TGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGC  
TGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
CGTGAAGAAGATGAAGAATACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC  
AGAGAAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGGCCTGAGCGAACTGGAT  
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
GGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGA  
TCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
GATTTCCAGTTTTACAAAGTGC GCGAGATCAACAATACTACCACCACGCCACGACGCC  
TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAAG  
CGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
CCACCGTGCGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAA

ES 2 757 623 T3

GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAG  
AAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTT  
CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
CAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCC  
CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACG  
AGATCATCGAGCAGATCAGCGAGTTCCTCAAGAGAGTGATCCTGGCCGACGCTAAT  
CTGGACAAAAGTGCTGTCCGCCTACAACAAGCACCCGGGATAAGCCCATCAGAGAGCA  
GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCCTGCCGCCTT  
CAAGTACTTTGACACCACCATCGACCCGGAAGAGGTACACCAGCACCAAAGAGGTGC  
TGGACGCCACCCTGATCCACCAGAGCATCACCCGGCCTGTACGAGACACGGATCGAC  
CTGTCTCAGCTGGGAGGGCGACAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
AAGAAAAAGAAA

> SpRNasa3-mCherry-NLS

ATGAAGCAGCTGGAGGAGTTACTTTCTACCTCTTTTCGACATCCAGTTTAAT  
GACCTGACCCTGCTGGAAACCGCCTTCACTCACACCTCCTACGCGAATGAGCACCCGC  
CTACTGAATGTGAGCCACAACGAGCGCCTGGAGTTTCTGGGGGATGCTGTCTTACAG  
CTGATCATCTCTGAATATCTGTTTGCCAAATACCCTAAGAAAACCGAAGGGGACATG  
TCAAAGCTGCGCTCCATGATAGTCAGGGAAGAGAGCCTGGCGGGCTTTAGTCGTTTT  
TGCTCATTTCGACGCTTATATCAAGCTGGGAAAAGGCGAAGAGAAGTCCGGCGGCAG  
GAGGCGCGATAACAATTCTGGGCGATCTCTTTGAAGCGTTTCTGGGCGCACTTCTACT  
GGACAAAGGGATCGACGCAGTCCGCCGCTTTCTGAAACAAGTGATGATCCCTCAGG  
TCGAAAAGGGAAACTTCGAGAGAGTGAAGGACTATAAAACATGTTTGCAGGAATTT  
CTCCAGACCAAGGGAGATGTAGCAATAGATTATCAGGTAATAAGTGAGAAAGGACC  
AGCTCACGCCAAACAATTGGAAGTTAGCATCGTTGTTAATGGCGCAGTGTTGTCGAA  
GGGCTTGGGTAATCAAAAAAACTGGCCGAGCAGGACGCTGCTAAAAACGCCCTCG  
CTCAGCTCAGCGAGGTAGGATCCGTGAGCAAGGGCGAGGAGGATAACATGGCCATC  
ATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGA  
5 GTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCA

ES 2 757 623 T3

AGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTC  
AGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACT  
TGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGAC  
GGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTA  
CAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGA  
AGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTG  
AAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTG  
AGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCCGGCGCCTACAAC  
GTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACA  
GTACGAACGCGCCGAGGGCCGCCACTCCACCGCGGCATGGACGAGCTGTACAAGA  
AGCGTCTGTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA

> 3xFLAG-NLS-SpCas9n-NLS (la mutación de nickasa D10A está subrayada)

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCAC  
GGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTC  
TGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATTCAAGG  
TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTG  
TTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAA  
GATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG  
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCTGGTGGAA  
AGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGG  
CCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGC  
ACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAGTTC  
CGGGGCCACTTCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAA  
GCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAA  
CGCCAGCGGCGTGGACGCCAAGGCCATCTGTCTGCCAGACTGAGCAAGAGCAGAC  
GGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGC  
AACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTG  
GCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAA  
CCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCAAGAACCT  
5 GTCCGACGCCATCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGG

ES 2 757 623 T3

CCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACC  
CTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC  
GACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG  
AGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCCGAGGAACTG  
CTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGG  
CAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG  
AAGATTTTACCCATTCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACC  
TTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGG  
ATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACCTCGAGGAAAGTGGTGGGA  
CAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC  
TGCCCAACGAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT  
ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGGAATGAGAAAGCCCGCCTTC  
CTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAA  
AGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT  
CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC  
GATCTGCTGAAAATTATCAAGGACAAGGACTTCTGGACAATGAGGAAAACGAGGA  
CATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA  
GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGA  
AGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC  
CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGC  
CAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGGA  
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGG  
CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATG  
TGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGC  
TGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
CGTGAAGAAGATGAAGAACTACTGGCAGGCTGCTGAACGCCAAGCTGATTACCC

ES 2 757 623 T3

AGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGCCTGAGCGAACTGGAT  
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
GGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGA  
TCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
GATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACCTACCACCACGCCCACGACGCC  
TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAG  
CGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
CCACCGTGCGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCAAGAGGAACAGCGATAA  
GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGAAAAAGGGCAAGTCCAAG  
AAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTT  
CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
CAAATATGTGAACTTCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCC  
CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACG  
AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT  
CTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCA  
GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT  
CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGC  
TGGACGCCACCCTGATCCACCAGAGCATACCGGCCTGTACGAGACACGGATCGAC  
CTGTCTCAGCTGGGAGGCGACAAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
AAGAAAAAGAAA

> hEMX1-molde de HR-HindIII-NheI

GAATGCTGCCCTCAGACCCGCTTCCTCCCTGTCCTTGTCTGTCCAAGGAGA  
ATGAGGTCTCACTGGTGGATTTCCGACTACCCTGAGGAGCTGGCACCTGAGGGACA  
5 AGGCCCCCACCTGCCAGCTCCAGCCTCTGATGAGGGGTGGGAGAGAGCTACATG

AGGTTGCTAAGAAAGCCTCCCCTGAAGGAGACCACACAGTGTGTGAGGTTGGAGTC  
 TCTAGCAGCGGGTTCTGTGCCCCAGGGATAGTCTGGCTGTCCAGGCACTGCTCTTG  
 ATATAAACACCACCTCCTAGTTATGAAACCATGCCATTCTGCCTCTCTGTATGGAA  
 AAGAGCATGGGGCTGGCCCGTGGGGTGGTGTCCACTT TAGGCCCTGTGGGAGATCA  
 TGGGAACCCACGCAGTGGGTCATAGGCTCTCTCATTTACTACTCACATCCACTCTGT  
 GAAGAAGCGATTATGATCTCTCCTCTAGAACTCGTAGAGTCCCATGTCTGCCGGCT  
 TCCAGAGCCTGCACTCCTCCACCTTGGCTTGGCTTTGCTGGGGCTAGAGGAGCTAGG  
 ATGCACAGCAGCTCTGTGACCCTTTGTTTGAGAGGAACAGGAAAACCACCCTTCTCT  
 CTGGCCCACTGTGTCTCTTCCCTGCCCTGCCATCCCCTTCTGTGAATGTTAGACCCAT  
 GGGAGCAGCTGGTCAGAGGGGACCCCGCCCTGGGGCCCTAACCCTATGTAGCCTC  
 AGTCTTCCCATCAGGCTCTCAGCTCAGCCTGAGTGTTGAGGCCCCAGTGGCTGCTCT  
 GGGGGCCTCCTGAGTTTCTCATCTGTGCCCTCCCTCCCTGGCCAGGTGAAGGTGT  
 GGTTCAGAACCGGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCC  
 TGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA  
 CGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACaagettgtagcGGTGG  
 GCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGGCCCTGCGT  
 GGGCCCAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTTGGGGAGGCCTGGAGTC  
 ATGGCCCCACAGGGCTTGAAGCCCCGGGGCCGCCATTGACAGAGGGACAAGCAATGG  
 GCTGGCTGAGGCCTGGGACCACTTGGCCTTCTCCTCGGAGAGCCTGCCTGCCTGGGC  
 GGGCCCCGCCGCCACCGCAGCCTCCCAGCTGCTCTCCGTGTCTCCAATCTCCCTTTTG  
 TTTTGATGCATTTCTGTTTTAATTTATTTTCCAGGCACCACTGTAGTTTAGTGATCCCC  
 AGTGTCCCCCTTCCCTATGGGAATAATAAAAGTCTCTCTCTTAATGACACGGGCATC  
 CAGCTCCAGCCCCAGAGCCTGGGGTGGTAGATTCCGGCTCTGAGGGCCAGTGGGGG  
 CTGGTAGAGCAAACGCGTTCAGGGCCTGGGAGCCTGGGGTGGGGTACTGGTGGAGG  
 GGGTCAAGGGTAATTCATTAACCTCTCTTTTGTGGGGGACCCTGGTCTCTACCTC  
 CAGCTCCACAGCAGGAGAAACAGGCTAGACATAGGGAAGGGCCATCCTGTATCTTG  
 AGGGAGGACAGGCCCAGGTCTTTCTTAACGTATTGAGAGGTGGGAATCAGGCCAG  
 GTAGTTCAATGGGAGAGGGAGAGTGCTTCCCTCTGCCTAGAGACTCTGGTGGCTTCT  
 CCAGTTGAGGAGAAACCAGAGGAAAGGGGAGGATTGGGGTCTGGGGGAGGGAACA  
 CCATTACAAAGGCTGACGGTTCAGTCCGAAGTCGTGGGCCACCAGGATGCTCA  
 CCTGTCCTTGGAGAACCGCTGGGCAGGTTGAGACTGCAGAGACAGGGCTTAAGGCT  
 GAGCCTGCAACCAGTCCCCAGTGA CT CAGGGCCTCCTCAGCCCAAGAAAGAGCAAC  
 GTGCCAGGGCCCCGCTGAGCTCTTGTGTTACCTG

5 > NLS-StCsn1-NLS

ES 2 757 623 T3

ATGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAA  
AAGTCCGACCTGGTACTTGGACTGGATATTGGTATCGGTTCCGGTGGGAGTCGGAATC  
CTCAACAAGGTCACGGGGGAGATCATTACAAGAACTCGCGGATCTTCCCCGCAGC  
TCAGGCTGAGAACAACCTTGGTGCGGAGAACG

AATAGGCAGGGCAGGCGACTGGCGAGGAGGAAGAAACACAGGAGAGTC  
CGATTGAACCGGCTGTTTCGAGGAGTCCGGTTTGATCACCGACTTTACGAAAATCTCG  
ATTAACCTTAATCCCTATCAGCTTCGGGTGAAAGGCCTGACAGACGAACTTTCGAAT  
GAGGAACTTTTCATCGCGCTGAAAAACATGGTCAAGCACAGAGGGATTTTCTACCTC  
GATGACGCCTCGGATGACGGAAATTCCTCAGTAGGAGATTATGCACAGATCGTGAA  
AGAGAACTCAAAGCAACTGGAAACAAAGACACCGGGGCAGATCCAACCTTGAAAGA  
TACCAGACATACGGACAGCTCAGAGGAGATTTTACGGTGGAGAAGGACGGTAAAAA  
GCACAGACTCATTAACGTATTTCCACGTCGGCGTACAGATCCGAAGCGCTCCGCAT  
CCTTCAGACTCAACAGGAGTTCAACCCGCAAATTACTGATGAGTTCATCAACCGCTA  
TTTGAAATCTTGACCGGAAAGCGCAAGTATTATCATGGGCCGGGTAATGAGAAAT  
CCAGAACAGATTACGGCCGATACAGAACTTCGGGGGAAACCTTGGATAACATCTTT  
GGTATTTTGATTGGAAAGTGCACCTTTTACCCGGACGAGTTTCGAGCGGCCAAGGCG  
TCATACACAGCACAAAGAGTTTAATCTCTTGAATGATTTGAACAACCTTGACGGTCCCC  
ACGGAGACAAAGAAGCTCTCAAAGAGCAAAAGAACCAAATCATCAACTACGTCA  
AGAACGAGAAGGCTATGGGGCCAGCGAAGCTGTTCAAGTATATCGCTAAACTTCTC  
AGCTGTGATGTGGCGGACATCAAAGGGTACCGAATCGACAAGTCGGGAAAAGCGGA  
AATTCACACGTTTGAAGCATATCGAAAGATGAAAACGTTGGAAACACTGGACATTG  
AGCAGATGGACCGGGAAACGCTCGACAAACTGGCATACGTGCTCACGTTGAATACT  
GAACGAGAGGGAATCCAAGAGGCCCTTGAACATGAGTTCGCCGATGGATCGTTCAG  
CCAGAAGCAGGTCGACGAACTTGTGCAATTCGCAAGGCGAATAGCTCCATCTTCG  
GGAAGGGATGGCACAACCTTTTCGGTCAAACCTCATGATGGAGTTGATCCAGAACCTT  
ATGAGACTTCGGAGGAGCAAATGACGATCTTGACGCGCTTGGGGAAACAGAAAACG  
ACAAGCTCATCGAACAACAACTAAGTACATTGATGAGAAATTGCTGACGGAAAGAAAT  
CTATAATCCGGTAGTAGCGAAATCGGTAAGACAAGCGATCAAAATCGTGAACGCGG

