



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



⑪ Número de publicación: **2 884 838**

⑮ Int. Cl.:

**C12N 15/11** (2006.01)  
**C12N 9/22** (2006.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑥ Fecha de presentación y número de la solicitud internacional: **05.04.2016 PCT/US2016/026028**

⑦ Fecha y número de publicación internacional: **13.10.2016 WO16164356**

⑨ Fecha de presentación y número de la solicitud europea: **05.04.2016 E 16720220 (9)**

⑩ Fecha y número de publicación de la concesión europea: **26.05.2021 EP 3280803**

---

④ Título: **ARN guía químicamente modificados para la regulación génica mediada por CRISPR/CAS**

⑩ Prioridad:

**06.04.2015 US 201562143729 P  
12.05.2015 US 201562160545 P**

④ Fecha de publicación y mención en BOPI de la traducción de la patente:  
**13.12.2021**

⑩ Titular/es:

**THE BOARD OF TRUSTEES OF THE LELAND  
STANFORD JUNIOR UNIVERSITY (50.0%)  
Office of the General Counsel, Building 170, Third  
Floor, Main Quad, P.O. Box 20386  
Stanford, CA 94305-2038, US y  
AGILENT TECHNOLOGIES, INC. (50.0%)**

⑩ Inventor/es:

**PORTEUS, MATTHEW H.;  
HENDEL, AYAL;  
CLARK, JOE;  
BAK, RASMUS O.;  
RYAN, DANIEL E.;  
DELLINGER, DOUGLAS J.;  
KAISER, ROBERT y  
MYERSON, JOEL**

⑩ Agente/Representante:

**UNGRÍA LÓPEZ, Javier**

**ES 2 884 838 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

ARN guía químicamente modificados para la regulación génica mediada por CRISPR/CAS

5 **Antecedentes de la invención**

La edición del genoma con nucleasas diseñadas es una tecnología revolucionaria para modificar esencialmente cualquier secuencia genómica de interés (Porteus, M.H. y Carroll, D., *Nature Biotechnology* 23, 967-973 (2005)). Esta tecnología explota las nucleasas diseñadas para generar roturas en la doble cadena (DSB, por sus siglas en inglés) específicas de sitio, seguidas de la resolución de las DSB mediante mecanismos de reparación celular endógenos. El resultado puede ser la mutación de un sitio específico a través de uniones terminales mutagénicas no homólogas (NHEJ, por sus siglas en inglés), creando inserciones o delecciones (ins/del) en el sitio de la rotura o el cambio preciso de una secuencia genómica mediante recombinación homóloga (RH) utilizando un molde donante introducido exógenamente (Hendel *et al.*, *Trends in Biotechnology* 33, 132-140 (2015)). Una adición importante reciente a esta plataforma es el sistema de repeticiones palindrómicas cortas, agrupadas y regularmente interespaciadas (CRISPR)/Cas que consiste en una nucleasa guiada por ARN (Cas) y un ARN guía corto (ARNsg) Jinek, M. *et al.*, *Science* 337, 816-821 (2012), Mali, P. *et al.*, *Science* 339, 823-826 (2013), Cong, L. *et al.*, *Science* 339, 819-823 (2013), Hsu *et al.*, *Cell* 157, 1262-1278 (2014)). El ARN guía está compuesto por dos ARN denominados ARN CRISPR (ARNcr) y ARNcr transactivador (ARNtracr), que se fusionan típicamente en un ARN guía único químérico (ARNsg).

Los ARNsg para la edición del genoma pueden consistir en 100 nucleótidos (nt) de los cuales 20 nt en el extremo 5' hibridan con una secuencia de ADN diana mediante el emparejamiento de bases de Watson-Crick y guían a la endonucleasa Cas para escindir el ADN genómico diana.

25 El sistema CRISPR/Cas también se ha adaptado para el control específico de secuencia de la expresión génica, por ejemplo, inhibición o activación de la expresión génica. Usando variantes particulares del polipéptido Cas9 que carecen de actividad endonucleasa, los genes diana pueden ser reprimidos o activados (Qi *et al.*, *Cell*, 2013, 152(5):1173-7783, Perez-Pinera *et al.*, *Nat Methods*, 2013, 10(10):973-976, Maeder *et al.*, *Nat Methods*, 2013, 30 10(10):977-979, Gilbert *et al.*, *Cell*, 2014, 159:647-661, O'Connell *et al.*, *Nature*, 2014, 516:263-266).

35 Desafortunadamente, la edición del genoma y la modulación de la expresión génica mediante el sistema CRISPR/Cas siguen siendo ineficaces, especialmente en células primarias. Como tales, existe la necesidad en la técnica de composiciones y métodos mejorados basados en el sistema CRISPR/Cas que puedan usarse para la regulación de genes, por ejemplo, editando el genoma, inhibiendo la expresión génica y activando la expresión génica. La presente invención satisface esta necesidad y también proporciona ventajas adicionales.

**Breve sumario de la invención**

40 La presente invención proporciona métodos *in vitro* para inducir (por ejemplo, inicio, modulación, potenciación, etc.) la regulación génica de un ácido nucleico diana en una célula como se define en las reivindicaciones. La invención incluye el uso de ARN guía único modificados (ARNsg) que mejoran la edición del genoma y/o la inhibición o activación de la expresión génica de un ácido nucleico diana en una célula primaria (por ejemplo, cultivada *in vitro* para su uso en terapia *ex vivo*) o en una célula en un sujeto tal como un ser humano (por ejemplo, para su uso en terapia *in vivo*). La presente invención se usa con células primarias y en cualquier locus génico que sea susceptible de tecnología de edición del genoma mediada por nucleasas.

45 En particular, la presente invención proporciona un método *in vitro* para inducir la regulación génica de un ácido nucleico diana en una célula primaria, comprendiendo el método introducir en la célula primaria:

50 (a) un ARN guía único (ARNsg) modificado que comprende una primera secuencia de nucleótidos que es complementaria al ácido nucleico diana y una segunda secuencia de nucleótidos que interactúa con un polipéptido de proteína asociada a CRISPR (Cas), en donde uno o más de los nucleótidos en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos son nucleótidos modificados, en donde los nucleótidos modificados se seleccionan de entre el grupo que consiste en un nucleótido de 2'-O-metil 3'-fosforotioato (MS), un nucleótido de 2'-O-metil 3'-tiofosfonoacetato (MSP) y una combinación de los mismos; y

55 (b) un polipéptido Cas, un ARNm que codifica un polipéptido Cas y/o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica un polipéptido Cas,

60 en donde el ARNsg modificado guía el polipéptido Cas hacia el ácido nucleico diana, y  
en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana con una actividad mejorada en relación con un ARNsg sin modificar correspondiente y  
en donde la actividad mejorada comprende un aumento o mejora en la eficiencia y/o la frecuencia de inducción, modulación, regulación o control de la edición del genoma y/o la expresión génica.

Otros objetos, características y ventajas de la presente invención resultarán evidentes para un experto en la técnica a partir de la siguiente descripción detallada y las figuras.

#### Breve descripción de los dibujos

5 Las Figuras 1A-G muestran que los ARNsg sintetizados y químicamente modificados facilitan altos niveles de escisión del ADN *in vitro* y altas frecuencias de ins/del en una línea celular humana (K562). La Figura 1A muestra la secuencia y el esquema de la estructura secundaria del ARNsg de *IL2RG* cargado en Cas9 y unido a su sitio diana genómico. Los nucleótidos con modificaciones químicas están marcados con señales blancas. La Figura 1B  
10 representa estructuras de modificaciones químicas incorporadas durante la síntesis química de ARNsg (Tabla 1 para secuencias). Cuando están presentes, modificaciones químicas que comprenden: 2'-O-metilo (M), 2'-O-metilo, 3'fosforotioato (MS) o 2'-O-metilo, 3'tioPACE (MSP) (mostrados en gris) se incorporaron en tres nucleótidos terminales en ambos extremos 5' y 3'. La Figura 1C muestra la alteración del gen por NHEJ mutagénico medido por secuenciación profunda de amplicones de PCR. La Figura 1D muestra la adición de genes por RH en los tres loci  
15 *IL2RG*, *HBB* y *CCR5* en células K562 inducidas por Cas9 en combinación con ARNsg sintéticos. Los ARNsg sintéticos se administraron a 1 µg (sombreado claro) o 20 µg (sombreado oscuro) por 1 millón de células. Cas9 se expresó a partir de un plásmido (2 µg) y para los experimentos de RH se incluyeron 5 µg de plásmido donante que codifica GFP. Como control positivo, se utilizaron 2 µg de plásmido de ARNsg que codifica tanto el ARNsg como la proteína Cas9 (barras grises). Las barras representan valores promedio +SEM, n=3. La Figura 1E muestra la  
20 especificidad de la escisión dirigida mediada por ARNsg sintéticos como se realiza en la Figura 1C para 20 µg de ARNsg. Las frecuencias de ins/del se midieron mediante la secuenciación profunda de amplicones de PCR de los loci genómicos diana y tres loci fuera de la diana predichos bioinformáticamente para cada gen. Las barras representan valores promedio +SEM, n=3. La Figura 1F muestra la administración escalonada de 15 µg de ARNm de Cas9 y 10 µg de ARNsg sintéticos de *IL2RG* en 1 millón de células K562 usando electroporación. Las barras representan el promedio de frecuencias de ins/del +SEM, n=3 medido por análisis TIDE de amplicones de PCR que abarcan los sitios diana de ARNsg, utilizando una muestra tratada de forma simulada como control de referencia. En la Figura 1G, la proteína Cas9 formó un complejo previamente con un exceso 2,5 molar de los ARNsg de *IL2RG* sintéticos indicados y se nucleofectó en 1 millón de células K562 en las cantidades indicadas. Las frecuencias de ins/del se midieron mediante análisis TIDE como se indica anteriormente y las barras representan las frecuencias promedio de inv/del +SEM, n=3.

30 Las Figuras 2A-2C muestran que los ARNsg modificados químicamente facilitan altas tasas de alteración génica en linfocitos T primarios humanos y células madre hematopoyéticas y progenitoras CD34+ (HSPC, por sus siglas en inglés). Se nucleofectaron 1 millón de linfocitos T humanos primarios con 10 µg de los ARNsg sintéticos indicados y 15 µg de ARNm de Cas9 o 1 µg de plásmido que codifica Cas9 (Figura 2A). Se incluyó 1 µg de plásmido que codifica tanto el ARNsg como la proteína Cas9 para la comparación. Las barras representan el promedio de las frecuencias de ins/del para tres donantes diferentes +SEM, n=6 medido por análisis TIDE de amplicones de PCR que abarcan los sitios diana de ARNsg y usando una muestra de tratamiento simulado como control de referencia. En la figura 2B, los linfocitos T estimulados se nucleofectaron como anteriormente, pero con 15 µg de proteína Cas9 en complejo previo con un exceso 2,5 molar de los ARNsg de *CCR5* sintéticos indicados. Las frecuencias de ins/del se midieron mediante análisis TIDE como se ha indicado anteriormente. Las barras representan el promedio de las frecuencias de ins/del para tres donantes diferentes +SEM, n=6. En la Figura 2C, se nucleofectaron 500.000 HSPC CD34+ de sangre periférica humana movilizada con 10 µg de los ARNsg sintéticos indicados dirigidos a *IL2RG* o *HBB* y 15 µg de ARNm de Cas9 o 1 µg de plásmido Cas9. Se incluyó 1 µg de plásmido de ARNsg que codifica tanto el ARNsg como la proteína Cas9 para la comparación. Las barras representan el promedio de frecuencias de ins/del +SEM, n=3 medido mediante ensayo de escisión con endonucleasa de T7. En la Figura 2D, se nucleofectaron 1 millón de linfocitos T estimulados o HSPC CD34+ de sangre periférica humanas movilizadas con 15 µg de ARNm de Cas9 y 10 µg de los ARNsg de *CCR5* sintéticos indicados. Cuando se usa en combinación, la cantidad de cada ARNsg fue de 5 µg. Las frecuencias de ins/del para muestras con ARNsg únicos se midieron mediante análisis TIDE como se ha indicado anteriormente y las frecuencias de ins/del para muestras con dos ARNsg se midieron mediante secuenciación de los productos de PCR clonados (véanse las FIGS. 18A-B). Las barras representan el promedio de frecuencias de ins/del +SEM, n=3.

55 La Figura 3 muestra la escisión de Cas9 de dianas e ADNbc dirigidas por ARNsg modificados químicamente *in vitro*. Las barras indican el rendimiento en porcentaje de los productos de escisión de los fragmentos de ADN diana (véase la Figura 4) tratados con proteína Cas9 y ARNsg. Se muestran los valores promedio +SEM para tres síntesis independientes de cada ARNsg.

60 La Figura 4 muestra la escisión de Cas9 de dianas e ADNbc dirigidas por ARNsg modificados químicamente *in vitro*. Los productos de escisión de la escisión bioquímica de ADNbc dianas se ensayaron en DNA 7500 LabChips en un Bioanalyzer 2200. Se muestran geles representativos para cada diana y se incluyen réplicas adicionales en los resultados representados en la Figura 3. Las muestras fueron las siguientes: (L): escalera, (1): ARNsg sin modificar + ADN diana (-tratado simulado con Cas9), (2): ADN diana + proteína Cas9 (-tratado simulado con ARNsg), (3): ARNsg sin modificar + ADN diana + proteína Cas9, (4): M ARNsg + ADN diana + proteína Cas9, (5): MS ARNsg + ADN diana + proteína Cas9 y (6): MSP ARNsg + ADN diana + proteína Cas9.

- La Figura 5 ilustra la especificidad de la escisión dirigida mediada por ARNsg sintéticos. La especificidad de la diana se evaluó como en la Figura IE usando secuenciación profunda Illumina, pero con las muestras de la Figura 1C nucleofectadas con 1 µg de ARNsg. Las frecuencias de ins/del se midieron mediante secuenciación profunda de amplicones de PCR de los tres loci fuera de la diana predichos bioinformáticamente para cada ARNsg. Las barras representan valores promedio +SEM, n=3, mostrados en una escala logarítmica. Véase, la Tabla 5 con los porcentajes de ins/del.
- 5 La Figura 6 muestra la titulación de MSP ARNsg dirigido al ARNm de IL2RG y Cas9 en células K562. Las frecuencias de ins/del medidas son promedios de tres réplicas y los valores se indican en un mapa de calor. El SEM de las réplicas (n = 3) no está indicado para mayor claridad, pero son todos menores del 4 % de los valores indicados medidos. Las frecuencias de ins/del se midieron mediante análisis TIDE de amplicones de PCR que abarcan los sitios diana de ARNsg y utilizando una muestra tratada de forma simulada como control de referencia.
- 10 La Figura 7 proporciona un diagrama experimental esquemático de la administración escalonada de ARNsg y ARNm de Cas9. Resumen esquemático del experimento que produce datos para la Figura 1G. Las células K562 se nucleofectaron en los puntos de tiempo indicados con ARNm de Cas9 y/o ARNsg dirigidos a IL2RG. El ADN genómico se extrajo 72 horas después de la nucleofeción del último componente y las frecuencias de ins/del se midieron mediante TIDE utilizando una muestra tratada de forma simulada como control de referencia.
- 15 La Figura 8 compara la proteína Cas9 de diferentes proveedores. Tres días después de que 1 millón de células K562 fueran nucleofectadas con 15 µg de proteína Cas9 previamente en complejo con un exceso 2,5 molar de MS ARNsg, se extrajo el ADN genómico y se midieron las frecuencias de ins/del mediante TIDE utilizando una muestra tratada de forma simulada como control de referencia. Las barras representan promedios +SEM, n=3.
- 20 La Figura 9A-9B muestran que la especificidad de la escisión dirigida está mediada por ARNsg de IL2RG sintéticos y plásmido, ARNm o proteína Cas9. La especificidad de la diana se evaluó en la Figura IE y la Figura 5 usando secuenciación profunda de Illumina y visualizada en una escala lineal (Figura 9A) y escala logarítmica (Fig. 9B). Se nucleofectaron 1 millón de células K562 con (i) 2 µg de plásmido Cas9 + 20 µg de ARNsg, (ii) 15 µg de ARNm de Cas9 + 10 µg de ARNsg, o (iii) 15 µg de proteína Cas9 en complejo previo con 7,6 µg de ARNsg (relación 30 molar proteína:ARNsg= 1:2,5). Los resultados del plásmido Cas9 son los mismos que se muestran en la Figura IE. Las barras representan el promedio de frecuencias de ins/del +SEM, n=3.
- 25 La Figura 10 muestra altas eficiencias de nucleofeción de ARN en linfocitos T humanos primarios. Se nucleofectaron linfocitos T estimulados de tres donantes diferentes con ARNm de GFP tres días después de la estimulación. La expresión de GFP se midió tres días después de la nucleofeción mediante citometría de flujo. La expresión de GFP en las células nucleofectadas (gris) se muestra en relación con las células transfectadas de forma simulada (negro).
- 30 La Figura 11 muestra que el aumento de las cantidades de ARNsg de CCR5 y ARNm de Cas9 en la nucleofeción de los linfocitos T produjo frecuencias de ins/del similares. Se nucleofectaron linfocitos T estimulados con las cantidades indicadas de ARNsg MSP CCR5 y ARNm de Cas9. Las frecuencias de ins/del se midieron mediante análisis TIDE de amplicones de PCR que abarcan los sitios diana de ARNsg y utilizando una muestra tratada de forma simulada como control de referencia. Se muestran las frecuencias promedio de ins/del +SEM, n=6.
- 35 La Figura 12 muestra frecuencias de ins/del similares en poblaciones de linfocitos T CD4+, CD8+ y totales. Los linfocitos T estimulados se nucleofectaron con ARNsg de CCR5 MSP y ARNm de Cas9 y posteriormente se clasificaron en subpoblaciones CD4+ y CD8+. Las frecuencias de ins/del se midieron mediante TIDE y se compararon con las frecuencias de ins/del en la población general. Las barras representan el promedio de 40 frecuencias de ins/del para un donante de linfocitos T +SEM, n=8.
- 45 La Figura 13 muestra que las frecuencias de ins/del en los linfocitos T humanos primarios son estables en el tiempo. Los linfocitos T estimulados se nucleofectaron con ARNsg de CCR5 MSP y ARNm de Cas9. El ADNg se extrajo de un subconjunto de células en los puntos de tiempo indicados y las frecuencias de ins/del se midieron mediante análisis TIDE de amplicones de PCR que abarcan el sitio diana del ARNsg y usando una muestra tratada de forma 50 simulada como control de referencia. Las frecuencias promedio de ins/del se muestran para tres donantes de linfocitos T diferentes +/-SEM, n=6.
- 55 La Figura 14 muestra frecuencias más bajas de muerte celular en linfocitos T nucleofectados con ARNsg sintéticos y ARNm de Cas9 en comparación con el plásmido Cas9. Se nucleofectaron 1 millón de linfocitos T estimulados con 10 µg de los ARNsg sintéticos indicados y 15 µg de ARNm de Cas9 o 1 µg de plásmido que codifica Cas9. Se incluyó 1 µg de plásmido que codifica tanto el ARNsg de CCR5 como la proteína Cas9 para comparación (plásmido de ARNsg). Tres días después de la nucleofeción, las células se tiñeron con la tinción de células VIVAS/MUERTAS. Las barras representan los porcentajes promedio de células muertas para tres donantes de linfocitos T diferentes +SEM, n=6.
- 60 La Figura 15 muestra los resultados de un ensayo de proliferación después de la nucleofeción de ARNsg sintéticos

en linfocitos T. Los linfocitos T estimulados de dos donantes diferentes se nucleofectaron el día 0 y se controló la proliferación celular usando el ensayo CellTiter Glo. El SEM de las réplicas no está indicado para mayor claridad, pero son todos menores del 15 % de los valores indicados.

- 5 La Figura 16 muestra la rotura de CCR5 en linfocitos T no estimulados. Se nucleofectaron linfocitos T humanos no estimulados de tres donantes diferentes el día del aislamiento con el ARNsg MS y el ARNm de Cas9. El ADNg se extrajo tres días después de la nucleofeción y las frecuencias de ins/del se midieron mediante TIDE utilizando una muestra tratada de forma simulada como control de referencia. Las barras representan el promedio +SEM, n=2.
- 10 La Figura 17 ilustra las frecuencias de ins/del en HSPC PB CD34+ movilizadas para *IL2RG* y *HBB*. Tres días después de la nucleofeción de HSPC CD34+, se extrajo el ADN genómico y se midieron las frecuencias de ins/del mediante el ensayo T7. Se muestra un gel representativo de tres réplicas biológicas para cada *IL2RG* y *HBB*. +SEM, n=2.
- 15 Las Figuras 18A-B muestran altas frecuencias de modificación del gen *CCR5* en linfocitos T humanos primarios y HSPC CD34+ usando dos ARNsg. Los linfocitos T estimulados de tres donantes diferentes y las HSPC CD34+ movilizadas con PB se nucleofectaron por triplicado con el ARNsg 'D' y 'Q' junto con ARNm Cas9. El ADNg se extrajo tres días después de la nucleofeción y la región modificada de *CCR5* se amplificó mediante PCR usando un par de cebadores que generaron un amplicón de 2,1 kb para los alelos sin modificar (Figura 18A). Los amplicones de PCR se subclonaron en un plásmido para su transformación y se secuenciaron colonias individuales que representan alelos individuales, referenciados respecto a la secuencia genómica esperada y clasificados de acuerdo con el genotipo alélico (Figura 18B).
- 20 La Figura 19 muestra que los ARNsg modificados con MS funcionan mejor que los ARNsg sin modificar en las HSPC CD34+. Las HSPC CD34+ se nucleofectaron con 30 µg de proteína Cas9 en complejo con un exceso 2,5 molar de los ARNsg de *HBB* sintéticos indicados. Las frecuencias de Ins/del se midieron mediante análisis TIDE cuatro días después de la nucleofeción. Las barras representan las frecuencias promedio de ins/del para tres donantes diferentes +SEM, n=3. \*\* p < 0,01, Prueba t de Student.
- 25 La Figura 20 muestra que los ARNsg modificados pueden usarse para una edición eficiente del genoma multiplexado. 1 millón de células K562 se nucleofectaron con 15 µg de ARNm de Cas9 y 5 µg de ARNsg de *CCR5*, *HBB* o *IL2RG* modificado con MS o los tres ARNsg (multiplexados) (3x5 µg). Las frecuencias de Ins/del se midieron mediante análisis TIDE en cada uno de los tres loci tres días después de la nucleofeción. Las barras representan las frecuencias promedio de ins/del +SEM, n=3. \*\*\* p < 0,001, n.s. = p ≥ 0,05, Prueba t de Student.
- 30 La Figura 21 muestra productos de PCR que abarcan el sitio diana *CCR5* después de la recombinación homóloga usando ARNsg de *CCR5* químicamente modificados, ARNm de Cas9 y un ssODN de *CCR5*.

#### Descripción detallada

##### I. Introducción

En el presente documento se proporcionan métodos para la edición del genoma basado en CRISPR/Cas y/o la modulación de la expresión génica en una célula *in vitro* (por ejemplo, una célula primaria para su uso en terapia *ex vivo*). En particular, los métodos proporcionados en el presente documento utilizan ARN de guía único modificados químicamente (ARNsg) que tienen una actividad mejorada durante la regulación génica (por ejemplo, editando el genoma, inhibición de la expresión génica y activación de la expresión génica) en comparación con los ARNsg sin modificar correspondientes. En determinados aspectos, la presente invención proporciona métodos *in vitro* para la regulación génica de un ácido nucleico diana en una célula mediante la introducción de un ARNsg químicamente modificado que hibrida con el ácido nucleico diana junto con una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento del mismo, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. En otros aspectos determinados, la presente divulgación proporciona métodos para prevenir o tratar una enfermedad genética en un sujeto mediante la administración de una cantidad suficiente del ARNsg químicamente modificado para corregir una mutación genética asociada con la enfermedad.

##### II. Aspectos generales

La práctica de la presente invención emplea, a no ser que se indique de otro modo, técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, genómica y ADN recombinante, que están dentro de la experiencia de la materia. Véase Sambrook, Fritsch y Maniatis, Molecular Cloning: A Laboratory Manual, 2<sup>a</sup> edición, 1989), Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds., (1987)), la serie Methods in Enzymology (Academic Press, Inc.); PCR 2: A Practical Approach (M. J. MacPherson, B. D. Hames y G. R. Taylor eds. (1995)), Harlow y Lane, eds. (1988) Antibodies, A Laboratory Manual, and Animal Cell Culture (R. I. Freshney, ed. (1987)).

Los oligonucleótidos que no están disponibles comercialmente se pueden sintetizar químicamente, por ejemplo, según el método de triéster de fosforamidita en fase sólida descrito por primera vez por Beaucage y Caruthers, Tetrahedron Lett. 22:1859-1862 (1981), utilizando un sintetizador automático, como se describe en Van Devanter et. al., Nucleic Acids Res. 12:6159-6168 (1984). La purificación de oligonucleótidos se realiza utilizando cualquier estrategia reconocida en la técnica, por ejemplo, electroforesis en gel de acrilamida nativa o cromatografía líquida de alto rendimiento (HPLC) de intercambio aniónico como se describe en Pearson y Reanier, J. Chrom. 255: 137-149 (1983).

10 **III. Definiciones**

A menos que se indique específicamente lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el que entienden comúnmente los expertos habituales en la materia a la que pertenece la presente invención. Además, cualquier método o material similar o equivalente a un método o material descrito en el presente documento puede usarse en la práctica de la presente invención. Para los fines de la presente invención, se definen los siguientes términos.

Los términos "un", "uno/a", o "el/la" como se usa en el presente documento no solo incluye aspectos con un miembro, sino que también incluye aspectos con más de un miembro. Por ejemplo, la forma singular "un", "uno/a", y "el/la" incluyen las referencias plurales salvo que el contexto indique claramente otra cosa. Por lo tanto, por ejemplo, la referencia a "una célula" incluye una pluralidad de dichas células y la referencia a "el agente" incluye una referencia a uno o más agentes conocidos por los expertos en la técnica, y así sucesivamente.

25 La expresión "célula primaria" se refiere a una célula aislada directamente de un organismo multicelular. Las células primarias típicamente han experimentado muy pocas duplicaciones de población y, por lo tanto, son más representativas del componente funcional principal del tejido del que derivan en comparación con las líneas celulares continuas (tumorales o inmortalizadas artificialmente). En algunos casos, las células primarias son células que se han aislado y luego se han utilizado inmediatamente. En otros casos, las células primarias no pueden dividirse indefinidamente y, por lo tanto, no pueden cultivarse durante largos períodos de tiempo *in vitro*.

30 La expresión "edición del genoma" se refiere a un tipo de ingeniería genética en la que el ADN se inserta, se reemplaza o se elimina de un ADN diana, por ejemplo, el genoma de una célula, utilizando una o más nucleasas y/o nickasas. Las nucleasas crean roturas específicas de doble cadena (DSB) en ubicaciones deseadas en el genoma y aprovechan los mecanismos endógenos de la célula para reparar la rotura inducida por reparación dirigida por homología (HDR, por sus siglas en inglés) (por ejemplo, recombinación homóloga) o por unión de extremos no homólogos (NHEJ). Las nickasas crean roturas específicas de una sola cadena en las ubicaciones deseadas del genoma. En un ejemplo no limitante, se pueden usar dos nickasas para crear dos roturas de una sola cadena en cadenas opuestas de un ADN diana, generando así un final romo o pegajoso. Se puede introducir cualquier nucleasa adecuada en una célula para inducir la edición del genoma de una secuencia de ADN diana que incluye, pero sin limitaciones, proteína asociada a CRISPR

45 (Cas) nucleasas, nucleasas de dedo de cinc (ZFN), nucleasas efectoras de tipo activador de la transcripción (TALEN), meganucleasas, otras endonucleasas o exonucleasas, variantes de las mismas, fragmentos de las mismas y combinaciones de las mismas. En realizaciones particulares, la edición del genoma mediada por nucleasas de una secuencia de ADN diana puede ser "inducida" o "modulada" (por ejemplo, mejorada) usando los ARN guía únicos (ARNsg) modificados descritos en el presente documento en combinación con nucleasas Cas (por ejemplo, polipéptidos Cas9 o ARNm de Cas9), por ejemplo, para mejorar la eficiencia de la edición precisa del genoma mediante la reparación dirigida por homología (HDR).

50 La expresión "reparación dirigida por homología" o "HDR" se refiere a un mecanismo en las células para reparar con precisión y exactitud las roturas del ADN de doble cadena utilizando un molde homólogo para guiar la reparación. La forma más común de HDR es la recombinación homóloga (RH), un tipo de recombinación genética en la que las secuencias de nucleótidos se intercambian entre dos moléculas de ADN similares o idénticas.

55 La expresión "unión de extremos no homólogos" o "NHEJ" se refiere a una ruta que repara roturas de ADN de doble cadena en las que los extremos rotos se ligan directamente sin necesidad de un molde homólogo.

60 La expresión "ácido nucleico", "nucleótido", o "polinucleótido" se refiere a ácidos desoxirribonucleicos (ADN), ácidos ribonucleicos (ARN) y polímeros de los mismos en forma monocatenaria, bicatenaria o multicatenaria. El término incluye, pero sin limitaciones, ADN o ARN monocatenario, bicatenario o multicatenario, ADN genómico, ADNc, híbridos de ADN-ARN o un polímero que comprende bases púricas y/o pirimidínicas u otras bases nucleotídicas naturales, químicamente modificadas, bioquímicamente modificadas, no naturales, sintéticas o derivatizadas. En algunas realizaciones, un ácido nucleico puede comprender una mezcla de ADN, ARN y análogos de los mismos. A menos que se limite específicamente, el término abarca ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares a las del ácido nucleico de referencia y se metabolizan de manera similar a los nucleótidos de origen natural. A menos que se indique otra cosa, una secuencia

- de ácido nucleico particular también abarca implícitamente variantes modificadas conservadoramente de la misma (por ejemplo, sustituciones de codones degenerados), alelos, ortólogos, polimorfismos de un solo nucleótido (SNP) y secuencias complementarias, así como la secuencia indicada explícitamente. En concreto, se pueden conseguir sustituciones de codones degeneradas generando secuencias en las que la tercera posición de uno o más (o todos) codones seleccionados está sustituida con restos de base mixta y/o desoxiinosina (Batzer *et al.*, Nucleic Acids Res. 19: 5081 (1991); Ohtsuka *et al.*, J. Biol. Chem. 260:2605-2608 (1985); y Rossolini *et al.*, Mol. Cell. Probes 8:91-98 (1994)). La expresión ácido nucleico se usa de forma indistinta con gen, ADNc y ARNm codificado por un gen.
- 5 La expresión "análogo nucleotídico" o "nucleótido modificado" se refiere a un nucleótido que contiene una o más modificaciones químicas (por ejemplo, sustituciones), en o sobre la base nitrogenada del nucleósido (por ejemplo, citosina (C), timina (T) o uracilo (U), adenina (A) o guanina (G)), en o sobre el resto de azúcar del nucleósido (por ejemplo, ribosa, desoxirribosa, ribosa modificada, desoxirribosa modificada, análogo de azúcar de seis miembros o análogo de azúcar de cadena abierta), o el fosfato.
- 10 15 El término/expresión "gen" o "secuencia de nucleótidos que codifica un polipéptido" significa el segmento de ADN implicado en la producción de una cadena polipeptídica. El segmento de ADN puede incluir regiones que preceden y siguen a la región codificante (líder y tráiler) implicadas en la transcripción/traducción del producto génico y la regulación de la transcripción/traducción, así como secuencias intermedias (intrones) entre segmentos de codificación individuales (exones).
- 20 25 30 35 40 45 50 55 60 65 Los términos "polipéptido", "péptido", y "proteína" se usan de forma intercambiable en el presente documento para hacer referencia a un polímero de restos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en donde uno o más restos de aminoácidos es un mimético químico artificial de un aminoácido de origen natural correspondiente, así como de polímeros de aminoácidos de origen natural y de polímeros de aminoácidos de origen no natural. Como se usa en el presente documento, los términos abarcan cadenas de aminoácidos de cualquier longitud, incluyendo proteínas de longitud completa, en donde los restos de aminoácidos están unidos por enlaces covalentes.
- El término "variante" se refiere a una forma de organismo, cepa, gen, polinucleótido, polipéptido, o característica que se desvía de lo que ocurre en la naturaleza.
- 30 35 40 45 50 55 60 65 El término "complementariedad" se refiere a la capacidad de un ácido nucleico para formar enlace(s) de hidrógeno con otra secuencia de ácido nucleico mediante tipos de Watson-Crick tradicionales u otros tipos no tradicionales. Un porcentaje de complementariedad indica el porcentaje de restos en una molécula de ácido nucleico que puede formar enlaces de hidrógeno (por ejemplo, emparejamiento de bases de Watson-Crick) con una segunda secuencia de ácido nucleico (por ejemplo, siendo 5, 6, 7, 8, 9, 10 de cada 10 un 50 %, 60 %, 70 %, 80 %, 90 % y 100 % complementarias). "Perfectamente complementario" significa que todos los restos contiguos de una secuencia de ácido nucleico formarán un puente de hidrógeno con el mismo número de restos contiguos en una segunda secuencia de ácido nucleico. "Sustancialmente complementaria", como se usa en el presente documento, se refiere a un grado de complementariedad que es al menos del 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 %, 99 % o 100 % en una región de 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 o más nucleótidos, o se refiere a dos ácidos nucleicos que se hibridan en condiciones rigurosas.
- 30 35 40 45 50 55 60 65 La expresión "condiciones rigurosas" para la hibridación se refiere a las condiciones en las que un ácido nucleico que tiene complementariedad con una secuencia diana hibrida predominantemente con la secuencia diana y sustancialmente no hibrida con secuencias no diana. Las condiciones rigurosas generalmente dependen de la secuencia y varían según una serie de factores. En general, cuanto más larga sea la secuencia, mayor es la temperatura a la que la secuencia se hibrida específicamente con su secuencia diana. Los ejemplos no limitantes de condiciones rigurosas se describen con detalle en Tijssen (1993), Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Part 1, Segundo capítulo "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y.
- 30 35 40 45 50 55 60 65 El término "hibridación" se refiere a una reacción en la que uno o más polinucleótidos reaccionan para formar un complejo que se estabiliza mediante enlaces de hidrógeno entre las bases de los restos de nucleótidos. El enlace de hidrógeno puede ocurrir por emparejamiento de bases de Watson Crick, unión de Hoogstein o de cualquier otra manera específica de secuencia. El complejo puede comprender dos cadenas que forman una estructura dúplex, tres o más cadenas que forman un complejo multicatenario, una sola cadena autohibridante o cualquier combinación de estas.
- 30 35 40 45 50 55 60 65 Un "vector de expresión recombinante" es una construcción de ácido nucleico, generada de forma recombinante o sintética, con una serie de elementos de ácido nucleico especificados que permiten la transcripción de una secuencia polinucleotídica particular en una célula hospedadora. Un vector de expresión puede ser parte de un plásmido, genoma viral o fragmento de ácido nucleico. Típicamente, un vector de expresión incluye un polinucleótido a transcribir, unido operativamente a un promotor. "Unido operativamente" en este contexto significa dos o más elementos genéticos, tal como una secuencia codificante de polinucleótidos y un promotor, colocados en posiciones relativas que permiten el correcto funcionamiento biológico de los elementos, tal como el promotor que dirige la