ES 2 757 623 T3

CGATCAAGGAATATGGTGACTTTGATAACATCGTAATTGAAATGGCTAGAGAGACG  
AACGAAGATGACGAGAAAAAGGCAATCCAGAAGATCCAGAAGGCCAACAAAGGATG  
AAAAAGATGCAGCGATGCTTAAAGCGGCCAACCAATACAATGGAAAGGCGGAGCT  
GCCCCATTCAGTGTTCACGGTCATAAACAGTTGGCGACCAAGATCCGACTCTGGCA  
TCAGCAGGGTGAGCGGTGTCTCTACACCGGAAAGACTATCTCCATCCATGACTTGAT  
TAACAATTCGAACCAGTTTGAAGTGGATCATATTCTGCCCTGTCAATCACCTTTGA  
CGACTCGCTTGCGAACAAGGTGCTCGTGTACGCAACGGCAAATCAGGAGAAAGGCC  
AGCGGACTCCGTATCAGGCGCTCGACTCAATGGACGATGCGTGGTCATTCCGGGAG  
CTGAAGGCGTTTCGTACGCGAGAGCAAGACACTGAGCAACAAAAAGAAAGAGTATCT  
GCTGACAGAGGAGGACATCTCGAAATTCGATGTCAGGAAGAAGTTCATCGAGCGGA  
ATCTTGTCGACACTCGCTACGCTTCCAGAGTAGTACTGAACGCGCTCCAGGAACACT  
TTAGAGCGCACAAAATTGACACGAAGGTGTCAGTGGTGAGAGGGCAGTTCACATCC  
CAACTCCGCCGACATTGGGGCATCGAAAAGACGCGGGACACATATCACCATCATGC  
GGTGGACGCGCTGATTATTGCCGCTTCGTCCCAGTTGAATCTCTGGAAAAAGCAGAA  
GAACACGCTGGTGTCTGATTTCGGAGGATCAGCTTTTGGACATCGAAACCGGGGAGC  
TGATTTCCGACGATGAATACAAAGAATCGGTGTTTAAAGGCACCATATCAGCATTTCG  
TGGACACGCTGAAGAGCAAAGAGTTTGAAGGACAGCATCCTCTTTTCGTACCAAGTG  
GACTCGAAGTTTAAATCGCAAGATTTACAGACGCCACAATCTACGCGACGAGGCAGGC  
GAAGGTGGGCAAAGATAAAGCAGATGAAACCTACGTCTTTGGTAAAATCAAGGACA  
TCTACACTCAGGACGGGTACGATGCGTTCATGAAAATCTACAAGAAGGATAAGTCG  
AAGTTTCTCATGTACCGCCACGATCCACAGACTTTCGAAAAAGTCATTGAGCCTATT  
TTGGAGAACTACCCTAACAAGCAAATCAACGAGAAAGGGAAAGAAGTCCCGTGCAA  
CCCCTTTCTGAAGTACAAGGAAGAGCACGGTTATATCCGCAAATACTCGAAGAAAG  
GAAATGGGCCTGAGATTAAGTCGCTTAAGTATTACGACTCAAAGTTGGGTAACCAC  
ATCGACATTACCCCGAAAGACTCCAACAACAAAGTCGTGTTGCAGTCCGTCTCGCCC  
TGGCGAGCAGATGTGTATTTAATAAGACGACCGGCAAATATGAGATCCTTGGACTC  
AAATACGCAGACCTTCAATTCGAAAAGGGGACGGGCACTTATAAGATTTACAAGA  
GAAGTACAACGACATCAAGAAAAAGGAAGGGGTCGATTCAGATTCGGAGTTCAAAT  
TCACCCTCTACAAAAACGACCTCCTGCTTGTGAAGGACACAGAAACGAAGGAGCAG  
CAGCTCTTTCGGTTCCTCTCACGCACGATGCCCAAACAAAAACATTACGTGAACTT  
AAACCTTACGATAAGCAAAAAGTTTGAAGGGGGAGAGGCACTGATCAAAGTATTGGG  
TAACGTAGCCAATAGCGGACAGTGTAAGAAAGGGCTGGGAAAGTCCAATATCTCGA  
TCTATAAAGTACGAACAGATGTATTGGGAAACCAGCATATCATCAAAAATGAGGGG  
GATAAACCCAAACTCGATTTCAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCT  
AAGAAAAAGAAATAA

5 > U6-St\_ARNtracr (7-97)

ES 2 757 623 T3

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT  
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAA  
TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAATTATGTTTT  
AAAATGGACTATCATATGCTTACCCTAACTTGAAAAGTATTTTCGATTTCTTGGCTTTAT  
ATATCTTGTGGAAAGGACGAAACACCGTTACTTAAATCTTGCAGAAGCTACAAAGA  
TAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTCGTTATT  
TAA

> EMX1\_TALEN\_izquierda

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
GGAGTCCCAGCAGCCGTAGATTTGAGAACTTTGGGATATTCACAGCAGCAGCAGGA  
AAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGG  
GTCATGGGTTTACACATGCCACATCGTAGCCTTGTCGCAGCACCCCTGCAGCCCTTG  
GCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGCGTTGCCGGAAGCCACACAT  
GAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCAGCGCTTGAGGCCCT  
GTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGT  
TGCTGAAGATCGCGAAGCGGGGAGGAGTCACGGCGGTTCGAGGCGGTGCACGCGTGG  
CGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCCCAGAGCAGGTCGTGGCAAT  
TGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTG  
TGCTGTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAAC  
CACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCA  
AGCGCACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAA  
AGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGG  
TTGACCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACGGGGGAAAGCAGGCACT  
CGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAG  
5 AGCAGGTCGTGGCAATTGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTC

ES 2 757 623 T3

CAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGT  
GGCAATTGCGAGCAACATCGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGC  
TGCCTGTGCTGTGCCAAGCGCACGGACTTACACCCGAACAAGTCGTGGCAATTGCG  
AGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCT  
GTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACG  
GGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCG  
CACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCA  
GGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGGTTGA  
CCCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCAGGCACTCGAA  
ACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCAGAGCA  
GGTCGTGGCAATTGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTCCAGA  
GGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGTGGCA  
ATTGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCC  
TGTGCTGTGCCAAGCGCACGGACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCAA  
CAACGGGGGCAGACCCGCACTGGAGTCAATCGTGGCCAGCTTTCGAGGCCGGACC  
CCGCGCTGGCCGCACTACTAATGATCATCTTGTAGCGCTGGCCTGCCTCGGCGGAC  
GACCCGCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTAAG  
CGGACCAACAGAAGGATCCCGAGAGGACATCACATCGAGTGGCAGGTTCCCAACT  
CGTGAAGAGTGAACCTGAGGAGAAAAAGTCGGAGCTGCGGCACAAATTGAAATACG  
TACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGACTCAAGACAGA  
ATCCTTGAGATGAAGGTAATGGAGTTCTTTATGAAGGTTTATGGATACCGAGGGAAG  
CATCTCGGTGGATCACGAAAACCCGACGGAGCAATCTATACGGTGGGGAGCCCGAT  
TGATTACGGAGTGATCGTCGACACGAAAGCCTACAGCGGTGGGTACAATCTTCCCAT  
CGGGCAGGCAGATGAGATGCAACGTTATGTCGAAGAAAATCAGACCAGGAACAAA  
CACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTAAG  
TTTTGTTTGTCTCTGGGCATTTCAAAGGCAACTATAAGGCCAGCTCACACGGTTG  
AATCACATTACGAACTGCAATGGTGGGTTTTGTCCGTAGAGGAACTGCTCATTGGT  
GGAGAAATGATCAAAGCGGGAACCTGACACTGGAAGAAGTCAGACGCAAGTTTAA  
CAATGGCGAGATCAATTTCCGCTCA

> EMX1\_TALEN\_derecha

ES 2 757 623 T3

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTA  
CAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCAC  
GGAGTCCCAGCAGCCGTAGATTTGAGAACTTTGGGATATTCACAGCAGCAGCAGGA  
AAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGG  
GTCATGGGTTTACACATGCCACATCGTAGCCTTGTGCGCAGCACCCCTGCAGCCCTTG  
GCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGGCGTTGCCGGAAGCCACACAT  
GAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGGCCCT  
GTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGT  
TGCTGAAGATCGCGAAGCGGGGAGGAGTCACGGCGGTTCGAGGCGGTGCACGCGTGG  
CGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCCCAGAGCAGGTCGTGGCAAT  
TGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTG  
TGCTGTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCAAC  
CACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCA  
AGCGCACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACCACGGGGGA  
AAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGG  
GTTGACCCCAGAGCAGGTCGTGGCAATTGCGAGCAACATCGGGGGAAAGCAGGCAC  
TCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCCAG  
AGCAGGTCGTGGCAATTGCGAGCAACCACGGGGGAAAGCAGGCACTCGAAACCGTC  
CAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGT  
GGCAATTGCGAGCCATGACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGC  
TGCCTGTGCTGTGCCAAGCGCACGGACTTACACCCGAACAAGTCGTGGCAATTGCG  
AGCCATGACGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCT  
GTGCCAAGCGCACGGACTTACGCCAGAGCAGGTCGTGGCAATTGCGAGCCATGACG  
GGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCG  
CACGGACTAACCCAGAGCAGGTCGTGGCAATTGCGAGCAACGGAGGGGGAAAGC  
AGGCACTCGAAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGGTTG  
ACCCAGAGCAGGTCGTGGCAATTGCGAGCAACGGAGGGGGAAAGCAGGCACTCG  
AAACCGTCCAGAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGCCTGACCCCAGAG  
CAGGTCGTGGCAATTGCGAGCCATGACGGGGGAAAGCAGGCACTCGAAACCGTCCA  
GAGGTTGCTGCCTGTGCTGTGCCAAGCGCACGGACTGACACCAGAGCAGGTCGTGG  
CAATTGCGAGCAACGGAGGGGGAAAGCAGGCACTCGAAACCGTCCAGAGGTTGCTG

CCTGTGCTGTGCCAAGCGCACGGACTCACGCCTGAGCAGGTAGTGGCTATTGCATCC  
AACGGAGGGGGCAGACCCGCACTGGAGTCAATCGTGGCCAGCTTTCGAGGCCGGA  
CCCCGCGCTGGCCGCACTACTAATGATCATCTTGTAGCGCTGGCCTGCCTCGGCGG  
ACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTA  
AGCGGACCAACAGAAGGATTCCCGAGAGGAATCACATCGAGTGGCAGGTTCCCAAC  
TCGTGAAGAGTGAAC TTGAGGAGAAAAGTCGGAGCTGCGGCACAAATTGAAATAC  
GTACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGACTCAAGACAG  
AATCCTTGAGATGAAGGTAATGGAGTTC TTTATGAAGGTTTATGGATACCGAGGGAA  
GCATCTCGGTGGATCACGAAAACCCGACGGAGCAATCTATACGGTGGGGAGCCCGA  
TTGATTACGGAGTGATCGTCGACACGAAAGCCTACAGCGGTGGGTACAATCTTCCCA  
TCGGGCAGGCAGATGAGATGCAACGTTATGTCGAAGAAAATCAGACCAGGAACAAA  
CACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGACCGAGTTTAAAG  
TTTTTGTGTCTCTGGGCATTTCAAAGGCAACTATAAGGCCAGCTCACACGGTTG  
AATCACATTACGAACTGCAATGGTTCGGTTTTGTCCGTAGAGGAACTGCTCATTGGT  
GGAGAAATGATCAAAGCGGGA ACTCTGACACTGGAAGAAGTCAGACGCAAGTTTAA  
CAATGGCGAGATCAATTTCCGCTCA

**Efectores transcripcionales de Cas9.** Las células HEK 293FT se cotransfectaron con la proteína de fusión de Cas9 mutante y un ARN de guía sintético (ARNgs) usando Lipofectamine 2000 (Life Technologies) 24 horas después de la siembra en una placa de 24 pocillos. Unas 72 horas después de la transfección, se purificó el ARN total (RNeasy Plus, Qiagen). Se retrotranscribió 1 ug de ARN en ADNc (qScript, Quanta BioSciences). La PCR en tiempo real cuantitativa se hizo de acuerdo con el protocolo del fabricante (Life Technologies) y se realizó por triplicado usando ensayos TaqMan para hKlf4 (Hs00358836\_ml), hSox2 (Hs01053049\_s1) y el control endógeno de GAPDH (Hs02758991\_g1).

5 El plásmido de activador de hSpCas9 se clonó en un vector lentivírico bajo la expresión del promotor hEF1a (pLenti-EF1a-Cas9-NLS-VP64). El plásmido represor de hSpCas9 se clonó en el mismo vector (pLenti-EF1a-SID4x-NLS-Cas9-NLS). Las secuencias de guía (20 pb) dirigidas al locus *KLF4* son: GCGCGCTCCACAACTCAC, GCAAAAATAGACAATCAGCA, GAAGGATCTCGGCCAATTTG. Las secuencias espaciadoras para los ARN de guía dirigidos al locus *SOX2* son: GCTGCCGGGTTTTGCATGAA, CCGGGCCCGCAGCAAATTC, GGGGCTGTCAGGGAATAAAT.

>hSpCas9(D10A,H840A)-conector-NLS-VP64

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLF  
 DSGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDDSFFHRLEESFLVEEDKK  
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG  
 DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGE  
 KKNGLFGNLIASLGLTPNFKSNFDLAEDAQLSKDQYDDDLNLLAQIGDQYADLFL  
 AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF  
 DQSKNGYAGYIDGGASQEEFYKFKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPH  
 QIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETI  
 TPWNFEVVVDKGASQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTE  
 GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL  
 GTYHDLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQL  
 KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA  
 QVSGQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQTT  
 QKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL  
 DINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRKGSDNVPSEEVVKKMKNYWRQL  
 LNAKLITQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE  
 NDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIHKYPKL  
 ESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKPLIET  
 NGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKK  
 DWDPKKYGGFDSPTVAYSVLVVAKVEKGGKSKLKSVKELLGITIMERSSEKPNPIDFLE  
 AKGYKEVKKDLIHKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY  
 EKLKGSPEQNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLKVLSAYNKHDKPI  
 REQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQITGLYETRIDLQ  
 LGGDSAGGGGSGGGGSGGGGSGPKKKRKYAAAGSGRADALDDFDLMLGSDALDDF  
 DLMLGSDALDDFDLMLGSDALDDFDLMLIN

>SID4X-NLS-FLAG-conector-hSpCas9(D10A,H840A)-NLS

MSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLE  
 RREREAEHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNIQMLL  
 EAADYLERREREAEHGYASMLPSRSAGPKKKRKYAAADYKDDDDKSRLEPGEKPYKC  
 PECGKSFSQSGALTRHQRTHTRDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNT

5

DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFF  
 HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLAL  
 AHMIKFRGHFLIEGDLNPDNSVDKLFQVLVQTYNQLFEENPINASGVDAKAILSARLSK  
 SRLENLIAQLPGEKKNLFGNLIASLGLTPNFKSNFDLAEDAQLQSKDQYDDDDLNL  
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL  
 VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL  
 LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNS  
 RFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT  
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS  
 VEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT  
 YAHLFDDKVMKQLKRRRYTGWGRLSRKLLINGIRDKQSGKTILDFLKSDFANRNFML  
 IHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHK  
 PENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYY  
 LQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE  
 VVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVA  
 QILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNA  
 VVGTAIIKKYPKLESEFVYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT  
 LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTQVQGGFSKESIL  
 PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKVKSKKLSVKELLGITIM  
 ERSSFENPIDFLEAKGYKEVKKDLIILKPKYSLFELENGRKRMLASAGELQKGNELALP  
 SKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLKD  
 VLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQ  
 SITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

**Ejemplo 3**

5 *Cas9 lentivírica inducible*

Preparación de lentivirus. Después de la clonación de pCasES10 (que contiene una estructura de plásmido de transferencia lentivírica), se sembraron HEK293FT a bajo pase (p=5) en un matraz T-75 hasta un 50 % de confluencia el día antes de la transfección en DMEM con suero bovino fetal al 10 % y sin antibióticos. Después de 20 horas, el medio se cambió a medio OptiMEM (sin suero) y la transfección se hizo 4 horas después. Las células se transfectaron con 10 ug de plásmido de transferencia lentivírica (pCasES10) y los siguientes plásmidos de empaquetado: 5 ug de pMD2.G (seudotipo VSV-g) y 7,5 ug de psPAX2 (gag/pol/rev/tat). La transfección se hizo en 4 ml de OptiMEM con un agente de suministro de lípidos catiónicos (50 ul de Lipofectamine 2000 y 100 ul de reactivo Plus). Después de 6 horas, el medio se cambió a DMEM sin antibiótico con suero bovino fetal al 10 %.