- transcripción de la secuencia codificante. El término "promotor" se usa en el presente documento para hacer referencia a una serie de secuencias de control de ácido nucleico que dirigen la transcripción de un ácido nucleico. Como se usa en el presente documento, un promotor incluye las secuencias de ácidos nucleicos necesarias próximas al sitio de inicio de la transcripción, tal como, en el caso de un promotor de la polimerasa de tipo II, un elemento TATA. Un promotor incluye también opcionalmente un potenciador distal o elementos represores, que se pueden localizar tanto como varios miles de pares de bases desde el sitio de inicio de la transcripción. Otros elementos que pueden estar presentes en un vector de expresión incluyen aquellos que mejoran la transcripción (por ejemplo, potenciadores) y terminan la transcripción (por ejemplo, terminadores), así como aquellos que confieren cierta afinidad o antigenicidad de unión a la proteína recombinante producida a partir del vector de expresión.
- "Recombinante" se refiere a un polinucleótido, polipéptido, célula, tejido u organismo modificado genéticamente. Por ejemplo, un polinucleótido recombinante (o una copia o complementario de un polinucleótido recombinante) es uno que ha sido manipulado usando métodos bien conocidos. Un casete de expresión recombinante que comprende un promotor unido operativamente a un segundo polinucleótido (por ejemplo, una secuencia codificante) puede incluir un promotor que es heterólogo del segundo polinucleótido como resultado de la manipulación humana (por ejemplo, mediante métodos descritos en Sambrook *et al.*, Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, (1989) o Current Protocols in Molecular Biology Volumes 1-3, John Wiley & Sons, Inc. (1994-1998)). Un casete de expresión recombinante (o vector de expresión) típicamente comprende polinucleótidos en combinaciones que no se encuentran en la naturaleza. Por ejemplo, los sitios de restricción manipulados por humanos o las secuencias del vector plasmídico pueden flanquear o separar el promotor de otras secuencias. Una proteína recombinante es aquella que se expresa a partir de un polinucleótido recombinante y las células, tejidos y organismos recombinantes son aquellos que comprenden secuencias recombinantes (polinucleótidos y/o polipéptidos).
- La expresión "polimorfismo de un solo nucleótido" o "SNP" se refiere a un cambio de un solo nucleótido con un polinucleótido, incluso dentro de un alelo. Esto puede incluir el reemplazo de un nucleótido por otro, así como la delección o inserción de un solo nucleótido. Lo más típicamente, los SNP son marcadores bialélicos, aunque también pueden existir marcadores trialélicos y tetraalélicos. A modo de ejemplo no limitante, una molécula de ácido nucleico que comprende SNP A \ C puede incluir una C o A en la posición polimórfica.
- Los términos "cultivo", "cultivar", "crecer", "crecimiento", "mantener", "que mantiene" "expandir", "en expansión", *etc.*, cuando hacen referencia al propio cultivo celular o al proceso de cultivo, se pueden usar indistintamente para significar que una célula (por ejemplo, una célula primaria) se mantiene fuera de su ambiente normal en condiciones controladas, por ejemplo, en condiciones adecuadas para la supervivencia. Se permite que las células cultivadas sobrevivan y el cultivo puede dar como resultado crecimiento celular, estasis, diferenciación o división. El término no implica que todas las células del cultivo sobreviven, crecen o se dividen, ya que algunas pueden morir o envejecer naturalmente. Las células se cultivan típicamente en medios, que se pueden cambiar durante el curso del cultivo.
- Los términos "sujeto", "paciente", e "individuo" se utilizan en el presente documento indistintamente para incluir un ser humano o un animal. Por ejemplo, el sujeto puede ser un mamífero, un primate (por ejemplo, un mono), un animal de ganado (por ejemplo, un caballo, una vaca, una oveja, un cerdo o una cabra), un animal de compañía (por ejemplo, un perro, un gato), un animal de ensayo de laboratorio (por ejemplo, un ratón, una rata, una cobaya, un ave), un animal de importancia veterinaria o un animal de importancia económica.
- Como se usa en el presente documento, el término "administrar" incluye administración oral, contacto tópico, administración como suppositorio, administración intravenosa, intraperitoneal, intramuscular, intralesional, intratecal, intranasal o subcutánea a un sujeto. La administración es por cualquier vía, incluyendo parenteral y transmucosa (por ejemplo, bucal, sublingual, palatal, gingival, nasal, vaginal, rectal o transdérmica). La administración parenteral incluye, por ejemplo, intravenosa, intramuscular, intraarteriolar, intradérmica, subcutánea, intraperitoneal, intraventricular e intracranial. Otros modos de administración incluyen, pero sin limitaciones, el uso de formulaciones liposomales, infusión intravenosa, parches transdérmicos, *etc.*
- El término "tratar" se refiere a un abordaje para obtener resultados beneficiosos o deseados que incluyen, entre otros, un beneficio terapéutico y/o un beneficio profiláctico. Por beneficio terapéutico se entiende cualquier mejora o efecto terapéuticamente relevante en una o más enfermedades, afecciones o síntomas en tratamiento. Para beneficio profiláctico, las composiciones pueden administrarse a un sujeto con riesgo de desarrollar una enfermedad particular, afección o síntoma, o a un sujeto que informa uno o más de los síntomas fisiológicos de una enfermedad, aunque la enfermedad, afección o síntoma puede no haberse manifestado todavía.
- La expresión "cantidad eficaz" o "cantidad suficiente" se refiere a la cantidad de un agente (por ejemplo, nucleasa Cas, ARN guía único modificado, *etc.*) que sea suficiente para lograr resultados beneficiosos o deseados. La cantidad terapéuticamente eficaz puede variar dependiendo de uno o más de: el sujeto y la enfermedad que se está tratando, el peso y la edad del sujeto, la gravedad de la patología, de la forma de administración y similares, que un experto en la técnica puede determinar con facilidad. La cantidad específica puede variar dependiendo de uno o más de: el agente particular elegido, el tipo de célula diana, la ubicación de la célula diana en el sujeto, el régimen de dosificación a seguir, si se administra junto con otros agentes, el momento de la administración y el sistema físico de

administración en el que se realiza.

La expresión "vehículo farmacéuticamente aceptable" se refiere a una sustancia que ayuda a la administración de un agente (por ejemplo, nucleasa Cas, ARN guía único modificado, etc.) a una célula, un organismo o un sujeto.

5 "Vehículo farmacéuticamente aceptable" se refiere a un vehículo o excipiente que puede incluirse en una composición o formulación y que no causa ningún efecto toxicológico adverso significativo en el paciente. Los ejemplos no limitantes de vehículos farmacéuticamente aceptables incluyen agua, NaCl, soluciones salinas normales, Ringer lactato, sacarosa normal, glucosa normal, aglutinantes, cargas, disgregantes, lubricantes, recubrimientos, edulcorantes, aromas y colores, y similares. Un experto en la materia reconocerá que son útiles 10 otros vehículos farmacéuticos en la presente invención.

La expresión "estabilidad creciente", con respecto a los componentes del sistema CRISPR, se refiere a modificaciones que estabilizan la estructura de cualquier componente molecular del sistema CRISPR. La expresión incluye modificaciones que disminuyen, inhiben, bajan o reducen la degradación de cualquier componente molecular 15 del sistema CRISPR.

La expresión "especificidad creciente", con respecto a los componentes del sistema CRISPR, se refiere a modificaciones que aumentan la actividad específica (por ejemplo, la actividad en la diana) de cualquier componente molecular del sistema CRISPR. La expresión incluye modificaciones que disminuyen, inhiben, bajan o reducen la actividad no específica (por ejemplo, la actividad fuera de la diana) de cualquier componente molecular del sistema 20 CRISPR.

La expresión "disminución de la toxicidad" con respecto a los componentes del sistema CRISPR, se refiere a modificaciones que disminuyen, inhiben, bajan o reducen el efecto tóxico de cualquier componente molecular del sistema 25 CRISPR en una célula, organismo, sujeto y similares.

La expresión "actividad potenciada", con respecto a los componentes del sistema CRISPR y en el contexto de la regulación génica, se refiere a un aumento o mejora de la eficiencia y/o la frecuencia de inducción, modulación, regulación o control de la edición del genoma y/o la expresión génica.

30 35 El término "aproximadamente" en relación con un valor numérico de referencia también puede incluir un intervalo de valores más o menos 10 % de dicho valor. Por ejemplo, la cantidad "aproximadamente 10" incluye cantidades de 9 a 11, incluyendo los números de referencia 9, 10 y 11. El término "aproximadamente" en relación con un valor numérico de referencia también puede incluir un intervalo de valores más o menos 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 % de ese valor.

#### IV. Descripción de las realizaciones

40 La presente invención proporciona métodos *in vitro* para inducir la regulación génica de un ácido nucleico diana en una célula. La invención incluye el uso de ARN guía únicos (ARNsg) modificados que mejoran la edición del genoma y/o la inhibición o activación de la expresión génica de un ácido nucleico diana en una célula primaria (por ejemplo, cultivada *in vitro* para su uso en terapia *ex vivo*). La presente divulgación también se refiere a métodos para prevenir o tratar una enfermedad en un sujeto mejorando la edición precisa del genoma para corregir una mutación en un gen diana asociado con la enfermedad. La presente invención se usa con células primarias y en cualquier locus génico 45 que sea susceptible de tecnología de edición del genoma mediada por nucleasas.

50 En un primer aspecto, la presente invención proporciona un método *in vitro* para la inducción (por ejemplo, inicio, modulación, potenciación, etc.) de la regulación génica de un ácido nucleico diana en una célula primaria, comprendiendo el método:

55 En algunas realizaciones, la actividad mejorada comprende una mayor estabilidad del ARNsg modificado y/o una mayor especificidad del ARNsg modificado para el ácido nucleico diana.

En algunas realizaciones, el ácido nucleico diana comprende un ADN diana o un ARN diana. La regulación génica 60 de un ácido nucleico diana abarca cualquier mecanismo utilizado por las células para aumentar o disminuir la producción de un producto génico específico (por ejemplo, proteína o ARN) mediante la introducción en la célula primaria:

(a) un ARN guía único (ARNsg) modificado que comprende una primera secuencia de nucleótidos que es complementaria al ácido nucleico diana y una segunda secuencia de nucleótidos que interactúa con un polipéptido de proteína asociada a CRISPR (Cas), en donde uno o más de los nucleótidos en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos son nucleótidos modificados, en donde los nucleótidos modificados se seleccionan de entre el grupo que consiste en un nucleótido de 2'-O-metil 3'-fosforotioato (MS), un nucleótido de 2'-O-metil 3'-tiofosfatoacetato (MSP) y una combinación de los mismos; y 65 (b) un polipéptido Cas, un ARNm que codifica un polipéptido Cas y/o un vector de expresión recombinante que

comprende una secuencia de nucleótidos que codifica un polipéptido Cas,

- en donde el ARNsg modificado guía el polipéptido Cas hacia el ácido nucleico diana, y en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana con una actividad mejorada en relación con un ARNsg sin modificar correspondiente y en donde la actividad mejorada comprende un aumento o mejora en la eficiencia y/o la frecuencia de inducción, modulación, regulación o control de la edición del genoma y/o la expresión génica. ácido nucleico diana e incluye la edición del genoma del ácido nucleico diana o la modulación (por ejemplo, inhibición o activación) de la expresión génica del ácido nucleico diana. En algunos casos, la regulación génica comprende la edición del genoma del ADN diana. La edición del genoma puede ser reparación dirigida por homólogos (HDR) o unión de extremos no homólogos (NHEJ) del ADN diana. En otros casos, la regulación génica comprende modular (por ejemplo, inhibir o activar) la expresión génica del ADN diana o del ARN diana usando un polipéptido Cas deficiente en endonucleasas.
- 15 En algunas realizaciones, el método comprende además introducir un molde de reparación de un donante recombinante en la célula primaria. En determinados casos, el molde de reparación del donante recombinante comprende dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de una secuencia de nucleótidos correspondiente al ADN diana para someterse a la edición del genoma. En otros casos, el molde de reparación del donante recombinante comprende un molde sintético de oligodesoxinucleótido monocatenario (ssODN) que comprende una secuencia de nucleótidos que codifica una mutación para corregir un polimorfismo de un solo nucleótido (SNP) y dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de la secuencia de nucleótidos que codifica la mutación.
- 20 25 En algunas realizaciones, la célula primaria se aísla de un organismo multicelular antes de introducir el ARNsg modificado y el polipéptido Cas en la célula primaria. El organismo multicelular puede ser una planta, un protista multicelular, un hongo multicelular o un animal, tal como un mamífero (por ejemplo, un ser humano). En determinados casos, la célula principal se selecciona de entre el grupo que consiste en una célula madre, una célula inmunitaria y una combinación de las mismas. Los ejemplos no limitantes de células madre incluyen células madre y progenitoras hematopoyéticas (HSPC) tales como HSPC CD34+, células madre mesenquimatosas, células madre neuronales, células madre de órganos y combinaciones de las mismas. Los ejemplos no limitantes de células inmunes incluyen linfocitos T (por ejemplo, linfocitos T CD3+, linfocitos T CD4+, linfocitos T CD8+, células infiltrantes de tumores (TIL), linfocitos T de memoria, linfocitos T progenitores de memoria, linfocitos T efectores), linfocitos citolíticos naturales, monocitos, células mononucleares de sangre periférica (PBMC, por sus siglas en inglés), linfocitos de sangre periférica (PBL) y combinaciones de los mismos. En otras realizaciones, la célula primaria o una progenie de la misma (por ejemplo, una célula derivada de la célula primaria) se devuelve (por ejemplo, administrada a través de cualquier sistema y vía de administración aceptables) al organismo multicelular (por ejemplo, ser humano) después de introducir el ARNsg modificado y el polipéptido Cas en la célula primaria.
- 30 35 40 45 50 55 60 65 66 En algunas realizaciones, la célula primaria comprende una población de células primarias. En algunos casos, el ARNsg modificado induce la regulación génica del ácido nucleico diana en al menos aproximadamente un 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o 95 % de la población de células primarias. En otros casos, la población de células primarias comprende al menos aproximadamente  $10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8$  o  $10^9$  células primarias. En determinados casos, la regulación génica comprende la edición del genoma (por ejemplo, HDR o NHEJ) del ADN diana en la población de células primarias. En otros casos determinados, la regulación génica comprende modular (por ejemplo, inhibir o activar) la expresión génica del ADN diana o el ARN diana en la población de células primarias usando un polipéptido Cas deficiente en endonucleasa. Como un ejemplo no limitante, el ARNsg modificado puede inducir HDR (por ejemplo, frecuencias de ins/del) en al menos aproximadamente el 30 %, 35 %, 40 %, 45 %, 50 %, 55 % o 60 % de una población de linfocitos T primarios después de introducir un ARNsg modificado en los linfocitos T primarios con un polipéptido Cas (por ejemplo, como un complejo de ribonucleoproteína (RNP)) o un ARNm que codifica un polipéptido Cas. Como otro ejemplo no limitante, el ARNsg modificado puede inducir HDR (por ejemplo, frecuencias de ins/del) en al menos aproximadamente el 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 % o 60 % de una población de células madre y progenitoras hematopoyéticas primarias (HSPC) después de introducir un ARNsg modificado en las HSPC primarias con un polipéptido Cas (por ejemplo, como un complejo de RNP) o un ARNm que codifica un polipéptido Cas.
- 60 65 66 El uno o más nucleótidos modificados en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos del ARNsg modificado comprenden nucleótidos de 2'-O-metil 3'-fosforotioato (MS), nucleótidos de 2'-O-metil 3'-tioPACE (MSP) o combinaciones de los mismos. En algunos casos, el ARNsg modificado incluye uno o más nucleótidos de MS en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos. En algunos casos, el ARNsg modificado incluye uno o más nucleótidos de MSP en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos. En algunos casos, el ARNsg modificado incluye uno o más nucleótidos de MS y uno o más nucleótidos de MSP en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos. En algunos casos, el ARNsg modificado no incluye M nucleótidos en la primera secuencia de nucleótidos y/o la segunda

secuencia de nucleótidos. En algunos casos, el ARNsg modificado incluye uno o más nucleótidos de MS y/o uno o más nucleótidos de MSP como los nucleótidos modificados en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos. En otros casos, el ARNsg modificado incluye uno o más nucleótidos de MS y/o uno o más nucleótidos de MSP en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos y puede incluir además uno o más nucleótidos de M en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos.

5 En algunas realizaciones, la primera secuencia de nucleótidos del ARNsg modificado tiene una longitud de aproximadamente 20 nucleótidos. En algunos casos, al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de los nucleótidos en la primera secuencia de nucleótidos son nucleótidos modificados. En determinados casos, aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 de los nucleótidos en la primera secuencia de nucleótidos (por ejemplo, una primera secuencia de nucleótidos de aproximadamente 20 nucleótidos de longitud) son nucleótidos modificados. En otros casos, todos los nucleótidos de la primera secuencia de nucleótidos (por ejemplo, una primera secuencia de nucleótidos de aproximadamente 20 nucleótidos de longitud) son nucleótidos modificados. En algunos casos, los nucleótidos modificados están ubicados en el extremo 5' (por ejemplo, el nucleótido terminal en el extremo 5') o cerca del extremo 5' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del nucleótido terminal en el extremo 5') de la primera secuencia de nucleótidos y/o en posiciones internas dentro de la primera secuencia de nucleótidos. En otros casos, de aproximadamente el 10 % a 20

En algunas realizaciones, la segunda secuencia de nucleótidos del ARNsg modificado tiene una longitud de aproximadamente 80 nucleótidos. En algunos casos, al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de los nucleótidos en la segunda secuencia de nucleótidos son nucleótidos modificados. En determinados casos, aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 de los nucleótidos en la segunda secuencia de nucleótidos (por ejemplo, una segunda secuencia de nucleótidos de aproximadamente 80 nucleótidos de longitud) son nucleótidos modificados. En otros casos, todos los nucleótidos de la segunda secuencia de nucleótidos (por ejemplo, una segunda secuencia de nucleótidos de aproximadamente 80 nucleótidos de longitud) son nucleótidos modificados. En algunos casos, los nucleótidos modificados están ubicados en el extremo 3' (por ejemplo, el nucleótido terminal en el extremo 3') o cerca del extremo 3' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del extremo 3') de la segunda secuencia de nucleótidos y/o en posiciones internas dentro de la segunda secuencia de nucleótidos. En otros casos, de aproximadamente el 1 % a 35

30 En algunas realizaciones, el ARNsg modificado comprende uno, dos o tres nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 5' (por ejemplo, el nucleótido terminal en el extremo 5') o cerca del extremo 5' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del nucleótido terminal en el extremo 5') de la primera secuencia de nucleótidos y uno, dos o tres nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 3' (por ejemplo, el nucleótido terminal en el extremo 3') o cerca del extremo 3' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del extremo 3') de la segunda secuencia de nucleótidos.

40 En algunos casos, el ARNsg modificado comprende un nucleótido modificado en el extremo 5' (por ejemplo, el nucleótido terminal en el extremo 5') o cerca del extremo 5' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del nucleótido terminal en el extremo 5') de la primera secuencia de nucleótidos y un nucleótido modificado en el extremo 3' (por ejemplo, el nucleótido terminal en el extremo 3') o cerca del extremo 3' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del extremo 3') de la segunda secuencia de nucleótidos.

45 En otros casos, el ARNsg modificado comprende dos nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 5' (por ejemplo, el nucleótido terminal en el extremo 5') o cerca del extremo 5' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del nucleótido terminal en el extremo 5') de la primera secuencia de nucleótidos y dos nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 3' (por ejemplo, el nucleótido terminal en el extremo 3') o cerca del extremo 3' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del extremo 3') de la segunda secuencia de nucleótidos.

55 En otros casos más, el ARNsg modificado comprende tres nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 5' (por ejemplo, el nucleótido terminal en el extremo 5') o cerca del extremo 5' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del nucleótido terminal en el extremo 5') de la primera secuencia de nucleótidos y tres nucleótidos modificados consecutivos o no consecutivos que comienzan en el extremo 3' (por ejemplo, el nucleótido terminal en el extremo 3') o cerca del extremo 3' (por ejemplo, dentro de 1, 2, 3, 4 o 5 nucleótidos del extremo 3') de la segunda secuencia de nucleótidos.

60 En realizaciones particulares, el ARNsg modificado comprende tres nucleótidos modificados consecutivos en el extremo 5' de la primera secuencia de nucleótidos y tres nucleótidos modificados consecutivos en el extremo 3' de la segunda secuencia de nucleótidos.

- En algunas realizaciones, el ARNsg modificado se sintetiza químicamente. En otras realizaciones, el método para inducir la regulación génica comprende la regulación génica multiplexada (por ejemplo, edición del genoma o modulación de la expresión génica) de una única secuencia de ácido nucleico diana o diferentes secuencias de ácido nucleico diana usando una pluralidad de ARNsg modificados. En realizaciones particulares, la regulación génica multiplexada es más eficiente y/o consistente en relación con el uso de los correspondientes ARNsg sin modificar. En determinados casos, la pluralidad de ARNsg modificados comprende al menos dos, tres, cuatro, cinco, diez, quince, veinte o más ARNsg modificados diferentes, en donde cada ARNsg modificado está dirigido a un ácido nucleico diana diferente. En otros casos, la pluralidad de ARNsg modificados comprende al menos dos, tres, cuatro, cinco, diez, quince, veinte o más ARNsg modificados diferentes, en donde cada ARNsg modificado está dirigido al mismo ácido nucleico diana.
- En determinadas realizaciones, el polipéptido Cas es una variante del polipéptido Cas o un fragmento de polipéptido Cas. En realizaciones particulares, el polipéptido Cas es un polipéptido Cas9, una variante del mismo o un fragmento del mismo. En otras realizaciones, la etapa de introducir en la célula primaria comprende electroporación (por ejemplo, mediante nucleofección) la célula primaria.
- En un segundo aspecto de la divulgación, la presente divulgación proporciona un método para prevenir o tratar una enfermedad genética en un sujeto, comprendiendo el método: administrar al sujeto un ARN guía único (ARNsg) modificado en una cantidad suficiente para corregir una mutación en un gen diana asociado con la enfermedad genética, en donde el ARNsg modificado comprende una primera secuencia de nucleótidos que es complementaria al gen diana y una segunda secuencia de nucleótidos que interactúa con un polipéptido de proteína (Cas) asociada a CRISPR, y en donde uno o más de los nucleótidos en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos son nucleótidos modificados.
- En algunas realizaciones, la enfermedad genética se selecciona de entre el grupo que consiste en una inmunodeficiencia combinada grave ligada al cromosoma X, anemia de células falciformes, talasemia, hemofilia, neoplasia, cáncer, degeneración macular relacionada con la edad, esquizofrenia, trastornos de repetición de trinucleótidos, síndrome del cromosoma X frágil, trastornos relacionados con priones, esclerosis lateral amiotrófica, drogadicción, autismo, enfermedad de Alzheimer, enfermedad de Parkinson, fibrosis quística, enfermedades o trastornos de la sangre y de la coagulación, inflamación, enfermedades o trastornos relacionados con el sistema inmunológico, enfermedades metabólicas, enfermedades y trastornos del hígado, enfermedades y trastornos renales, enfermedades y trastornos musculares/esqueléticos, enfermedades y trastornos neurológicos y neuronales, enfermedades y trastornos cardiovasculares, enfermedades y trastornos pulmonares, enfermedades y trastornos oculares e infecciones víricas (por ejemplo, infección por VIH).
- En algunas realizaciones, el método comprende además administrar al sujeto un polipéptido Cas, un ARNm que codifica un polipéptido Cas y/o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica un polipéptido Cas.
- En algunas realizaciones, el método comprende además administrar al sujeto un molde de reparación de donante recombinante. En determinados casos, el molde de reparación de donante recombinante comprende dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del gen diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de una secuencia de nucleótidos correspondiente al gen diana que se somete a edición del genoma. En otros casos, el molde de reparación de donante recombinante comprende un molde sintético de oligodesoxinucleótido monocatenario (ssODN) que comprende una secuencia de nucleótidos que codifica una mutación para corregir un polimorfismo de un solo nucleótido (SNP) en el gen diana y dos secuencias de nucleótidos que comprenden dos porciones homólogas no solapantes del gen diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de la secuencia de nucleótidos que codifica la mutación.
- En determinadas realizaciones, la administración del ARNsg modificado mejora el efecto del polipéptido Cas para corregir la mutación en el gen diana en comparación con la administración del ARNsg sin modificar correspondiente. Las realizaciones no limitantes relacionadas con el ARNsg modificado usado en el método para prevenir o tratar una enfermedad genética en un sujeto se han descrito anteriormente.
- En algunas realizaciones, el ARNsg modificado, polipéptido Cas y/o el molde de reparación de donante recombinante se administra al sujeto con un vehículo farmacéuticamente aceptable.
- En algunas realizaciones, el ARNsg modificado, el polipéptido Cas y/o el molde de reparación de donante recombinante se administra al sujeto a través de un sistema de administración seleccionado de entre el grupo que consiste en una nanopartícula, un liposoma, una micela, un virosoma, un complejo de ácido nucleico y una combinación de los mismos. En determinados casos, el complejo de ácido nucleico comprende el ARNsg modificado en complejo con el polipéptido Cas.
- En algunas realizaciones, el ARNsg modificado, el polipéptido Cas y/o el molde de reparación de donante recombinante se administra al sujeto a través de una vía de administración seleccionada de entre el grupo que

consiste en oral, intravenosa, intraperitoneal, intramuscular, intradérmica, subcutánea, intraarteriolar, intraventricular, intracraneal, intralesional, intratecal, tópica, transmucosal, intranasal y una combinación de las mismas.

5 Se contempla que el ARNsg químicamente modificado descrito en el presente documento puede usarse con cualquier tecnología asociada a CRISPR, incluyendo, pero sin limitaciones, CRISPRi para la inhibición de la expresión génica, CRISPRa para la activación de la expresión génica, herramientas de imágenes CRISPR para la visualización dinámica de loci genómicos y reconocimiento y escisión de ARN mediado por CRISPR. En algunos casos, el ARNsg se puede utilizar para la formación de imágenes y/o la administración de pequeñas moléculas o proteínas en células *in vitro* o células *in vivo*. Por consiguiente, los ARNsg se pueden utilizar en aplicaciones de 10 investigación y terapéuticas.

#### A. Sistema CRISPR/Cas

15 El sistema CRISPR/Cas de modificación del genoma incluye una nucleasa Cas (por ejemplo, nucleasa Cas9) o una variante o fragmento de la misma, un ARN dirigido al ADN (por ejemplo, ARNsg modificado) que contiene una secuencia guía que dirige la nucleasa Cas al ADN genómico diana y una secuencia armazón que interactúa con la nucleasa Cas (por ejemplo, ARNtracr) y, opcionalmente, un molde de reparación de donante. En algunos casos, se puede usar una variante de una nucleasa Cas, tal como una Cas9 mutante, que contiene una o más de las 20 siguientes mutaciones: D10A, H840A, D839A y H863A, o una Cas9 nickasa. En otros casos, se puede usar un fragmento de una nucleasa Cas o una variante de la misma con las propiedades deseadas (por ejemplo, capaz de 25 generar roturas de cadena simple o doble y/o modular la expresión génica). El molde de reparación de donante puede incluir una secuencia de nucleótidos que codifica un polipéptido indicador, tal como una proteína fluorescente o un marcador de resistencia a antibióticos, y brazos de homología que son homólogos al ADN diana y flanquean el sitio de modificación del gen. Como alternativa, el molde de reparación de donante puede ser un oligodesoxinucleótido monocatenario (ssODN).

##### 1. Cas nucleasas y variantes de las mismas

30 El sistema de nucleasa CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas)/Cas (proteína asociada a CRISPR) es un sistema de nucleasa diseñado basado en un sistema bacteriano que se puede utilizar para la ingeniería del genoma. Se basa en parte en la respuesta inmunológica adaptativa de muchas bacterias y arqueas. Cuando un virus o plásmido invade una bacteria, los segmentos del ADN del invasor se convierten en ARN de CRISPR (ARNcr) mediante la respuesta "inmunológica". A continuación, el ARNcr se asocia, a través de una región de complementariedad parcial, con otro tipo de ARN llamado ARNtracr para guiar la nucleasa 35 Cas (por ejemplo, Cas9) a una región homóloga al ARNcr en el ADN diana llamado "protoespaciador". La nucleasa Cas (por ejemplo, Cas9) escinde el ADN para generar extremos romos en la rotura de la doble cadena en los sitios especificados por una secuencia guía de 20 nucleótidos contenida en el transcripto de ARNcr. La nucleasa Cas (por ejemplo, Cas9) requiere tanto el ARNcr como el ARNtracr para el reconocimiento y la escisión del ADN específico de sitio. Este sistema ahora se ha diseñado de tal manera que el ARNcr y el ARNtracr se pueden combinar en una 40 molécula (el "ARN guía único" o "ARNsg"), y la porción equivalente de ARNcr del ARN guía único se puede diseñar para guiar la Cas (por ejemplo, Cas9) nucleasa para apuntar a cualquier secuencia deseada (véase, por ejemplo, Jinek *et al.* (2012) *Science*, 337:816-821; Jinek *et al.* (2013) *eLife*, 2:e00471; Segal (2013) *eLife*, 2:e00563). Por lo tanto, el sistema CRISPR/Cas se puede diseñar para crear una rotura de doble cadena en una diana deseada en un 45 genoma de una célula y aprovechar los mecanismos endógenos de la célula para reparar la rotura inducida mediante reparación dirigida por homología (HDR) o unión de extremos no homólogos (NHEJ).

En algunas realizaciones, la nucleasa Cas tiene actividad de escisión del ADN. La nucleasa Cas puede dirigir la escisión de una o ambas cadenas en una ubicación en una secuencia de ADN diana. Por ejemplo, la nucleasa Cas puede ser una nickasa que tiene uno o más dominios catalíticos inactivados que escinde una sola cadena de una 50 secuencia de ADN diana.

Los ejemplos no limitantes de nucleasas 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, variantes de las mismas, fragmentos de las mismas, mutantes de las mismas y derivados de las mismas. Hay tres tipos principales de nucleasas Cas (tipo I, tipo II y tipo III) y 10 subtipos, incluidas 5 proteínas de tipo I, 3 de tipo II y 2 de tipo III (véase, por ejemplo, Hochstrasser y Doudna, *Trends Biochem Sci*, 2015:40(1):58-66). Las nucleasas Cas de tipo II incluyen Cas1, Cas2, Csn2 y Cas9. Estas nucleasas Cas son conocidas para los expertos en la técnica. Por ejemplo, la secuencia de aminoácidos del polipéptido Cas9 de tipo salvaje de *Streptococcus pyogenes*, se expone, por ejemplo, en NBCI Ref. Seq. n.º NP\_269215 y la secuencia de aminoácidos del polipéptido Cas9 de tipo salvaje de *Streptococcus thermophilus* se expone, por ejemplo, en NBCI Ref. Seq. N.º WP\_011681470. Se desvelan endonucleasas relacionadas con CRISPR que son útiles en la presente invención, por ejemplo, en las publicaciones de solicitud de Estados Unidos n.º 2014/0068797, 2014/0302563 y 2014/0356959.

65 Las nucleasas Cas, por ejemplo, los polipéptidos Cas9, pueden derivarse de diversas especies bacterianas, que

incluyen, pero sin limitaciones, *Veillonella atypical*, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus kitaharae*, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Finegoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subsp. *Torquens*, *Ilyobacter polytropus*, *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus cellulolyticus*, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratiractor salsugininis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subsp. *Succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum rubrum*, *Candidatus Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subsp. *Jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subsp. *Multocida*, *Sutterella wadsworthensis*, *proteobacteria*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes* y *Francisella novicida*.

"Cas9" se refiere a una proteína nucleasa de unión a ADN de doble cadena guiada por ARN o proteína nickasa. La nucleasa Cas9 de tipo salvaje tiene dos dominios funcionales, por ejemplo, RuvC y HNH, que cortan diferentes cadenas de ADN. Cas9 puede inducir roturas de doble cadena en el ADN genómico (ADN diana) cuando ambos dominios funcionales están activos. La enzima Cas9 puede comprender uno o más dominios catalíticos de una proteína Cas9 derivada de bacterias pertenecientes al grupo que consiste en *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flavivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratiractor* y *Campylobacter*. En algunas realizaciones, los dos dominios catalíticos derivan de diferentes especies de bacterias.

Las variantes útiles de la nucleasa Cas9 pueden incluir un solo dominio catalítico inactivo, tal como una enzima RuvC<sup>-</sup> o HNH<sup>-</sup> o una nickasa. Una nickasa Cas9 tiene solo un dominio funcional activo y puede cortar solo una cadena del ADN diana, creando así una rotura o muesca de una sola cadena. En algunas realizaciones, la nucleasa Cas9 mutante que tiene al menos una mutación D10A es una nickasa Cas9. En otras realizaciones, la nucleasa Cas9 mutante que tiene al menos una mutación H840A es una nickasa Cas9. Otros ejemplos de mutaciones presentes en una nickasa Cas9 incluyen, sin limitación, N854A y N863A. Se puede introducir una rotura de doble cadena usando una nickasa Cas9 si se usan al menos dos ARN dirigidos al ADN que están dirigidos a cadenas de ADN opuestas. Una rotura de doble cadena inducida por doble muesca puede ser reparada por NHEJ o HDR (Ran *et al.*, 2013, *Cell*, 154:1380-1389). Esta estrategia de edición de genes favorece la HDR y disminuye la frecuencia de mutaciones ins/del en sitios de ADN fuera de la diana. Se describen ejemplos no limitantes de nucleasas Cas9 o nickasas en, por ejemplo, las patentes de Estados Unidos N.º 8,895,308; 8,889,418; y 8,865,406 y las publicaciones de solicitud de Estados Unidos 2014/0356959, 2014/0273226 y 2014/0186919. La nucleasa Cas9 o nickasa se puede optimizar por codones para la célula diana o el organismo diana.

En algunas realizaciones, la nucleasa Cas puede ser un polipéptido Cas9 que contiene dos mutaciones silenciadoras de los dominios de nucleasa RuvC1 y HNH (D10A y H840A), que se conoce como dCas9 (Jinek *et al.*, *Science*, 2012, 337:816-821; Qi *et al.*, *Cell*, 152(5): 1173-1183). En una realización, el polipéptido dCas9 de *Streptococcus pyogenes* comprende al menos una mutación en la posición D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, A987 o cualquier combinación de las mismas. Las descripciones de dichos polipéptidos dCas9 y variantes de los mismos se proporcionan en, por ejemplo, la publicación de patente internacional n.º WO 2013/176772. La enzima dCas9 puede contener una mutación en D10, E762, H983 o D986, así como una mutación en H840 o N863. En algunos casos, la enzima dCas9 contiene una mutación D10A o D10N. También, la enzima dCas9 puede incluir un H840A, H840Y o H840N. En algunas realizaciones, la enzima dCas9 de la presente invención comprende sustituciones D10A y H840A; D10A y H840Y; D10A y H840N; D10N y H840A; D10N y H840Y; o D10N y H840N. Las sustituciones pueden ser sustituciones conservadoras o no conservadoras para hacer que el polipéptido Cas9 sea catalíticamente inactivo y capaz de unirse al ADN diana.

En determinadas realizaciones, el polipéptido dCas9 es catalíticamente inactivo, tal como defectuoso en la actividad nucleasa. En algunos casos, la enzima dCas9 o una variante o fragmento de la misma puede bloquear la transcripción de una secuencia diana y, en algunos casos, bloquear la ARN polimerasa. En otros casos, la enzima dCas9 o una variante o fragmento de la misma puede activar la transcripción de una secuencia diana.

Para los métodos de edición del genoma, la nucleasa Cas puede ser una proteína de fusión Cas9 tal como un polipéptido que comprende el dominio catalítico de la enzima de restricción de tipo IIS, FokI, unida a dCas9. La proteína de fusión FokI-dCas9 (fCas9) puede usar dos ARN guía para unirse a una sola cadena de ADN diana para generar una rotura de doble cadena.

Para la regulación génica (por ejemplo, modulación de la transcripción del ADN diana), una proteína Cas deficiente en nucleasa, tal como, pero sin limitación, dCas9, puede usarse para activación transcripcional o represión

transcripcional. Se describen métodos para inactivar la expresión génica utilizando una proteína Cas nula para nucleasa, por ejemplo, en Larson *et al.*, Nat. Protoc., 2013, 8(11):2180-2196.

- 5 En algunas realizaciones, una secuencia de nucleótidos que codifica la nucleasa Cas está presente en un vector de expresión recombinante. En determinados casos, el vector de expresión es una construcción vírica, por ejemplo, una construcción de virus adenoasociado recombinante, una construcción adenoviral recombinante, una construcción lentiviral recombinante, etc. Por ejemplo, los vectores víricos pueden basarse en el virus vaccinia, poliovirus, adenovirus, virus adenoasociado, SV40, virus del herpes simple, virus de inmunodeficiencia humana y similares. Un vector retroviral puede basarse en el virus de la leucemia murina, el virus de la necrosis del bazo y vectores procedentes de retrovirus tales como el virus del sarcoma de Rous, el virus del sarcoma de Harvey, el virus de la leucosis aviar, un lentivirus, el virus de la inmunodeficiencia humana, el virus del sarcoma mieloproliferativo, el virus del tumor mamario y similares. Los expertos en la técnica conocen vectores de expresión útiles y están disponibles comercialmente. Los siguientes vectores se proporcionan a modo de ejemplo para células hospedadoras eucariotas: pXT1, pSG5, pSVK3, pBPV, pMSG y pSVLSV40. Sin embargo, se puede utilizar cualquier otro vector si es compatible con la célula hospedadora. Por ejemplo, los vectores de expresión útiles que contienen una secuencia de nucleótidos que codifica una enzima Cas9 están disponibles comercialmente en, por ejemplo, Addgene, Life Technologies, Sigma-Aldrich y Origene.
- 10 Dependiendo de la célula diana/sistema de expresión utilizado, cualquiera de una serie de elementos de control de la transcripción y la traducción, incluyendo promotores, potenciadores de la transcripción, terminadores de la transcripción y similares, etc. pueden usarse en el vector de expresión. Los promotores útiles pueden derivar de virus o de cualquier organismo, por ejemplo, organismos procariotas o eucariotas. Los promotores adecuados incluyen, pero sin limitaciones, el promotor temprano de SV40, promotor de la repetición terminal larga (LTR, por sus siglas en inglés) del virus del tumor mamario de ratón; promotor tardío principal de adenovirus (Ad MLP, por sus siglas en inglés); un promotor del virus del herpes simple (VHS), un promotor de citomegalovirus (CMV) tal como la región promotora temprana inmediata del CMV (CMVIE, por sus siglas en inglés), un promotor del virus del sarcoma de Rous (RSV, por sus siglas en inglés), un promotor nuclear pequeño U6 humano (U6), un promotor U6 mejorado, un promotor H1 humano (H1), etc.
- 15 20 25 30 35
- 30 La nucleasa Cas y variantes o fragmentos de la misma se pueden introducir en una célula (por ejemplo, una célula *in vitro* como una célula primaria para terapia *ex vivo* o una célula *in vivo* tal como en un paciente) como un polipéptido Cas o una variante o fragmento del mismo, un ARNm que codifica un polipéptido Cas o una variante o fragmento del mismo, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica un polipéptido Cas o una variante o fragmento del mismo

## 2. ARN guía único modificado (ARNsg)

- 40 Los ARNsg modificados para su uso en el sistema CRISPR/Cas de modificación del genoma normalmente incluyen una secuencia guía (por ejemplo, ARNcr) que es complementaria a una secuencia de ácido nucleico diana y una secuencia de armazón (por ejemplo, ARNtracr) que interactúa con una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento del mismo. Los presentes inventores han descubierto que los ARNsg modificados que contienen una o más modificaciones químicas pueden aumentar la actividad, la estabilidad y la especificidad, y/o disminuir la toxicidad de los ARNsg modificados en comparación con los ARNsg sin modificar correspondientes cuando se utilizan para la regulación génica basada en CRISPR (por ejemplo, edición del genoma o modulación de la expresión génica) en células primarias (por ejemplo, linfocitos T o células madre y progenitoras hematopoyéticas). Las ventajas de los ARNsg modificados sobre la técnica anterior incluyen, pero sin limitaciones, mayor facilidad de administración en las células diana, tales como células primarias, así como una mayor estabilidad, aumento de la duración de la actividad y reducción de la toxicidad en las células diana. En algunos casos, los ARNsg modificados como parte del sistema CRISPR/Cas proporcionan frecuencias más altas de regulación génica en la diana en comparación con otros sistemas. En otros casos, los ARNsg modificados proporcionan una actividad y/o especificidad mejoradas en comparación con sus equivalentes de secuencia sin modificar.

- 45 50 55 60
- 60 En determinados casos, el ARNsg modificado forma un complejo con una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma para formar un sistema de administración basado en ribonucleoproteína (RNP) para su introducción en una célula (por ejemplo, una célula *in vitro*, tal como una célula primaria para terapia *ex vivo*, o una célula *in vivo* tal como en un paciente). En otros casos, el ARNsg modificado se introduce en una célula (por ejemplo, una célula *in vitro*, tal como una célula primaria para terapia *ex vivo* o una célula *in vivo* tal como en un paciente) con un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. En otros casos más, el ARNsg modificado se introduce en una célula (por ejemplo, una célula *in vitro* tal como una célula primaria para terapia *ex vivo*, o una célula *in vivo*, tal como en un paciente) con un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma.

- 65 En algunos casos, se puede usar una pluralidad de ARNsg modificados para la regulación de genes basada en CRISPR multiplexada eficiente (por ejemplo, edición del genoma o modulación de la expresión génica) en células diana, tales como células primarias. La pluralidad de ARNsg modificados puede incluir al menos aproximadamente

- 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más ARNsg modificados que hibridan a la misma secuencia de ácido nucleico diana o a diferentes secuencias de ácido nucleico diana. La pluralidad de ARNsg modificados se puede introducir en una célula (por ejemplo, una célula *in vitro*, tal como una célula primaria para terapia *ex vivo*, o una célula *in vivo*, tal como en un paciente) en un complejo con una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o como una secuencia de nucleótidos (por ejemplo, ARNm o vector de expresión recombinante) que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma.
- 5 La secuencia de ácido nucleico del ARNsg modificado puede ser cualquier secuencia de polinucleótidos que tenga suficiente complementariedad con una secuencia de polinucleótidos diana (por ejemplo, secuencia de ADN diana) para hibridar con la secuencia diana y unión directa específica de secuencia de un complejo CRISPR a la secuencia diana. En algunas realizaciones, el grado de complementariedad entre una secuencia guía del ARNsg modificado y su correspondiente secuencia diana, cuando se alinea de manera óptima utilizando un algoritmo de alineación adecuado, es de aproximadamente el 50 % o más, 60 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97,5 %, 99 % o más. La alineación óptima se puede determinar con el uso de cualquier algoritmo adecuado para alinear secuencias, un ejemplo no limitante de los cuales incluyen el algoritmo Smith-Waterman, el algoritmo Needleman-Wunsch, algoritmos basados en la Transformada de Burrows-Wheeler (por ejemplo, el Alineador de Burrows Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, California), SOAP (disponible en [soap.genomics.org.cn](http://soap.genomics.org.cn)) y Maq (disponible en [maq.sourceforge.net](http://maq.sourceforge.net)). En algunas realizaciones, una secuencia guía tiene aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 o más nucleótidos de longitud. En algunos casos, una secuencia guía tiene aproximadamente 20 nucleótidos de longitud. En otros casos, una secuencia guía tiene aproximadamente 15 nucleótidos de longitud. En otros casos, una secuencia guía tiene aproximadamente 25 nucleótidos de longitud. La capacidad de una secuencia guía para dirigir la unión específica de secuencia de un complejo CRISPR a una secuencia diana puede evaluarse mediante cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema CRISPR suficientes para formar un complejo CRISPR, incluyendo la secuencia guía que se va a analizar, pueden proporcionarse a una célula hospedadora que tiene la secuencia diana correspondiente, tal como mediante transfección con vectores que codifican los componentes de la secuencia CRISPR, seguido de una evaluación de la escisión preferencial dentro de la secuencia diana. De manera similar, la escisión de una secuencia de 30 polinucleótidos diana puede evaluarse en un tubo de ensayo proporcionando la secuencia diana, los componentes de un complejo CRISPR, incluyendo la secuencia guía que se va a analizar y una secuencia guía de control diferente de la secuencia guía de ensayo, y comparar la unión o la velocidad de escisión en la secuencia diana entre las reacciones de la secuencia guía de ensayo y de control.
- 10 La secuencia de nucleótidos de un ARNsg modificado se puede seleccionar utilizando cualquiera de los programas de software basados en la web descritos anteriormente. Las consideraciones para seleccionar un ARN dirigido al ADN incluyen la secuencia PAM para la nucleasa Cas (por ejemplo, polipéptido Cas9) que se utilizará y estrategias para minimizar las modificaciones fuera de la diana. Herramientas, tales como la herramienta de diseño CRISPR, pueden proporcionar secuencias para preparar el ARNsg modificado, para evaluar la eficiencia de modificación de la diana y/o evaluar la escisión en sitios fuera de la diana. Otra consideración para seleccionar la secuencia de un ARNsg modificado incluye reducir el grado de estructura secundaria dentro de la secuencia guía. La estructura secundaria puede determinarse mediante cualquier algoritmo de plegamiento de polinucleótidos adecuado. Algunos programas se basan en el cálculo de la energía libre mínima de Gibbs. Los ejemplos de algoritmos adecuados incluyen mFold (Zuker y Stiegler, Nucleic Acids Res, 9 (1981), 133-148), paquete UNAFold (Markham *et al.*, Methods Mol Biol, 2008, 453: 3-31) y RNAfold forman el paquete ViennaRNA.
- 15 Uno o más nucleótidos de la secuencia guía y/o uno o más nucleótidos de la secuencia de armazón del ARNsg modificado pueden ser un nucleótido modificado. Por ejemplo, una secuencia guía que tiene aproximadamente 20 nucleótidos de longitud puede tener 1 o más, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos modificados. En algunos casos, la secuencia guía incluye al menos 2, 3, 4, 5, 6, 7, 8, 9, 10 o más nucleótidos modificados. En otros casos, la secuencia guía incluye al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20 o más nucleótidos modificados. Los nucleótidos modificados pueden ubicarse en cualquier posición de ácido nucleico de la secuencia guía. En otras palabras, los nucleótidos modificados pueden estar en o cerca del primer y/o último nucleótido de la secuencia guía, y/o en cualquier posición intermedia. Por ejemplo, para una secuencia guía de 20 nucleótidos de longitud, el uno o más nucleótidos modificados se pueden ubicar en la posición 1 del ácido nucleico, en la posición 2, en la posición 3, en la posición 4, en la posición 5, en la posición 6, en la posición 7, en la posición 8, en la posición 9, en la posición 10, en la posición 11, en la posición 12, en la posición 13, en la posición 14, en la posición 15, en la posición 16, en la posición 17, en la posición 18, en la posición 19 y/o en la posición 20 de la secuencia guía. En determinados casos, de aproximadamente un 10 % a aproximadamente un 30 %, por ejemplo, de aproximadamente el 10 % a aproximadamente el 25 %, de aproximadamente el 10 % a aproximadamente el 20 %, de aproximadamente el 10 % a aproximadamente el 15 %, de aproximadamente el 15 % a aproximadamente el 30 %, de aproximadamente el 20 % a aproximadamente el 30% o de aproximadamente el 25 % a aproximadamente el 30 % de la secuencia guía puede comprender nucleótidos modificados. En otros casos, de aproximadamente un 10 % a aproximadamente un 30 %, por ejemplo, aproximadamente un 10 %, aproximadamente un 11 %, aproximadamente un 12 %, aproximadamente un 13 %, aproximadamente un 14 %, aproximadamente un 15 %, aproximadamente un 16 %, aproximadamente un 17 %, aproximadamente un 18 %,

aproximadamente un 19 %, aproximadamente un 20 %, aproximadamente un 21 %, aproximadamente un 22 %, aproximadamente un 23 %, aproximadamente un 24 %, aproximadamente un 25 %, aproximadamente un 26 %, aproximadamente un 27 %, aproximadamente un 28 %, aproximadamente un 29 % o aproximadamente un 30 % de la secuencia guía puede comprender nucleótidos modificados.

5 En algunas realizaciones, la secuencia de armazón del ARNsg modificado contiene uno o más nucleótidos modificados. Por ejemplo, una secuencia de armazón que tiene aproximadamente 80 nucleótidos de longitud puede tener 1 o más, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 76, 77, 78, 79, 80 o más nucleótidos modificados. En algunos 10 casos, la secuencia de armazón incluye al menos 2, 3, 4, 5, 6, 7, 8, 9, 10 o más nucleótidos modificados. En otros casos, la secuencia de armazón incluye al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20 o más nucleótidos modificados. Los nucleótidos modificados pueden ubicarse en cualquier posición de ácido nucleico de la secuencia de armazón. Por ejemplo, los nucleótidos modificados pueden estar en o cerca del primer y/o último nucleótido de la secuencia de armazón, y/o en cualquier posición intermedia. Por ejemplo, para una secuencia de 15 andamio que tiene aproximadamente 80 nucleótidos de longitud, el uno o más nucleótidos modificados se pueden ubicar en la posición 1 del ácido nucleico, en la posición 2, en la posición 3, en la posición 4, en la posición 5, en la posición 6, en la posición 7, en la posición 8, en la posición 9, en la posición 10, en la posición 11, en la posición 12, en la posición 13, en la posición 14, en la posición 15, en la posición 16, en la posición 17, en la posición 18, en la posición 19, en la posición 20, en la posición 21, en la posición 22, en la posición 23, en la posición 24, en la posición 25, en la posición 26, en la posición 27, en la posición 28, en la posición 29, en la posición 30, en la posición 31, en la posición 32, en la posición 33, en la posición 34, en la posición 35, en la posición 36, en la posición 37, en la posición 38, en la posición 39, en la posición 40, en la posición 41, en la posición 42, en la posición 43, en la posición 44, en la posición 45, en la posición 46, en la posición 47, en la posición 48, en la posición 49, en la posición 50, en la posición 51, en la posición 52, en la posición 53, en la posición 54, en la posición 55, en la posición 56, en la posición 57, en la posición 58, en la posición 59, en la posición 60, en la posición 61, en la posición 62, en la posición 63, en la posición 64, en la posición 65, en la posición 66, en la posición 67, en la posición 68, en la posición 69, en la posición 70, en la posición 71, en la posición 72, en la posición 73, en la posición 74, en la posición 75, en la posición 76, en la posición 77, en la posición 78, en la posición 79 y/o en la posición 80 de la secuencia. En algunos casos, de aproximadamente un 1 % a 20 aproximadamente un 10 %, por ejemplo, de aproximadamente el 1 % a aproximadamente el 8 %, de aproximadamente el 1 % a aproximadamente el 5 %, de aproximadamente un 5 % a aproximadamente un 10 % o de aproximadamente un 3 % a aproximadamente un 7 % de la secuencia de armazón puede comprender nucleótidos modificados. En otros casos, de aproximadamente un 1 % a aproximadamente un 10 %, por ejemplo, 25 aproximadamente un 1 %, aproximadamente un 2 %, aproximadamente un 3 %, aproximadamente un 4 %, aproximadamente un 5 %, aproximadamente un 6 %, aproximadamente un 7 %, aproximadamente un 8 %, aproximadamente un 9 % o aproximadamente un 10 % de la secuencia de armazón puede comprender nucleótidos modificados.

40 Los nucleótidos modificados del ARNsg pueden incluir una modificación en el grupo ribosa (por ejemplo, azúcar), grupo fosfato, nucleobase o cualquier combinación de los mismos. En algunas realizaciones, la modificación en el grupo ribosa comprende una modificación en la posición 2' de la ribosa.

45 En algunas realizaciones, el nucleótido modificado incluye un ácido 2'fluoro-arabino nucleico, ADN triciclo (tc-ADN), ácido nucleico peptídico, ácido nucleico ciclohexeno (CeNA), ácido nucleico bloqueado (LNA, por sus siglas en inglés), ácido nucleico con puente de etileno (ENA, por sus siglas en inglés), un fosfodiamidato morfolino, o una combinación de los mismos.

50 Los nucleótidos o análogos de nucleótidos modificados pueden incluir ribonucleótidos modificados con azúcar y/o la cadena principal (es decir, incluir modificaciones en la cadena principal de fosfato-azúcar). Por ejemplo, los enlaces fosfodiéster de un ARN nativo o natural pueden modificarse para incluir al menos uno de un heteroátomo de nitrógeno o azufre. En algunos ribonucleótidos modificados en la cadena principal, el grupo fosfoéster que se conecta a los ribonucleótidos adyacentes puede ser reemplazado por un grupo modificado, por ejemplo, del grupo fosfato. En los ribonucleótidos modificados con azúcar preferidos, el resto 2' es un grupo seleccionado de entre H, OR, R, halo, SH, SR, NH<sub>2</sub>, NHR, NR<sub>2</sub> u ON, en donde R es alquilo C<sub>1</sub>-C<sub>6</sub>, alquenilo o alquinilo y halo es F, Cl, Br o I.

55 En algunas realizaciones, el nucleótido modificado contiene una modificación de azúcar. Los ejemplos no limitantes de modificaciones de azúcar incluyen 2'-desoxi-2'-fluoro-oligorribonucleótido (2'-fluoro-2'-desoxicitidina-5'-trifosfato, 2'-fluoro-2'-desoxiuridina-5'-trifosfato), oligorribonucleótido de 2'-desoxi-2'-desamina (2'-amino-2'-desoxicitidina-5'-trifosfato, 2'-amino-2'-desoxiuridina-5'-trifosfato), 2'-O-alquil oligorribonucleótido, 2'-desoxi-2'-C-alquil oligorribonucleótido (2'-O-metilcitidina-5'-trifosfato, 2'-metiluridina-5'-trifosfato), 2'-C-alquil oligorribonucleótido e isómeros de los mismos (2'-aracitidina-5'-trifosfato, 2'-arauridina-5'-trifosfato), azidotrifosfato (2'-azido-2'-desoxicitidina-5'-trifosfato, 2'-azido-2'-desoxiuridin-5'-trifosfato) y combinaciones de los mismos.

60 65 En algunas realizaciones, el ARNsg modificado contiene una o más modificaciones 2'-fluro, 2'-amino y/o 2'-tio. En algunos casos, la modificación es una 2'-fluoro-citidina, 2'-fluoro-uridina, 2'-fluoro-adenosina, 2'-fluoro-guanosina, 2'-

amino-citidina, 2'-amino-uridina, 2'-amino-adenosina, 2'-amino-guanosina, 2,6-diaminopurina, 4-tiouridina, 5-amino-alil-uridina, 5-bromo-uridina, 5-yodo-uridina, 5-metil-citidina, ribotimidina, 2-aminopurina, 2'-amino-butiril-pireno-uridina, 5-fluoro-citidina y/o 5-fluoro-uridina.

- 5 Hay más de 96 modificaciones de nucleósidos de origen natural que se encuentran en el ARN de mamíferos. Véase, por ejemplo, Limbach *et al.*, Nucleic Acids Research, 22(12):2183-2196 (1994). La preparación de nucleótidos y nucleótidos y nucleósidos modificados es bien conocida en la técnica y se describe en, por ejemplo, las patentes de Estados Unidos números 4.373.071.4.458.066, 4.500.707, 4.668.777, 4.973.679, 5.047.524, 5.132.418, 5.153.319, 5.262.530 y 5.700.642. Numerosos nucleósidos modificados y nucleótidos modificados que son adecuados para su uso como se describe en el presente documento están disponibles comercialmente. El nucleósido puede ser un análogo de un nucleósido de origen natural. En algunos casos, el análogo es dihidrouridina, metiladenosina, metilcitidina, metiluridina, metilpseudouridina, tiouridina, desoxicitodina y desoxiuridina.
- 10 En algunos casos, el ARNsg modificado descrito en el presente documento incluye un ribonucleótido modificado con nucleobase, es decir, un ribonucleótido que contiene al menos una nucleobase de origen no natural en lugar de una nucleobase de origen natural. Los ejemplos no limitantes de nucleobases modificadas que pueden incorporarse en nucleósidos modificados y nucleótidos modificados incluyen m5C (5-metilcitidina), m5U (5-metiluridina), m6A (N6-metiladenosina), s2U (2-tiouridina), Um (2'-0-metiluridina), m1A (1-metil adenosina), m2A (2-metiladenosina), Am (2-1-O-metiladenosina), ms2m6A (2-metiltio-N6-metiladenosina), i6A (N6-isopentenil adenosina), ms2i6A (2-metiltio-N6isopenteniladenosina), io6A (N6- (cis-hidroxiisopentenil) adenosina), ms2io6A (2-metiltio-N6-(cis-hidroxiisopentenil) adenosina), g6A (N6-glicinilcarbamoiladenosina), t6A (N6-treonil carbamoiladenosina), ms2t6A (2-metiltio-N6-treonilcarbamoiladenosina), m6t6A (N6-metil-N6-treonilcarbamoiladenosina), hn6A (N6-hidroxinorvalilcarbamoil adenosina), ms2hn6A (2-metiltio-N6-hidroxinorvalilcarbamoiladenosina), Ar(p) (2'-O-ribosiladenosina (fosfato)), I (inosina), m11 (1-metilinosina), m'1m (1,2'-O-dimetilinosina), m3C (3-metilcitidina), Cm (2T-O-metilcitidina), s2C (2-tiocitidina), ac4C (N4-acetilcitidina), f5C (5-fonilcitidina), m5Cm (5,2-O-dimetilcitidina), ac4Cm (N4acetil2TOMetilcitidina), k2C (lisidina), m1G (1-metilguanosina), m2G (N2-metilguanosina), m7G (7-metilguanosina), Gm (2'-O-metilguanosina), m22G (N2,N2-dimetilguanosina), m2Gm (N2,2'-O-dimetilguanosina), m22Gm (N2,N2,2'-O-trimetilguanosina), Gr(p) (2'-O-ribosilguanosina (fosfato)), yW (wibutosina), o2yW (peroxiwibutosina), OHyW (hidroxiwibutosina), OHyW\* (hidroxiwibutosina no modificada), imG (wiosina), mimG (metilguanosina), Q (queuosina), oQ (epoxiqueuosina), galQ (galtactosil-queuosina), manQ (manosil-queuosina), preQo (7-ciano-7-deazaguanosina), preQo (7-aminometil-7-deazaguanosina), G (arcaeosina), D (dihidrouridina), m5Um (5,2'-O-dimetiluridina), s4U (4-tiouridina), m5s2U (5-metil-2-tiouridina), s2Um (2-tio-2'-O-metiluridina), acp3U (3-(3-amino-3-carboxipropil)uridina), ho5U (5-hidroxuridina), mo5U (5-metoxuridina), cmo5U (ácido uridina 5-oxiacético), mcmo5U (éster metílico de ácido uridina 5-oxiacético), chm5U (5-(carboxihidroximetil)uridina), mchm5U (éster metílico de 5-(carboxihidroximetil)uridina), mcm5U (5-metoxicarbonil metiluridina), mcm5Um (S-metoxicarbonilmetil-2-O-metiluridina), mcm5s2U (5-metoxicarbonilmetil-2-tiouridina), nm5s2U (5-aminometil-2-tiouridina), mmn5U (5-metilaminometiluridina), mmn5s2U (5-metilaminometil-2-tiouridina), mmn5se2U (5-metilaminometil-2-selenouridina), ncm5U (5-carbamoilmetiluridina), ncm5Um (5-carbamoilmetil-2'-O-metiluridina), cmnm5U (5-carboximetilaminometiluridina), cnmm5Um (5-carboximetilaminometil-2-L-ometiluridina), cmnm5s2U (5-carboximetilaminometil-2-tiouridina), m62A (N6,N6-dimetiladenosina), Tm (2'-O-metilinosina), m4C (N4-metilcitidina), m4Cm (N4,2-O-dimetilcitidina), hm5C (5-hidroximetilcitidina), m3U (3-metiluridina), cm5U (5-carboximetiluridina), m6Am (N6, T-O-dimetiladenosina), rn62Am (N6, N6, O-2-trimetiladenosina), m2'7G (N2,7-dimetilguanosina), m2'2'7G (N2, N2,7-trimetilguanosina), m3Um (3,2T-O-dimetiluridina), m5D (5-metildihidrouridina), f5Cm (5-formil-2'-O-metilcitidina), m1Gm (1,2'-O-dimetilguanosina), m'Am (1,2-0-dimetil adenosina) irinometiluridina), tm5s2U (S-taurinometil-2-tiouridina)), imG-14 (4-desmetil guanosina), imG2 (isoguanosina) o ac6A (N6-acetiladenosina), hipoxantina, inosina, 8-oxo-adenina, derivados 7-sustituidos de las mismas, dihidrouracilo, pseudouracilo, 2-tiouracilo, 4-tiouracilo, 5-aminouracilo, 5-alquiluracilo (C<sub>1</sub>-C<sub>6</sub>), 5-metiluracilo, 5-alqueniluracilo (C<sub>2</sub>-C<sub>6</sub>), 5-alquiniluracilo (C<sub>2</sub>-C<sub>6</sub>), 5-(hidroximetil)uracilo, 5-clorouracilo, 5-fluorouracilo, 5-bromouracilo, 5-hidroxicitosina, 5-alquicitosina (C<sub>1</sub>-C<sub>6</sub>), 5-metilcitosina, 5-alquenilcitosina (C<sub>2</sub>-C<sub>6</sub>), 5-alquinilcitosina (C<sub>2</sub>-C<sub>6</sub>), 5-clorocitosina, 5-fluorocitosina, 5-bromocitosina, N<sup>2</sup>-dimetilguanina, 7-desazaguanina, 8-azaguanina, 7-deaza-guanina-7-sustituida, 7-deaza-7- (C<sub>2</sub>-C<sub>6</sub>)alquinilguanina, 7-deaza-guanina-8-sustituida, 8-hidroxiguanina, 6-tioguanina, 8-oxoguanina, 2-aminopurina, 2-amino-6-cloropurina, 2,4-diaminopurina, 2,6-diaminopurina, 8-azapurina, 7-deazapurina sustituida, purina 7-deaza-7-sustituida, purina 7-deaza-8-sustituida y combinaciones de las mismas.
- 45 55 En algunas realizaciones, la cadena principal de fosfato del ARNsg modificado está alterado. El ARNsg modificado puede incluir uno o más enlaces fosforotioato, fosforamidato (por ejemplo, N<sup>3</sup>'-P<sup>5</sup>'-fosforamidato (NP)), 2'-O-metoxietilo (2'MOE), 2'-O-metil-etilo (2'ME) y/o metilfosfonato.
- 60 55 En realizaciones particulares, uno o más de los nucleótidos modificados de la secuencia guía y/o uno o más de los nucleótidos modificados de la secuencia de armazón del ARNsg modificado incluyen un nucleótido 2'-O-metilo (M), un nucleótido 2'-O-metil 3'-fosforotioato (MS), un nucleótido 2'-O-metil 3'tioPACE (MSP) o una combinación de los mismos. En algunos casos, el ARNsg modificado incluye uno o más nucleótidos de MS. En otros casos, el ARNsg modificado incluye uno o más nucleótidos de MSP. En otros casos más, el ARNsg modificado incluye uno o más nucleótidos de MS y uno o más nucleótidos de MSP. En otros casos, el ARNsg modificado no incluye nucleótidos de M. En determinados casos, el ARNsg modificado incluye uno o más nucleótidos de MS y/o uno o más nucleótidos de MSP, y, además, incluye uno o más nucleótidos de M. En otros casos determinados, los nucleótidos de MS y/o los

nucleótidos de MSP son los únicos nucleótidos modificados presentes en el ARNsg modificado.

Cabe señalar que cualquiera de las modificaciones descritas en el presente documento puede combinarse e incorporarse en la secuencia guía y/o la secuencia de armazón del ARNsg modificado.

- 5 En algunos casos, el ARNsg modificado también incluye una modificación estructural, tal como un bucle del tallo, por ejemplo, bucle del tallo M2 o un tetrabucle.
- 10 El ARNsg modificado puede sintetizarse mediante cualquier método conocido por un experto en la materia. En algunas realizaciones, el ARNsg modificado se sintetiza químicamente. Los ARNsg modificados se pueden sintetizar usando nucleósidos fosforamiditas protegidos con 2'-O-tiono-carbamato. Los métodos se describen en, por ejemplo, Dellinger *et al.*, J. American Chemical Society 133, 11540-11556 (2011); Threlfall *et al.*, Organic & Biomolecular Chemistry 10, 746-754 (2012); y Dellinger *et al.*, J. American Chemical Society 125, 940-950 (2003).
- 15 Los ARNsg modificados químicamente se pueden utilizar con cualquier tecnología asociada a CRISPR, por ejemplo, y tecnología guiada por ARN. Tal como se describe en el presente documento, el ARNsg modificado puede servir como guía para cualquier nucleasa Cas o variante o fragmento de la misma, incluyendo cualquier polipéptido Cas9 manipulado o artificial. Los ARNsg modificados pueden dirigirse a moléculas de ADN y/o ARN en células primarias aisladas para terapia *ex vivo* o *in vivo* (por ejemplo, en un animal). Los métodos desvelados en el presente documento se pueden aplicar a la edición del genoma, la regulación génica, el procesamiento de imágenes y cualquier otra aplicación basada en CRISPR.

### 3. Molde de reparación de donante

- 25 En algunas realizaciones, la presente invención proporciona un molde de reparación de donante recombinante que comprende dos brazos de homología que son homólogos a porciones de una secuencia de ADN diana (por ejemplo, gen o locus diana) en cada lado de un sitio de escisión de nucleasa Cas (por ejemplo, nucleasa Cas9). En determinados casos, el molde de reparación de donante recombinante comprende un casete indicador que incluye una secuencia de nucleótidos que codifica un polipéptido indicador (por ejemplo, un polipéptido detectable, polipéptido fluorescente o un marcador seleccionable), y dos brazos de homología que flanquean el casete indicador y son homólogos a porciones del ADN diana en cada lado del sitio de escisión de la nucleasa Cas. El casete indicador puede comprender además una secuencia que codifica un péptido de autoescisión, una o más señales de localización nuclear y/o un polipéptido fluorescente, por ejemplo, GFP superlegador (sfGFP).
- 35 En algunas realizaciones, los brazos de homología tienen la misma longitud. En otras realizaciones, los brazos de homología tienen diferentes longitudes. Los brazos de homología pueden tener al menos aproximadamente 10 pares de bases (pb), por ejemplo, al menos aproximadamente 10 pb, 15 pb, 20 pb, 25 pb, 30 pb, 35 pb, 45 pb, 55 pb, 65 pb, 75 pb, 85 pb, 95 pb, 100 pb, 150 pb, 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 950 pb, 1000 pb, 1,1 kilobases (kb), 1,2 kb, 1,3 kb, 1,4 kb, 1,5 kb, 1,6 kb, 1,7 kb, 1,8 kb, 1,9 kb, 2,0 kb, 2,1 kb, 2,2 kb, 2,3 kb, 2,4 kb, 2,5 kb, 2,6 kb, 2,7 kb, 2,8 kb, 2,9 kb, 3,0 kb, 3,1 kb, 3,2 kb, 3,3 kb, 3,4 kb, 3,5 kb, 3,6 kb, 3,7 kb, 3,8 kb, 3,9 kb, 4,0 kb o más. Los brazos de homología pueden tener de aproximadamente 10 pb a aproximadamente 4 kb, por ejemplo, de aproximadamente 10 pb a aproximadamente 20 pb, de aproximadamente 10 pb a aproximadamente 50 pb, de aproximadamente 10 pb a aproximadamente 100 pb, de aproximadamente 10 pb a aproximadamente 200 pb, de aproximadamente 10 pb a aproximadamente 500 pb, de aproximadamente 10 pb a aproximadamente 2 kb, de aproximadamente 1 kb a aproximadamente 10 pb a aproximadamente 1 kb, de aproximadamente 1 kb a aproximadamente 2 kb, de aproximadamente 10 pb a aproximadamente 4 kb, de aproximadamente 100 pb a aproximadamente 500 pb, de aproximadamente 100 pb a aproximadamente 1 kb, de aproximadamente 100 pb a aproximadamente 2 kb, de aproximadamente 1 kb a aproximadamente 500 pb a aproximadamente 1 kb, de aproximadamente 500 pb a aproximadamente 2 kb, de aproximadamente 500 pb a aproximadamente 4 kb, de aproximadamente 500 pb a aproximadamente 1 kb a aproximadamente 2 kb, de aproximadamente 1 kb a aproximadamente 4 kb o de aproximadamente 2 kb a aproximadamente 4 kb.

55 El molde de reparación de donante se puede clonar en un vector de expresión. Pueden usarse vectores de expresión convencionales basados en virus y no virus conocidos por los expertos en la técnica.