15 Purificación de lentivirus. Los sobrenadantes víricos se recogieron después de 48 horas. Los sobrenadantes se limpiaron en primer lugar de desechos y se filtraron a través de un filtro de baja unión de proteínas de 0,45 um (PVDF). Después se centrifugaron en una ultracentrífuga durante 2 horas a 24 000 r.p.m. Los sedimentos víricos se resuspendieron en 50 ul de DMEM durante una noche a 4 °C. Después se dividieron en alícuotas y se congelaron inmediatamente a -80 °C.

25 Aislamiento clonal usando FACS. Para el aislamiento clonal de células embrionarias humanas HEK293FT y HUES64, las células se infectaron en suspensión con 1 ul o 5 ul de virus purificado. Veinticuatro horas después de la infección, se añadió doxiciclina 1 uM al medio de cultivo celular. Después de 24 o 48 horas más, las células se sometieron a clasificación celular asistida por fluorescencia (FACS) en un instrumento BD FACSAria Ilu para aislar las células individuales que expresaban de forma robusta EGFP (y, por tanto, Cas9) después de tratamiento con doxiciclina. Las células se sembraron a granel o en pocillos individuales para permitir la selección de poblaciones clonales con una Cas9 inducible integrada para uso adicional. La eficacia de clasificación fue siempre mayor >95 % y las células se visualizaron inmediatamente después de la siembra en placa para verificar la fluorescencia de EGFP.

La figura 15 representa los diseños del vector Tet Cas9.

La figura 16 representa un vector y la expresión de EGFP en células 293FT.

5

Secuencia de Cas9 inducible de pCasES020:

```
caacttgatagaaaagttggctccgaattcgcccttcaggtccgaggttctagacgagttfactccctatcagtgatagaga  
acgatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatc  
agtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgagttactccctatcagtgatagagaacgtatgtcgag  
gtagggcgtgtacgggtgggaggcctatataagcagagctcgtttagtgaaccgtcagatcgcaaagggcgaattcgacccaagttgtacag  
ccaccATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAA  
AGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA  
GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGT  
GGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCAGCAAGAAATCAAGGTGC  
TGGGCAACACCGACCGGCACAGCATCAAGAAGAACCCTGATCGGAGCCCTGCTGTTC  
GACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGAT  
ACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATG  
GCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGAAGA
```

GGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCT  
 ACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACC  
 GACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCGGG  
 GGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCT  
 GTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAACG  
 CCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGG  
 CTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGCAA  
 CCTGATTGCCCTGAGCCTGGGCCTGACCCCAACTTCAAGAGCAACTTCGACCTGGC  
 CGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACC  
 TGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGT  
 CCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCC  
 CCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCT  
 GCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTGA  
 CCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG  
 TTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCCGAGGAACTGCT  
 CGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCA  
 GCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAA  
 GATTTTTACCCATTCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTC  
 CGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTTCGCTGGAT  
 GACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTTCGAGGAAGTGGTGGAC  
 AAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCT  
 GCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT  
 ATAACGAGCTGACCAAAGTGAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC  
 CTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAA  
 AGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT  
 CCGTGGAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC  
 GATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGA  
 CATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA  
 GGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGA  
 AGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATC  
 CGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGC

ES 2 757 623 T3

CAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA  
TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAAT  
CTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGG  
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGG  
CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA  
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC  
GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG  
GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATG  
TGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGC  
TGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGT  
CGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC  
AGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGGCCTGAGCGAACTGGAT  
AAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT  
GGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGA  
TCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAG  
GATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACCTACCACCACGCCCACGACGCC  
TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAAG  
CGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA  
GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATG  
AACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCT  
GATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTG  
CCACCGTGCGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAG  
GTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAA  
GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC  
CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAG  
AAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAAGAAGCAGCTT  
CGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGG  
ACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAG  
AGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTC  
CAAATATGTGAACTTCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCC  
CGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTGGACG

AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT  
 CTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCA  
 GGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT  
 CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGC  
 TGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGAC  
 CTGTCTCAGCTGGGAGGCGACAAAAGGCCGGCGCCACGAAAAAGGCCGGCCAGG  
 CAAAAAGAAAAAGgaattctctagaGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACAT  
 GCGGTGACGTGAGGAGAATCCTGGCCAGTGAGCAAGGGCGAGGAGCTGTTACC  
 GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG  
 CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA  
 TCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCT  
 ACGGCGTGCAGTGCTTCAGCCGCTACCCCCACCACATGAAGCAGCACGACTTCTTCA  
 AGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC  
 GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG  
 CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGC  
 TGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAC  
 GGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT  
 CGCCGACCACTACCAGCAGAACACCCCCATCGGGCAGGGCCCCGTGCTGCTGCCCCG  
 ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGC  
 GATCACATGGTCCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCATGGAC  
 GAGCTGTACAAGCTCGAGGGAAGCGGAGCTACTAATTGAGCCTGCTGAAGCAGGC  
 TGGCGACGTGGAGGAGAACCCTGGACCTatgtctaggctggacaagagcaagtcataaacggagctctggaa  
 ttactcaatggtgcteggtatcgaaggcctgacgacaagaaaactcgtcaaaagctgggagftgagcagcctaccctgtaactggcactgga  
 agaacaagcgggcccctgctcgaatgcccctgccaatcgagatgctggacaggeatcatacccacttctgccccctggaaggcgagtcagtcg  
 aagacttctgaggacaacgccaagtataccgctgtgctctcctctcacatcggcaggggctaaagtgcctcggcacccgccaac  
 agagaaacagfacgaaacctggaaaatcagctcgcgttctgtgctcagcaaggttctcctgggagaacgcactgtacgctctgccc  
 gtggccacttfacactgggctgctgattggaggaaacaggagcatcaagtagcaaaaggaaagagagacacactaccaccgattctatg  
 cccccactctgagacaagcaattgagctgttcgaccggcaggagccgaacctgccttcttctggcctggaactaatcatatgtggcctg  
 gagaacagctaaagtgcgaaagcggggggccgaccgacgcccttgacgatttgacttagacatgctcccagecgateccttgacgatt  
 ttgacctgacatgctccccggtaattgtacaagtggtaattccggcaatcgcatacaagcttgcgataatcaacctctggattacaaaatt  
 tgtgaaagattgactggtattcttaactatgtgctcctttacgctatgtggatgacgctgcttaatgcctttgtatcatgctattgctcccgtatgg

cttfcattttctcctctgtataaatcctggfctgctctttatgaggagttggcccgtgtcaggcaacgtggcgtggtgactgtgtttg  
ctgacgcaacccccactggfaggccattgccaccacctgacgctctttccgggaacttctcttccccctccctattgccacggcggaaac  
tcatcgcgcctgcttcccgtctgtgacagggctcggctgttgggactgacaattccgtggtgttctggggaatcatgctcttc  
cttggctgctcgcctgtgtccacctggaffctgcgcgggacgctctctgctacgctccctcggccctcaatccagcggacctctctccg  
cggcctgctgcggctctgcggctcttccgctcttccgctcttccgctcagacagctcggatcctcttgggcccctccccgacgata  
ccgtcgacctgagacctagaaaaacatggagcaatcacaagtagcaatacagcagctaccaatgctgattgctcctggctagaagcaca  
agaggaggaggaggfagggtttccagtcacacctcaggtaccttfaagaccaatgactfacaaggcagctgtagatcttagccactttfaa  
agaaaagggggactggaaggctaactcccaacgaagacaagatatccttgaatctggtatctaccacacaaggctacttccct  
gattggcagaactcacaccagggccaggatcagataccactgaccttggatggtgtctacaagctagtaccagttgagcaagagaag  
gtagaagaagccaatgaaggagagaacaccgctgttacacctgtgagcctgcatggatggatgacctggagagagaagtattaga  
gtggagggtgacagccctagcatttcaatcactgcccagagactgcatccggactgtactgggtctctctgtttagaccagatctgagc  
ctgggagctctctgctaactagggaaaccactgcttaagcctcaataagcttgcctgagtgcttcaagtagtgtgtcccgtctgtgtgtg  
actctgtaactagagatccctcagaccttttagtcagtgtgaaaatctctagcagggcccgttfaaacccgctgatcagcctcactgtgc  
cttctagttgccagccatctgttcttccccctccccgctctctgacctggaagggtccactccactgtcttccataaaaaatgagg  
aaattgcatcgactgtctgagtaggtgtcattctattctgggggtgggtggggcaggacagcaagggggaggattgggaagacaatag  
caggcatgctggggatgcccgtggctctatggctctgagggcgaagaaaccagctggggcttaggggtatccccacgcgcctgta  
ggcgccatfaagcgcggcgggtgtgtgtgttacgagcgtgaccgctacacttgcacgcgcctagcggccgctcttctgcttcttct  
ccttcttctcggcagcttccggcttccccgctcaagctcaaaatcggggctccccctttaggggtccgatttagtctttacggcacctcgac  
ccccaaaaactgattagggtgatgggtcacgtagtggccatcgcctgatagacggttttcggccttgacgtggagccacgttcttfaat  
agtggactctgttccaactggaacaacactcaacctatctcggctcttctttgattataagggaatttgcgatttccgcttattggttaa  
aaatgagctgattfaacaaaaatfaacggaatfaattctgtggaatgtgtcagttagggtgtgaaagtccccaggcctccagcaggc  
agaagfatgcaaagcatgcatctcaattagtcagcaaccaggtgtgaaaagtccccaggctccagcaggcagaagtatgcaaagcatg  
catctcaattagtcagcaaccatagctcccggcctaactcggccatccccggcctaactcggccagttccggccattctcggcccatg  
tgactaattttttattatgacagggccaggccgctctgctctgagctattccagaagtagtagggagccttttggaggcctagggcttt  
gcaaaaagctccgggagctgtatataatcttggatctgatcagcagctgttgacaattaatcctggcatagtataatcggcatagtataat  
acgacaaggtgaggaactaaacctggccaagttgaccagtgccgttccgggtgctcaccgcgcgcgacgtcgggagcggctgagttc  
tggaccgaccgctcgggtctcgggactcgtggaggacgactcggcgggtgtgtcgggacgacgtgacctgttcatcagcgcg  
gtccaggaccaggtgtgctgggacaacacctggcctgggtgtgggtgctggcggcctggacgagctgtacggcaggtgtgaggtg  
gtccacgaactcgggacgctccggcggccatgaccgagatcggcagcagcctggggggggagttcgcctgctcggacc  
cggccggcaactcgtgacttctggtggcggagcaggaactgacacgtgctacgagatttccattccaccggcctctatgaaagggt  
gggctcggaaatgtttccgggacggcggctggatgatctccagcgcggggatctcatgctggagttcttcccacccaactgtttatt