En lugar de un molde de reparación de donante recombinante, se puede usar un molde donante de oligodesoxinucleótido monocatenario (ssODN) para la reparación mediada por recombinación homóloga. Un ssODN es útil para introducir modificaciones cortas dentro de un ADN diana. Por ejemplo, los ssODN son adecuados para corregir con precisión mutaciones genéticas tales como SNP. Los ssODN pueden contener dos secuencias homólogas flanqueantes en cada lado del sitio diana de la escisión de la nucleasa Cas y pueden estar orientadas en la dirección sentido o antisentido con respecto al ADN diana. Cada secuencia flanqueante puede tener al menos aproximadamente 10 pares de bases (pb), por ejemplo, al menos aproximadamente 10 pb, 15 pb, 20 pb, 25 pb, 30 pb, 35 pb, 40 pb, 45 pb, 50 pb, 55 pb, 60 pb, 65 pb, 70 pb, 75 pb, 80 pb, 85 pb, 90 pb, 95 pb, 100 pb, 150 pb, 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 950 pb, 1 kb, 2 kb, 4 kb o más. En algunas realizaciones, cada brazo de homología tiene de

aproximadamente 10 pb a aproximadamente 4 kb, por ejemplo, de aproximadamente 10 pb a aproximadamente 20 pb, de aproximadamente 10 pb a aproximadamente 50 pb, de aproximadamente 10 pb a aproximadamente 100 pb, de aproximadamente 10 pb a aproximadamente 200 pb, de aproximadamente 10 pb a aproximadamente 500 pb, de aproximadamente 10 pb a aproximadamente 1 kb, de aproximadamente 10 pb a aproximadamente 2 kb, de aproximadamente 10 pb a aproximadamente 4 kb, de aproximadamente 100 pb a aproximadamente 500 pb, de aproximadamente 100 pb a aproximadamente 1 kb, de aproximadamente 100 pb a aproximadamente 2 kb, de aproximadamente 500 pb a aproximadamente 1 kb, de aproximadamente 500 pb a aproximadamente 500 pb, de aproximadamente 500 pb a aproximadamente 2 kb, de aproximadamente 500 pb a aproximadamente 4 kb, de aproximadamente 500 pb a aproximadamente 1 kb a aproximadamente 2 kb, de aproximadamente 1 kb a aproximadamente 2 kb, de aproximadamente 1 kb a aproximadamente 4 kb o de aproximadamente 2 kb a aproximadamente 4 kb. El ssODN puede tener al menos aproximadamente 25 nucleótidos (nt) de longitud, por ejemplo, al menos aproximadamente 25 nt, 30 nt, 35 nt, 40 nt, 45 nt, 50 nt, 55 nt, 60 nt, 65 nt, 70 nt, 75 nt, 80 nt, 85 nt, 90 nt, 95 nt, 100 nt, 150 nt, 200 nt, 250 nt, 300 nt o más. En algunas realizaciones, el ssODN es de aproximadamente 25 a aproximadamente 50; de aproximadamente 50 a aproximadamente 100; de aproximadamente 100 a aproximadamente 150; de aproximadamente 150 a aproximadamente 200; de aproximadamente 200 a aproximadamente 250; de aproximadamente 250 a aproximadamente 300; o de aproximadamente 25 nt a aproximadamente 300 nt de longitud.

En algunas realizaciones, el molde de ssODN comprende al menos uno, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500 o más nucleótidos modificados descritos en el presente documento. En algunos casos, al menos un 1 %, 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 99 % de la secuencia de ssODN incluye un nucleótido modificado. En algunas realizaciones, los nucleótidos modificados se encuentran en uno o en ambos extremos terminales del ssODN. Los nucleótidos modificados pueden estar en el primero, segundo, tercero, cuarto, quinto, sexto, séptimo, octavo, noveno o décimo nucleótido terminal o cualquier combinación de los mismos. Por ejemplo, los nucleótidos modificados pueden estar en los tres nucleótidos terminales en ambos extremos del molde de ssODN. Adicionalmente, los nucleótidos modificados se pueden ubicar en el interior de los extremos terminales.

#### 4. ADN diana

En el sistema CRISPR/Cas, la secuencia de ADN diana puede ser complementaria a un fragmento de un ARN dirigido a ADN (por ejemplo, ARNsg modificado) y puede ser seguida inmediatamente por una secuencia motivo adyacente protoespaciador (PAM). El sitio del ADN diana puede encontrarse inmediatamente en 5' de una secuencia PAM, que es específico de la especie bacteriana de Cas9 utilizada. Por ejemplo, la secuencia PAM de Cas9 *derivada de Streptococcus pyogenes* es NGG; la secuencia PAM de Cas9 *derivada de Neisseria meningitidis* es NNNNGATT; la secuencia PAM de Cas9 *derivada de Streptococcus thermophilus* es NNAGAA; y la secuencia PAM de Cas9 *derivada de Treponema denticola* es NAAAAC. En algunas realizaciones, la secuencia PAM puede ser 5'-NGG, en donde N es cualquier nucleótido; 5'-NRG, en donde N es cualquier nucleótido y R es una purina; o 5'-NNGRR, en donde N es cualquier nucleótido y R es una purina. Para el sistema de *S. pyogenes*, la secuencia de ADN diana seleccionada debe preceder inmediatamente (por ejemplo, estar ubicada en 5') a 5' NGG PAM, en donde N es cualquier nucleótido, de manera que la secuencia guía del ARN que se dirige al ADN (por ejemplo, ARNsg modificado) se empareja con la cadena opuesta para mediar la escisión en aproximadamente 3 pares de bases cadena arriba de la secuencia de PAM.

En algunas realizaciones, el grado de complementariedad entre una secuencia guía del ARN dirigido al ADN (por ejemplo, ARNsg modificado) y su secuencia de ADN diana correspondiente, cuando se alinea de manera óptima utilizando un algoritmo de alineación adecuado, es de aproximadamente el 50 % o más, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más. La alineación óptima se puede determinar con el uso de cualquier algoritmo adecuado para alinear secuencias, un ejemplo no limitante de los cuales incluyen el algoritmo Smith-Waterman, el algoritmo Needleman-Wunsch, algoritmos basados en la Transformada de Burrows-Wheeler (por ejemplo, el Alineador de Burrows Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, Selangor, Malasia) y ELAND (Illumina, San Diego, CA).

El sitio de ADN diana se puede seleccionar en una secuencia genómica predefinida (gen) utilizando un software basado en la web como el software ZiFiT Targeter (Sander *et al.*, 2007, Nucleic Acids Res, 35:599-605; Sander *et al.*, 2010, Nucleic Acids Res, 38:462-468), E-CRISP (Heigwer *et al.*, 2014, Nat Methods, 11:122-123), RGEN Tools (Bae *et al.*, 2014, Bioinformatics, 30(10):1473-1475), CasFinder (Aach *et al.*, 2014, bioRxiv), DNA2.0 gNRA Design Tool (DNA2.0, Menlo Park, CA) y la herramienta de diseño CRISPR (Broad Institute, Cambridge, MA). Dichas herramientas analizan una secuencia genómica (por ejemplo, gen o locus de interés) e identifican el sitio diana adecuado para la edición de genes. Para evaluar las modificaciones de genes fuera de la diana para cada ARN dirigido al ADN (por ejemplo, ARNsg modificado), las predicciones computacionales de los sitios fuera de la diana se basan en el análisis de especificidad cuantitativa de la identidad de emparejamiento erróneo, posición y distribución de las bases.

## 5. Modulación de la expresión génica

El sistema CRISPR/Cas de regulación de la expresión génica, tal como inhibir la expresión génica o activar la expresión génica, puede incluir una variante o fragmento de la nucleasa Cas nativa o de tipo salvaje (por ejemplo, variante o fragmento del polipéptido Cas9) y un ARNsg dirigido al ADN o un ARNsg dirigido al ARN. Como un ejemplo no limitante, un complejo que comprende una variante o fragmento de Cas9 y un ARNsg que puede unirse a una secuencia de ADN diana complementaria a una porción del ARNsg puede bloquear u obstaculizar el inicio y/o elongación de la transcripción por la ARN polimerasa. Esto, a su vez, puede inhibir o reprimir la expresión génica del ADN diana. Como alternativa, un complejo que comprende una variante o fragmento de Cas9 diferente y un ARNsg que puede unirse a una secuencia de ADN diana complementaria a una porción del ARNsg puede inducir o activar la expresión génica del ADN diana.

Las descripciones detalladas de los métodos para realizar la interferencia CRISPR (CRISPRi) para inactivar o reducir la expresión génica se encuentran en, por ejemplo, Larson *et al.*, *Nature Protocols*, 2013, 8(11):2180-2196 y Qi *et al.*, *Cell*, 152, 2013, 1173-1183. En CRISPRi, el complejo variante ARNsg-Cas9 puede unirse a una cadena de ADN sin molde de una región codificante de proteína y bloquear la elongación de la transcripción. En algunos casos, cuando el complejo variante ARNsg-Cas9 se une a una región promotora de un gen, el complejo previene o dificulta el inicio de la transcripción.

Las descripciones detalladas de los métodos para realizar la activación de CRISPR para aumentar la expresión génica se encuentran en, por ejemplo, Cheng *et al.*, *Cell Research*, 2013, 23:1163-1171, Konerman *et al.*, *Nature*, 2015, 517:583-588 y la patente de Estados Unidos n.º 8,697,359.

Se encuentran descripciones detalladas de los métodos para la unión y/o escisión de ARN basados en CRISPR para modular la expresión génica en, por ejemplo, O'Connell *et al.*, *Nature*, 2014, 516: 263 -266 y la Publicación de Solicitud de Patente de Estados Unidos n.º 2014/0302563.

Para el control de la expresión génica basado en CRISPR, puede usarse una variante catalíticamente inactiva de la nucleasa Cas (por ejemplo, polipéptido Cas9) que carece de actividad endonucleolítica. En algunas realizaciones, la nucleasa Cas es una variante de Cas9 que contiene al menos dos mutaciones puntuales en los dominios nucleasa de tipo RuvC y HNH. En algunas realizaciones, la variante Cas9 tiene sustituciones de aminoácidos D10A y H840A, que se conoce como dCas9 (Jinek *et al.*, *Science*, 2012, 337:816-821; Qi *et al.*, *Cell*, 152(5):1173-1183). En algunos casos, el polipéptido dCas9 de *Streptococcus pyogenes* comprende al menos una mutación en la posición D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, A987 o cualquier combinación de las mismas. Las descripciones de dichos polipéptidos dCas9 y variantes de los mismos se proporcionan en, por ejemplo, la publicación de patente internacional n.º WO 2013/176772. La enzima dCas9 puede contener una mutación en D10, E762, H983 o D986, así como una mutación en H840 o N863. En algunos casos, la enzima dCas9 contiene una mutación D10A o D10N. También, la enzima dCas9 puede incluir un H840A, H840Y o H840N. En algunos casos, la enzima dCas9 comprende sustituciones D10A y H840A; D10A y H840Y; D10A y H840N; D10N y H840A; D10N y H840Y; o D10N y H840N. Las sustituciones pueden ser sustituciones conservadoras o no conservadoras para hacer que el polipéptido Cas9 sea catalíticamente inactivo y capaz de unirse al ADN diana.

En determinadas realizaciones, el polipéptido dCas9 es catalíticamente inactivo, tal como defectuoso en la actividad nucleasa. En algunos casos, la enzima dCas9 o una variante o fragmento de la misma puede bloquear la transcripción de una secuencia diana y, en algunos casos, bloquear la ARN polimerasa. En otros casos, la enzima dCas9 o una variante o fragmento de la misma puede activar la transcripción de una secuencia diana.

En determinadas realizaciones, la variante Cas9 que carece de actividad endonucleolítica (por ejemplo, dCas9) puede fusionarse con un dominio de represión transcripcional, por ejemplo, un dominio de caja asociada a Kruppel (KRAB), o un dominio de activación transcripcional, por ejemplo, un dominio de transactivación VP16. En algunas realizaciones, la variante Cas9 es un polipéptido de fusión que comprende dCas9 y un factor de transcripción, por ejemplo, el factor omega de ARN polimerasa, el factor de choque térmico 1 o un fragmento del mismo. En otras realizaciones, la variante Cas9 es un polipéptido de fusión que comprende dCas9 y una ADN metilasa, histona acetilasa o un fragmento de la misma.

Para el control basado en CRISPR de la expresión génica mediada por la unión del ARN y/o la escisión del ARN, puede usarse una variante de nucleasa Cas adecuada (por ejemplo, polipéptido Cas9) que tiene actividad endorribonucleasa, como se describe en, por ejemplo, O'Connell *et al.*, *Nature*, 2014, 516:263-266. Otras variantes de la nucleasa Cas útiles (por ejemplo, Cas9) se describen en, por ejemplo, la Publicación de Solicitud de Patente de Estados Unidos n.º 2014/0302563. Otras enzimas relacionadas con CRISPR que pueden escindir el ARN incluyen una endorribonucleasa Csy4, una enzima Cas6 relacionada con CRISPR, una enzima miembro de la familia Cas5, una enzima miembro de la familia Cas6, una endorribonucleasa del sistema CRISPR de tipo I, una endorribonucleasa del sistema CRISPR de tipo II, una endorribonucleasa del sistema CRISPR de tipo III y variantes de las mismas.

En algunas realizaciones de escisión de ARN basada en CRISPR, se usa un oligonucleótido de ADN que contiene una secuencia de PAM (por ejemplo, PAMmer) con el ARNsg modificado y la variante de nucleasa Cas (por ejemplo, Cas9) descrita en el presente documento para unirse y escindir un transcripto de ARN monocatenario. Las

descripciones detalladas de las secuencias de PAMmer adecuadas se encuentran en, por ejemplo, O'Connell *et al.*, *Nature*, 2014, 516:263-266.

5 La nucleasa Cas (por ejemplo, polipéptido Cas9) o variantes o fragmentos de la misma se pueden proporcionar como polipéptido, un ARNm que codifica el polipéptido o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica el polipéptido. Se pueden encontrar detalles adicionales en lo que antecede.

10 En algunas realizaciones, se usa una pluralidad de ARNsg modificados para apuntar a diferentes regiones de un gen diana para regular la expresión génica de ese gen diana. La pluralidad de ARNsg modificados puede proporcionar una modulación sinérgica (por ejemplo, inhibición o activación) de la expresión génica de un único gen diana en comparación con cada ARNsg modificado solo. En otras realizaciones, se utiliza una pluralidad de ARNsg modificados para regular la expresión génica de al menos dos genes diana.

#### B. Células primarias

15 La presente invención se usa para inducir la regulación génica de un ácido nucleico diana en cualquier célula primaria de interés. La célula primaria puede ser una célula aislada de cualquier organismo multicelular, por ejemplo, una célula vegetal (por ejemplo, una célula de arroz, una célula de trigo, una célula de tomate, una célula de *Arabidopsis thaliana*, una célula de *Zea mays*, y similares), una célula de un protista multicelular, una célula de un hongo multicelular, una célula animal, tal como una célula de un animal invertebrado (por ejemplo, mosca de la fruta, cnidario, equinodermo, nematodo, etc.) o una célula de un animal vertebrado (por ejemplo, peces, anfibios, reptiles, aves, mamíferos, etc.), una célula de un ser humano, una célula de un ser humano sano, una célula de un paciente humano, una célula de un paciente con cáncer, etc. En algunos casos, la célula primaria con regulación génica inducida se puede trasplantar a un sujeto (por ejemplo, paciente). Por ejemplo, la célula primaria puede derivarse del sujeto (por ejemplo, paciente) a tratar.

20 Cualquier tipo de célula primaria puede ser de interés, tal como una célula madre, por ejemplo, célula madre embrionaria, célula madre pluripotencial inducida, célula madre adulta (por ejemplo, célula madre mesenquimatosa, célula madre neuronal, célula madre hematopoyética, célula madre de órgano), una célula progenitora, una célula somática (por ejemplo, fibroblasto, hepatocito, célula cardíaca, hepatocito, célula pancreática, célula muscular, célula de la piel, célula sanguínea, célula neural, célula inmunitaria) y cualquier otra célula del cuerpo, por ejemplo, cuerpo humano. Las células pueden ser células primarias o cultivos de células primarias derivados de un sujeto, por ejemplo, un sujeto animal o un sujeto humano, y se deja crecer *in vitro* durante un número limitado de pases. En 25 algunas realizaciones, las células son células enfermas o derivadas de un sujeto con una enfermedad. Por ejemplo, las células pueden ser cancerosas o tumorales.

30 Las células primarias se pueden recolectar de un sujeto mediante cualquier método estándar. Por ejemplo, células de tejidos, tales como la piel, musculares, de médula ósea, de bazo, de hígado, de riñón, de páncreas, de pulmón, de intestino, de estómago, etc., puede recolectarse mediante una biopsia de tejido o un aspirado con aguja fina. Las 35 células sanguíneas y/o las células inmunitarias se pueden aislar de sangre completa, plasma o suero. En algunos casos, las células primarias adecuadas incluyen células mononucleares de sangre periférica (PBMC), linfocitos de sangre periférica (PBL) y otros subconjuntos de células sanguíneas tales como, pero sin limitaciones, linfocito T, un linfocito citolítico natural, un monocito, un linfocito T citolítico natural, una célula precursora de monocitos, una célula madre y progenitora hematopoyética (HSPC), tales como HSPC CD34+, o una célula madre no pluripotencial. En 40 algunos casos, la célula puede ser cualquier célula inmunitaria, incluyendo, pero sin limitaciones, cualquier linfocito T, tales como células infiltrantes de tumores (TIL), linfocitos T CD3+, linfocitos T CD4+, linfocitos T CD8+ o cualquier otro tipo de linfocito T. El linfocito T también puede incluir linfocitos T de memoria, linfocitos T madre de memoria o linfocitos T efectores. Los linfocitos T también pueden estar sesgados hacia poblaciones y fenotipos particulares. Por ejemplo, los linfocitos T se pueden sesgar para que comprendan fenotípicamente CD45RO(-), CCR7(+), CD45RA(+), CD62L(+), CD27(+), CD28(+) y/o IL-7R $\alpha$ (+). Se pueden seleccionar células adecuadas que comprendan uno o más marcadores seleccionados de una lista que comprende CD45RO(-), CCR7(+), CD45RA(+), CD62L(+), CD27(+), CD28(+) y/o IL-7R $\alpha$ (+). Las células madre pluripotenciales inducidas se pueden generar a partir de células diferenciadas de acuerdo con los protocolos estándar descritos en, por ejemplo, las Patentes de Estados Unidos n.º 7,682,828, 8,058,065, 8,530,238, 8,871,504, 8,900,871 y 8,791,248.

45 55 En algunas realizaciones, la célula primaria es *in vitro*. En otras realizaciones, la célula primaria es *ex vivo*.

#### C. Terapia *ex vivo*

60 Los métodos descritos en el presente documento pueden usarse en terapia *ex vivo*. La terapia *ex vivo* puede comprender administrar una composición (por ejemplo, una célula) generada o modificada fuera de un organismo a un sujeto (por ejemplo, un paciente). En algunas realizaciones, la composición (por ejemplo, una célula) puede generarse o modificarse mediante los métodos desvelados en el presente documento. Por ejemplo, la terapia *ex vivo* puede comprender administrar una célula primaria generada o modificada fuera de un organismo a un sujeto (por ejemplo, un paciente), en donde la célula primaria ha sido cultivada *in vitro* de acuerdo con los métodos de la 65 presente invención que incluye poner en contacto el ácido nucleico diana en la célula primaria con uno o más ARNsg

modificados descritos en el presente documento y una nucleasa Cas (por ejemplo, polipéptido Cas9) o variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma.

- 5 En algunas realizaciones, la composición (por ejemplo, una célula) se puede derivar del sujeto (por ejemplo, un paciente) que se va a tratar mediante terapia *ex vivo*. En algunas realizaciones, la terapia *ex vivo* puede incluir terapia basada en células, tal como inmunoterapia adoptiva.
- 10 En algunas realizaciones, la composición utilizada en la terapia *ex vivo* puede ser una célula. La célula puede ser una célula primaria, que incluyen, pero sin limitaciones, células mononucleares de sangre periférica (PBMC, por sus siglas en inglés), linfocitos de sangre periférica (PBL) y otros subconjuntos de células sanguíneas. La célula primaria puede ser una célula inmunitaria. La célula primaria puede ser un linfocito T (por ejemplo, linfocitos T CD3+, linfocitos T CD4+ y/o linfocitos T CD8+), un linfocito citolítico natural, un monocito, un linfocito T citolítico natural, una célula precursora de monocitos, una célula madre hematopoyética o una célula madre no pluripotencial, una célula madre o una célula progenitora. La célula primaria puede ser una célula madre o progenitora hematopoyética (HSPC), tales como HSPC CD34+. La célula primaria puede ser una célula humana. La célula primaria se puede aislar, seleccionar y/o cultivar. La célula primaria se puede expandir *ex vivo*. La célula primaria se puede expandir *in vivo*. La célula primaria puede ser CD45RO(-), CCR7(+), CD45RA(+), CD62L(+), CD27(+), CD28(+) y/o IL-7R $\alpha$ (+).
- 15 20 La célula primaria puede ser autóloga de un sujeto que la necesite. La célula primaria puede no ser autóloga para un sujeto que la necesite. La celda primaria puede ser un reactivo compatible con las buenas prácticas de fabricación (GMP). La célula primaria puede ser parte de una terapia combinada para tratar enfermedades, incluyendo cáncer, infecciones, trastornos autoinmunes o enfermedad de injerto contra huésped (EICH), en un sujeto que lo necesite.
- 25 30 Como ejemplo no limitativo de terapia *ex vivo*, una célula primaria puede aislarse de un organismo multicelular (por ejemplo, una planta, protista multicelular, hongo multicelular, animal invertebrado, animal vertebrado, etc.) antes de poner en contacto un ácido nucleico diana dentro de la célula primaria con una nucleasa Cas y un ARNsg modificado. Después de poner en contacto el ácido nucleico diana con la nucleasa Cas y el ARNsg modificado, la célula primaria o su progenie (por ejemplo, una célula derivada de la célula primaria) se puede devolver al organismo multicelular.

#### **D. Métodos para introducir ácidos nucleicos en células diana**

- 35 Los métodos para introducir polipéptidos y ácidos nucleicos en una célula diana (célula hospedadora) son conocidos en la técnica y se puede usar cualquier método conocido para introducir una nucleasa o un ácido nucleico (por ejemplo, una secuencia de nucleótidos que codifica la nucleasa, un ARN dirigido al ADN (por ejemplo, un ARN guía único modificado), un molde de reparación de donante para la reparación dirigida por homología (HDR), etc.) en una célula, por ejemplo, una célula primaria, tal como una célula madre, una célula progenitora o una célula diferenciada. Los ejemplos no limitantes de métodos adecuados incluyen electroporación, infección vírica o por bacteriófago, 40 transfección, conjugación, fusión de protoplastos, lipofección, precipitación con fosfato de calcio, transfección mediada por polietilenimina (PEI), transfección mediada por DEAE-dextrano, transfección mediada por liposomas, tecnología de pistola de partículas, precipitación con fosfato de calcio, microinyección directa, administración de ácido nucleico mediado por nanopartículas y similares.
- 45 55 En algunas realizaciones, los componentes de la regulación génica mediada por CRISPR/Cas pueden introducirse en una célula utilizando un sistema de administración. En determinados casos, el sistema de administración comprende una nanopartícula, una micropartícula (por ejemplo, un micropolímero de polímero), un liposoma, una micela, un virosoma, una partícula vírica, un complejo de ácido nucleico, un agente de transfección, un agente de electroporación (por ejemplo, usando un sistema de transfección NEON), un agente de nucleofección, un agente de lipofección y/o un sistema tampón que incluye un componente de nucleasa (como un polipéptido o codificado por una construcción de expresión) y uno o más componentes de ácido nucleico, tales como un ARN dirigido al ADN (por ejemplo, un ARN guía único modificado) y/o un molde de reparación de donante. Por ejemplo, los componentes se pueden mezclar con un agente de lipofección de modo que se encapsulen o empaqueten en emulsiones catiónicas submicrónicas de aceite en agua. Como alternativa, los componentes se pueden administrar sin un sistema de administración, por ejemplo, como solución acuosa.

60 Los métodos para preparar liposomas y encapsular polipéptidos y ácidos nucleicos en liposomas se describen en, por ejemplo, *Methods and Protocols, Volumen 1: Pharmaceutical Nanocarriers: Methods and Protocols*. (ed. Weissig). Humana Press, 2009 y Heyes *et al.* (2005) *J Controlled Release* 107:276-87. Los métodos para preparar micropartículas y encapsular polipéptidos y ácidos nucleicos se describen en, por ejemplo, *Functional Polymer Colloids and Microparticles volume 4 (Microspheres, microcapsules & liposomes)*. (eds. Arshady y Guyot). Citus Books, 2002 y *Microparticulate Systems for the Delivery of Proteins and Vaccines*. (eds. Cohen y Bernstein). CRC Press, 1996.

#### **E. Métodos para evaluar la eficiencia de la edición del genoma**

Para probar funcionalmente la presencia de la modificación de edición del genoma correcta, el ADN diana puede analizarse mediante métodos estándar conocidos por los expertos en la técnica. Por ejemplo, las mutaciones ins/del se pueden identificar mediante secuenciación utilizando el kit de detección de mutaciones SURVEYOR® (Integrated DNA Technologies, Coralville, IA) o el kit de identificación Guide-it™ Indel (Clontech, Mountain View, CA).

5 La reparación dirigida por homología (HDR) puede detectarse mediante métodos basados en PCR y en combinación con secuenciación o análisis RFLP. Los ejemplos no limitantes de kits basados en PCR incluyen el kit de detección de mutaciones Guide-it (Clontech) y el kit de detección de escisión genómica GeneArt® (Life Technologies, Carlsbad, CA). También se puede utilizar la secuenciación profunda, particularmente para una gran cantidad de muestras o sitios potenciales en la diana/fuera de la diana.

10 En determinadas realizaciones, la eficiencia (por ejemplo, la especificidad) de la edición del genoma corresponde al número o porcentaje de eventos de escisión del genoma en la diana en relación con el número o porcentaje de todos los eventos de escisión del genoma, incluidos los eventos dentro y fuera de la diana.

15 En algunas realizaciones, los ARNsg modificados descritos en el presente documento son capaces de mejorar la edición del genoma de una secuencia de ADN diana en una célula, tal como una célula primaria, en relación con los ARNsg sin modificar correspondientes. La edición del genoma puede comprender la reparación dirigida por homología (HDR) (por ejemplo, inserciones, delecciones o mutaciones puntuales) o unión de extremos no homólogos (NHEJ).

20 En determinadas realizaciones, la eficiencia de la edición del genoma mediada por nucleasa de una secuencia de ADN diana en una célula aumenta al menos aproximadamente 0,5 veces, 0,6 veces, 0,7 veces, 0,8 veces, 0,9 veces, 1 veces, 1,1 veces, 1,2 veces, 1,3 veces, 1,4 veces, 1,5 veces, 2 veces, 2,5 veces, 3 veces, 3,5 veces, 4 veces, 4,5 veces, 5 veces, 5,5 veces, 6 veces, 6,5 veces, 7 veces, 7,5 veces, 8 veces, 8,5 veces, 9 veces, 9,5 veces, 25 10 veces, 15 veces, 20 veces, 25 veces, 30 veces, 35 veces, 40 veces, 45 veces, 50 veces o más en presencia de un ARNsg modificado descrito en el presente documento en comparación con la secuencia de ARNsg no modificada correspondiente.

#### **F. Métodos para la regulación génica de un ácido nucleico diana en una célula**

30 En el presente documento se proporciona un método *in vitro* para inducir la regulación génica, por ejemplo, editar y/o modular el genoma (por ejemplo, inhibir o activar) la expresión génica, de un ácido nucleico diana en una célula primaria.

35 El método *in vitro* para inducir la edición del genoma incluye introducir en una célula primaria el ARNsg modificado descrito en el presente documento y una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. El ARNsg modificado guía la nucleasa Cas (por ejemplo, polipéptido Cas9) o variante o fragmento de la misma hacia el ácido nucleico diana (por ejemplo, ADN diana). El ARNsg modificado tiene actividad, estabilidad y/o especificidad mejoradas para el ADN diana en comparación con una secuencia de ARNsg no modificada correspondiente. En algunos casos, la edición del genoma es la unión de extremos no homólogos (NHEJ) del ADN diana. En otros casos, la edición del genoma es la reparación dirigida por homólogos (HDR) del ADN diana. En algunas realizaciones de HDR, se añade a la célula 40 un molde de reparación de donante recombinante. El molde de reparación de donante recombinante puede incluir dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de una secuencia de nucleótidos correspondiente al ADN diana. En algunas realizaciones, el molde de reparación de donante recombinante comprende un molde sintético de oligodesoxinucleótido monocatenario (ssODN) que comprende una secuencia de nucleótidos que codifica una mutación para corregir un polimorfismo de un solo nucleótido (SNP) y dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de la secuencia de nucleótidos que codifica la mutación. El ARNsg modificado y/o una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o un fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o 45 un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma se introducen en la célula usando cualquier método tal como mediante electroporación.

50 El método *in vitro* para modular (por ejemplo, inhibir o activar) la expresión génica de un ácido nucleico diana, por ejemplo, un ADN diana, en una célula primaria incluye introducir (por ejemplo, electroporación) en la célula el ARNsg modificado descrito en el presente documento y una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. En algunas 55 realizaciones, la variante de nucleasa Cas (por ejemplo, Cas9) es un polipéptido Cas deficiente en endonucleasa (por ejemplo, dCas9). En algunos casos, la variante Cas9 puede tener dos o más sustituciones de aminoácidos en

comparación con el polipéptido Cas9 de tipo salvaje. En otros casos, la variante Cas9 no puede escindir el ADN bicatenario. La variante de nucleasa Cas puede ser un polipéptido de fusión Cas (por ejemplo, dCas9). En algunas realizaciones, el polipéptido de fusión incluye un dominio de represión transcripcional, un dominio de activación transcripcional, factor de transcripción, enzima modificadora de histonas (por ejemplo, histona desacetilasa, histona metil transferasa, histona acetiltransferasa), una enzima modificadora del ADN (por ejemplo, ADN metiltransferasa) y similares.

5 El método *in vitro* para modular (por ejemplo, inhibir o activar) la expresión génica de un ácido nucleico diana, por ejemplo, un ARN diana, en una célula primaria incluye introducir (por ejemplo, electroporación) en la célula el ARNsg modificado descrito en el presente documento y una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. En algunas 10 realizaciones, la variante de Cas nucleasa (por ejemplo, Cas9) tiene actividad endonucleolítica reducida o carece de la misma. La variante de nucleasa Cas puede contener dos o más sustituciones de aminoácidos de modo que el polipéptido no pueda escindir el ADN bicatenario. La variante de Cas nucleasa (por ejemplo, Cas9) puede tener actividad endorribonucleasa y puede escindir el ARN diana.

15 20 **G. Métodos para prevenir o tratar una enfermedad genética en un sujeto**

La divulgación del ARNsg modificado en el presente documento se puede usar para modular la eficiencia de la regulación génica. Por ejemplo, el ARNsg modificado puede inducir la regulación génica con una actividad mejorada en relación con un ARNsg sin modificar correspondiente. En algunos casos, la actividad mejorada comprende una 25 mayor estabilidad del ARNsg modificado y/o una mayor especificidad del ARNsg modificado para un ácido nucleico diana. Como ejemplo adicional, el ARNsg modificado puede inducir la regulación génica con una disminución de la toxicidad celular en relación con un ARNsg sin modificar correspondiente.

30 Los ARNsg modificados se pueden aplicar a terapias de enfermedades genéticas basadas en nucleasas dirigidas. Los enfoques actuales para corregir con precisión las mutaciones genéticas en el genoma de las células primarias de los pacientes han sido muy ineficaces (menos del 1 por ciento de las células se pueden editar con precisión). La divulgación de ARNsg modificados en el presente documento puede mejorar la actividad de edición del genoma y aumentar la eficacia de las terapias basadas en la edición del genoma. En realizaciones particulares de la divulgación, los ARNsg modificados pueden usarse para la edición génica de genes *in vivo* en sujetos con una 35 enfermedad genética. Los ARNsg modificados se pueden administrar a un sujeto mediante cualquier vía de administración adecuada y a dosis o cantidades suficientes para potenciar el efecto (por ejemplo, mejorar la eficacia de edición del genoma) de la terapia basada en nucleasas.

40 La divulgación en el presente documento es un método para prevenir o tratar una enfermedad genética en un sujeto que lo necesita corrigiendo una mutación genética asociada con la enfermedad. El método incluye administrar al sujeto un ARNsg modificado descrito en el presente documento en una cantidad que sea suficiente para corregir la mutación. También se desvela en el presente documento el uso de un ARNsg modificado descrito en el presente documento en la fabricación de un medicamento para prevenir o tratar una enfermedad genética en un sujeto que lo necesita corrigiendo una mutación genética asociada con la enfermedad. El ARNsg modificado puede estar 45 contenido en una composición que también incluye una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, un ARNm que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma, o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica una nucleasa Cas (por ejemplo, polipéptido Cas9) o una variante o fragmento de la misma. En algunos casos, el ARNsg modificado se incluye en un sistema de administración descrito anteriormente.

50 55 60 Las enfermedades genéticas que pueden corregirse con el método incluyen, pero sin limitaciones, Inmunodeficiencia combinada grave ligada al cromosoma X, anemia de células falciformes, talasemia, hemofilia, neoplasia, cáncer, degeneración macular relacionada con la edad, esquizofrenia, trastornos de repetición de trinucleótidos, síndrome del cromosoma X frágil, trastornos relacionados con priones, esclerosis lateral amiotrófica, drogadicción, autismo, enfermedad de Alzheimer, enfermedad de Parkinson, fibrosis quística, enfermedades o trastornos de la sangre y de la coagulación, inflamación, enfermedades o trastornos relacionados con el sistema inmunológico, enfermedades metabólicas, enfermedades y trastornos del hígado, enfermedades y trastornos renales, enfermedades y trastornos musculares/esqueléticos (por ejemplo, distrofia muscular, distrofia muscular de Duchenne), enfermedades y trastornos neurológicos y neuronales, enfermedades y trastornos cardiovasculares, enfermedades y trastornos pulmonares, enfermedades y trastornos oculares, infecciones víricas (por ejemplo, infección por VIH) y similares.

**V. Ejemplos**

Los siguientes ejemplos se ofrecen para ilustrar, no para limitar, la invención reivindicada.

**Ejemplo 1. Los ARN guía modificados químicamente mejoran la edición del genoma por CRISPR/Cas en células primarias humanas.**

La edición del genoma mediada por CRISPR/Cas se basa en ARN guía que dirigen la escisión del ADN específica de sitio facilitada por la endonucleasa Cas. En el presente documento se informa del uso de ARN guía únicos (ARNsg) sintetizados químicamente y se muestra que la modificación de los ARNsg por alteraciones químicas mejora dramáticamente la edición del genoma en linfocitos T primarios humanos y células madre y progenitoras hematopoyéticas CD34+. Este enfoque es una forma simple y altamente efectiva de agilizar el desarrollo de la edición del genoma con el potencial de acelerar una amplia gama de aplicaciones biotecnológicas y terapéuticas de la tecnología CRISPR/Cas.

La edición del genoma con nucleasas diseñadas es una tecnología revolucionaria para modificar esencialmente cualquier secuencia genómica de interés (Porteus, M.H. y Carroll, D., *Nature biotechnology* 23, 967-973 (2005)). Esta tecnología explota las nucleasas diseñadas para generar roturas en la doble cadena (DSB, por sus siglas en inglés) específicas de sitio, seguidas de la resolución de las DSB mediante mecanismos de reparación celular endógenos. El resultado puede ser la mutación de un sitio específico a través de uniones terminales mutagénicas no homólogas (NHEJ, por sus siglas en inglés), creando inserciones o delecciones (ins/del) en el sitio de la rotura o el cambio preciso de una secuencia genómica mediante recombinación homóloga (RH) utilizando un molde donante introducido exógenamente (Hendel *et al.*, *Trends in Biotechnology* 33, 132-140 (2015)). Una adición importante reciente a esta tecnología es el sistema de repeticiones palindrómicas agrupadas y regularmente interespaciadas (CRISPR)/Cas que consiste en una nucleasa guiada por ARN (Cas) y un ARN guía corto (ARNsg) Jinek, M. *et al.*, *Science* 337, 816-821 (2012), Mali, P. *et al.*, *Science* 339, 823-826 (2013), Cong, L. *et al.*, *Science* 339, 819-823 (2013), Hsu *et al.*, *Cell* 157, 1262-1278 (2014)). El ARN guía está compuesto por dos ARN denominados ARN CRISPR (ARNcr) y ARNcr transactivador (ARNtracr), que, para fines de edición de genes, se fusionan típicamente en un ARN guía único (ARNsg) químérico. Los ARNsg consisten en 100 nucleótidos (nt) de los cuales 20 nt en el extremo 5' pueden hibridar con una secuencia de ADN diana por medio del emparejamiento de bases de Watson-Crick y guiar a la endonucleasa Cas para escindir el ADN genómico diana (Figura 1A). El ARNsg se puede administrar a las células como ARN, por ejemplo, preparado mediante transcripción *in vitro*, o utilizando un vector de ADN con el ARNsg expresado a partir de un promotor de ARN polimerasa III. Si bien la edición del genoma con el sistema CRISPR/Cas es altamente eficiente en líneas celulares humanas, la edición del genoma CRISPR/Cas en células humanas primarias generalmente supone un reto mayor. Las razones de esta disminución de la actividad siguen siendo esquivas, pero factores que contribuyen pueden implicar diferencias en las tasas de transfección, la actividad promotora, la actividad exonucleasa, la respuesta inmunológica innata al interferón cuando se administran ácidos nucleicos y la fidelidad de la reparación del ADN. En el presente documento se demuestra que los ARNsg sintetizados químicamente pueden inducir altos niveles de edición del genoma y se demuestra además que las alteraciones químicas de los ARNsg pueden mejorar dramáticamente la edición del genoma tanto en linfocitos T primarios humanos como en células madre y progenitoras hematopoyéticas (HSPC) CD34+. El aumento en la edición del genoma en estos tipos de células que utilizan ARNsg modificados químicamente se mejora aún más mediante la administración de Cas9 como ARNm o proteína en lugar de a través de un plásmido de expresión de ADN, generando así un sistema de administración simple y completo basado en ARN o ribonucleoproteína (RNP) para el sistema CRISPR/Cas.

## Resultados

Un avance reciente en la tecnología de síntesis de ARN hizo que fuera práctico sintetizar químicamente ARN de más de 100 nt de longitud (Dellinger, D.J. *et al.*, *Journal of the American Chemical Society* 133, 11540-11556 (2011)). Para probar la utilidad de los ARNsg sintetizados químicamente para la edición del genoma, se sintetizaron ARNsg de longitud completa de 100 nt usando un sintetizador ABI 394 y fosforamiditas de nucleósidos protegidas con 2'-O-tionocarbamato de acuerdo con los procedimientos descritos anteriormente (Dellinger, D.J. *et al.*, *Journal of the American Chemical Society* 133, 11540-11556 (2011)). También se sintetizaron ARNsg con diversas modificaciones químicas en ambos extremos para evaluar sus efectos sobre la eficacia (Figuras 1A y 1B, y Tabla 1). Cuando están presentes, las modificaciones químicas que comprenden 2'-O-metilo (M), 3'-fosforotioato de 2'-O-metilo (MS) o 2'-O-metilo 3'tioPACE (MSP) se incorporaron en tres nucleótidos terminales en ambos extremos 5' y 3'. Estas tres modificaciones se seleccionaron para la evaluación debido a su estabilidad previamente informada a las fosfordiesterasas de suero y veneno de serpiente, así como a su amplia gama de efectos informados sobre las propiedades inmunoestimuladoras de los ácidos nucleicos (Deleavey, G.F. y Damha, M.J., *Journal of the American Chemical Society* 125, 940-950 (2003), Hendel *et al.*, *Cell Reports* 7, 293-305 (2014)). Se seleccionaron tres ARNsg que previamente se ha indicado que producen altas tasas de edición de genes en líneas celulares de cada uno de los siguientes genes humanos (Hendel *et al.*, *Cell Reports* 7, 293-305 (2014), Cradick *et al.*, *Nucleic Acids Research* 41, 9584-9592 (2013)): (i) *IL2RG*, mutaciones en las que son responsables de la inmunodeficiencia primaria congénita SCID-X1, (ii) *HBB*, mutaciones en las que son responsables de la anemia de células falciformes y la talasemia, y (iii) *CCR5*, que codifica un correceptor del VIH y actualmente se está investigando como diana para la edición de genes terapéuticos en ensayos clínicos anti-VIH (Tebas, P. *et al.*, *The New England Journal of Medicine* 370, 901-910 (2014)). Primero se probó la escisión del ADN diana en presencia de proteína Cas9 recombinante purificada y cada uno de los ARNsg sintéticos *in vitro*. Todos los ARNsg modificados y sintetizados químicamente escindieron sus dianas de ADN de manera eficaz en presencia de la proteína Cas9 (Figuras 3 y 4).

A continuación, se examinaron si los ARNsg sintetizados podrían inducir ins/del dirigidas indicativas de NHEJ

mutagénico y alteración de genes en líneas celulares humanas. Se administró cada ARNsg junto con el plásmido de ADN que codifica Cas9 en células K562 mediante nucleofección y se analizaron las frecuencias de ins/del. La administración de 1 µg de ARNsg sintético sin modificar dirigido al locus *IL2RG* generó frecuencias de ins/del dirigidas del 2,4 %, demostrando su funcionalidad (Figura 1c, barras de sombreado claro). Para el ARNsg modificado con M, observamos un pequeño aumento en las frecuencias de ins/del hasta el 13,5 %, lo que sugiere una mejora modesta en la estabilidad sobre el ARNsg sin modificar. Sorprendentemente, la misma cantidad de ARNsg modificados químicamente aumentó la frecuencia de ins/del al 68,0 % y al 75,7 % para los ARNsg modificados con MS y con MSP, respectivamente. Incrementar la cantidad de ARNsg modificados 20 veces aumentó aún más la frecuencia de ins/del en todos los casos, llevando los ARNsg MS y MSP al 75,3 % y al 83,3 %, respectivamente (Figura 1C, barras de sombreado oscuro), que era comparable a las frecuencias obtenidas al expresar el sistema CRISPR/Cas a partir de un plásmido. Se obtuvieron resultados muy similares para las dianas *HBB* y *CCR5* (Figura 1C), lo que demuestra la capacidad general de los ARNsg modificados químicamente para dirigir altos niveles de mutaciones génicas dirigidas en células humanas. A continuación, los inventores determinaron si los ARNsg sintéticos podrían estimular el direccionamiento génico a través de HR. Diseñaron vectores de direccionamiento para cada uno de los tres loci con brazos de ~ 0,8 kb de homología 5' y 3' del sitio de corte de CRISPR. Entre los brazos de homología, se incluyó un casete de expresión de GFP que se puede integrar de manera estable tras una RH exitosa en el locus diana (Lombard. *et al.*, *Nature Methods*, 8, 861-869 (2011); Voit *et al.*, *Nucleic Acids Research* 42, 1365-1378 (2014)). En los tres loci diana, los ARNsg MS y MSP estimularon niveles significativamente más altos de RH que los ARNsg sin modificar y modificados con M (Figura 1D). A niveles más altos de ARNsg (20 µg) se midieron tasas de HR del 20,6 %, 25,5 % y 50,0 % para los ARNsg MSP en *IL2RG*, *HBB* y *CCR5*, respectivamente. Estas frecuencias son comparables o superiores a las obtenidas al expresar el sistema CRISPR/Cas completamente a partir de un plásmido.

Para investigar si los ARNsg modificados químicamente afectan a la actividad fuera de la diana, se utilizaron secuenciación profunda para medir las frecuencias de mutación fuera de la diana en tres loci diferentes para cada ARNsg. Ocho de estos sitios fuera de la diana se predijeron mediante herramientas de predicción *in silico* (Hsu *et al.*, *Nature Biotechnology* 31, 827-832 (2013), Cradick *et al.*, *Molecular therapy*. Los ácidos nucleicos, 3, e214 (2014)) y para el ARNsg de HBB se incluyó un sitio fuera de la diana para el que el ARNsg ha demostrado previamente tener altos niveles de actividad fuera de la diana (Cradick *et al.*, *Nucleic Acid Research*, 42, 9584-9592, (2013)). Para cuatro de los ocho sitios predichos, se descubrió una actividad fuera de la diana cercana a la basal para todos los ARNsg sintetizados químicamente a pesar de detectar niveles altos de actividad en la diana para los ARNsg modificados (Figuras 1E y 5, Tabla 5). La modificación química de los ARNsg tendió a resultar en una mayor actividad fuera de la diana en los otros cuatro sitios predichos, pero los niveles de actividad fueron variables. En algunos casos, la proporción de frecuencias de ins/del en la diana y fuera de la diana se mejoró con los ARNsg modificados, lo que sugiere una especificidad relativa mejorada. Los detalles de estas conclusiones se resumen a continuación. Solo se detectó actividad para el ARNsg sin modificar en dos sitios fuera de la diana (*IL2RG* "fuera de la diana 2" y *HBB* "fuera de la diana 1" en los que el ARNsg de *HBB* ha demostrado previamente una actividad significativa fuera de la diana), lo que permitió comparar las proporciones dentro:fuera de la diana entre ARNsg modificados y sin modificar. En el sitio de *IL2RG*, la proporción dentro:fuera de la diana fue 5,8 veces mejor para el ARNsg sin modificar en comparación con el ARNsg MSP (Figura 1E y Tabla 5), mientras que para el sitio *HBB* esta proporción fue 2,6 veces y 1,5 veces mejor (para 1 µg y 20 µg) para el ARNsg MSP en comparación con el ARNsg sin modificar (Figuras 1E y 5, Tabla 5). Al comparar las proporciones dentro:fuera de la diana de los ARNsg modificados con las proporciones dentro:fuera de la diana del plásmido de ARNsg, el plásmido de ARNsg tenía una mejor proporción en *CCR5* "fuera de la diana 2" y el *IL2RG* "fuera de la diana 2", mientras que para el *HBB* "fuera de la diana 1", el ARNsg MSP tenía la mejor proporción dentro:fuera de la diana (Figura 1E y Tabla 5). A excepción del sitio de *HBB* fuera de la diana 1, las frecuencias fuera de la diana más altas medidas fueron con los ARNsg *IL2RG* en "fuera de la diana 2", lo que arroja frecuencias de ins/del del 1,0 % y 7,8 % cuando se usa 1 µg y 20 µg de ARNsg MS, respectivamente. Curiosamente, la actividad fuera de la diana del ARNsg *IL2RG* MSP en el mismo sitio fue 2,7 veces y 2,8 veces menor, respectivamente, en comparación con el ARNsg MS, a pesar de tener una mayor actividad en la diana. Análogamente, la actividad fuera de la diana del ARNsg MSP fue mejor en comparación con el ARNsg MS en el sitio *HBB* fuera de la diana-1, en donde el ARNsg *HBB* ha demostrado previamente una actividad significativa fuera de la diana. En el *CCR5* "fuera de la diana 2" se observó lo contrario, el ARNsg MSP tenía una mayor actividad fuera de la diana en comparación con el ARNsg MS. Tomado en conjunto, estos resultados sugieren que típicamente los ARNsg modificados químicamente conservan una alta especificidad. Las diferencias observadas en las proporciones dentro:fuera de la diana sugieren la posibilidad de que las alteraciones químicas del ARNsg puedan tener el potencial de modular las actividades dentro:fuera de la diana, sin embargo, el impacto de una alteración química dada parece depender de la secuencia y también puede depender de otros factores como el tipo de célula y las condiciones de administración. Si estas observaciones son generalizables a otros ARNsg que se dirigen a diferentes loci en diferentes especies, será necesario realizar más estudios.

Para explorar más a fondo el rendimiento de los ARNsg modificados químicamente en líneas celulares humanas, se recurrió a una plataforma de administración de "todo el ARN" que coadministra los ARNsg con el ARNm que codifica Cas9. Se midieron las frecuencias de ins/del en el locus *IL2RG* usando cantidades variables de ARNm de Cas9 (1-15 µg) junto con cantidades variables del ARNsg MSP (1-20 µg) (Figura 6). Se observaron tasas igualmente altas de ins/del entre el 81-90 % para todas las concentraciones probadas, excepto por una modesta disminución al 70 % de ins/del cuando se usa 1 µg de cada uno de los ARNm de Cas9 y ARNsg MSP, lo que demuestra la alta eficiencia de

este enfoque de "todo el ARN". Para explorar si las modificaciones químicas tuvieron un efecto sobre la semivida de la actividad del ARNsg, se llevó a cabo un experimento en el que se coadministró o se administró secuencialmente ARNm de Cas9 y los diversos ARNsg sintetizados dirigidos a *IL2RG* mediante nucleofeción (Figura 7). La coadministración de ARNm de Cas9 y el ARNsg MS o MSP dio como resultado altas frecuencias de edición del 87 % y 84 % de ins/del, respectivamente (Figura 1F). Mientras que el ARNsg m dio lugar al 66 % de ins/del, el ARNsg sin modificar dio lugar a un modesto 7,0 %, lo que significa claramente la importancia de las modificaciones del ARNsg cuando se coadministran con el ARNm de Cas9. Curiosamente, la administración de ARNm de Cas9 primero y los ARNsg 4 u 8 horas más tarde produjo niveles altos y comparables de ins/del (83,1 % -92,4 %) en los cuatro ARNsg sin diferencia en las eficiencias entre los ARNsg sin modificar y modificados químicamente (Figura 1F). Por el contrario, cuando se administra el ARNsg primero seguido por la administración de ARNm de Cas9 4, 8, 12 o 24 horas después, las frecuencias de ins/del obtenidas con el ARNsg sin modificar cayeron a niveles cercanos al basal ya en el punto de tiempo de 4 horas. Para el ARNsg M también se observó una disminución en las frecuencias de ins/del a niveles cercanos a los basales, pero esta caída se retrasó ligeramente, lo que sugiere una mejora modesta en la estabilidad sobre el ARNsg sin modificar. Para el ARNsg MS, no se observaron una disminución en las frecuencias de ins/del hasta el punto de tiempo de 24 horas, en el que las tasas se redujeron al 53 %. Sorprendentemente, para el ARNsg MSP no se detectó una disminución significativa en la actividad incluso después de un retraso de 24 horas entre el ARNsg y la administración de Cas9. Estos resultados son consistentes con el modelo de los inventores en el que las modificaciones finales de ARNsg mejoran la estabilidad intracelular, permitiendo así una mayor eficacia de la edición del genoma cuando el ARNm de Cas9 y los ARNsg se coadministran en células humanas. El hecho de que las frecuencias de ins/del dirigidas por el ARNsg sin modificar no se redujeron cuando se administró Cas9 por primera vez, lo que sugiere que la proteína Cas9 protege a los ARNsg de la degradación. Para investigar esta hipótesis, los inventores formaron un complejo del ARNsg sin modificar o con MS IL2RG con la proteína Cas9 recombinante antes de la electroporación de este RNP activo en células K562 en dos cantidades diferentes (Figuras 1G y 8). Con la baja cantidad de RNP, se observaron frecuencias de ins/del 3,8 veces más altas cuando se usó el ARNsg MS en comparación con el ARNsg sin modificar (35,7 % frente al 9,5 %) y para la alta cantidad de RNP este aumento fue de 2,3 veces (81,0 % frente al 35,9 %). Esta proporción entre ARNsg MS y sin modificar fue, sin embargo, significativamente menor que la observada al coadministrar los ARNsg con ARNm de Cas9 (compárese con la Figura 1F "Coadministración"), lo que indica que la proteína Cas9 protege parcialmente el ARNsg sin modificar de la degradación. No obstante, las modificaciones en el ARNsg aún mejoran la eficiencia de edición de genes cuando se administra el ARNsg en complejo previo con la proteína Cas9. Seguidamente, los inventores compararon la actividad fuera de la diana de los ARNsg sin modificar y MS IL2RG en los tres sitios fuera de la diana previamente investigados cuando se administraron con el plásmido DE Cas9, ARNm de Cas9 o en complejo previo con la proteína Cas9. Con el ARNsg sin modificar, los inventores detectaron niveles bajos de ins/del (<0,37 %) en las tres fuentes Cas9 y los tres sitios fuera de la diana (Figuras 9A-B). Para el ARNsg MS, los inventores observaron mejores proporciones dentro:fuera de la diana en los tres sitios fuera de la diana al administrar el plásmido de Cas9 en comparación con el ARNm de Cas9 (2,6-3,0 veces). De forma notable, para la administración de Cas9 RNP con ARNsg MS, se detectaron proporciones dentro:fuera de la diana significativamente mejores en comparación con el plásmido de Cas9 y el ARNm para los tres sitios, con actividad fuera de la diana cercana a la basar en dos de los sitios. En resumen, los ARNsg modificados químicamente demuestran una ventaja significativa sobre el ARNsg sin modificar para la edición de genes en líneas celulares humanas cuando se coadministran con el ARNm de Cas9 o se administran como RNP

A continuación, los inventores probaron los ARNsg modificados químicamente en células primarias. La alteración del gen *CCR5* utilizando nucleasas de dedo de cinc (ZFN) en linfocitos T CD4+ se está explorando actualmente en ensayos clínicos como un tratamiento anti-VIH (Tebas *et al.* *New Eng Jour of Med*, 370, 901-910 (2014)). Sin embargo, un estudio reciente descubrió que el genoma de los linfocitos T primarios humanos es intrínsecamente difícil de editar con el sistema CRISPR/plásmido Cas9 usando un solo ARNsg y que los ARNsg que dan lugar a altas frecuencias de modificación de alelos en líneas celulares tienen una eficacia sustancialmente reducida en los linfocitos T, incluso cuando se enriquece para células transfectadas (Mandal *et al.*, *Cell Stem Cell*, 15, 643-652 (2014)). Los inventores probaron los ARNsg de *CCR5* modificados químicamente en linfocitos T primarios humano coadministrados con ARNm codificante de Cas9. Cabe destacar que, al usar ARNm de GFP, los inventores observaron de forma consistente eficiencias de nucleofeción superiores al 98 % en linfocitos T medidas mediante citometría de flujo (Figura 10), obviando la necesidad de enriquecer las células transfectadas. La nucleofeción del plásmido de ARNsg que codifica el ARNsg y Cas9 no dio lugar a frecuencias de modificación de alelos por encima de las basales (Figura 2A). Sin embargo, la cotransfección del ARNsg MSP con el plásmido de expresión de ADN para Cas9 fue capaz de rescatar la actividad al 9,3 % de las frecuencias de ins/del. La nucleofeción del ARNm de Cas9 con el ARNsg sin modificar o M no dio lugar a tasas de modificación de alelos por encima de las basales. Por el contrario, la nucleofeción del ARNm Cas9 con el ARNsg MS o MSP generó frecuencias notables del 48,7 % y del 47,9 % de ins/del, respectivamente. El incremento de las cantidades de ARNsg o ARNm de Cas9 no dio frecuencias de ins/del más altas (Figura 11). Para las células nucleofectadas con ARNm de Cas9 y ARNsg MSP, se encontraron frecuencias de modificación de alelos comparables en las poblaciones de linfocitos T CD4+, CD8+ y totales (Figura 12). Cabe destacar que, las altas frecuencias de modificación observadas fueron estables cuando se midieron para el ARNsg MSP durante un transcurso de tiempo de 21 días (Figura 13), y solo se observaron pequeños impactos en la viabilidad y proliferación celular después de la nucleofeción con los ARNsg modificados, en contraste con la nucleofeción plasmídica, lo que produjo muerte celular significativa y disminución del potencial de proliferación (Figuras 14 y 15). También se probaron el ARNsg MS en linfocitos T no estimulados, que han demostrado ser más

difíciles de editar que los linfocitos T estimulados (Yi *et al.*, Molecular Therapy Nucleic Acids, 3, e198 (2014)). A diferencia de los linfocitos T estimulados, las frecuencias de ins/del en los linfocitos T no estimulados oscilaron entre el 6,6 % y el 22,2 %, lo que muestra mayor variabilidad del donante (Figura 16). Aunque es menor que en los linfocitos T estimulados, estas frecuencias de edición en linfocitos T no estimulados pueden todavía tener utilidad, en

5 particular en los linfocitos T diseñados en los que la activación y el cultivo prolongado puede afectar a la posterior funcionalidad biológica de las células. A continuación, los inventores probaron la administración de RNP utilizando ARNsg MS y sin modificar en linfocitos T primarios estimulados (Figura 2B). De manera similar a los resultados obtenidos en las células K562, los inventores observaron una mejora de 2,4 veces en las frecuencias de ins/del del ARNsg MS sobre el ARNsg sin modificar (30,7 % frente a 12,8 %), lo que respalda aún más el uso de ARNsg químicamente modificados cuando se administran complejos previos con la proteína Cas9.

10 La terapia génica en las HSPC se ha explorado ampliamente para tratar trastornos genéticos o adquiridos del sistema hematopoyético. Los inventores probaron si los ARNsg modificados químicamente dirigidos a *IL2RG* y *HBB* son funcionales en las HSPC CD34+ aisladas de sangre periférica movilizada. De nuevo, el plásmido de ARNsg que 15 expresa tanto ARNsg como Cas9 no dio lugar a frecuencias de ins/del detectables, pero la cotransfección del ARNsg MSP con el plásmido de expresión de ADN de Cas9 rescató frecuencias de ins/del del 3,2 % y 5,2 % para *IL2RG* y *HBB*, respectivamente (Figuras 2C y 17). Similar a las observaciones de los inventores en linfocitos T, no se detectaron ins/del en ninguno de los loci utilizando los ARNsg sin modificar o M cuando se cotransfectaron con el ARNm de Cas9, pero los ARNsg MS y MSP dieron lugar a altas frecuencias de ins/del: para *IL2RG* se detectó un 20 17,5 % y un 17,7 % de frecuencias de ins/del para los ARNsg MS y MSP, respectivamente, y para *HBB*, las frecuencias fueron incluso mayores al 23,4 % y 22,0 %, respectivamente. Otros estudios podrían dilucidar si las modificaciones genéticas de las HSPC mediante ARNsg modificados químicamente afectan a su capacidad de multipotencialidad o si las eficiencias de edición difieren entre las células madre re poblantes a largo plazo y las 25 células progenitoras comprometidas con el linaje.

25 Un estudio reciente mostró que el uso simultáneo de dos ARNsg podría mejorar la alteración de genes en los linfocitos T primarios humanos y en las HSPC CD34+ (Mandal *et al.*, Cell Stem Cell, 15, 643-652 (2014)). Los inventores sintetizaron ARNsg MS y MSP de CCR5 que tenían las secuencias informadas en ese estudio (denominadas "D" y "Q") que separaban 205 pb. Los inventores los probaron en linfocitos T y en HSPC CD34+ 30 coadministrados con ARNm de Cas9. Cuando los dos ARNsg se utilizaron individualmente, se cuantificaron las frecuencias de modificación de los alelos mediante análisis TIDE. Cuando se usaron ambos, se cuantificaron las frecuencias de modificación de alelos mediante la secuenciación de productos de PCR clonados, lo que también permitió cuantificar el espectro completo de eventos de edición, incluyendo la alta incidencia notificada previamente de delecciones de secuencia entre los dos sitios diana de ARNsg. Cuando se usa individualmente en linfocitos T primarios, el ARNsg "D" dio lugar al 56,0 % y 56,3 % de ins/del para el ARNsg MS y MSP, respectivamente, y el ARNsg "Q" dio lugar a 62,6 % y 69,6 %, respectivamente (Figuras 2D y 18A). Cuando se usan en combinación, las 35 frecuencias de modificación de los alelos aumentaron, ya que se observaron 73,9 % y 93,1 % para los ARNsg MS y MSP, respectivamente, de los cuales la mayoría de los eventos de modificación fueron delecciones entre los dos sitios diana de ARNsg (Figura 18B). En HSPC CD34+, las observaciones de los inventores fueron similares, aunque 40 las frecuencias generales fueron más bajas. Por lo tanto, para el ARNsg "D" se observaron frecuencias de modificación de alelos del 9,8 % y 11,2 % para el ARNsg MS y MSP, respectivamente, y 17,8 % y 19,2 % para el ARNsg "Q" (Figuras 2D y 18B). Cuando se usa en combinación, las frecuencias aumentaron al 37,8 % y al 43,0 % para los ARNsg MS y MSP, respectivamente. Los inventores concluyeron que el uso de dos ARNsg modificados químicamente es una forma muy eficaz de facilitar la alteración de genes en linfocitos T humanos primarios y HSPC 45 CD34+.

50 La Figura 19 proporciona datos experimentales adicionales que muestran que los ARNsg modificados con MS funcionan mejor que los ARNsg sin modificar en las HSPC CD34+. La Figura 20 muestra que los ARNsg modificados pueden usarse para una edición eficiente del genoma multiplexado.

55 En este estudio, los inventores mostraron que los ARNsg sintetizados químicamente se pueden utilizar de forma eficaz para la edición del genoma dirigida y se demostró que los ARNsg modificados químicamente mejoran significativamente las eficiencias de edición del genoma en los linfocitos T primarios humanos y las HSPC CD34+. Los ARNsg modificados y sintetizados químicamente ofrecen ventajas sobre los ARNsg expresados o transcritos *in vitro*, entre ellos: (1) mayor eficacia, (2) producción robusta y escalable de ARNsg de alta pureza para aplicaciones biotecnológicas y terapéuticas, (3) mayor flexibilidad en el diseño de ARNsg en contraste con las limitaciones sobre los primeros nucleótidos transcritos impuestas por los promotores U6 o T7, que normalmente se usan para la expresión en plásmidos o la transcripción *in vitro* de ARNsg, respectivamente y (4) habilitación de una plataforma CRISPR de "ARN solo" o RNP altamente activa con menor citotoxicidad en células primarias que los sistemas basados en plásmidos de ADN. La simplificación del sistema CRISPR/Cas a un sistema CRISPR puramente de ARN o RNP se presta a la formulación en diferentes vectores de nanopartículas para su administración *in vivo*, de modo que la nucleasa no se expresará continuamente, como sería cuando se administra como parte de un vector viral. Adicionalmente, los inventores anticiparon que los ARNsg modificados químicamente mejorarán la edición del genoma multiplexado, así como los experimentos de pocillos múltiples de alto rendimiento. También anticiparon que 60 los ARNsg modificados refinarán una amplia gama de tecnologías asociadas con CRISPR, como los sistemas CRISPRi/CRISPRa para la inhibición y activación de la expresión génica (Gilbert, L.A. *et al.*, Cell 159, 647-661 65

(2014)) la herramienta de imágenes CRISPR para la visualización dinámica de loci genómicos (Chen, B. *et al.*, *Cell* 155, 1479-1491 (2013)) y reconocimiento y escisión de ARN mediada por CRISPR (O'Connell, M.R. *et al.*, *Nature* 516, 263-266 (2014)). La tecnología se puede desarrollar aún más para mejorar la administración intracelular a las células o tejidos diana y permitir la conjugación con varias moléculas para la formación de imágenes y estudios bioquímicos. Estudios futuros podrían investigar una mayor variedad de modificaciones químicas, explorar diferentes ubicaciones de las modificaciones para el diseño racional de ARNsg optimizados, así como el mecanismo para la actividad mejorada de ARNsg modificados. En conclusión, se cree que los ARNsg modificados químicamente como los presentados en el presente documento mejorarán significativamente una amplia gama de aplicaciones biotecnológicas y terapéuticas CRISPR/Cas.

5 10 **Materiales y métodos**

### 1. Síntesis de ARNsg

15 Todos los oligómeros de ARN se sintetizaron en un sintetizador ABI 394 (Life Technologies, Carlsbad, CA, EE.UU.) utilizando fosforamiditas de nucleósidos protegidas con 2'-O-tionocarbamato (Sigma-Aldrich, St. Louis, MO, EE.UU. o Thermo Fisher, Waltham, MA, EE.UU) de acuerdo con los procedimientos descritos previamente (Dellinger *et al.*, *Journal of the American Chemical Society* 133, 11540-11556 (2011)). Las 2'-O.metyl fosforamiditas se adquirieron en Thermo Scientific, Grand Island, NY y se incorporaron en oligómeros de ARN en las mismas condiciones que las 20 fosforamiditas protegidas con 2'-O-tionocarbamato. Los nucleósidos de éster de 2'-O-metyl-3'-O-(diisopropilamino)fosfinoacético-1,1-dimetilcianoetilo-5'-O-dmetoxitritilo utilizados para la síntesis de ARN modificados con tiofosfonoacetato (tioPACE) se sintetizaron esencialmente de acuerdo con métodos publicados (Threlfall *et al.*, *Organic & biomolecular chemistry*, 10, 746-754 (2012); Dellinger *et al.*, *Journal of the American Chemical Society*, 125, 940-950 (2003)). Para oligómeros que contienen fosforotioato, la etapa de oxidación del yodo 25 después de la reacción de acoplamiento se reemplazó por una etapa de sulfuración usando una solución 0,05 M de 3-((N,N-dimetilaminometiliden)amino)-3H-1,2,4-ditiazol-5-tiona en una mezcla de piridina-acetonitrilo (3:2) durante 6 minutos. A menos que se indique lo contrario, los reactivos para la síntesis de ARN en fase sólida se adquirieron en Glen Research (Sterling, VA, Estados Unidos).

30 35 Todos los oligonucleótidos se purificaron usando HPLC de fase inversa y se analizaron mediante LC-MS usando un sistema LC Agilent serie 1290 Infinity acoplado a un espectrómetro de masas Agilent 6520 Q-TOF (Agilent Technologies, Santa Clara, CA, EE.UU.). En la tabla 1 se muestran las secuencias de todos los ARNsg utilizados y las masas obtenidas de la desconvolución de la serie de picos de estado de carga encontrados. La desconvolución se realizó utilizando el software Mass Hunter Qualitative Analysis (versión B.06.00) (Agilent).

Tabla 1. Descripción general de todos los ARNsg utilizados en el presente estudio.

SEQ ID NO:	Nombre	Secuencia	Calc. Mas. (Da)	Obs. Mas. (Da)
SEQ ID NO: 4	ARNsg de HBB sin modificar	CUUGCCCCACAGGGCAGUAAGUUUUAGAGCUAGAAAAG AAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAAAAG UGGCACCGAGUCGGUGCUUU	32187,42	32187,84
SEQ ID NO: 5	ARNsg M de HBB	<u>CUUGCCCCACAGGGCAGUAAGUUUUAGAGCUAGAAAAG</u> AAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAAAAG UGGCACCGAGUCGGUGCUUU	32271,54	32271,35
SEQ ID NO: 6	ARNsg MS de HBB	<u>C•U•U•GCC</u> CCCCACAGGGCAGUAAGUUUUAGAGCUAGAAA AGCAAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32367,42	32367,31
SEQ ID NO: 7	ARNsg MSP de HBB	<u>C•U•U•GCC</u> CCCCACAGGGCAGUAAGUUUUAGAGCUAGAAA AGCAAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32619,93	32619,39
SEQ ID NO: 8	ARNsg de IL2G sin modificar	UGGUAAUGAUGGCUUCAACAGUUUUAGAGCUAGAAAAG CAAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAAAA UGGCACCGAGUCGGUGCUUU	32214,40	32214,37
SEQ ID NO: 9	ARNsg M de IL2G	<u>UGGUAAUGAUGGCUUCAACAGUUUUAGAGCUAGAAAAG</u> CAAGUUAAAAUAAGGCUAGUCGUUAUCACUUGAAAAA UGGCACCGAGUCGGUGCUUU	32298,52	32297,01

(continuación)

SEQ ID NO:	Nombre	Secuencia	Calc. Mas. (Da)	Obs. Mas. (Da)
SEQ ID NO: 10	ARNsg MS de IL2G	<u>U•G•G</u> •UAUAUGAUGGCUUCAACAGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32394,4	32395,43
SEQ ID NO: 11	ARNsg MSP de IL2G	<u>U•G♦G</u> •UAUAUGAUGGCUUCAACAGUUUUAGAGCUAGAAA UAGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA	32646,91	32645,39
		AAAGUGGCACCGAGUCGGUGCU <u>♦U♦U♦U</u>		
SEQ ID NO: 12	ARNsg de CCR5 sin modificar	GGCAGCAUAGUGAGCCCAGAGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA GUGGCACCGAGUCGGUGCUUUU	32290,51	32289,07
SEQ ID NO: 13	ARNsg M de CCR5	GGCAGCAUAGUGAGCCCAGAGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA GUGGCACCGAGUCGGUGCUUUU	32374,63	32375,3
SEQ ID NO: 14	ARNsg MS de CCR5	<u>G•G•C</u> •AGCAUAGUGAGCCCAGAGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32470,51	32469,92
SEQ ID NO: 15	ARNsg MSP de CCR5	<u>G♦G♦C</u> •AGCAUAGUGAGCCCAGAGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>♦U♦U♦U</u>	32723,02	32721,96
SEQ ID NO: 16	ARNsg "D" MS de CCR5	<u>U•C•A</u> •CUAUGCUGCCGCCAGUGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32281,32	32282,52
SEQ ID NO: 17	ARNsg MSP "D" de CCR5	<u>U♦C♦A</u> •CUAUGCUGCCGCCAGUGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>♦U♦U♦U</u>	32533,83	32533,55
SEQ ID NO: 18	ARNsg MS "Q" de CCR5	<u>G•C•U</u> •GUGUUUGCGUCUCUCCGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>•U•U•U</u>	32252,22	32253,21
SEQ ID NO: 19	ARNsg MSP "Q" de CCR5	<u>G♦C♦U</u> •GUGUUUGCGUCUCUCCGUUUUAGAGCUAGAAA AGCAAGUUAAAAAUAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <u>♦U♦U♦U</u>	32504,73	32504,63

Se indican las secuencias de ARNsg así como los pesos moleculares calculados y observados. Los nucleótidos con modificaciones 2'-O-metilo están subrayados. Las modificaciones en la cadena principal de fosfato se indican con • (MS) y ♦ (MSP).