gcagcftafaatggttacaataaagcaatagcatcacaaffcacaataaagcattttttcactgacttagttagtgggttgccaaactcat  
 caatgfatcftatcatgtctgtataccgtcgacctctagctagagctfggcgfaatcatggcatagctgtttcctgtgtgaaattgttatccgctca  
 caattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatagagtgaactaacacattaattgcgttgcgctca  
 ctgcccgtttccagtcgggaaacctgtctgtccagctgcattaatgaatcggccaacgcgcggggagagcgggttgcgtattgggctct  
 ctccgcttctctgctactgactcgtcgtcggctcggctcggctcggcgagcgggtatcagctcactcaagcgggtaatacggfatacca  
 cagaatcaggggataacgcaggaaagaacatgtgagcaaaagccagcaaaagccaggaaccgtaaaagccgcgttgcgtggcggtt  
 ttccataggctccgccccctgacgagcatcacaanaatcgacgtcaagtcagaggtggcgaaacccgacaggactataaagatacca  
 ggcggttccctggaagctccctcgtgctctctgttccgacctgcegttaccggatacctgtccgcttctccctcgggaagcgtg  
 gcgctttctcatagctcagctgttaggtatctcagttcgggtgtaggtcgttcctcaagctgggtgtgtgcacgaacccccgttgcagccc  
 gaccgctgcgcttaccggtaactatcgtcttgagccaacccggtaagacacgacttatcggcactggcagcagccactgtaacaggat  
 tagcagagcggatgtgtagggcgtgtacagagttcttgaagtggtggcctaactacggctacactagaagaacagtafttggfctctg  
 ctctgtgaagccagttaccctcggaaaaagagttggtagctcttgcacggcaaaacaaccaccgctggtagcgggtgttttttgggttgc  
 gcagcagattacgcgcagaaaaaaggatctcaagaagatccttgccttctacggggtctgacgctcagtggaacgaaaactcaggt  
 aagggttttggctafgagatatacaaaaggatctaccagatcttttaaaftaaaaatgaagtttaataatcaatcaaaagtatafatagtaa  
 acttggctgacagttaccaatgcttaatcagtgaggcacctatctcagcagatctgtctatttctcatccatagttgcctgactccccgtcgtg  
 agataactacgafacgggagggcttaccatctggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagattatcagc  
 aataaaccagccagccggaaggccgagcgcagaagtggctcctgcaactttatccgctccatccagcttataattgttgcgggaagcta  
 gagtaagtagtccagtaataagtttgcgcaacggttggcaattgctacagcagcgtggtgacgctcgtcgttggatagcctcattca  
 gctccgggttcccaacgataagggagttacatgatccccatgtgtgcaaaaaagcgggttagctcttccggctcctcagctgttgcagaa  
 gtaagttggccgaggttatacactcatggttatggcagcactgcataattctctactgcatccatccgtaagatgctttctgtgactgggta  
 gtactcaaccaagtcattctgagaatagtgatgcggcgaccgagttgctcttcccccggcgaataacgggataataccgcgccacatagca  
 gaactftaaaagtgtcatcattggaaaaagttctcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcagtaaccact  
 cgtgcaccaactgatctcagcatctttactttaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgcccgaaaaaagggaat  
 aaggcgacacggaaatgtgaatactcactcttctttcaatattatgaagcatttatcagggttattgtctcatgagcggatacatattg  
 aatgtatttagaaaaataacaaatagggttccgcacattccccgaaaagtgccacctgacgtcagcggatcgggagatctcccgatc  
 ccctatgggtgactctcagtaaatctgctctgatgccafagftaagccagtatctgctccctgctgtgtgttggaggctcgtgagtagtgc  
 gcgagcaaaatfaagctacaacaaggcaaggcttaccgacaattgcatgaagaatctgctfagggttaggcgttttgcgctcttccgcat  
 gtacgggcccagatatacgcgttgacattgattatgactagttatfaatagtaaatcaatfacggggtcattagttcatagcccatatagagttcc  
 gcgttacataacttacggtaaatggcccgtggctgaccgccaacgacccccgcccactgacgtcaataatgacgtatgttcccatagta  
 acgccaatagggaacttccactgacgtcaatgggtggagatttacggtaaaactgccacttggcagtacatcaagtgtatcatatgccaagta  
 cggccccctattgacgtcaatgacggtaaatggcccctggcattatgccagttacatgaccttattgggacttctacttggcagttacatcta

ES 2 757 623 T3

cgfaffagtcacgctaffaccatggfgatcggttttggcagtfacatcaatgggctggatagcggttgactcacggggatttccaagtctc  
acccattgacgtcaatgggagfittgtttggcaccaaaatcaacgggactttccaaaatgtcgtacaactccgccccattgacgcaaatgg  
gggtagggcggtfacgggtgggaggtctatataaagcagcgcttttgcctgactgggtctctctggtagaccagatctgagcctgggagctc  
tctggctaactagggaaaccactgctaaagcctcaataaagcttgccttgagtcttcaagtagtgtgtcccgtctgtgtgactctggtaa  
ctagagatccctcagacccttttagtcagtggtgaaaatcttagcagtgccgcccgaacagggacttgaagcgaaaaggaaaccagag  
gagctctctcagcagggactcggcttctgaagcgcgcacggcaagagggcgagggcgggcactggtgagtagccaaaaafttgac  
tagcggaggctagaaggagagagatgggtgagagcgtcagtattaagcgggggagaattagatcgcgatgggaaaaaftcggtaa  
ggccagggggaaagaaaaataaaftaaacatatagtatgggcaagcagggagctagaacgattcgcagttaactctggcctgtaga  
aacatcagaaggtgtagaacaactgggacagctacaacctccctcagacaggtacagaagaacttagatcattatataatagtag  
caaccctctattgtgcatcaaagatagagataaaagacaccaaggaagctttagacaagatagaggaagagcaaaacaaaagtaaga  
ccaccgcacagcaagcggcctgatctcagacctggaggaggagatatgagggacaattggagaagtgaattatataatataaagta  
gtaaaaattgaaccattagtagtagcaccaccaagcaagagaagagtggtgagagagaaaaagagcagtggggaataggagcttt  
gtccttgggtcttgggagcagcaggaagcactatggggcagcgtcaatgacgctgacggtacaggccagacaatfctgtctgtatag  
tgacgagcagaacaafftctgagggctattgagggcgaacagcatctgtgcaactcacagcttggggcatcaagcagctccaggcaa  
gaaactctgctgtgaaaagatacctaaaggatcaacagctcctggggatttgggtgctctggaaaactcatttcaccactgctgtcctt  
ggaatgctagttggagtaataatctctggaacagatttgaatcacacgacctggatggagtgaggacagagaaftaacaattacacaagc  
ttaatacactccttaattgaagaatcgaaaaccagcaagaaaagaatgaacaagaatfctggaatagataaatgggcaagtttgggaatt  
ggfttaacatacaaafttggctgtggtatataaaftatcataatgatagtaggagcttggtaggttaagaatagttttgtctgactttctatag  
tgaatagagttaggcagggatattcaccattatcgtttcagaccacctcccaaccccaggggacccgacaggcccgaaggaatagaag  
aagaaggtggagagagagacagacagatccattcgattagtgaacggatcggcactgctgccaattctgagacaaatggcagt  
attcatccacaatfttaaaagaaaagggggattgggggtacagtgagggaagaatagtagacataatagcaacagacatacaaac  
aaaagaattcaaaaacaaattacaaaatftcaaaatftcgggtttattacagggacagcagagatccagtttggtaattaa

**REIVINDICACIONES**

1. Un complejo de repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR)-Cas, que comprende:
- 5
- una Cas9, en la que la Cas9 tiene dos o más señales de localización nuclear;
  - una secuencia de guía unida a una secuencia de acoplamiento tracr, y
  - una secuencia tracr hibridable con la totalidad o una parte de la secuencia de acoplamiento tracr;
- 10
- en el que la secuencia de guía se diseña para que tenga complementariedad con una secuencia diana en una célula eucariota y pueda dirigir la unión específica de secuencia del complejo de CRISPR a la secuencia diana en la célula eucariota.
2. Un sistema de vector de repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR)-Cas, que comprende:
- 15
- a) un primer elemento regulador unido de forma funcional a una o más secuencias que codifican (1) una secuencia de guía unida a una secuencia de acoplamiento tracr y (2) una secuencia tracr hibridable con la totalidad o una parte de la secuencia de acoplamiento tracr;
- 20
- b) un segundo elemento regulador unido de forma funcional a una secuencia que codifica una Cas9,
- en el que la Cas9 tiene dos o más señales de localización nuclear;
- en el que los componentes (a) y (b) están ubicados en el mismo vector o diferentes vectores del sistema y en el que la secuencia de guía se diseña para que tenga complementariedad con una secuencia diana en una célula eucariota.
- 25
3. El sistema de vector de CRISPR-Cas de acuerdo con la reivindicación 2, en el que la secuencia que codifica la Cas9 tiene los codones optimizados para la expresión en células eucariotas.
- 30
4. El sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 2 o 3, que comprende dos o más secuencias de guía.
5. El sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 2-4, en el que dicho mismo o diferentes vectores son uno o más vectores víricos.
- 35
6. El sistema de vector de CRISPR-Cas de acuerdo con reivindicación 5, en el que dicho uno o más vectores víricos son uno o más vectores retrovíricos, adenovíricos, alfavíricos adenoasociados o víricos del herpes simple.
- 40
7. El sistema de vector de CRISPR-Cas de acuerdo con la reivindicación 6, en el que el retrovirus es un lentivirus.
8. El sistema de vector de CRISPR-Cas de acuerdo con la reivindicación 6, en el que el alfavirus es un virus Sindbis, virus del bosque Semliki o virus de la encefalitis equina venezolana.
- 45
9. El sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 2-8, en el que los componentes (a) y (b) están ubicados en el mismo vector.
10. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-9, en el que la secuencia de acoplamiento tracr se une con la secuencia tracr para formar un ARN quimérico (ARNqui).
- 50
11. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-10, en el que la Cas9 dirige la escisión de ambas hebras de un polinucleótido diana que contiene la secuencia diana.
- 55
12. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-10, en el que la Cas9 se muta con respecto a una Cas9 de tipo silvestre correspondiente de modo que la Cas9 mutada carezca de la capacidad de escindir una o ambas hebras de un polinucleótido diana que contiene la secuencia diana.
- 60
13. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-10 y 12, en el que la Cas9 comprende una mutación seleccionada del grupo que consiste en D10A, H840A, N854A y N863A con referencia a la numeración de posiciones de una Cas9 de *S. Pyogenes*.

14. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-13, en el que la Cas9 comprende una o más señales de localización nuclear en el extremo amino y una o más señales de localización nuclear en el extremo carboxi.
- 5 15. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-14, en el que la Cas9 es una Cas9 de *S. Pyogenes*.
16. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-14, en el que la Cas9 es una de *thermophilus* Cas9.
- 10 17. El complejo de CRISPR-Cas o sistema de vector de CRISPR-Cas de acuerdo con una cualquiera de las reivindicaciones 1-16 para su uso en tratamiento.
- 15 18. Uso del complejo de CRISPR o sistema de vector de CRISPR-cas de acuerdo con una cualquiera de las reivindicaciones 1-16 para manipulación genómica, en el que dicho uso no es un método de tratamiento del organismo animal o humano por terapia, y en el que dicho uso no es un proceso para modificar la identidad genética de la línea germinal de los seres humanos.