## 2. Ensayos de escisión *in vitro*

Se prepararon dianas direccionables por PAM de 4 kb mediante amplificación por PCR preparativa de secuencias humanas transmitidas por plásmidos. En un volumen de reacción de 20  $\mu$ l, 50 fmoles de ADN diana linealizado en presencia de ARNsg 50 nM, proteína Cas9 purificada recombinante 39 nM (Agilent) y MgCl<sub>2</sub> 10 mM a pH 7,6 se incubaron a 37 °C durante 30 minutos. Al finalizar, se añadieron 0,5  $\mu$ l de RNase It (Agilent) y se continuó la incubación a 37 °C durante 5 minutos y luego a 70 °C durante 15 minutos. Posteriormente se añadieron 0,5  $\mu$ l de proteinasa K (calidad Mol. Bio., NEB) y se incubó a 37 °C durante 15 minutos. Se cargaron alícuotas en un LabChip de ADN 7500 y se analizaron en un Bioanalyzer 2200. Las frecuencias de escisión se calcularon mediante la fórmula: (a/b) x 100 donde "a" es la suma de las intensidades de banda de los dos productos de escisión y b es la suma de las intensidades de banda del ADN escindido y no escindido.

## 10 3. Plásmidos

Los vectores de expresión de ARNsg se construyeron clonando secuencias diana de oligonucleótidos de 20 pb en px330 (plásmido Addgene n.º 42230) que contiene un casete de expresión de SpCas9 optimizado para codones humanos y un promotor U6 humano que dirige la expresión del ARNsg químérico (véase la Tabla 1 para las secuencias de ARNsg).

Los tres vectores de direccionamiento de plásmidos llevan aproximadamente 2 brazos de homología de 800 pb, que se generaron mediante amplificación por PCR de los loci correspondientes utilizando ADN genómico aislado de células K562. A continuación, los brazos de homología se clonaron en un vector de ~ 2.900 pares de bases basado en pBluescript SK + usando métodos de clonación estándar. Entre los brazos de homología, tanto los donantes *HBB* como *CCR5* contienen el promotor EF1 $\alpha$  que dirige la expresión de GFP. El donante de *IL2RG* carece de un promotor y depende de la actividad endógena del gen *IL2RG* para impulsar la expresión de GFP. La secuencia de aminoácidos del vector dirigido a *IL2RG* se expone en la SEQ ID NO:1. La secuencia de aminoácidos del vector dirigido a *HBB* se expone en la SEQ ID NO: 2. La secuencia de aminoácidos del vector dirigido a *CCR5* se expone en la SEQ ID NO:3.

## 4. Cultivo celular y nucleofecciones

Las células K562 y linfocitos T se cultivaron a 37 °C, con 5 % de CO<sub>2</sub> y niveles de oxígeno ambiental. Se cultivaron células madre/progenitoras hematopoyéticas (HSPC) CD34+ a 37 °C, 5 % de CO<sub>2</sub> y 5 % de O<sub>2</sub>. Las células K562 se mantuvieron en RPMI 1640 (HyClone) suplementado con suero de crecimiento bovino al 10 %, 100 mg/ml de estreptomicina, 100 unidades/ml de penicilina y L-glutamina 2 mM. Las células K562 se nucleofectaron usando el Nucleofector 2b de Lonza (programa T-016) y un tampón de nucleofección que contenía KH<sub>2</sub>PO<sub>4</sub> 100 mM, NaHCO<sub>3</sub> 15 mM, MgCl<sub>2</sub> 12 mM x 6H<sub>2</sub>O, ATP 8 mM, glucosa 2 mM (pH 7,4). Condiciones de nucleofección: 100  $\mu$ l de solución de nucleofección, 10<sup>6</sup> células, 1 a 20  $\mu$ g de ARNsg químicamente modificado, 1 a 15  $\mu$ g de ARNm de Cas9 (ARNm de Cas9, 5 meC,  $\Psi$ , Código de producto: L-6125, TriLink BioTechnologies, San Diego, CA, EE.UU.), 2  $\mu$ g de plásmido que codifica ARNsg/Cas9, o 5  $\mu$ g de plásmido donante para RH. Se aislaron linfocitos T CD3+ de capas leucocitarias obtenidas del Centro de sangre de la Escuela de Medicina de Stanford utilizando un kit de aislamiento de células Pan T Cell humanas (Miltenyi Biotec, San Diego, CA, EE.UU.). Las células CD3+ se mantuvieron en X-VIVO 15 (Lonza, Walkersville, MD, EE.UU.) suplementado con suero humano al 5 % (Sigma-Aldrich, St. Louis, MO, EE.UU.), 100 UI/ml de IL-2 recombinante humana (Peprotech, Rocky Hill, NJ, EE.UU.) y 10 ng/ml de IL-7 recombinante humana (BD Biosciences, San Jose, CA, EE.UU.). Antes de la nucleofección, los linfocitos T se activaron durante tres días con anticuerpos anti-CD3 inmovilizados (clon: OKT3, eBioscience, San Diego, CA, EE.UU.) y anticuerpos anti-CD28 solubles (clon: CD28.2, eBioscience). Para los linfocitos T CD3+ no activados, las células se nucleofectaron inmediatamente después del aislamiento. Los linfocitos T se nucleofectaron usando el Nucleofector 2b de Lonza (programa U-014) y el Kit Nucleofector de linfocitos T humanos (VPA-1002, Lonza). Condiciones de nucleofección: 100  $\mu$ l de solución de nucleofección, 10<sup>6</sup> células, 10 a 20  $\mu$ g de ARNsg químicamente modificado, 15 a 30  $\mu$ g de Cas9 (o 15  $\mu$ g de ARNm de eGFP, TriLink BioTechnologies, San Diego, CA, EE.UU.), 1  $\mu$ g de plásmido que codifica ARNsg/Cas9. Se adquirieron HSPC CD34+ de sangre periférica humana movilizada de AllCells y se descongelaron de acuerdo con las instrucciones del fabricante. Las HSPC CD34+ se mantuvieron en X-VIVO 15 (Lonza) suplementado con SCF (100 ng/ml), TPO (100 ng/ml), ligando Flt3 (100 ng/ml), IL-6 (100 ng/ml) y StemRegenin1 (0,75 mM). Las HSPC CD34+ se nucleofectaron usando el Nucleofector 4D de Lonza (programa EO-100) y el kit de nucleofector 4D de células primarias P3 (V4XP-3024). Condiciones de nucleofección: 100  $\mu$ l de solución de nucleofección, 5x10<sup>5</sup> células, 10  $\mu$ g de ARNsg químicamente modificado, 15  $\mu$ g de ARNm de Cas9, 1  $\mu$ g de plásmido. Para experimentos con Cas9 RNP, la proteína Cas9 se adquirió en PNA Bio (Thousand Oaks, CA, EE.UU.) o Life Technologies (Carlsbad, CA, EE.UU.). Para todos los experimentos de RNP excepto para la Figura 8, se utilizó proteína Cas9 de PNA Bio. La proteína Cas9 se hizo formar complejo con ARNsg en una proporción molar de Cas9:ARNsg de 1:2,5 durante 10 minutos a 25 °C. Las RNP se nucleofectaron en células K562 o linfocitos T como se ha descrito anteriormente con 10<sup>6</sup> células en 100  $\mu$ l de las respectivas soluciones de nucleofección. Para los experimentos de ARNsg duales, la cantidad total de ARNsg fue de 10  $\mu$ g (10  $\mu$ g cuando se usan individualmente y 2 x 5  $\mu$ g cuando se usan juntos). Tanto para los linfocitos T como para las HSPC CD34+, los ARNsg se nucleofectaron con 15  $\mu$ g de ARNm de Cas9 en 10<sup>6</sup> células. Las nucleofecciones de linfocitos T se realizaron como anteriormente, mientras que la nucleofección de HSPC CD34+ fue similar a las nucleofecciones de células T usando el Nucleofector 2b de Lonza (programa U-014) y el Kit de Nucleofector de linfocitos T humanos (VPA-1002, Lonza). Directamente después de la nucleofección, las HSPC CD34+ se incubaron a 30 °C durante 24 horas, después de lo

cual se transfirieron a 37 °C hasta la recolección del ADN genómico.

#### 5. Citometría de flujo y clasificación de células activadas por fluorescencia (FACS)

- 5 Para experimentos de RH, las células se analizaron 2-3 semanas después de la nucleofección dependiendo del tipo de célula cuando no quedaba expresión de eGFP del plásmido episomal. La expresión de eGFP se midió en un citómetro de flujo Accuri C6 (BD Biosciences, San Jose, CA, EE.UU.). La muerte celular se midió con el kit de tinción Fixable Red Dead Cell Stain Kit VIDAS/MUERTAS (Life Technologies, Carlsbad, CA, EE.UU.) de acuerdo con las instrucciones del fabricante y las células se analizaron en el citómetro de flujo Accuri C6. Para clasificar los linfocitos T CD3+ en poblaciones CD4+ y CD8+, las células se tiñeron con una mezcla de PE-Cy7 anti-CD4 humano (clon: RPA-T4, Tonbo Biosciences, San Diego, CA, EE.UU.) y un anti-CD8a humano de APC (clon: RPA-T8, Tonbo Biosciences) y las dos poblaciones se clasificaron en un FACS Aria II SORP.

#### 6. Medición de las frecuencias de modificación de alelos usando un ensayo TIDE y T7

- 15 Para los ensayos TIDE y T7, el ADNg se extrajo de las células tres días después de la nucleofección (si no se indica lo contrario) utilizando la solución de extracción de ADN QuickExtract (Epicentre, Madison, WI, EE.UU.) siguiendo las instrucciones del fabricante. Los amplicones de PCR que abarcan los sitios diana genómicos de ARNsg se generaron utilizando la mezcla iProof High-Fidelity Master Mix (Bio-Rad, Hercules, CA, EE.UU.) con los siguientes pares de cebadores:

IL2RG\_dir (SEQ ID NO: 84): 5'-T CACACAGCACATATTGCCACACCCCTTG-3',

IL2RG\_INV (SEQ ID NO: 85): 5'-TGCCCACATGATTGTAATGCCAGTGG-3',

25 HBB\_dir (SEQ ID NO: 86): 5'-CCAACCTCTAACGCCAGTGCCAGAAAGAG-3',

HBB\_inv (SEQ ID NO: 87) 5'-AGTCAGTGCCTATCAGAAACCCAAGAG-3',

30 CCR5\_dir (SEQ ID NO: 88): 5'-GCACAGGGTGGAAACAAGATGG-3',

CCR5\_inv (SEQ ID NO: 89): 5'-CACCAACCCAAAGGTGACCGT-3'.

- 35 Para los ensayos de T7, los amplicones de PCR se purificaron y se desnaturizaron 200 ng y se rehíbridaron en un termociclador y se digirieron con Endonucleasa I de T7 (New England Biolabs, Waltham, MA, EE.UU.) de acuerdo con el protocolo del fabricante. El ADN digerido se procesó en un gel de poliacrilamida TBE al 4-20 %, se tiñó con colorante de ácido nucleico Diamond (Promega, Madison, WI, EE. UU.) y se visualizó en un ChemiDoc XRS+ (Bio-Rad). Las intensidades de banda se analizaron utilizando el software Image Lab (Bio-Rad) y las frecuencias de modificación de alelos se calcularon con la fórmula:  $100 \times (1 - (1 - \text{fracción escindida})^{0,5})$ . Para analizar las 40 frecuencias de modificación de alelos usando TIDE (seguimiento de ins/del por descomposición, Brinkman *et al.*, Nucleic Acids Research 42, e168 (2014)), los productos de PCR purificados se secuenciaron por Sanger usando ambos cebadores y cada cromatograma de secuencia se analizó con el software TIDE en línea disponible en el sitio web tide.nki.nl. Los análisis se realizaron utilizando una secuencia de referencia de una muestra transfectada de forma simulada. Los parámetros se establecieron en un tamaño máximo de ins/del de 10 nucleótidos y la ventana de 45 descomposición para cubrir la ventana más grande posible con trazas de alta calidad. Todos los análisis TIDE por debajo de la sensibilidad de detección del 3,5 % se establecieron en 0 %. Para la secuenciación de fragmentos de PCR clonados con TOPO se generó un amplicón de 2,1 kb (tamaño SV) que abarcaba el sitio o sitios de escisión usando la mezcla i Proof High-Fidelity Master Mix y los cebadores 5'-GGCTGTTGTCATCTGACCTTCCC-3' (SEQ ID NO: 90) y 5'-TGAAACTGAGCTTGCTCGCTG-3' (SEQ ID NO: 91) con 25 ciclos, incluyendo una temperatura de hibridación de 72 °C y un tiempo de elongación de 2 minutos. Los productos de reacción de la PCR se subclonaron en n plásmido usando directamente el kit de clonación Zero Blunt TOPO PCR (Life Technologies) de acuerdo con el protocolo del fabricante. Las reacciones TOPO se transformaron en células competentes XL-1 Blue, se sembraron en placas de agar con kanamicina y las colonias individuales se secuenciaron directamente de las placas por McLab (South San Francisco, CA, EE.UU.) mediante amplificación con círculo rodante seguido de secuenciación usando el cebador 5'-GCACAGGGTGGAAACAAGATGG-3' (SEQ ID NO: 92).

#### 7. Ensayo de proliferación

- 60 Para medir la proliferación de linfocitos T después de la nucleofección, se usó el ensayo CellTiter-Glo 2.0 (Promega, Madison, WI, EE. UU.). Directamente después de la nucleofección, los linfocitos T se transfirieron a múltiples placas de 96 pocillos con fondo en U de 96 pocillos a  $3 \times 10^4$  células/pocillo. Directamente después de la nucleofección y a intervalos de 24 horas, las células se transfirieron a placas blancas de 96 pocillos en 100 µl de medio y se añadieron 100 µl de CellTiter-Glo 2.0 según las directrices del fabricante. La luminiscencia se leyó en un Tecan Infinite 200 PRO (Tecan, Mannedorf, Suiza) utilizando un tiempo de integración de 1 segundo.

**8. Secuenciación profunda para cuantificar la eficiencia y la especificidad de la modificación del genoma**

Para cada experimento de direccionamiento de genes, se extrajo ADN genómico de las diversas células K562 tratadas con CRISPR y de control 48 horas después de la transfección. Las regiones genómicas que flanquean la diana de CRISPR y tres fuera de la diana (Tabla 2) se amplificaron mediante dos rondas de PCR para vincular códigos de barras (específicos del tratamiento) y adaptadores de secuenciación de Illumina (Tabla 3). Los amplicones de PCR con código de barras se agruparon equimolarmente, se purificaron mediante una columna de giro y se secuenciaron en la plataforma del secuenciador de ADN Illumina MiSeq. Véase, "9. Generación de amplicones CRISPR dentro y fuera de la diana previsto para secuenciación profunda" en la sección siguiente para obtener más detalles.

La Tabla 2 proporciona una lista de loci dentro y fuera de la diana interrogados por secuenciación profunda de amplicones de PCR. Para los experimentos de CRISPR dirigidos a CCR5, HBB e IL2RG, la secuencia diana genómica pretendida ("EN") y tres secuencias diana FUERA predichas computacionalmente ("FUERA 1-3") se presentan con su ubicación genómica (ensamblaje de construcción del genoma humano hg38). Las secuencias seleccionadas fuera de la diana fueron las que pronosticaron las máximas puntuaciones tanto por el CÓSMIDO (Cradick *et al.*, Molecular therapy. Nucleic acids 3, e214 (2014) and Optimized CRISPR Design (Hsu *et al.*, Nature Biotechnology, 31, 827-832 (2013)) herramientas web (región MIT), excepto HBB-FUERA3 que solo se predijo que tendría una actividad significativa por CÓSMIDO. La actividad diana prevista aumenta con el aumento de los valores de puntuación CÓSMIDO y la disminución de los valores de puntuación del "diseño MIT". Los sitios PAM y los emparejamientos erróneos en las secuencias fuera de la diana se indican mediante texto en negrita y de color rojo, respectivamente.

Tabla 2. Lista de loci dentro y fuera de la diana interrogados mediante secuenciación profunda de amplicones de PCR.

SEQ ID NO:	ID de la diana	Secuencia del sitio diana (5' → 3')	Ubicación genómica		Puntuación diana	
			Cadena	CÓSMIDO	Diseño MIT	
ARN guía dirigido a <i>CCR5</i>						
SEQ ID NO: 20	EN	<b>GGCAGCATAGTAGGCCAGAG</b> <b>G</b>	Chr3:46373153-46373175	-	0	56
SEQ ID NO: 21	OFF1	<u>ATCATCATAGTAGGCCAGAG</u> <b>G</b>	Chr3:15440658-15440680	+	0,44	2,4
SEQ ID NO: 22	FUERA2	<u>ACCAGCAGAGTAGGCCAGAGG</u> <b>G</b>	Chr4:3744369-3744391	+	0,52	2,6
SEQ ID NO: 23	FUERA3	<u>AGGAGCAGAGTAGGCCAGAGA</u> <b>G</b>	Chr15:92469456-92469478	+	0,54	2,6
ARN guía dirigido a <i>HBB</i>						
SEQ ID NO: 24	EN	<b>CTTGCCCCACAGGGCAGTAACGG</b>	Chr11:5226968-5226990	+	0	65
SEQ ID NO: 25	OFF1	<u>TCAGCCCCACAGGGCAGTAAGG</u> <b>G</b>	Chr9:101833584-101833606	+	0,4	2,3
SEQ ID NO: 26	FUERA2	<b>CCTCTCCCACAGGGCAGTAAAGG</b>	Chr17:68628098-68628120	-	0,49	2,4
SEQ ID NO: 27	FUERA3	<u>TTTTCCCCAAAGGGCAGTAATAG</u>	Chr13:10916598-8-109166010	+	0,79	ND
ARN guía dirigido a <i>IL2RG</i>						
SEQ ID NO: 28	EN	<b>TGGTAATGATGGCTAACATGG</b>	ChrX:7111519-7111541	+	0	49
SEQ ID NO: 29	OFF1	<u>TGGGAAGGATGGCTAACACAG</u>	Chr7:151485304-151485326	-	0,4	3,9
SEQ ID NO: 30	FUERA2	<u>TGGTGAGGATGGCTAACACGG</u>	Chr1:167730172-167730194	-	0,42	3,7
SEQ ID NO: 31	FUERA3	<b>TGGTAATGATGACTTCAACATAG</b>	Chr3:72764801-72764823	-	0,8	49,2

La Tabla 3 proporciona una lista de cebadores de oligonucleótidos utilizados para la generación de amplicones dentro y fuera de la diana para cuantificar las frecuencias de ins/del mediante secuenciación profunda. Las secuencias de hibridación específicas de genes de los cebadores de amplicones específicos de genes y los códigos de barras de los cebadores de códigos de barras de Illumina se indican con texto subrayado y en negrita, respectivamente.

5

Tabla 3. Lista de cebadores de oligonucleótidos utilizados para la generación de amplicones dentro y fuera de la diana para cuantificar las frecuencias de ins/del mediante secuenciación profunda.

SEQ ID NO:	Nombre del cebador	Secuencia del cebador (5' → 3')
<b>Cebadores de amplicones específicos de genes</b>		
SEQ ID NO: 32	CCR5_ON-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CCAAACA</u> <u>CAGCATGGACGACA</u>
SEQ ID NO: 33	CCR5_FUERA-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTGGCTG</u> <u>AAGAGCATGACTGACA</u>
SEQ ID NO: 34	CCR5_OFF1-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CGGGAA</u> <u>GCAGTCTGGACTTAGA</u>
SEQ ID NO: 35	CCR5_FUERA1-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTCCTGC</u> <u>CATTAAATCCACCAAA</u>
SEQ ID NO: 36	CCR5_OFF2-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CAGCGAG</u> <u>TCGAGTTAGGTG</u>
SEQ ID NO: 37	CCR5_FUERA2-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTGGCTA</u> <u>CCTACCCCAGGTTCT</u>
SEQ ID NO: 38	CCR5_OFF3-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CTCTCAC</u> <u>CAACACTGCCGAAT</u>
SEQ ID NO: 39	CCR5_FUERA3-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTTGCA</u> <u>TATAGTGCTCCCCACT</u>
SEQ ID NO: 40	HBB_EN-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CTCTGTCT</u> <u>CCACATGCCAGT</u>
SEQ ID NO: 41	HBB_EN-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTCAGG</u> <u>GCAGAGCCATCTATTG</u>
SEQ ID NO: 42	HBB_OFF1-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CTCCGT</u> <u>TCTCCACCCAATAG</u>
SEQ ID NO: 43	HBB_FUERA1-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTGATT</u> <u>CCAGGCTATGCTTCCA</u>
SEQ ID NO: 44	HBB_FUERA2-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CGTTGGC</u> <u>AGGGAGACTTACCA</u>
SEQ ID NO: 45	HBB_FUERA2-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGAT <u>CTCCAT</u> <u>GGTACGACTGTTCTCA</u>
SEQ ID NO: 46	HBB_FUERA3-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CTGGGGC</u> <u>CTTCAAGTGTCTT</u>

(continuación)

SEQ ID NO:	Nombre del cebador	Secuencia del cebador (5'→ 3')
SEQ ID NO: 47	HBB_FUERA3-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT <u>GTGTG</u> <u>CTCCTATGCCTGGTT</u>
SEQ ID NO: 48	IL2RG_EN-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CCATTGG</u> <u>GCGTCAGAATTGTC</u>
SEQ ID NO: 49	IL2RG_EN-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT <u>AGGTT</u> <u>CTTTCCACCGGAAGC</u>
SEQ ID NO: 50	IL2RG_FUERA1-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CTCCGGA</u> <u>AGTTATTCAAGTCTGA</u>
SEQ ID NO: 51	IL2RG_FUERA1-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT <u>GTCTG</u> <u>GCATCAGAGCACAAA</u>
SEQ ID NO: 52	IL2RG_FUERA2-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CCCTGGG</u> <u>CCATATCAAGAGAC</u>
SEQ ID NO: 53	IL2RG_FUERA2-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT <u>TTTT</u> <u>GGGGTGATTTGTG</u>
SEQ ID NO: 54	IL2RG_FUERA3-dir	CGACAGGTTCAGAGTTCTACAGTCCGACGAT <u>CCACAAC</u> <u>AGTTGACCCAGGAA</u>
SEQ ID NO: 55	IL2RG_FUERA3-inv	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT <u>CCCAA</u> <u>CCCAGGTCTCTGAAC</u>
<b>Cebadores de códigos de barras Illumina</b>		
SEQ ID NO: 56	P5-BC_A-dir	AATGATA <u>CGCGACCACCGAGATCTACACTAGAGCTC</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 57	P5-BC_B-dir	AATGATA <u>CGCGACCACCGAGATCTACACTCATAGCG</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 58	P5-BC_C-dir	AATGATA <u>CGCGACCACCGAGATCTACACGTAGCACT</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 59	P5-BC_D-dir	AATGATA <u>CGCGACCACCGAGATCTACACGCTCATAG</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 60	P5-BC_E-dir	AATGATA <u>CGCGACCACCGAGATCTACACATGCATCG</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
		CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 61	P5-BC_F-dir	AATGATA <u>CGCGACCACCGAGATCTACACAGTCGATC</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 62	P5-BC_G_dir	AATGATA <u>CGCGACCACCGAGATCTACACCACTGTGA</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 63	P5-BC_H_dir	AATGATA <u>CGCGACCACCGAGATCTACACCGAGTATC</u> CGACAGGTT <u>CAGAGTTCTACAGTCCGACGATC</u>
SEQ ID NO: 64	P7-BC-A_inv	CAAGCAGAAGACGGCATACGAGATT <u>AGAGCTCGTGAC</u> TGGAGIT <u>TCAGACGTGTGCTCTCCGATCT</u>

(continuación)

SEQ ID NO:	Nombre del cebador	Secuencia del cebador (5'→ 3')
SEQ ID NO: 65	P7-BC_B inv	CAAGCAGAACGGCATACGAGATT <b>CATAGCGGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 66	P7-BC_C inv	CAAGCAGAACGGCATACGAGAT <b>GTCAGACTGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 67	P7-BC_D inv	CAAGCAGAACGGCATACGAGAT <b>GTCATAGGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 68	P7-BC_E_inv	CAAGCAGAACGGCATACGAGAT <b>ATGCACTCGGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 69	P7-BC_F_inv	CAAGCAGAACGGCATACGAGAT <b>AGTCGATCGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 70	P7-BC_G_inv	CAAGCAGAACGGCATACGAGAT <b>CACTGTGAGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT
SEQ ID NO: 71	P7-BC_H_inv	CAAGCAGAACGGCATACGAGAT <b>CGAGTATCGTGAC</b> TGGAGTTAGACGTGTGCTCTCCGATCT

Las secuencias de hibridación específicas de genes de los cebadores de amplicones específicos de genes y los códigos de barras de los cebadores de códigos de barras de Illumina se indican con texto subrayado y en negrita, respectivamente.

5 Para el análisis de los datos de secuenciación, las lecturas de diferentes tratamientos se agruparon con sus correspondientes códigos de barras de tratamiento y se mapearon en el genoma utilizando BWA-MEM (Li *et al.*, Bioinformatics 26, 589-595 (2010)) (bwa-0.7.10) con parámetros predeterminados. Las asignaciones de extremos emparejados inconsistentes (> 1 Kbp de diferencia) se filtraron del análisis junto con las lecturas alineadas secundarias y asignadas de baja calidad según lo definido por los parámetros predeterminados en BWA-MEM. Para cada una de las regiones dentro y fuera de la diana, se calculó el % de lecturas con ins/del alrededor del sitio de corte previsto, donde se estima que el sitio de corte previsto se encuentra entre la 3<sup>a</sup> y la 4<sup>a</sup> posición aguas arriba del sitio PAM. El % de ins/del para cada posición en el genoma se calculó usando  $ins/del[i] = (I[i] + D[i]) / C[i]$ , donde  $D[i]$  e  $I[i]$  indican el número de lecturas con una delección o una inserción de cualquier tamaño en la posición *i*, respectivamente y  $C[i]$  indica el número de lecturas asignadas a cualquier intervalo genómico que contenga la posición *i*. El % de ins/del para cada diana se calculó mediante  $ins/del[c]$ , donde *c* es el sitio de corte esperado que es la cuarta posición aguas arriba del sitio PAM. En casos con secuencia de homonucleótidos en la posición 4, "AA" en IL2Rg y "CCC" en CCR5 y un evento de inserción, el alineador BWA\_MEM no puede resolver la posición del nucleótido insertado. Esto se corrigió tomando la posición relevante en el informe de alineación en lugar de la posición 4, en concreto, tomar las posiciones 3 y 6 para IL2Rg y CCR5 fuera de la diana, respectivamente.

#### 9. Generación de amplicones CRISPR dentro y fuera de la diana previstos para secuenciación profunda

25 Para cada locus de gen diana, se transfecaron  $10^6$  células K562 con 1  $\mu$ g o 20  $\mu$ g de los ARN guía sintéticos (sin modificar, M, MS o MSP) y 2  $\mu$ g del plásmido de expresión Cas9 (PX330), 2  $\mu$ g de plásmido de ARNsg que codifica tanto ARNsg como la proteína Cas9 (control positivo) o 2  $\mu$ g de PX330 solamente (solo Cas9, control negativo). Adicionalmente, para experimentos dirigidos al locus del gen IL2RG, se transfecaron  $10^6$  células K562 con 15  $\mu$ g de ARNm de Cas9 y 10  $\mu$ g de ARN guía sintético (sin modificar o MS), 15  $\mu$ g de proteína Cas9 se formaron en complejo con 7,6  $\mu$ g de ARN guía sintético (sin modificar o MS). El ADN genómico de estas muestras y las muestras de transfección simulada (segundo control negativo) se extrajo 72 horas después de la transfección utilizando la solución de extracción de ADN QuickExtract™ (Epicentre, Madison, WI) de acuerdo con las especificaciones del fabricante. Se utilizaron 40 ng de ADN genómico como molde para la amplificación por PCR de los loci en la diana y tres loci fuera de la diana predichos computacionalmente (Tabla 2), utilizando PfuUltra II HS 2x Master Mix (Agilent Technologies, Santa Clara, CA) y cebadores específicos de genes que etiquetan los extremos del amplicón con cebadores de secuenciación utilizados en la secuenciación profunda mediante la plataforma MiSeq (Illumina, San Diego, CA) (Tabla 3). Se llevó a cabo una segunda reacción de PCR en los amplicones dentro y fuera de la diana (Tabla 4) para agregar adaptadores de secuenciación de Illumina adicionales (es decir, P5, P7) y códigos de barras duales de 8 pb personalizados, identificando de forma única los correspondientes tratamientos de transfección. Despues de la segunda PCR, los amplicones con códigos de barras se cuantificaron mediante Agilent D1000 TapeStation, se agruparon en concentraciones equimolares antes de la purificación con un kit de purificación

QIAquick PCR (Qiagen, Valencia, CA) de acuerdo con las instrucciones del fabricante. La biblioteca purificada se secuenció en un secuenciador de ADN Illumina MiSeq, 2 x 20<sup>1</sup> ciclos con indexación dual por un servicio de secuenciación de ADN NGS (Seqmatic, Fremont, CA).

- 5 La Tabla 4 proporciona una lista de amplicones dentro y fuera de la diana *CCR5*, *HBB* e *IL2RG* generados para el análisis de secuenciación profunda de frecuencias ins/del. Los tamaños de los amplicones (de ADN genómico sin editar) varían de 183-220 pb, con un mínimo de 50 pb desde el sitio diana hasta la secuencia de hibridación del cebador específico del gen. Las secuencias de hibridación utilizadas para la generación de amplicones a partir de ADN genómico y sitios diana CRISPR putativos se indican en texto subrayados y en negrita, respectivamente.

10

Tabla 4. Lista de amplicones dentro y fuera de la diana *CCR5*, *HBB* e *IL2RG* generados para el análisis de secuenciación profunda de frecuencias ins/del.

SEQ ID NO:	ID de la diana	Secuencia de amplicones diana (5' → 3')
ARN guía dirigido a <i>CCR5</i>		
SEQ ID NO: 72	EN	<u>CAAACACAGCATGGACGACAGCCAGGTACCTATCGATTGTCAGGAGGATGATGAAG</u> AAGATTCCAGAGAAGAAGCCTATAAAATAGAGCCCTGTCAAGAGTTGACACATTGTA TTTCAAAGTCCCCTGGGC <b>GGCAGCATACTGAGCCCAGAAGGGGACAGTAAGAAG</b> GAAAAACAGGTCAGAGATGCCAGGTTGAGCAGGTAGAT <b>GTCAGTCATGCTCT</b> <b>TCAGCC</b>
SEQ ID NO: 73	OFF1	<u>GGGGAAAGCAGTCTGGACTTAGAAAGGAAATAGGTGGTCTGTCATAGGGCTTCAT</u> TAGAGTAAACTCATAGAGTCACACTGTTCATCATCATAGT <b>GAGCCCAGAGAGCCA</b> CTGCCAGCAGCATGTCACACCACCTACCCTAGTGTAGGTAATAGGTCTACGCTAG GACCCCGTCTGGCTCTAGGCCATCATGAGAT <b>TTGGTGGATTAAATGGCAGG</b>
SEQ ID NO: 74	FUERA2	<u>AGCGAGTCGAGTTAGGTGGAGCAGAGGGGCCACCAGCAGAGCGAGTCGAGT</u> CCAGGGGGAGCAGAGGGCCAC <b>ACCAGCAGAGTGAGCCCAGAGGGTTAAAGAA</b> GGGGCGGTCTCTACGGTATGGTAGAGTCAGGGAACTAGGAAAGGACAGAGCAG <b>AACCTGGGTAGGTAGCC</b>
SEQ ID NO: 75	FUERA3	<u>TCTCACCAACACTGCCGAATGTCATCTCTCATCTTATCTCTATT</u> CTTGCTTCTGTCTTCAGGGCTCTCCCTGGCATTCA <b>CCAGGAGCA</b> <b>GAGTGAGCCCAGAGAGCTGAGTGGTATCCCTCTTGGTCCCT</b> GAGCCCTGACCTGGAGCAATGCTGTGAGACAGCAGGAAAGGAGGGG <b>AGTGTGGAGTGGGAGCACTATATGCCA</b>
ARN guía dirigido a <i>HBB</i>		
SEQ ID NO: 76	EN	<u>TCTGTCTCCACATGCCAG</u> TTCTATTGGTCTCCTTAAACCTGTCTTG TAACCTTGATACCAACCTGCCAGGGCCTCACCAACTTCATCCA CGTTCAC <b>CTTGCCCCACAGGCAGTAACGGCAGACTTCTCCTCAG</b> GAGTCAGATGCACCATGGTCTGTGTTGAGGTTGCTAGTGAACACAG TTGTGTCAGAACATGTAAGCAATAGATGGCTCTGCCCTG
SEQ ID NO: 77	OFF1	<u>TCCCCTCTCCACCAATAGCTATGGAAAGGGGAAGATCCCAGAGA</u> ACTTGGATAGGAAAGGTGAAGTCAGAGCAGTGCT <b>CAGCCCCACA</b> <b>GGGCAGTAAGGGCAGCCTCCTCAAATACCAAGATCCCAAATCTG</b> GCTGTGCTTCAATTGGAGTTGGACATACTGCTAAACTATAATTTC TTAGGCCGTACCTAAATATTATGGAAGCATAGCCTGGAAATC
SEQ ID NO: 78	FUERA2	<u>GTTGGCAGGGAGACTTACCA</u> GCTTCCGTATCTCCCTCCACATGGAG GCAGGACACGCTCTGGCCTGCCACCCCTCCACTAGC <b>CCTCTCCCAC</b> <b>AGGGCAGTAAGGTGAGTCTGGAGAAAGAACCGGTCAGACTTAG</b> TTCAGCTCCACCCCTTCCCTGGAGTGAGTCTTCCAAGACAGAG CATGTTTTCTACCCCTCAGT <b>GAGAACAGTCGTACCATGGG</b>

(continuación)

SEQ ID NO:	ID de la diana	Secuencia de amplicones diana (5' → 3')
SEQ ID NO: 79	FUERA3	<u>TGGGGCCTCAAGTGTCTTCCCAAGAGTCAGAGTGAACCAGAACCA</u> AGAACCATGTTGAGTTGCCAGATGTAACCAGGCCTACAGGTACCTG GGAGAAACACGTGTACATTTC <u>CCCAAAGGGCA</u> GTAA <u>TAGCATCC</u> TAGGCTCAAAACATTCATAGAAACCATTTC <u>CAAATGCAAAAGTCCA</u> ACACAGTTAGAA <u>ATAACCAGGC</u> ATAGGAGCACAC
		ARN guía dirigido a <i>IL2RG</i>
SEQ ID NO: 80	EN	<u>CATTGGCGTCAGAATTGTCGTGTCAGCCCCACTCCCAGCAGGGGC</u> AGCTGCAGGAATAAGAGGGATGTGAAT <u><b>TGGTAATGATGGCTTCAAC</b></u> <u><b>ATGGCGCTGCTCTCATTCCCTGGGTGATGCTGTCTGTAGGA</b></u> ACCTGGGTCCCTCACCCACTACCCCTCCCCACCCACAGTTCCCTTG TCATA <u>GCTTCCGGTGGAAAGAACCT</u>
SEQ ID NO: 81	OFF1	<u>TCCGGAAGTTATTCAAGTCTGATTTCCTTCCTCCCTGTCAGGGAAA</u> AGAAGTTGTGACAATTGCTTGGATCCTTAAGCTTAAGT <u><b>GGGAAGG</b></u> <u><b>ATGGCTTCAACACAGAACATCTGTTCATGCTGTTTATCCGTCAG</b></u> TAAA <u>ACTGTTACTTCTTTATGTACTAAAGTTCTAGCACTTA</u> ACTA ATATTAGCTTTGTGCTGTGATGCCAGAC
SEQ ID NO: 82	FUERA2	<u>CCTGGGCCATATCAAGAGACTCTGCCTCAAAAAAGAAAAGAAAGAA</u> AGAAAAAAGAAAAAAAAGAACATCATTAAAATCCCTGAGGAGC ATTTAGAGTATTGGTGGCACAAACAGATTCTGCATGATT <u><b>GGTGAG</b></u> <u><b>GATGGCTTCAACACAGGCAGCTTATTCCCTTTAACAGAGTCAGCA</b></u> GCATCAAGGCATGAGGATCTTGG <u>CACAAACATCACCCAAAGA</u>
SEQ ID NO: 83	FUERA3	<u>CACAAACAGTTGACCCAGGAACAGGGGGACCTCCCACCATTCCCAT</u> CCC <u>ACTGTTGATCAGATCCAAGAATCCACAATATTGAGAGTGAATG</u> AAA <u>AGTGTCAAGCTGGTAATGATGACTCAACATAGTCAGAACTCTT</u> TGGGGTGTCCAAACATCATGGTGCATATGTATTAC <u>CTGGGAGTCTT</u> GTTAAAAA <u>AGACTCCTGTTCAAGAGACCTGGGTTGGG</u>

Los tamaños de los amplicones (de ADN genómico sin editar) varían de 183-220 pb, con un mínimo de 50 pb desde el sitio diana hasta la secuencia de hibridación del cebador específico del gen. Las secuencias de hibridación utilizadas para la generación de amplicones a partir de ADN genómico y sitios diana CRISPR putativos se indican en texto subrayados y en negrita, respectivamente.