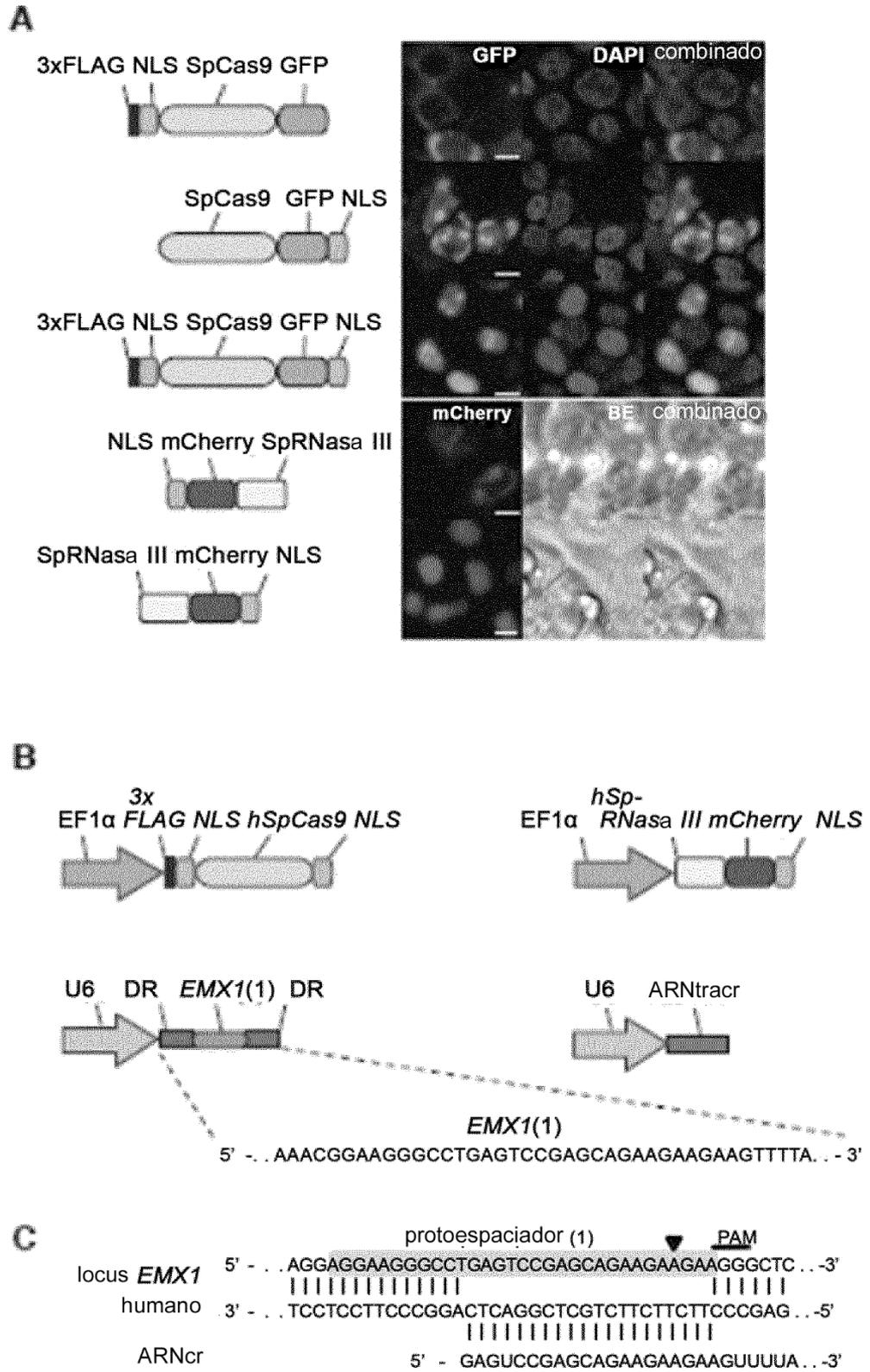


FIG. 1A-C





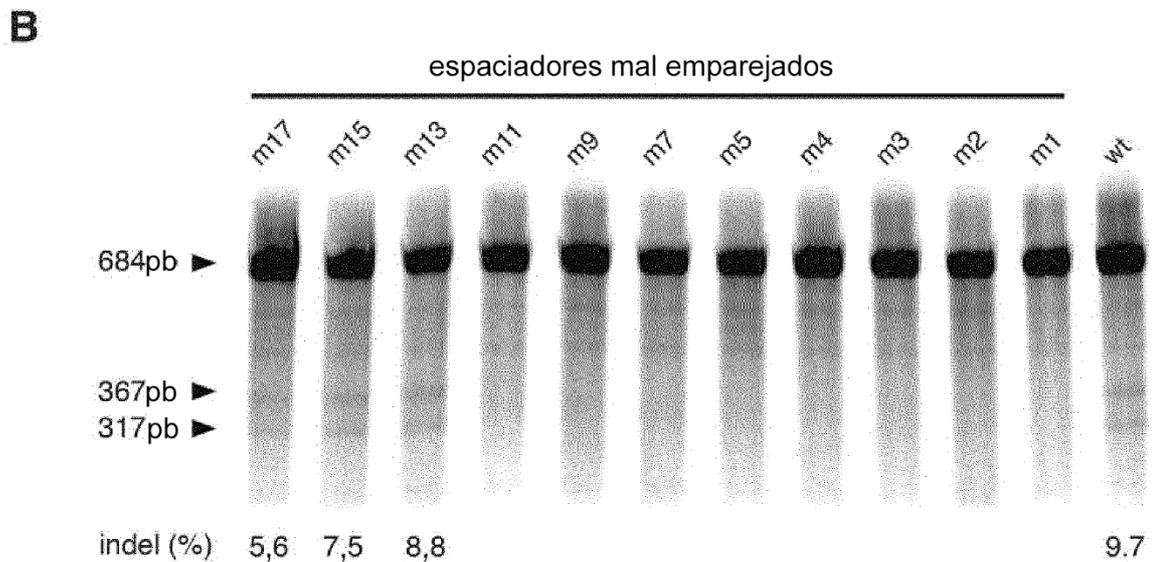
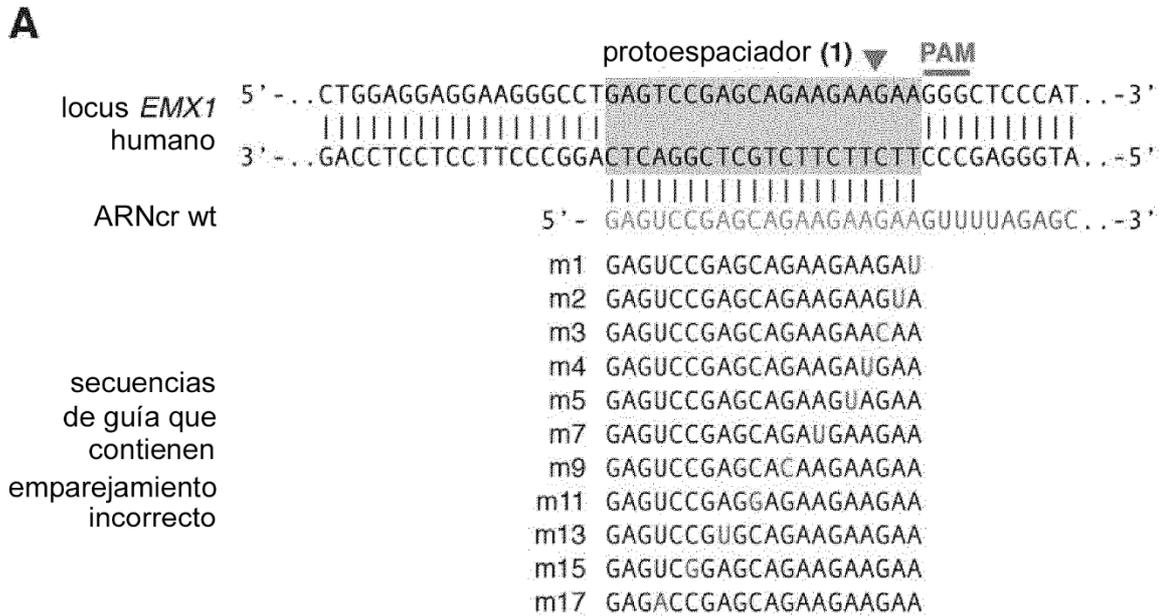
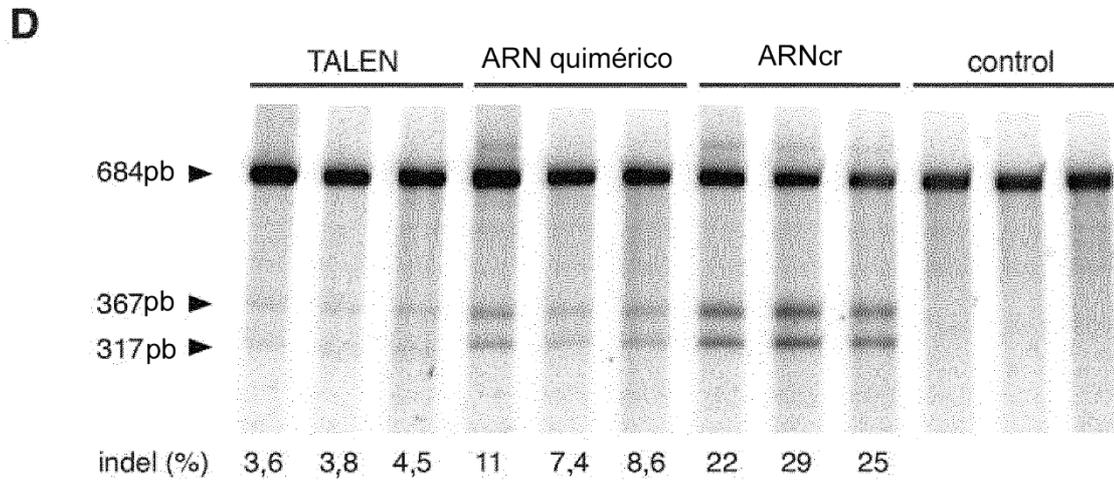


FIG. 3A-C



**FIG. 3D**





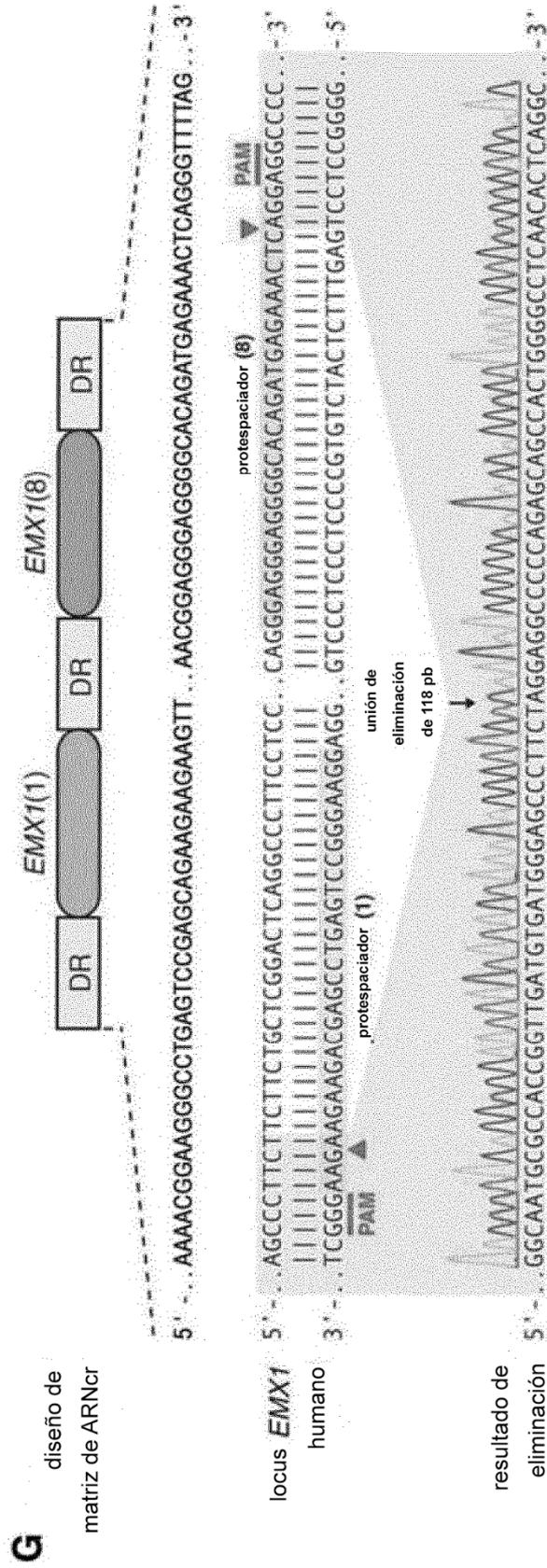


FIG. 4G

locus de CRISPR de tipo II de *Streptococcus pyogenes* SF370

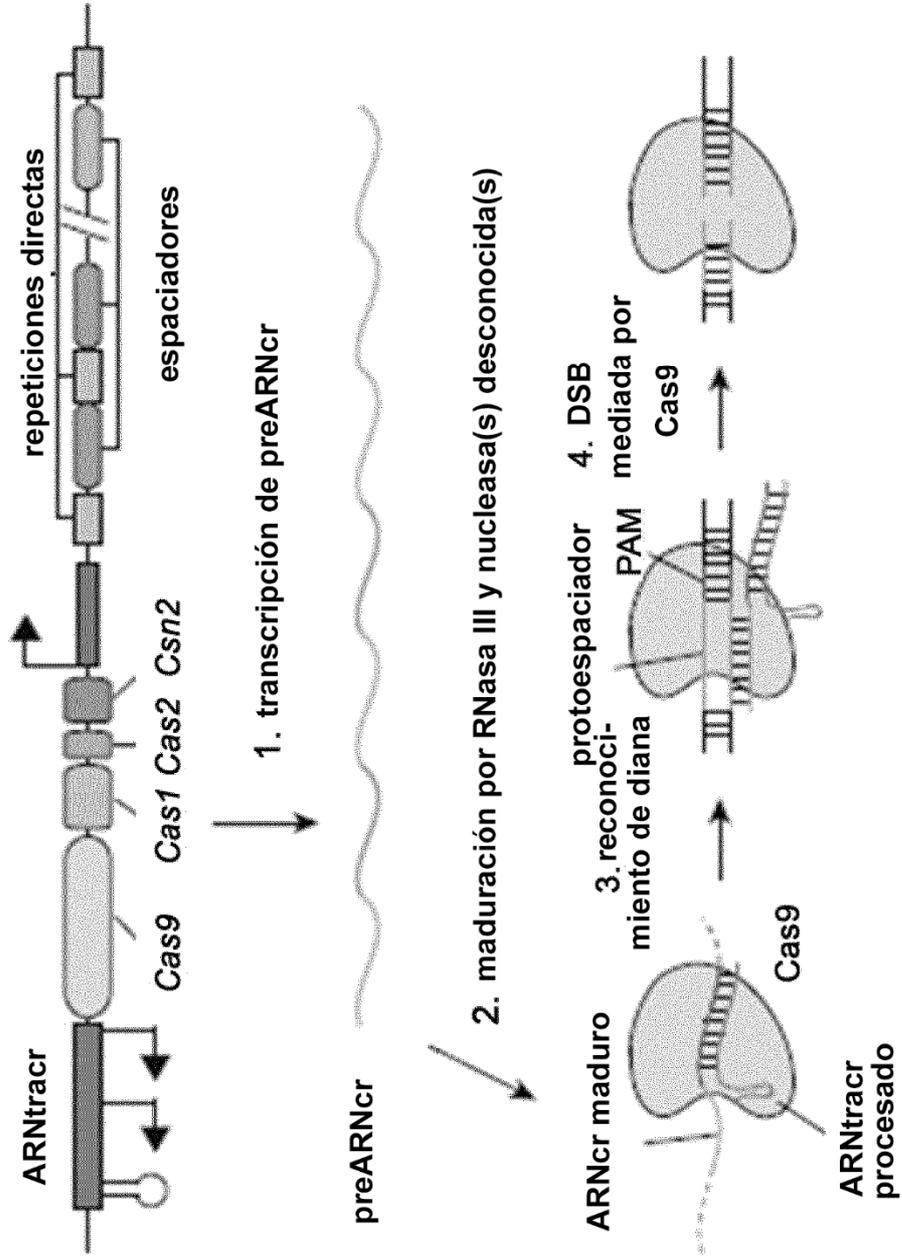
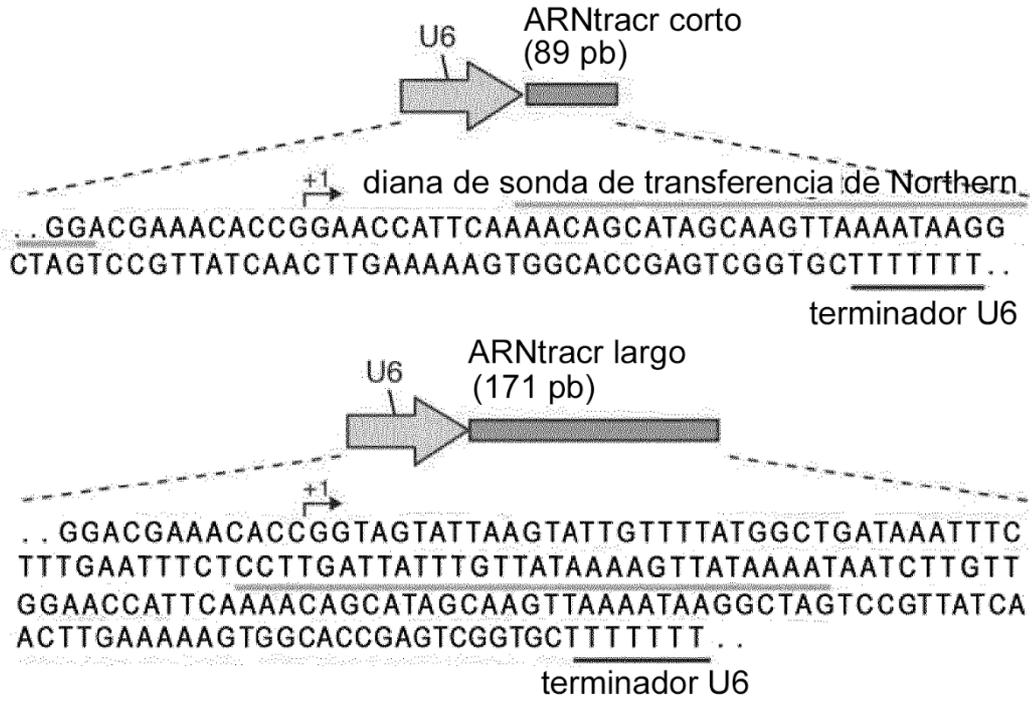


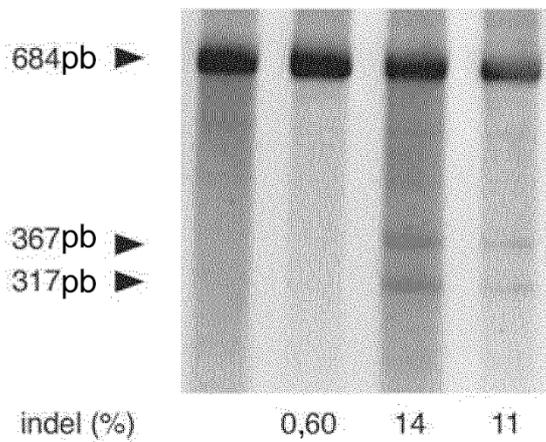
FIG. 5

**A**



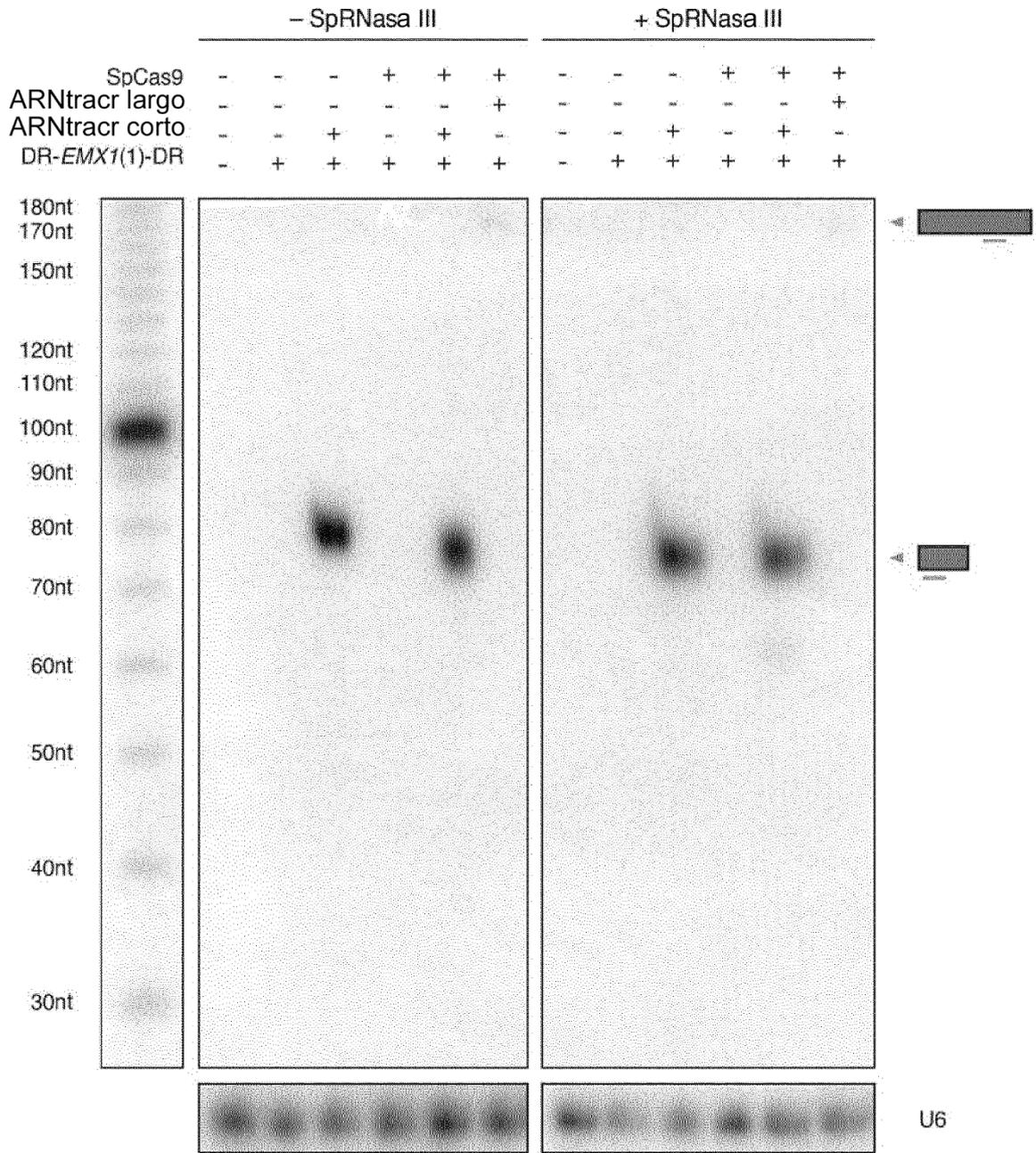
**B**

SpCas9	+	+	+	+
ARNtracr largo	+	+	-	-
ARNtracr corto	-	-	+	+
SpRNasa III	+	+	+	+
DR-EMX1(1)-DR	+	+	+	+



**FIG. 6A-B**

**C**



**FIG. 6C**

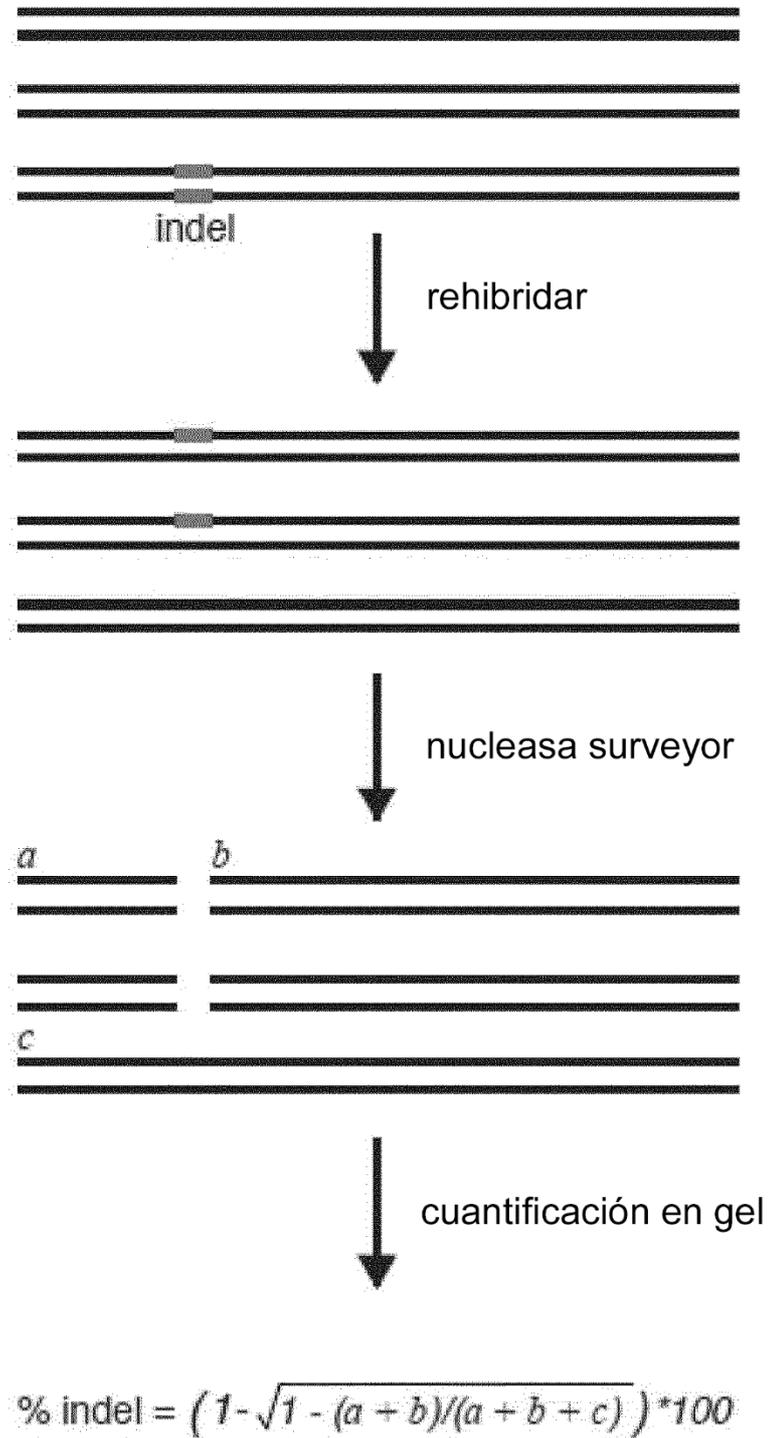
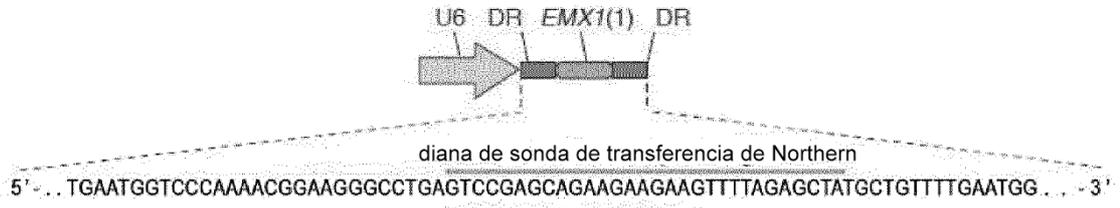
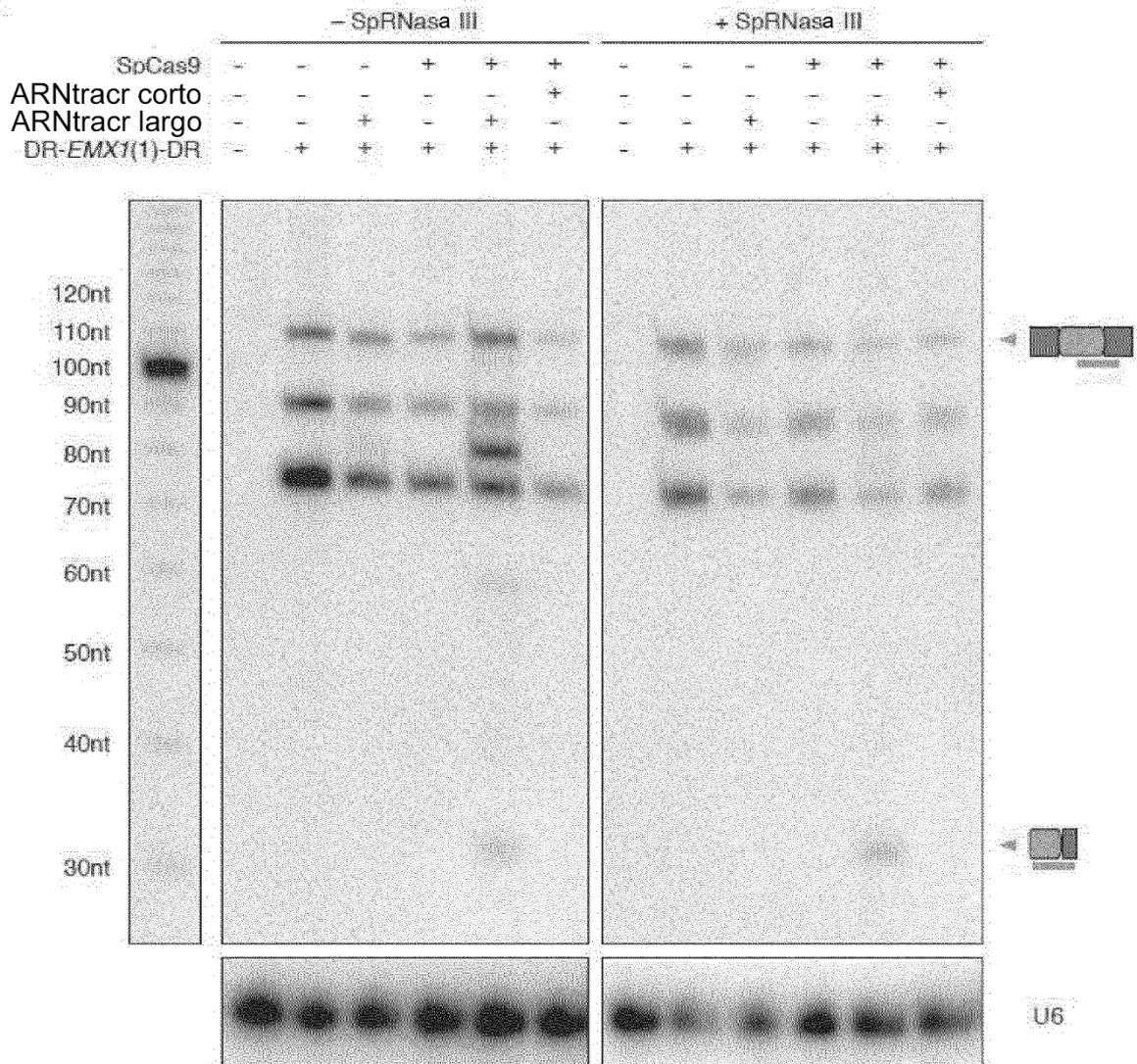