5 5 Los tamaños de los amplicones (de ADN genómico sin editar) varían de 183-220 pb, con un mínimo de 50 pb desde el sitio diana hasta la secuencia de hibridación del cebador específico del gen. Las secuencias de hibridación utilizadas para la generación de amplicones a partir de ADN genómico y sitios diana CRISPR putativos se indican en texto subrayados y en negrita, respectivamente.

#### 10. Actividad dentro y fuera de la diana de ARNsg modificados químicamente

10 La Tabla 5 muestra los números en los que se basan las figuras 1C, 1E y 5. Especificidad de la escisión dirigida mediada por ARNsg sintéticos como se realiza en la Figura 1C con 2 µg de plásmido de Cas9 y 1 µg de ARNsg (panel superior) o 20 µg de ARNsg (panel inferior). Las frecuencias de ins/del se midieron mediante la secuenciación profunda de amplicones de PCR de los loci genómicos diana y tres loci fuera de la diana predichos bioinformáticamente para cada gen. Los valores promedio se muestran +/-SEM, n=3.

15

Tabla 5

2 µg de plásmido de Cas9 y 1 µg de ARNsg						
		Simulado	Sin modificar	M	EM	MSP
<b>IL2RG</b>	<b>En la diana</b>	0,16 ±0,03	2,43 ±0,23	13,47 ±0,54	68 ±1,06	75,73 ±1,30
	<b>Fuera de la diana 1</b>	0,00	0,00	0,00	0,00	0,00
	<b>Fuera de la diana 2</b>	0,00	0,00	0,10 ±0,00	1,00 ±0,10	0,37 ±0,03
	<b>Fuera de la diana 3</b>	0,00	0,00	0,00	0,00	0,00
<b>HBB</b>	<b>En la diana</b>	0,15 ±0,05	1,3 ±0,08	12,4 ±0,28	48,17 ±0,39	38,68 ±0,85
	<b>Fuera de la diana 1</b>	0,01 ±0,01	0,84 ±0,07	7,01 ±0,36	27,31 ±0,60	9,51 ±0,25

(continuación)

2 µg de plásmido de Cas9 y 1 µg de ARNsg						
		Simulado	Sin modificar	M	EM	MSP
	<b>Fuera de la diana 2</b>	0,00	0,00	0,01 ±0,00	0,00	0,00
	<b>Fuera de la diana 3</b>	0,00	0,00	0,00	0,00	0,00
<b>CCR5</b>	<b>En la diana</b>	0,02 ±0,01	4,26 ±0,18	3,41 ±0,23	24,6 ±1,25	22,78 ±1,76
	<b>Fuera de la diana 1</b>	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00	0,02 ±0,00
	<b>Fuera de la diana 2</b>	0,01 ±0,01	0,01 ±0,00	0,01 ±0,00	0,08 ±0,03	0,54 ±0,10
	<b>Fuera de la diana 3</b>	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00

2 µg de plásmido de Cas9 y 20 µg de ARNsg

		Sin modificar	M	EM	MSP	Plásmido de ARNsg
<b>IL2RG</b>	<b>En la diana</b>	23,4 ±	48,13 ± 0,4	75,3 ±5,1	83,27 ±0,7	70,53 ±0,01
	<b>Fuera de la diana 1</b>	0,00	0,03 ±0,03	0,13 ±0,03	0,00	0,00
	<b>Fuera de la diana 2</b>	0,13 ±0,03	±0,06	7,83 ±0,58	2,77 ±0,24	0,07 ±0,00
	<b>Fuera de la diana 3</b>	0,00	0,10 ±0,00	0,30 ±0,00	0,10 ±0,00	0,17 ±0,00
<b>HBB</b>	<b>En la diana</b>	19,42 ±0,27	40,99 ±1,4	65,91 ±0,62	60,71 ±0,25	31,11 ±0,04
	<b>Fuera de la diana 1</b>	9,2 ±0,38	33,56 ±2,3	55,1 ±0,97	19,08 ±0,49	25,44 ±0,04
	<b>Fuera de la diana 2</b>	0,00	0,00	0,00	0,00	Aceite
	<b>Fuera de la diana 3</b>	0,00	0,00	0,00	0,00	0,00
<b>CCR5</b>	<b>En la diana</b>	11,04 ±2,5	14,87 ±2,1	56,3 ±3,3	52,19 ±4,5	45,86 ±0,03
	<b>Fuera de la diana 1</b>	0,02 ±0,01	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00	0,01 ±0,00
	<b>Fuera de la diana 2</b>	0,01 ±0,00	0,02 ±0,00	1,75 ±0,18	5,32 ±0,68	0,02 ±0,00
	<b>Fuera de la diana 3</b>	0,01 ±0,00	0,01 ±0,00	0,02 ±0,01	0,02 ±0,00	0,01 ±0,00

5 **Ejemplo 2. Recombinación homóloga basada en CRISPR/Cas9 utilizando ARN guía modificados químicamente y un molde de oligodesoxinucleótido monocatenario sintético (ssODN).**

Se nucleofectaron linfocitos T primarios humanos estimulados de tres donantes diferentes con 10 µg de ARNsg de CCR5 (sin modificar o MS), 15 µg de ARNm de Cas9 y 2,81 µg de un ssODN de CCR5 de 183 nt (con o sin enlaces de fosforotioato ("PS") entre los tres nucleótidos terminales en ambos extremos). El ssODN contenía un sitio de restricción HindIII central que no estaba presente en la secuencia SV de CCR5. El ADN genómico (ADNg) se extrajo tres días después de que se generaran nucleofecciones y productos de PCR que abarcaban el sitio diana (fuera de la secuencia homóloga a la ssODN) y se digirieron con HindIII. Los fragmentos de restricción se analizaron en un gel de agarosa TBE al 2 % y se calcularon las frecuencias de HDR.

- 10 15 La Figura 21 representa los geles de agarosa del experimento de HDR. La frecuencia de HDR se aumentó cuando se usaron ARNsg modificados. Además, las frecuencias de HDR aumentaron aún más cuando el ssODN contenía enlaces de fosforotioato entre los tres nucleótidos terminales en ambos extremos.
- 20 Aunque la invención anterior se ha descrito con algo de detalle a modo de ilustración y ejemplos a efectos de claridad de comprensión, un experto en la técnica apreciará que se pueden poner en práctica ciertos cambios y modificaciones dentro del alcance de las reivindicaciones adjuntas.

**Listado informal de secuencias**

SEQ ID NO: 1

Secuencia del vector de direccionamiento de IL2RG

ctaaattgttaagcgtaatattttgttaaaattcgcgtaaattttgttaatcagctcatttttaaccaata  
 ggcgaaatcgcaaaatccctataaatcaaaaagaaatagaccgagatagggttagtgttccagttggaa  
 caagagtccactattaaagaacgtggactccaacgtcaaagggcggaaaaccgtctatcagggcgatggccact  
 acgtgaaccatcaccctaatacaagttttgggtcgaggtgccgtaaagcactaaatcggaaaccctaaagggag  
 cccccgatttagagcttgacggggaaagccggcaacgtggcgagaaagaaggaaaggaaagcggaaaggagcggg  
 cgcttagggcgtggcaagtgtacgcgtcacgcgtcgacaccacaccacccggcgcttaatgcgcgcgtaca  
 gggcgctccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgccggcctttcgctattacg  
 ccagctggcgaaaggggatgtgctgcaggcgatataagggttaacgcggatgggtttccagtcacgcacgttg  
 taaaacgacggccagtgcgcgtataacgactcactataggcgaaattgggtacctaactataacggcccta  
 aggttagcgatataatgggagaaacaccacagaacagcagatgggttatattcttgcgttagagaggggaga  
 aattgaagctgattctgagggttcaagtcgtgggtgacttagaggggtgacgataccattgactgaggtggggagg  
 caggaagagaagcagatggggaaagatgggaaagcttgcgttagtattttttccatcttgcgttagatatt  
 tttgtattataatggcgttgcacacttcgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 tgggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 aaaaacagtagagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ggggtgggggttcagaacttcattatagaaggtaatgatgggtggggatgggtggggatgggtggggatgggt  
 tagtgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 tagtttcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ccacaccctctgtaaagccctgggttataagggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ggggtgggggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 gaggggttagtgggtggggatggggaccgcggatgggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 agaattaccctgttatccctaaggtaatgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 tggtttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ccacagggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 GCACCCCTGAACGGCGTGGAGTCAGCTGGTGGCGGGAGAGGGCACCCCGAGCAGGGCGCATGACCAACA  
 AGATGAAGAGCACCAAAAGGCCTGACCTCAGCCCCTACCTGCTGAGCCACGTGATGGCTACGGCTTCTACC  
 ACTTCGGCACCTACCCAGCGCTACGAGAACCCCTTCTGCACGCCATAACAACGGCGCTACACCAACACCC  
 GCATCGAGAAGTACGAGGACGGCGCGTGCACGTGAGCTCAGCTACCGCTACGAGGGCGCGTGTGATCG  
 GCGACTTCAGGTGGTGGGACCGGGCTTCCCGAGGACAGCGTGATCTTCACCGACAAGATCATCCGCAGCAACG  
 CCACCGTGGAGCACCTGCACCCCATGGCGATAACGTGCTGGGAGCTTCACCGACAAGATCATCCGCAGCAACG  
 ACGGCGGCTACTACAGCTCGTGGTGGACAGCCACATGCACCTCAAGAGCGCCATCCACCCAGCATCCTGCAGA  
 ACGGGGGCCCATGGTGCCTTCCGCCGCGTGGAGGAGCTGCACAGCAACACCGAGCTGGGATCGTGGAGTAC  
 AGCACGCCTCAAGACCCCCATGCCTTCGCCAGATCTCGAGCTCGAtgactcgaggcgccgcgtatcag  
 cctcgactgtgccttctagttqccagccatctgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ccactccactgtccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ggggtqgggtqggqcaqqaqcaqcaaggqqaqgatqggaaqacaataqcaqcaqcatqctqggqatqcggtqggqct  
 ctatggcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 tcacatcccttatttcgtcagctgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 atgaagacaccacagctgggtggaaatctggactggagggggctgggtgagaagggtggctgtgggaaaggggccg  
 tacagagatctgggtgcctgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ggggagggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 gtctcactatgttgcctaggctggctcaaaactcctggcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 caggattacagacatgagccaccgtgtggcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ttttcctcccaatcttagatttcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 agaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 caacctcactctgcattattggtatgagaaggacggaggggggatgagaaggatgggtgggtggatcag  
 ccaagagagaggtagcaagtctccaggatgggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 cattaccttttccgcaccgcatacatgagctccagctttgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 tggcgtatcatggcatagctgtttcctgtgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 gaagcataaagtgttaaaggctgggtgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 cttccactcggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 ttggcgcttccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 actcaaaggcggttaatcggttatccacagaatcagggtataacgcaggaaagaacatgtgagcaaaaaggccag  
 aaaaggccaggaaaccgtaaaaaggcccggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt  
 aaaatcgcgtcaagtcaaggtggcgaaaccgcacaggactataaagataccaggcgttcccttgcgttgcgtt  
 ccctcggtcgcttcctgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt

ES 2 884 838 T3

cgctttctcatagctacgctgttaggtatctcagttcggttaggtcgctccaagctgggctgtgcacg  
aaccggccgttcagccgaccgctgcgccttacggtaactatcgtttgagtcacccgtaagacacgact  
tatcgccactggcagcagccactggtaacaggattagcagagcggatgttaggcggctacagagttctga  
agtggggcttaactacggctacactagaaggacagtatggatctgcgtctgtgaaggcagttacccctcg  
aaaaagagttggtagcttgcataccggctggtagcgggtgggggggggggggggggggggggggggggggg  
agattacgcgcagaaaaaaaggatctcaagaagatccttgcataccgggtctgacgctcgtggaaacg  
aaaactcacgttaagggatttgcatacgatattcaaaaaggatctcacctagatcctttaattaaaaat  
gaagtttaaatcaatctaaagtatataatgagtaactggctgacagttaccaatgcttaatcgtgaggcac  
ctatctcagcgtatgtctatttcgttcatccatagttgcctgactcccgctgttagataactacgatcgg  
agggcttaccatctggccccagtgctgcaatgataccgcgagacccacgcaccggctccagattatcagcaa  
taaaccagccagccggaaaggcggagcgcagaagtggctgcacattatccgcctccatcgcattatcatt  
gttgcgggaaagctagagtaagttagttcgccagttaatagttgcgcacgttgcgcattgtctacaggcatcg  
tgggtcacgctcgtcggtatggcttattcgcgtccgggtcccaacgcataaggcgcattacatgtatccc  
ccatgttgcaaaaaagcggttagctccgcgtccgcattcgcattgtcagaagtaagttggccgcagtttat  
cactcatggttatggcagcactgcataattctctactgtcatgcgcattcgtaagatgcatttgcattgt  
agtactcaaccaagtattctgagaatagtgtatgcggcaccgcagttgcatttgcgcattgtcaataacgggata  
ataccgcgcacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaagga  
tcttaccgcgttgcattgcacccactcgtgcacccactgatcttcagcatcttacttca  
ccagcggttctgggtgagaaaaacagaaggaaaaatgcgcaaaaaaggaaataaggcgacacggaaatgtt  
gaatactcataactcttcattttcaatattatttgaaagcatttcagggatttgcattgtctcatgagcggatcatat  
ttgaatgtatttagaaaaataacaaatagggttcgcgcacattccccgaaaagtgccac

Los brazos de homología de IL2RG están en negrita y en cursiva. El intrón quimérico está subrayado. La secuencia de GFP está en mayúscula y en cursiva. La secuencia de polí(A) de BGH está subrayada doble.

SEQ ID NO: 2  
Secuencia del vector de direccionamiento de HBB

ctaaaattgttaagcgttaatattttgttaaaattcgcggttaaattttgttaaatcagctcatttttaaccata  
ggccgaatcgccaaaatcccttataaatcaaaagaatagaccgagatagggttagtgttccagttggaa  
caagagtccactattaaagaacgtggactccaacgtcaaaggcgaaaaaccgtctatcagggcatggccact  
acgtgaaccatcacctaatacgatttttgggtogaggtgcgtaaagactaaatcggaacctaaggag  
cccccgatttagagcttgacgggaaagccggcaacgtggcgagaaaggaaaggaaagcgaaaggagccgg  
cgctagggcgtggcaagtgtacggctacgctgcgttaaccaccacaccgcgttaatgcgcgttaca  
gggcgttccattcgccattcaggctgcgtcaactgttggaaaggcgatcggtgcggccttgcgttattacg  
ccagctggcgaaaggggatgtgtcaaggcgattaaagtggtaacgcggcagggtttccactacgacgttgc  
taaaacgacggccagtgagcgcgttaatacgactactataggcgatgggttacttaactataacggttca  
aggtagcattaaaggcagaaacagtttagatgtccctagttaccccttatttgacaccactgattaccc  
attgatagtcacacttgggtgttaagtgacttttattttttagttactgttactgcattaagaggctcttagt  
ttttatctcttgcgtttccaaaacctaataagtaactaatgcacagagcacattgattgttattttatctattt  
tagacataattttagcatgcgtgagcaaaattaagaaaaacaacaatgaatgcataatataatgttatatgt  
tgtgttatatacacacatataatataatatttttcttaccagaagggtttatccaaaataaggaga  
agatatgcttagaaccgaggtagagtttcatccattctgtccgttaagtatttgcatattctggagacgcagg  
aagagatccatctacatatccaaagctgaattatgttagacaaaactctccacttttagtgcataacttctt  
atttgcataaagaaaattggaaaacgatcttcaatatgcttaccaagctgtgattccaaatattacgttaat  
acacttgc当地aggaggatgttttagtgcataatttgcgttactgtgtatggccaaagagatatacttagaggga  
gggctgagggttgcaggactccatctgttgcacacttgcgttactgcgttactgcgttactgcgttacttag  
acctcaccctgtggagccacaccctagggttgcgttacttgcgttactgcgttactgcgttactgcgttacttag  
ggcataaaaagtgcaggccatcttgcgttacttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
acagacaccatggtgcacctgactcctgaggagaagtgcgttactgcgttactgcgttactgcgttacttag  
agtggcagagcgcacatgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
gaaggtggcgccccgtaaactggaaagtgtgttgcgttactggctccgcctttcccgagggtggggagaac  
cgtatataagtgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
tgtgtgttcccgccccgttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
agtacgtgattctgtatcccggacttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
cttcgcctcgtgcttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
cctgtctcgctgttgcataactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
aagatagtcttgcataatgcggccaagatctgcacactgttactgcgttactgcgttactgcgttacttag  
cccgtgcgtccagcgcacatgttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
ctcaagctggccggctgttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
cccgtgcgtccagcgttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
cgcggcgttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
catgtgactccacggacttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
tagttggggggagggtttatgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
ggcacttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
tgttcaaaagttttcttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
CCGCCATGAAGATCGAGTCCCCCATCACCGCACCTGAACGGCGTGGAGTCGGCTGGTGGGGGGGGAGAG  
GCACCCCCGAGCAGGGCCGATGACCAACAAGATGAAGAGCACAAAGGCGCCCTGACCTTCAGCCCCCTACCTGC  
TGAGCCACGTGATGGCTACGGCTTCTACCACCTCGGCACCTACCCCGAGGGCTACGAGAACCCCTCCTGCACG  
CCATCAACAACGGCGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGTGCTGCACGTGAGCTCA  
GCTACCGCTACGAGGCGCCCGTGTACGGCACTCAAGGTGGTGGCACCGGCTTCCCCGAGGACAGCGTGA  
TCTTCACCGACAAGATCATCCGCAGCAACGCCACCGTGGAGCACCTGCACCCATGGCGATAACGTGCTGGTGG  
GCAGCTTCGCCGCACCTTCAGCCTGCACGGCGCTACTACAGCTTCGTGGACAGCCACATGCACTTCA  
AGAGCGCCATCCACCCAGCATCCGTCAAGAACGGGGCCCATGTTGCCTTCCGCCGTGGAGGAGCTGCACA  
GCAACACCGAGCTGGCATCGTGGAGTACCGACGCCCTCAAGACCCCCATGCCCTCGCCAGATCTGAGTCT  
agctcgagggcgcggccgtatcgccctgcgttactgcgttactgcgttactgcgttactgcgttacttag  
ccgtgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
gtctqagtaggtgtcattcttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
gcaggcatqctgggtqcggtggqctatggcttgcgttactgcgttactgcgttactgcgttactgcgttacttag  
gaacgtggatgaagtgggtggtagggccctggcagggtggtagtgtatcaagggttacaagacaggtttaaggagacaa  
tagaaactgggcgttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
atttccacccttgcgttactgcgttactgcgttactgcgttactgcgttactgcgttactgcgttacttag  
ctccctgtgttatggcaaccctaaggtgaaggctcatggcaagaaaagtgcgttactgcgttactgcgttacttag  
tggctcacctggacaacactcaaggccaccccttgcgttactgcgttactgcgttactgcgttacttag  
ctgagaacttcagggttagtctatggacgcgttgcgttactgcgttactgcgttactgcgttacttag

Los brazos de homología de HBB están en negrita y en cursiva. El promotor EF1 $\alpha$  está subrayado. La secuencia de GFP está en mayúscula y en cursiva. La secuencia de poli(A) de BGH está subrayada doble.

SEQ ID NO: 3

## Secuencia del vector de direccionamiento de CCR5

ctaaatgttaagcgtaatatttgttaaattcgcttaaattttgttaatcagtcatttttaaccaata  
 ggccgaaatcgccaaaatccottataaatcaaaagaatagaccgagatagggtttagtgttccagttggaa  
 caagagtccactattaaagaacgtggactccaacgtcaaagggcaaaaaaaaaccgtctatcagggcagtgccact  
 acgtgaaccatcaccctaatcaagttttgggtcgaggtgccgtaaagcactaaatcggaaaccctaaaggag  
 cccccgatttagagcttgcacgggaaagccggcgaacgtggcagaaaaggaaaggaaaggcgaaggagcggg  
 cgcttagggcgtggcaagttagcggtcacgtgcccgtaaaccaccacaccacccggcgttaatcgccgtaca  
 gggcgctccattcgccattcaggctgcccactgtggaaaggcgatcggtgcgggccttcgtattacg  
 ccagctggcgaagggggatgtgtcaaggcgatataagggttacgcaggtttccagtcacgacgttg  
 taaaacgacggccagtggcgtaaacgactcactataggcgaatttggtaactaactataacggctcta  
 aggttagcgattaaatagaacagttagggcatccaggatgtgcctcgaggccctctaattatactggcttg  
 tcatagtgcattgttgcattgtggctaaactctagcgtcaataaaaaatgttaagactgagttgcagccggcatgg  
 ggctcatgcctgtatccagcattctaggaggctgaggcaggaggatcggttagccaggagttcgagaccag  
 cctggcaacatagtgtatcttgcattataaaaaataacaaaatttagcttggtgggtggccctgttagtccc  
 cagccacttggaggggtgagggtgagaggattgttagccggatggccaggctgcagtgagccatgatcgt  
 ccactgcactccaggctggcgcacagagttagcggccatgtctcacaacaacaacaacaacaaaaggctgagc  
 tgcaccatgcattgtgaccctaggatcttgcattgtcaagcattcactccatgtgtctatagagcacaaga  
 ttttatttggtagatggtgatcatgaaattccccaaacagagccaaacgcctccatctagtggacaggaaag  
 agcagcaaacccttcccttactacaaaacttcattgttgcattgtggccaaaaagagagatgttcaatgttag  
 atataaaagatcacttttattatgcacagggtggaaagatggattatcaagtgtcaagtccatctatgaca  
 tcaattattatagcgcccttagaattaccctgttatccctaagatctggcattccgtccggcgtcagtgcc  
 cgcacatcgcccacagtcccgagaagtgg  
 ggggtaaactggggaaagtgtatgtactggctccgcattttcccgagggtgggggggggggggggggggg  
 gcaatgtgtccgtgacgttgcacccgggttgcacccgggttgcggccagaacacaggtaatgtccgtgt  
 cggggccctggcctttacgggttatggcccttgcgtgcattgttgcattacttccacccgtgcactgtgatt  
 ctgtatcccggcttgggttggaaagtgggtggagagttcgaggcccttgcgttgcattaggagcccttcgc  
 gcttgagttgaggccctggccttggcgtctggggccggccgtgcgaatctgtggcacttcgcgcctgtc  
 gcttcgataagtcttagccattaaaattttgtatgttgcacccgttgcacgccttttgcattgttgc  
 gtaatgcggccaagatctgcacactgttgcatttgcgtttttggccggccggccggccggccggcc  
 cagccacatgttgcggccggccgtgcggccaccgagaatcgacggggtagtgcattgcattgc  
 cggccctgtctggcctggcctcgccgtatcgccccccctggccggcaaggctggccggccgtcc  
 ccagttgcgtgagccggaaagatggccgttccggccgttgcaggagctcaaaatggaggacgcggcc  
 ggagagccggccgggtgagtcacccacacaaaggcccttccgtctcggccgttgcattgtgactcc  
 acggagttaccggccgtccaggcacccgttgcattgttgcgttgcggccaccgttgcatttgc  
 gaggggtttatgcgtatggagttccacactgtgttgcgttgcggccactgttgcattgt  
 taatttgcatttgcgttgcatttgcgttgcatttgcgttgcatttgcatttgcatttgc  
 ttttttgcatttgcgttgcatttgcgttgcatttgcgttgcatttgcatttgcatttgc  
 GAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAAGTTCAAGCGTG  
 TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAACGGCAAGCTGCC  
 GTGCCCTGGCCCACCCCTCGTGACCACCTCGGCTACGGCCTGATGTGCTTCGCCCGTACCCCCGACCACATGAAG  
 CAGCACGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCCACCATCTTCTTCAAGGACGA  
 CGGC  
 AACTACAAGACCCCGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC  
 TTCAAGGAGGACGGCAACATCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACACGTCTATATCATGGCC  
 GACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCC  
 GACCACTTACCAAGCAGAACACCCCCATGGCGACGGCCCGTGTGCTGCCGACAACCAACTACCTGAGCTACCAG  
 TCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGCCTGCTGGAGTTCTGACCGCCGCCGGGATC  
 ACTCTCGGCATGGACGAGCTGTACAAGttaactcgaggccgcgttgcattcaacccctgttgcatttgc  
 agattgtacttgtatt  
 gctatttgcatttgcgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 tggccctgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 accacccgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 ctttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 ttttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 aatccagccgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 acgagtccgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 tgccaggccatctgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc  
 taataaaatgaggaaattgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc

aqcaagggggqaqqattqqqaqacaataqcaqqcatqctqqqqatqcqqgtqqqctctatggcttctgaggcggaa  
 agaacgtttaaacactagtctatgacatcaattattatacatcgagccctgccaaaaatcaatgtgaagcaaa  
tcgcagccgcctccgcctcgactactggttcatcttgggttgtggcaacatgctggcatcc  
tcatcctgataaaactgcaaaggctgaagagcatgactgacatctacctgctcaacctggccatctctgacctgt  
tttccttcttactgtcccttctggctcaactatgctgcccagtgggactttggaaatacatgtcaac  
tcttgcacaggctctatTTTataggcttctctggaatcttcatcatcctctgacaatcgataggta  
tggctgtcgccatgctgtgtttcttaaaagccaggacggtcaccttgggtggtacaagtgtgatcactt  
gggtggggctgtgtttcgctctccaggaaatcatcttaccagatctcaaaaagaaggcttattacac  
gcagctatcatTTTccatacagtcgtatcaattctggaagaattccagacattaaagatagtcatctgggc  
tggtcctgcccgtgtgtcatggtcatctgctactcggaatcctaaaaactctgcttcgggtcgaaatgaga  
agaagaggcacaggcgtgaggcttcatcttaccatcatgattgtttatTTTcttctggctccataaca  
ttgtccttcctgaacacacTTTccaggaaattcttgcgttgcataattgcagtagctctaaccagggtggacc  
ctatgcaggtgactgcatttattacctttccgcacccgacatagatgagctccagctttgtccctt  
 agtgagggttaattgcgcgttgcgtatcatggtcatagctgtttctgtgtgaaattgttattccgctacaa  
 ttccacacaacatacgagccggaaagcataaagtgtaaagctgggtgcataatgagtgagctaacattaa  
 ttgcgttgcgttactgcccgttccagtcgtggaaacctgtcgccagctgcattatgaatcgccaaacgcg  
 cggggagaggcgggttgcgtattggcgcttccgttccgcgttccgcgttccgcgttccgcgttccgc  
 tgcggcggacttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 acatgtgagcaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgcgtggctttccataggctcc  
 gccccctgcgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 aggcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 ttctcccttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 ccaagctggctgtgcacgaaccccccgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 ccaacccggtaagacacgacttatgcgcactggcagccactgttgcataaggattacagagcgaggtatgt  
 gcggtgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 tgctgaagccagttacccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 gtttttttttgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 ggtctgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 agatccctttaattaaaaatgaagtttaatcaatctaaagtatatactgagtaacttggctgacagttacc  
 aatgcctaattcgtgaggcacccatctcagcgatctgttcatccatagttgcctgactcccgctcg  
 tgttagataactacgatacgggaggcttaccatctggcccccagtgtgtcaatgataccgcgagaccacgc  
 cggctccagattatcagcaataaaaccagccagccggaaaggcccagcgcagaatggcttccgcgttccgc  
 cctccatccagtttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 ttgcattgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 caaggcgagttacatgtatccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 gtaagttggccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 gatgtttctgtgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 gcccggcgtaatacggataataccgcgcacatagcagaactttaaaagtgttccgcgttccgcgttccgc  
 cggggcgaaaactctcaaggatcttaccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 cttagcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgcgttccgc  
 taagggcgacacggaaatgttgaataactcataacttccctttcaatattattgaagcatttcatgggtt  
 gtctcatgagcgatataattgaatgtttagaaaaataaacaataggggttccgcgcacatttccccgaa  
 aagtgcac

Los brazos de homología de HBB están en negrita y en cursiva. El promotor EF1α está subrayado. La secuencia de GFP está en mayúscula y en cursiva. La secuencia de WPRE está subrayada de forma discontinua. La secuencia de polí(A) de BGH está subrayada doble.