FIG. 7

**A**



**B**



**FIG. 8A-B**

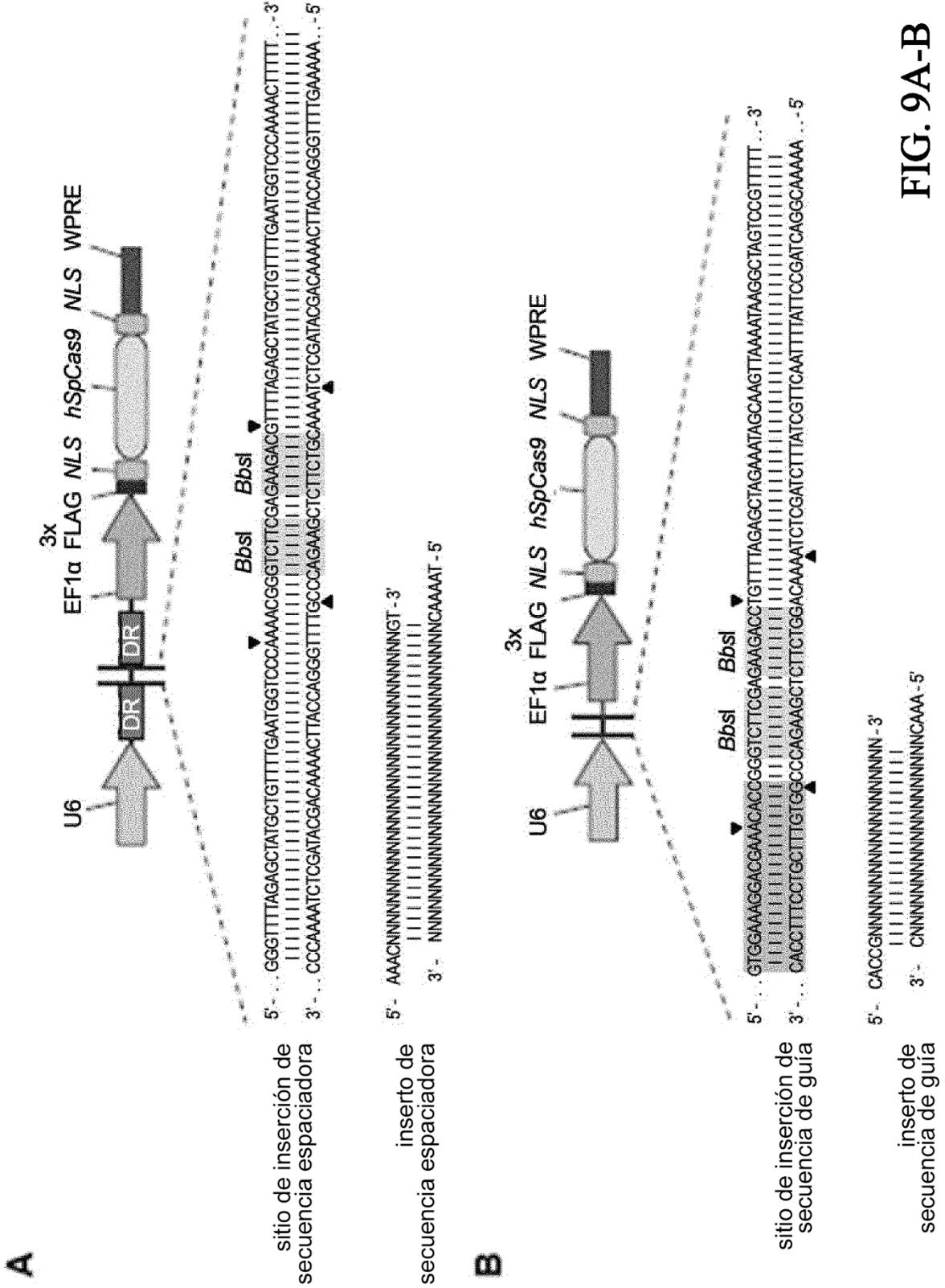


FIG. 9A-B

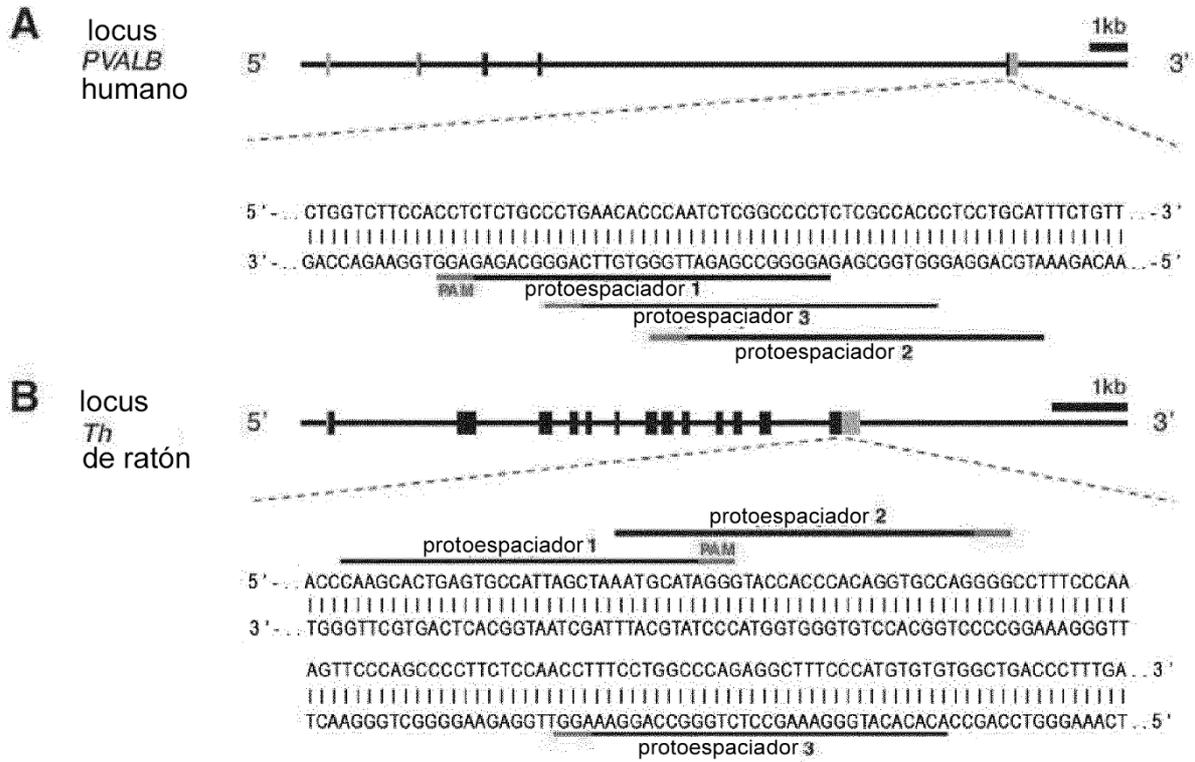


FIG. 10A-B

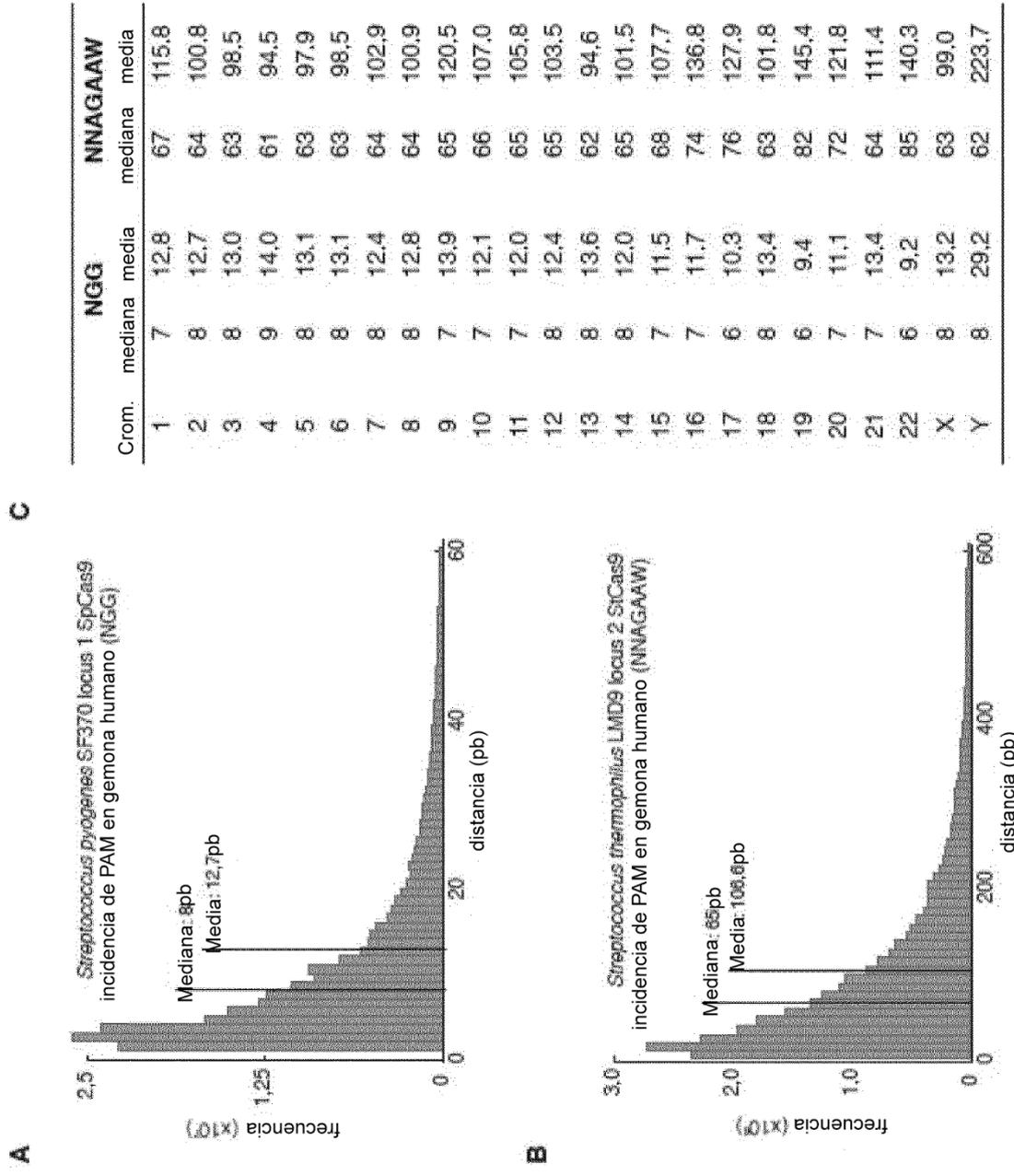


FIG. 11A-C

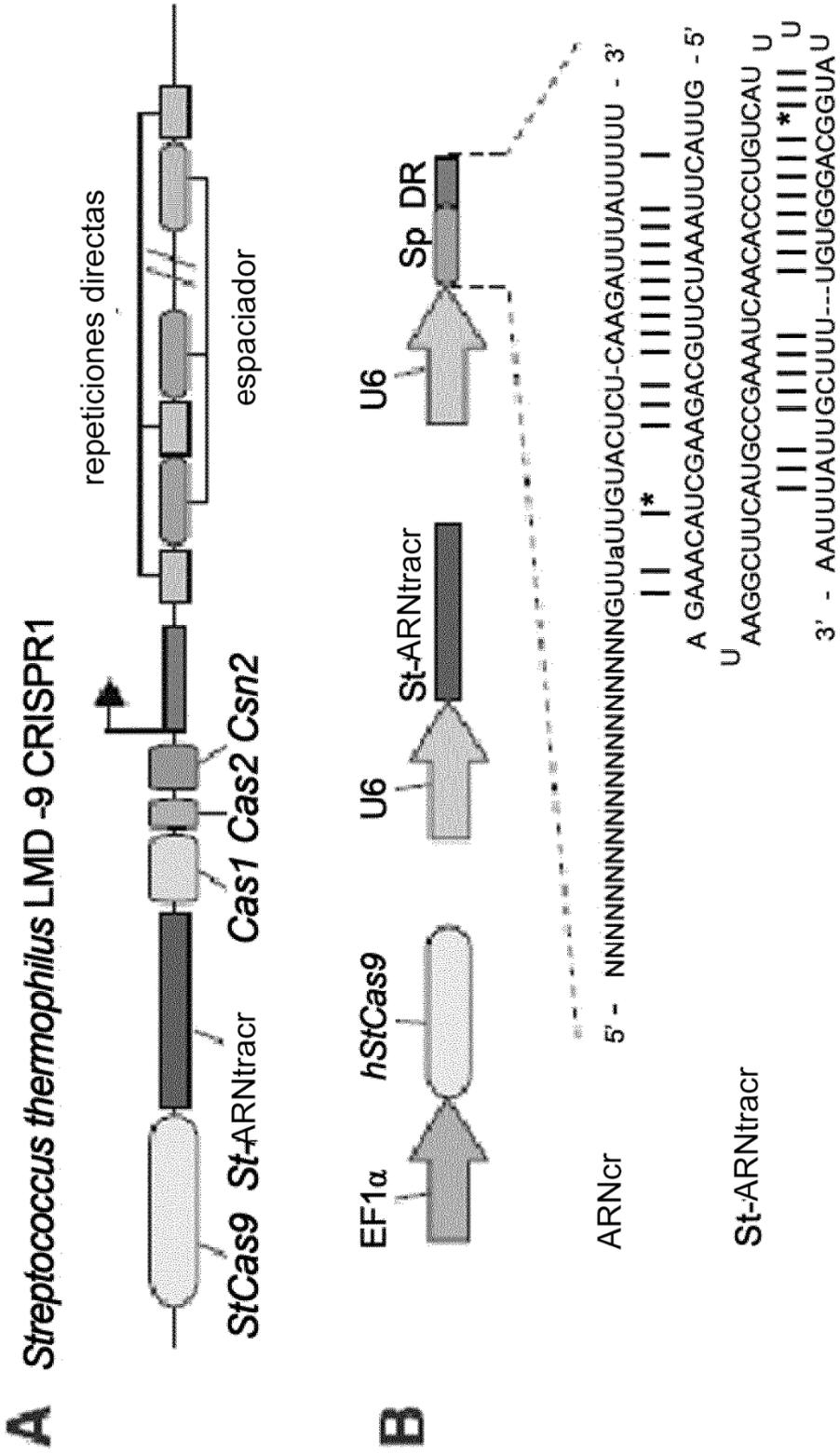
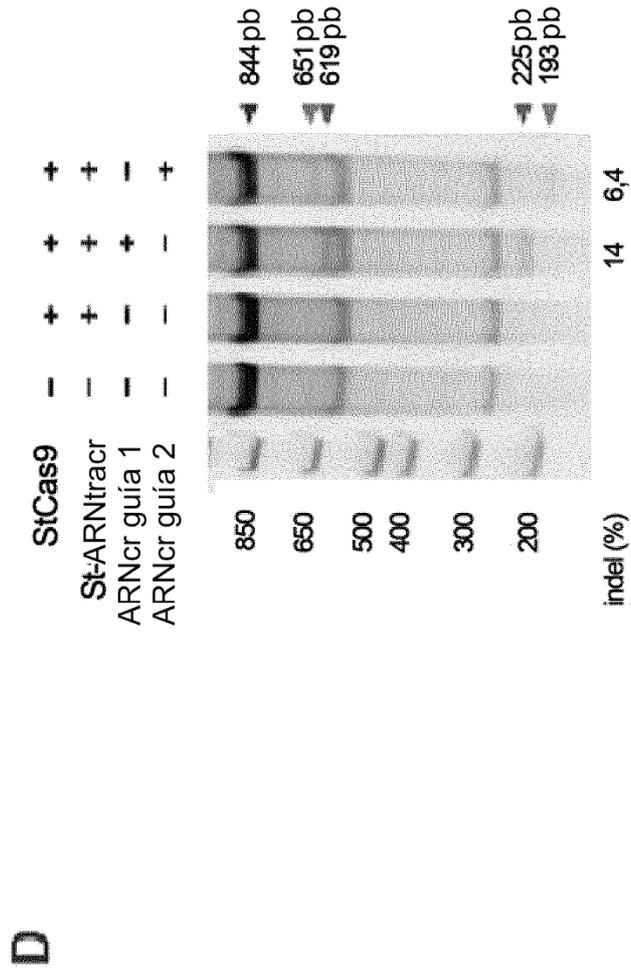
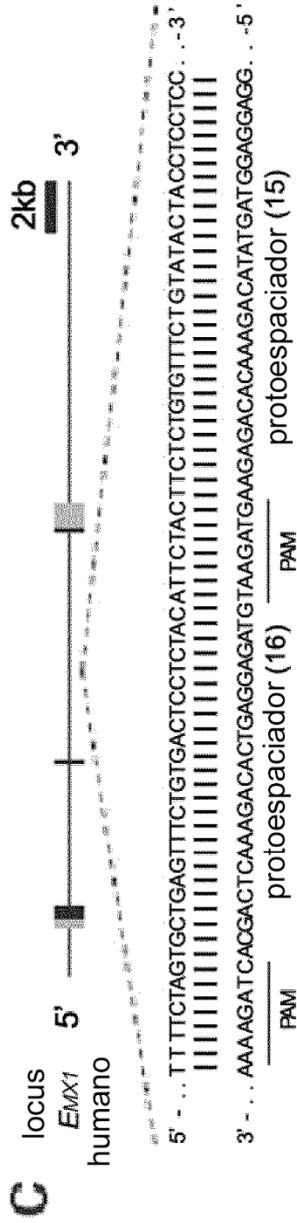
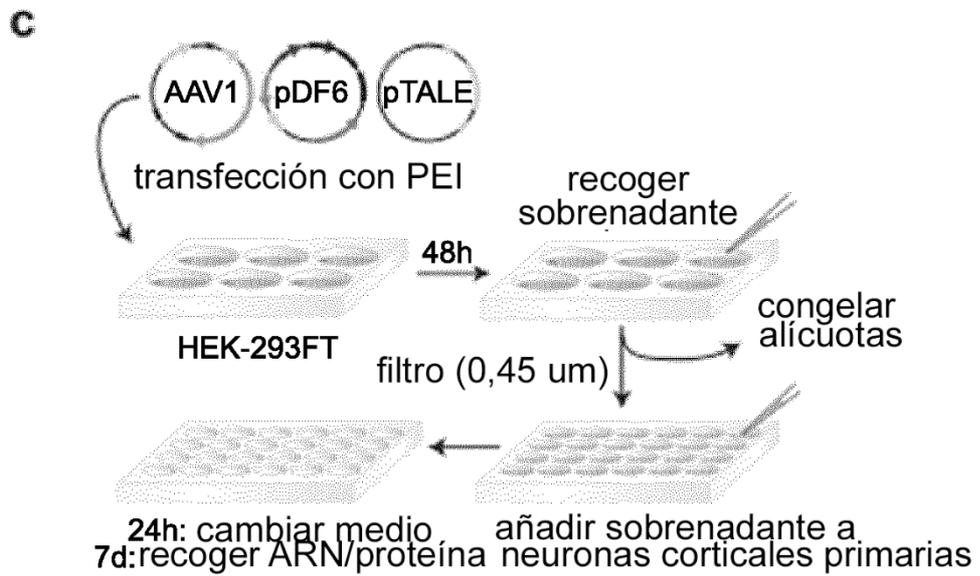
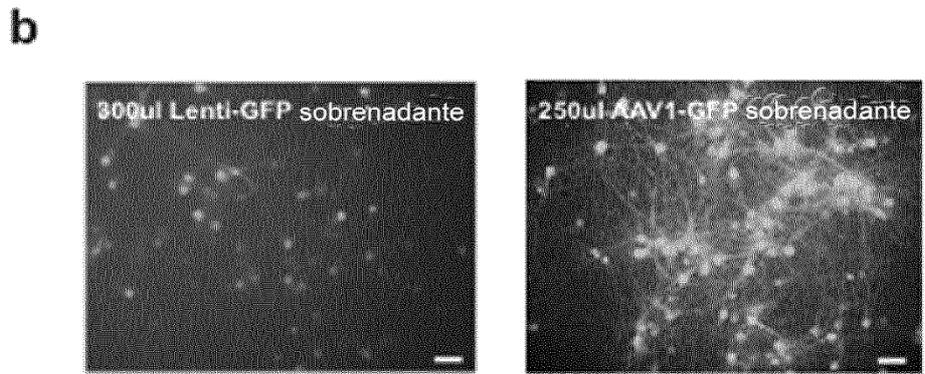
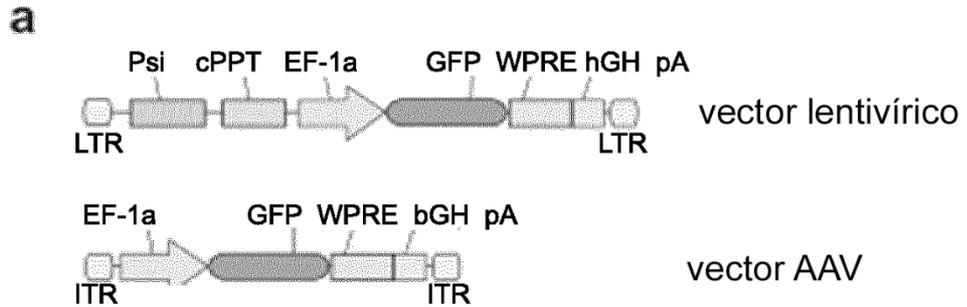


FIG. 12A-B



**FIG. 12C-D**



**FIG. 13A-C**





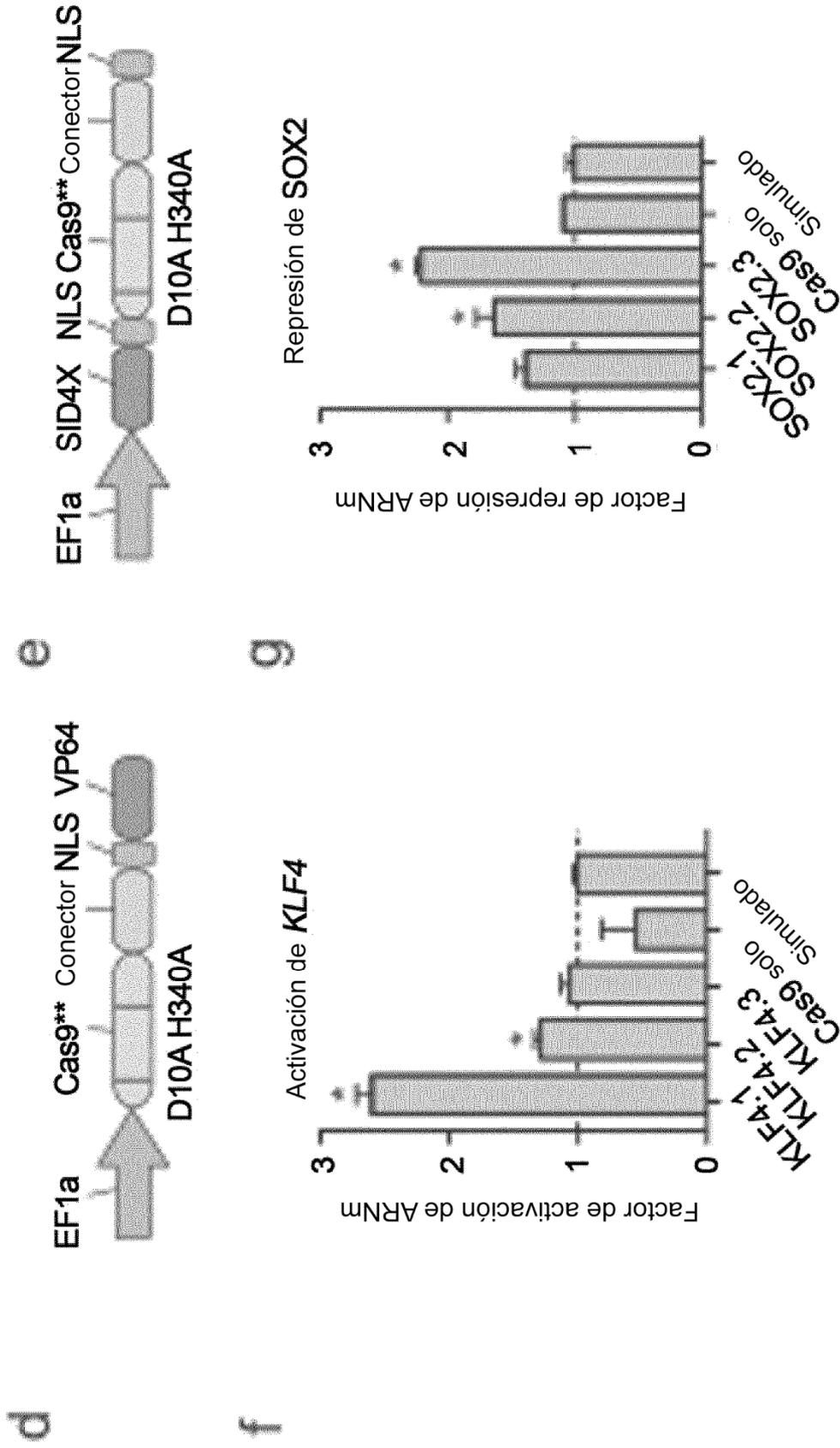


FIG. 14D-G

Casetes lentivíricos transgénicos de Cas9 ES

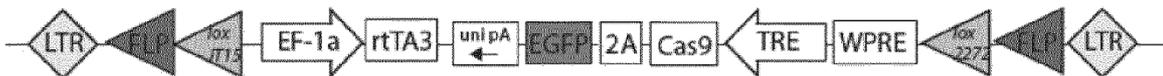
1) Bicistrónico de rtTA3 y Cas9



2) Autoregulador de retroalimentación de rtTA3



3) Casete inverso de Cas9-EGFP con poliA unidireccional



4) Virus diferentes para rtTA3 y Cas9

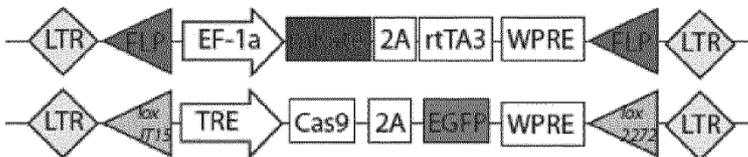


FIG. 15

Cas9 autoinductora sensible a Dox en HEK293FT mediante lentivirus purificado

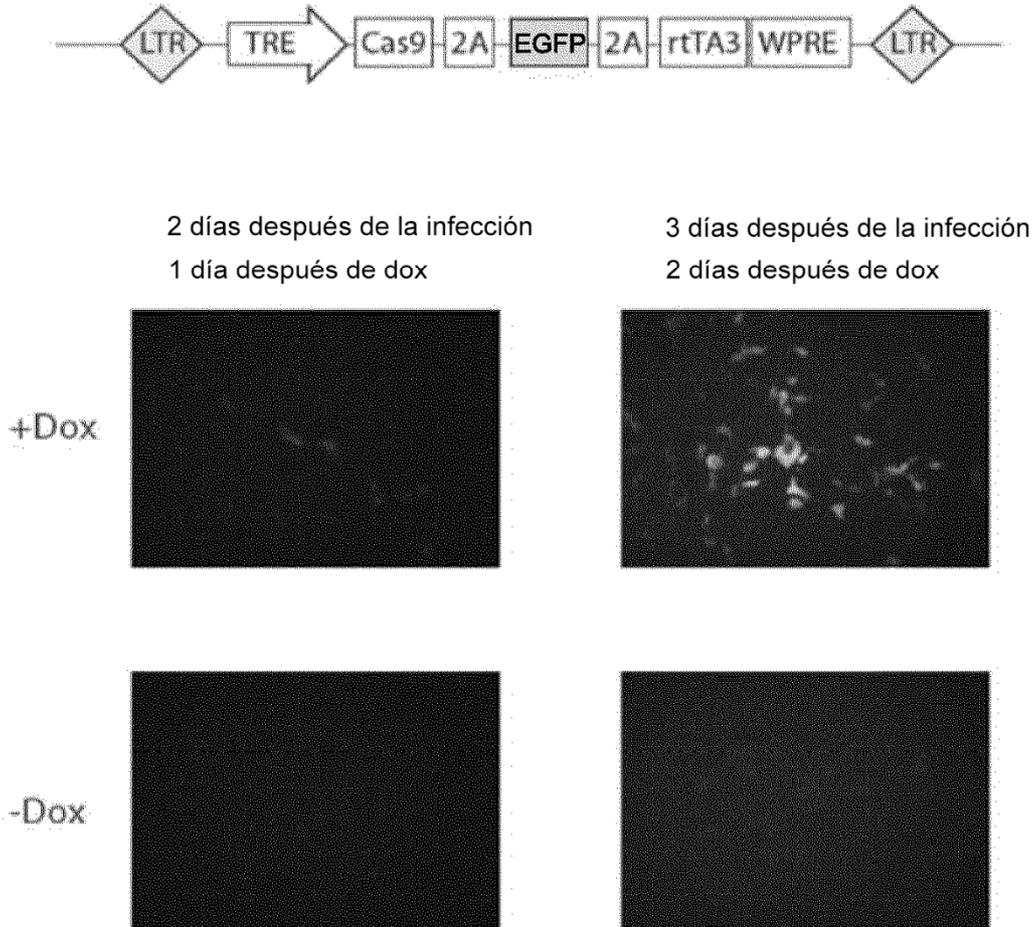


FIG. 16