SEQ ID NOS: 4-19 Véase, Tabla 1

SEQ ID NOS: 20-31 Véase, Tabla 2

SEQ ID NOS: 32-71 Véase, Tabla 3

SEQ ID NOS: 72-83 Véase, Tabla 4

SEQ ID NO: 84 IL2RG\_dir: 5'-TCACACAGCACATATTGCCACACCCTCTG-3',

SEQ ID NO: 85 IL2RG\_INV: 5'-TGCCCACATGATTGTAATGCCAGTGG-3'

SEQ ID NO: 86 HBB\_dir: 5'-CCAACTCCTAAGCCAGTGCAGAAGAG-3'

SEQ ID NO: 87 HBB\_inv: 5'-AGTCAGTGCCTATCAGAAACCCAAGAG-3'

SEQ ID NO: 88 CCR5\_dir: 5'-GCACAGGGTGGAACAAAGATGG-3'

5 SEQ ID NO: 89 CCR5\_inv: 5'-CACCACCCCAAAGGTGACCGT-3'

SEQ ID NO: 90 5'-GGCTGTTGTCATCTATGACCTTCCC-3'

10 SEQ ID NO: 91 5'-TGTAAACTGAGCTTGCTCGCTCG-3'

SEQ ID NO: 92 5'-GCACAGGGTGGAACAAAGATGG-3'

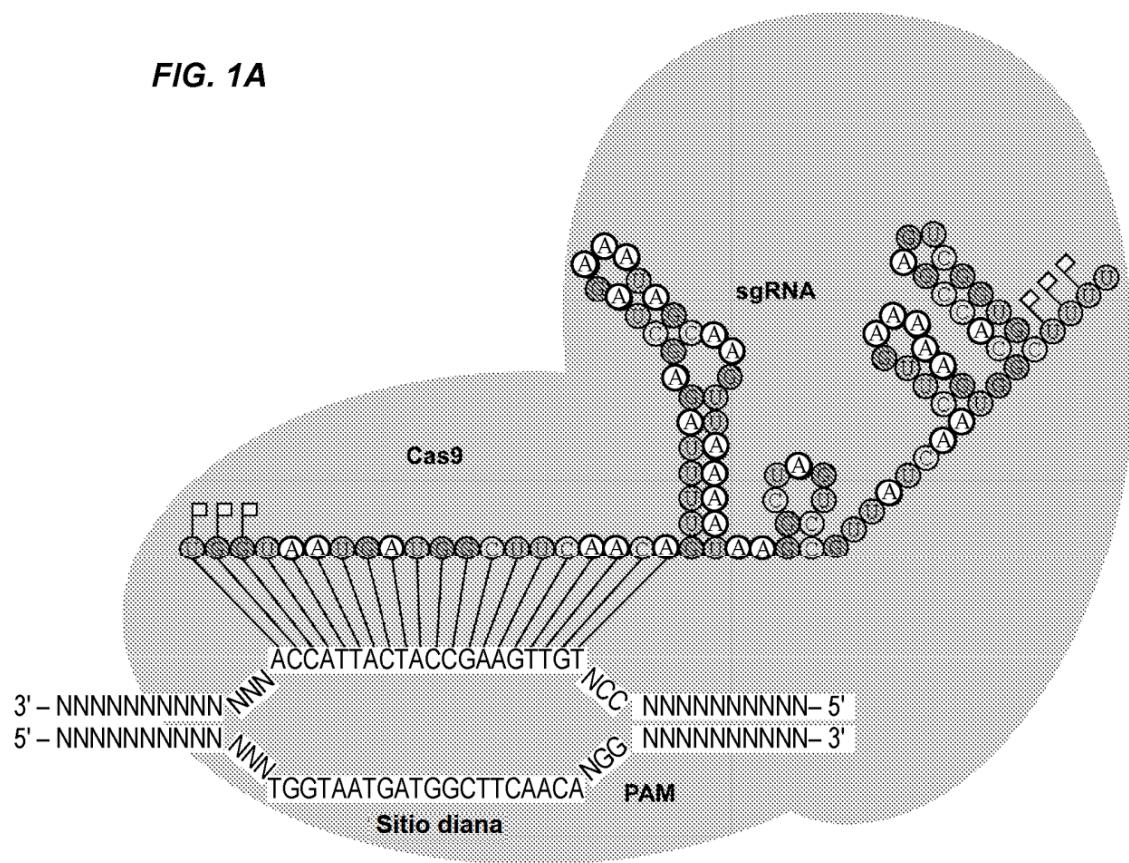
## REIVINDICACIONES

1. Un método *in vitro* para inducir la regulación génica de un ácido nucleico diana en una célula primaria, comprendiendo el método: introducir en la célula primaria:
- 5 (a) un ARN guía único (ARNsg) modificado que comprende una primera secuencia de nucleótidos que es complementaria al ácido nucleico diana y una segunda secuencia de nucleótidos que interactúa con un polipéptido de proteína asociada a CRISPR (Cas), en donde uno o más de los nucleótidos en la primera secuencia de nucleótidos y/o la segunda secuencia de nucleótidos son nucleótidos modificados, en donde los nucleótidos modificados se seleccionan de entre el grupo que consiste en un nucleótido de 2'-O-metil 3'-fosforotioato (MS), un nucleótido de 2'-O-metil 3'-tiofosfatoacetato (MSP) y una combinación de los mismos; y
- 10 (b) un polipéptido Cas, un ARNm que codifica un polipéptido Cas y/o un vector de expresión recombinante que comprende una secuencia de nucleótidos que codifica un polipéptido Cas,
- 15 en donde el ARNsg modificado guía el polipéptido Cas hacia el ácido nucleico diana, y en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana con una actividad mejorada en relación con un ARNsg sin modificar correspondiente y en donde la actividad mejorada comprende un aumento o mejora en la eficiencia y/o la frecuencia de inducción, modulación, regulación o control de la edición del genoma y/o la expresión génica.
- 20 2. El método de la reivindicación 1, en donde la actividad mejorada comprende una mayor estabilidad del ARNsg modificado y/o una mayor especificidad del ARNsg modificado por el ácido nucleico diana.
- 25 3. El método de la reivindicación 1 o 2, en donde el ácido nucleico diana comprende un ADN diana o un ARN diana.
4. El método de la reivindicación 3, en donde la regulación génica comprende la edición del genoma del ADN diana.
- 30 5. El método de la reivindicación 4, en donde la edición del genoma comprende la reparación dirigida por homólogos (HDR) o la unión de extremos no homólogos (NHEJ) del ADN diana.
6. El método de la reivindicación 4 o 5, que comprende además la introducción de un molde de reparación de donante recombinante en la célula primaria.
- 35 7. El método de la reivindicación 6, en donde el molde de reparación de donante recombinante comprende dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de una secuencia de nucleótidos correspondiente al ADN diana para someterse a la edición del genoma o en donde el molde de reparación de donante recombinante comprende un molde de oligodesoxinucleótido monocatenario sintético (ssODN) que comprende una secuencia de nucleótidos que codifica una mutación para corregir un polimorfismo de un solo nucleótido (SNP) y dos secuencias de nucleótidos que comprenden dos porciones homólogas no superpuestas del ADN diana, en donde las secuencias de nucleótidos están ubicadas en los extremos 5' y 3' de la secuencia de nucleótidos que codifica la mutación.
- 40 8. El método de la reivindicación 3, en donde la regulación génica comprende la inhibición o activación de la expresión génica del ADN diana o el ARN diana usando un polipéptido Cas que es deficiente en endonucleasas.
- 45 9. El método de una cualquiera de las reivindicaciones 1 a 8, en donde la célula primaria se aísla de un organismo multicelular antes de introducir el ARNsg modificado y el polipéptido Cas en la célula primaria.
- 50 10. El método de la reivindicación 9, en donde el organismo multicelular es una planta, un protista multicelular, un hongo multicelular o un animal.
11. El método de una cualquiera de las reivindicaciones 1 a 10, en donde la célula primaria se selecciona de entre el grupo que consiste en una célula madre, una célula inmunitaria y una combinación de las mismas.
- 55 12. El método de la reivindicación 11, en donde la célula madre se selecciona del grupo que consiste en una célula madre y progenitora hematopoyética (HSPC), una célula madre mesenquimatosa, una célula madre neuronal, una célula madre de un órgano y una combinación de las mismas, o en donde la célula inmunitaria se selecciona del grupo que consiste en un linfocito T, un linfocito citolítico natural, un monocito, una célula mononuclear de sangre periférica (PBMC), un linfocito de sangre periférica (PBL) y una combinación de los mismos.
- 60 13. El método de una cualquiera de las reivindicaciones 1 a 12, en donde la célula primaria comprende una población de células primarias.
- 65 14. El método de la reivindicación 13, en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana en al menos aproximadamente el 30 % de la población de células primarias, o en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana en al menos aproximadamente el 40 % de la población de

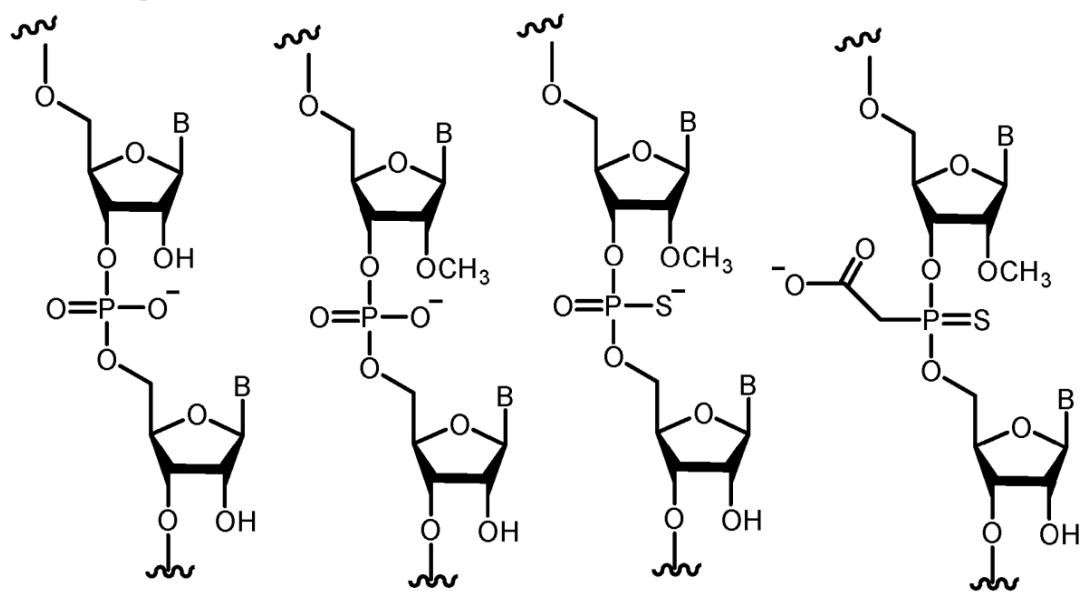
células primarias o en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana en al menos aproximadamente el 50 % de la población de células primarias o en donde el ARNsg modificado induce la regulación génica del ácido nucleico diana en al menos aproximadamente el 60 % de la población de celdas primarias.

- 5 15. El método de una cualquiera de las reivindicaciones 1 a 14, en donde la primera secuencia de nucleótidos tiene aproximadamente 20 nucleótidos de longitud y/o en donde al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de los nucleótidos en la primera secuencia de nucleótidos son nucleótidos modificados y/o en donde los nucleótidos modificados están ubicados en o cerca del extremo 5' de la primera secuencia de nucleótidos y/o en posiciones internas dentro de la primera secuencia de nucleótidos y/o en donde desde aproximadamente el 10 %
- 10 hasta aproximadamente el 30 % de los nucleótidos en la primera secuencia de nucleótidos son nucleótidos modificados y/o en donde la segunda secuencia de nucleótidos tiene una longitud de aproximadamente 80 nucleótidos y/o en donde al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de los nucleótidos en la segunda secuencia de nucleótidos son nucleótidos modificados y/o en donde los nucleótidos modificados están ubicados en o cerca del extremo 3' de la segunda secuencia de nucleótidos y/o en posiciones internas dentro de la
- 15 segunda secuencia de nucleótidos y/o en donde desde aproximadamente el 1 % hasta aproximadamente el 10 % de los nucleótidos en la segunda secuencia de nucleótidos son nucleótidos modificados y/o en donde el ARNsg modificado comprende uno, dos o tres nucleótidos modificados consecutivos o no consecutivos en o cerca del extremo 5' de la primera secuencia de nucleótidos y uno, dos o tres nucleótidos modificados consecutivos o no consecutivos en o cerca del extremo 3' de la segunda secuencia de nucleótidos y/o en donde el ARNsg modificado
- 20 comprende tres nucleótidos modificados consecutivos en el extremo 5' de la primera secuencia de nucleótidos y tres nucleótidos modificados consecutivos en el extremo 3' de la segunda secuencia de nucleótidos y/o en donde el ARNsg modificado se sintetiza químicamente y/o en donde el ARNsg modificado comprende al menos dos ARNsg modificados diferentes, en donde cada ARNsg modificado está dirigido a un ácido nucleico diana diferente y/o en donde el polipéptido Cas es un ARNm que codifica el polipéptido Cas y/o en donde el polipéptido Cas es un polipéptido Cas9, una variante del mismo o un fragmento del mismo y/o en donde la etapa de introducir en la célula
- 25 primaria comprende electroporar la célula primaria.

**FIG. 1A**



**FIG. 1B**



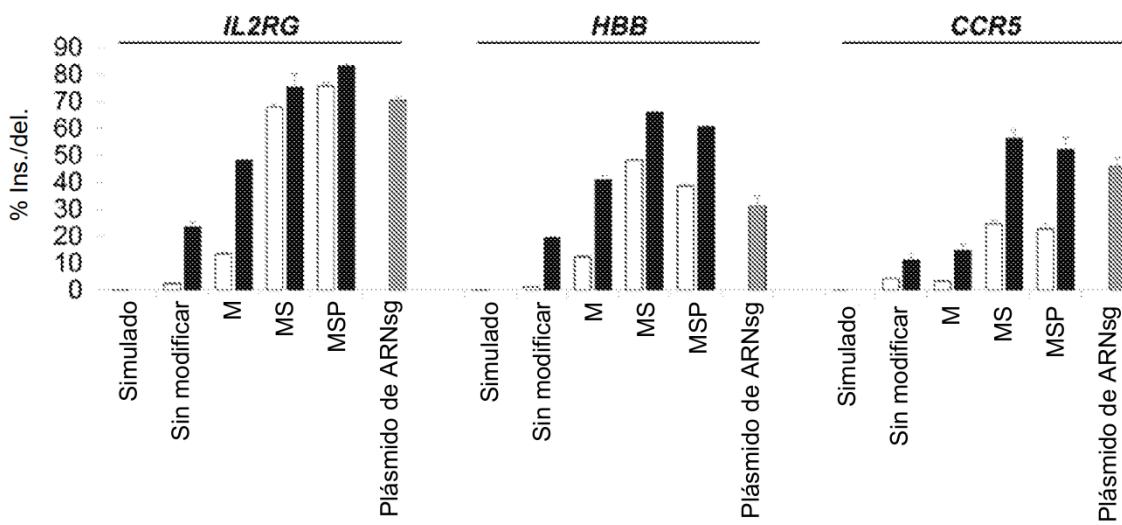
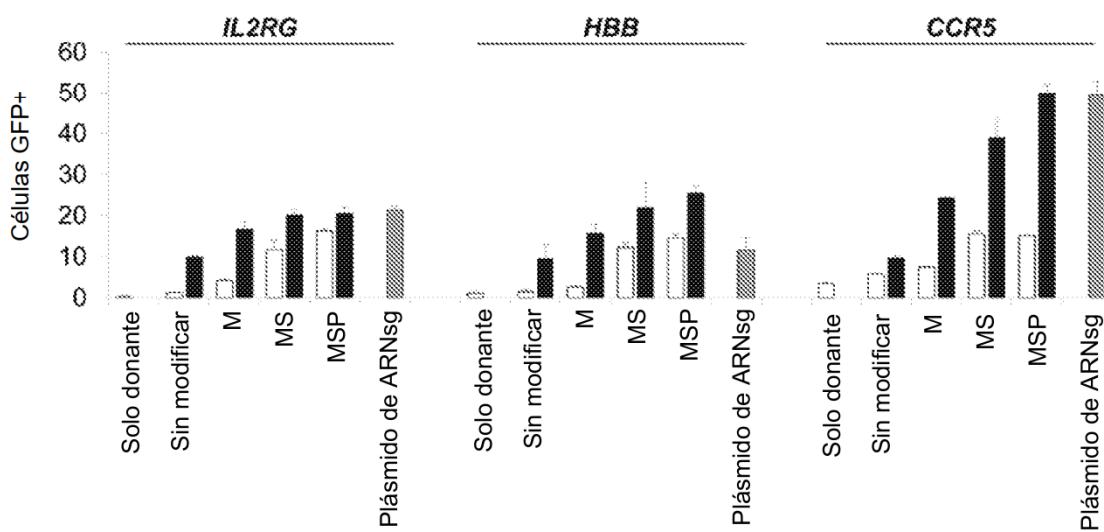
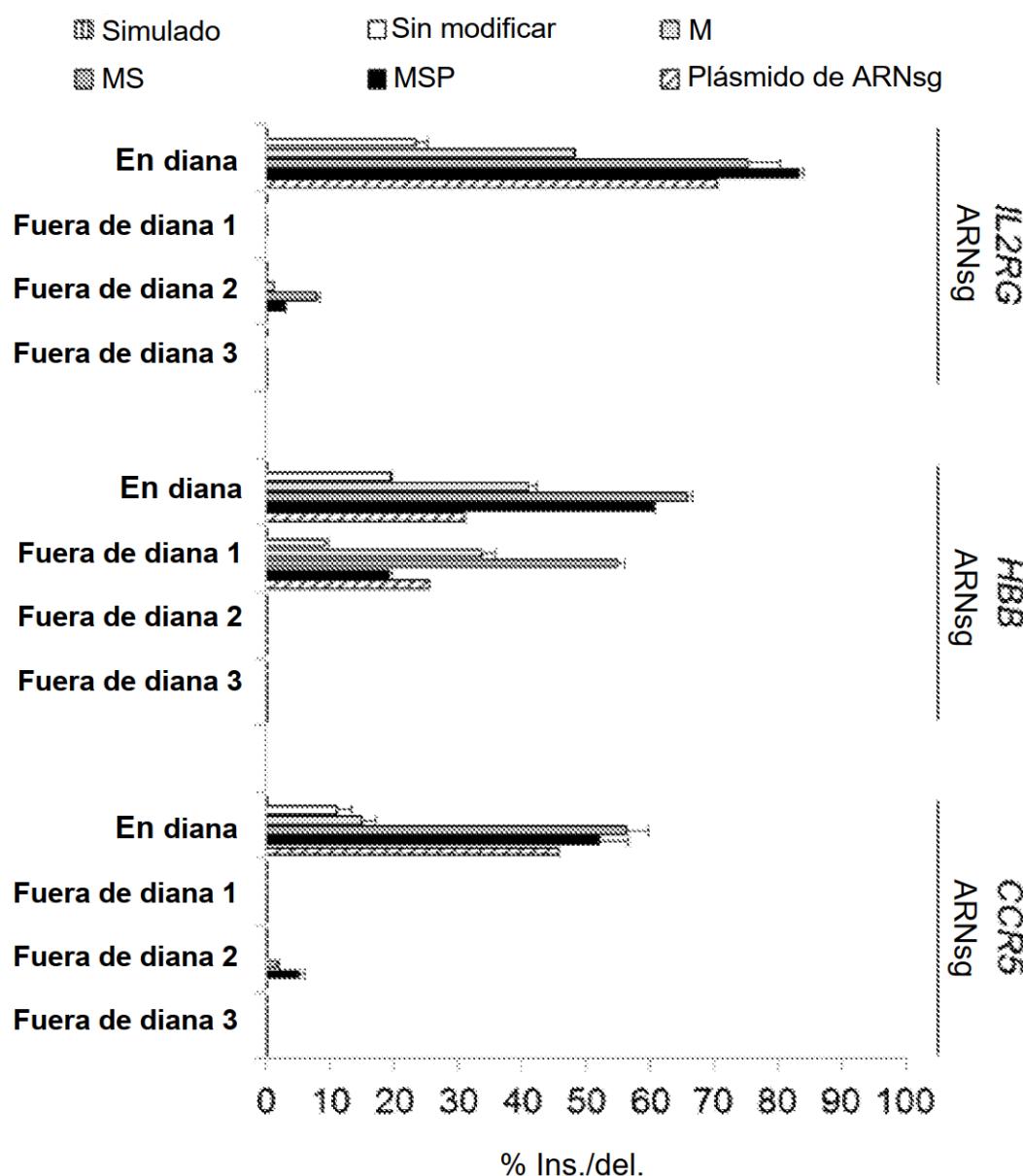
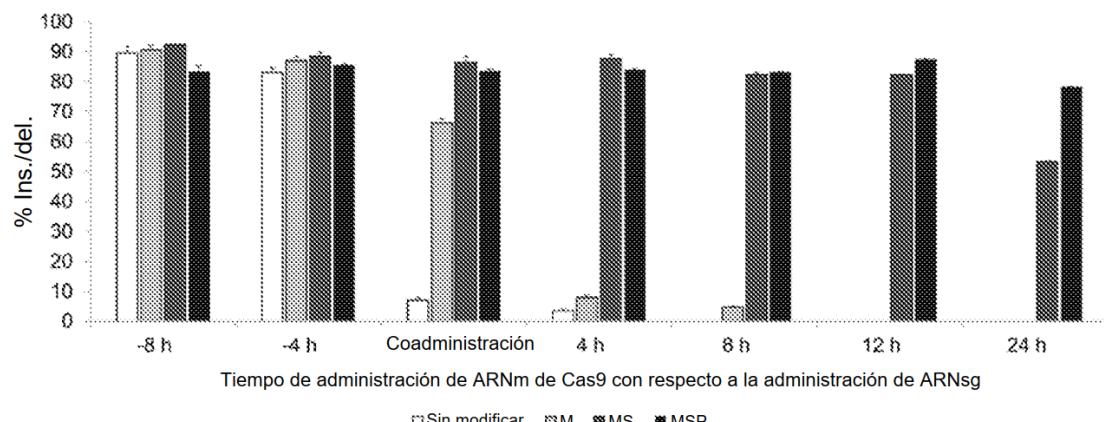
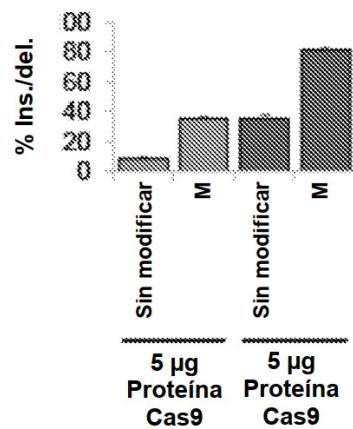
**FIG. 1C****FIG. 1D**

FIG. 1E



**FIG. 1F****FIG. 1G**

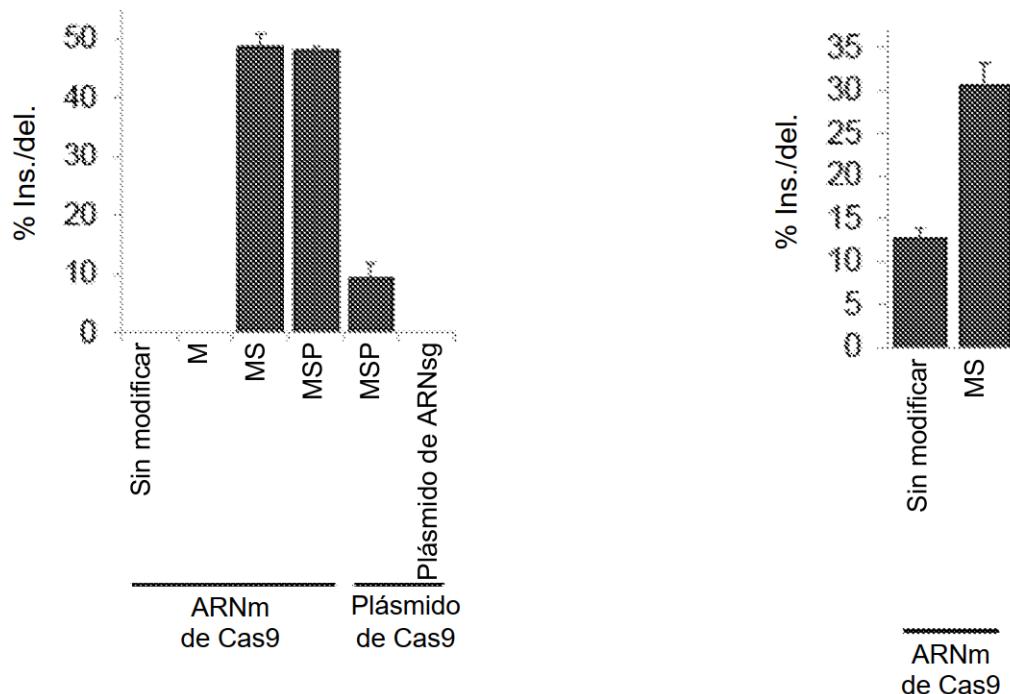
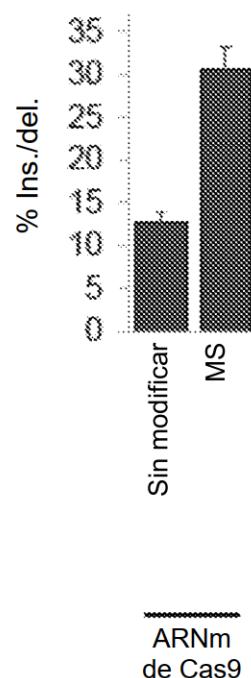
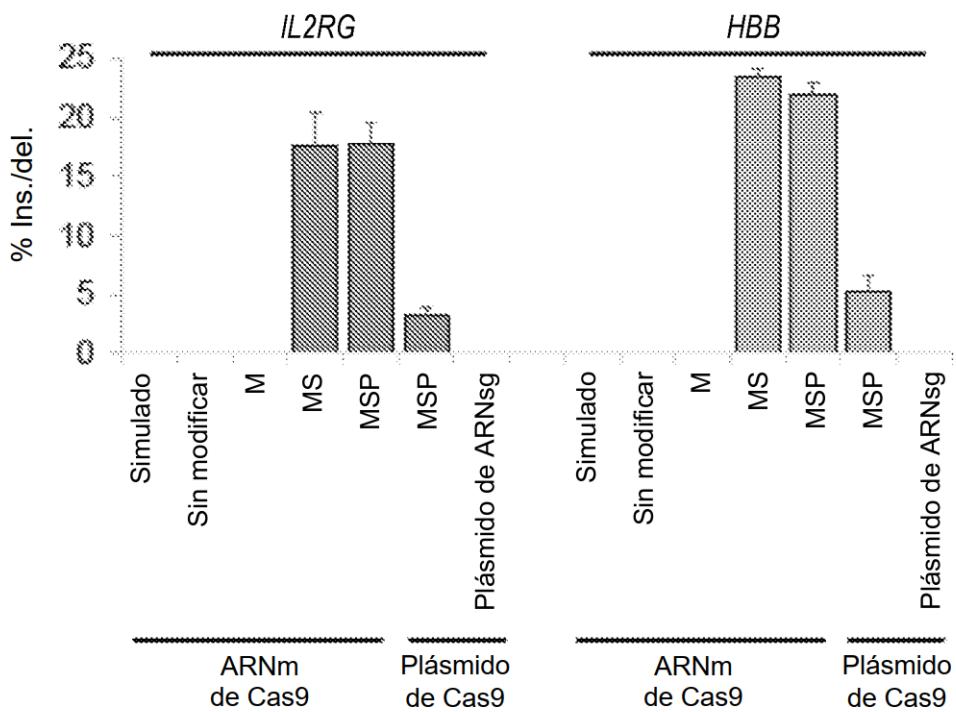
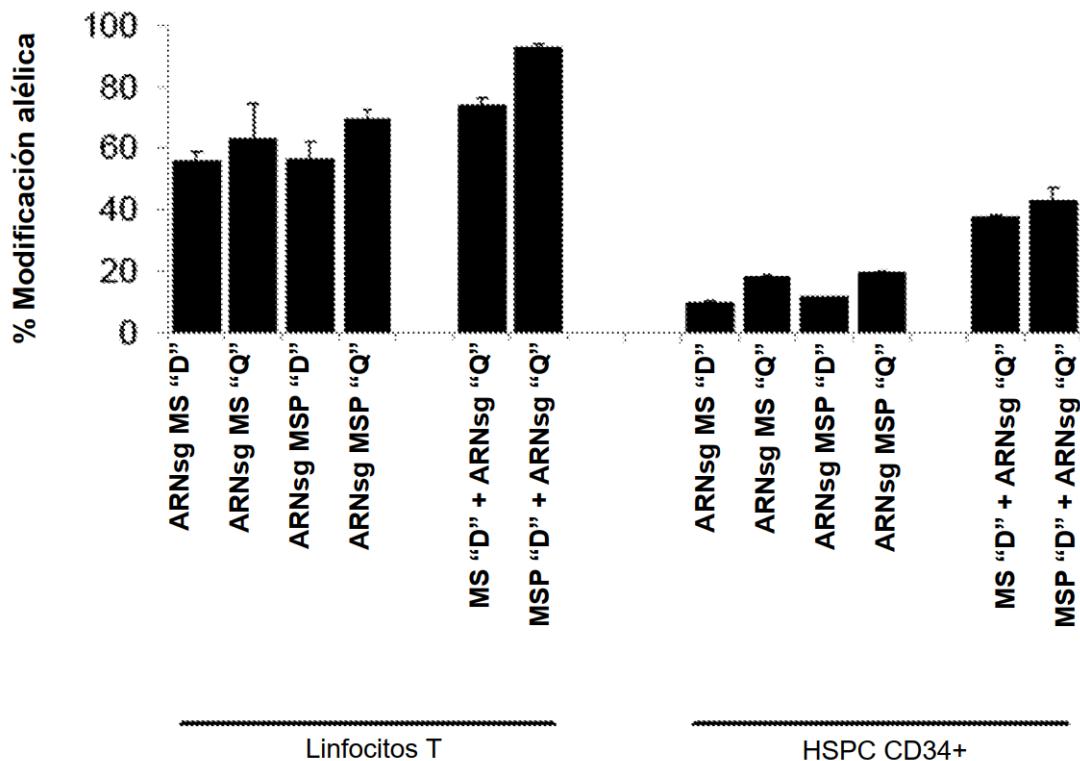
**FIG. 2A****FIG. 2B****FIG. 2C**

FIG. 2D



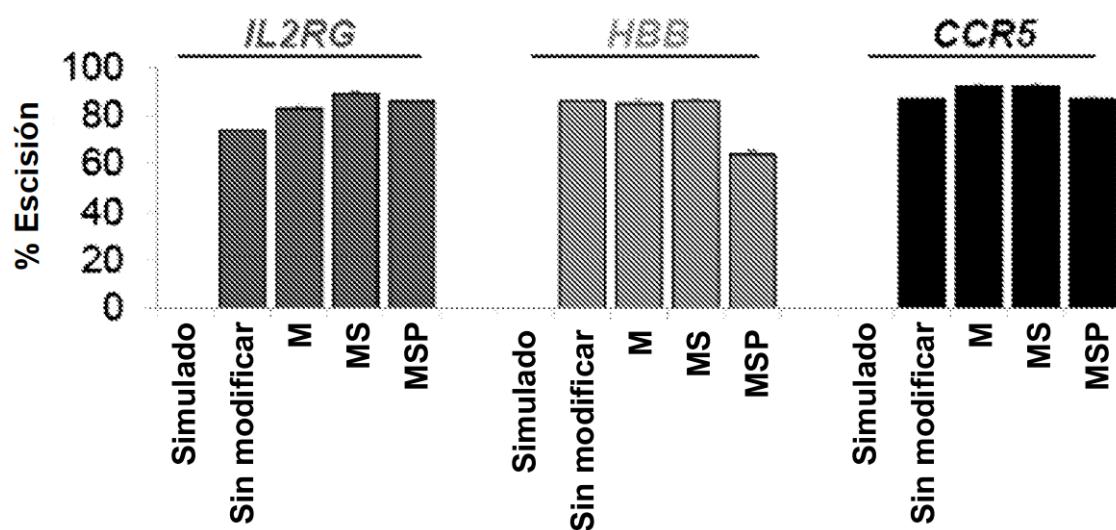


FIG. 3

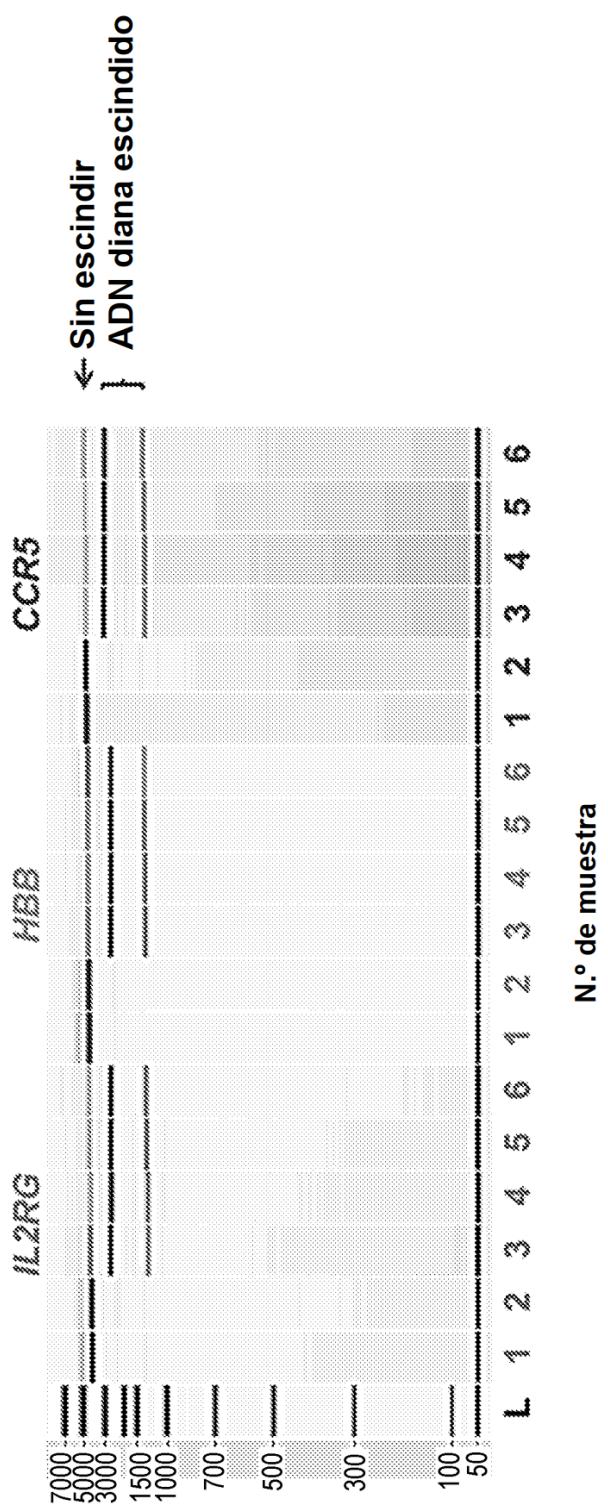


FIG. 4

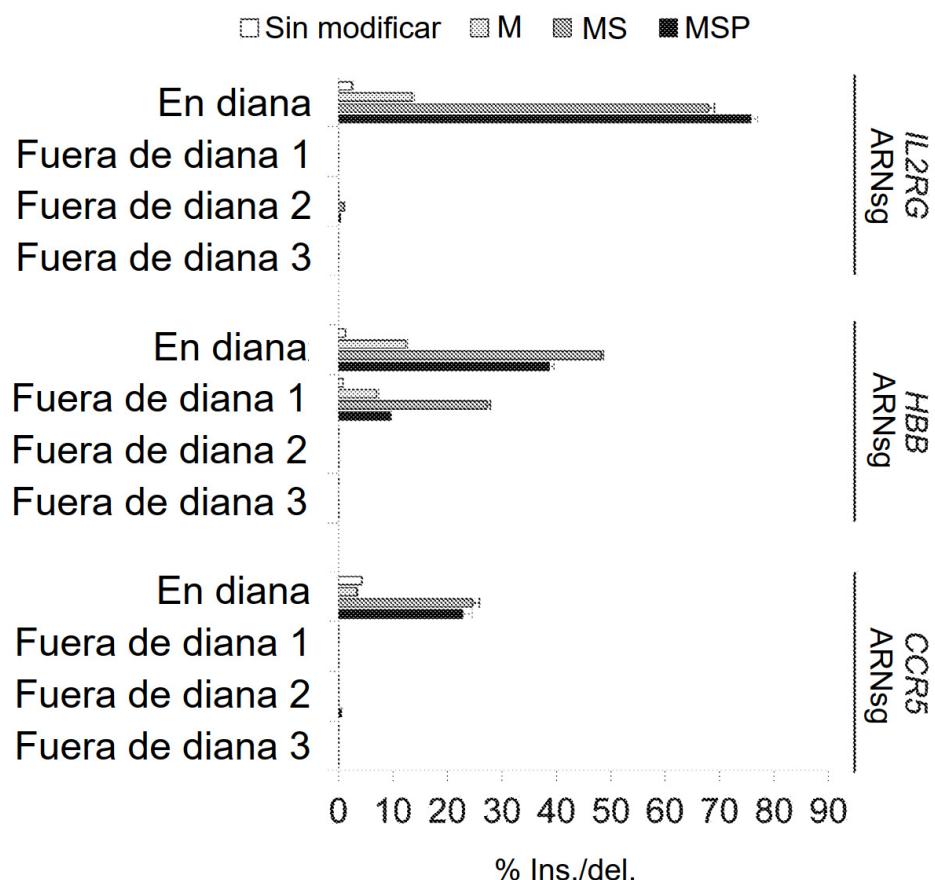


FIG. 5

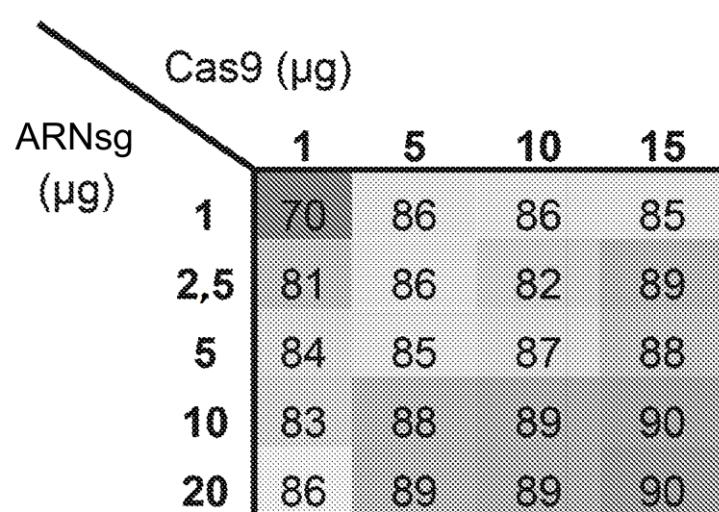


Figure 6 is a line graph showing the relationship between Cas9 concentration (μg) on the y-axis and ARNsg concentration (μg) on the x-axis. The data is represented by a 5x5 grid of numbers, with the first cell containing the value 70.

ARNsg (μg)	1	5	10	15	
Cas9 (μg)	1	70	86	86	85
	2,5	81	86	82	89
	5	84	85	87	88
	10	83	88	89	90
	20	86	89	89	90

*FIG. 6*

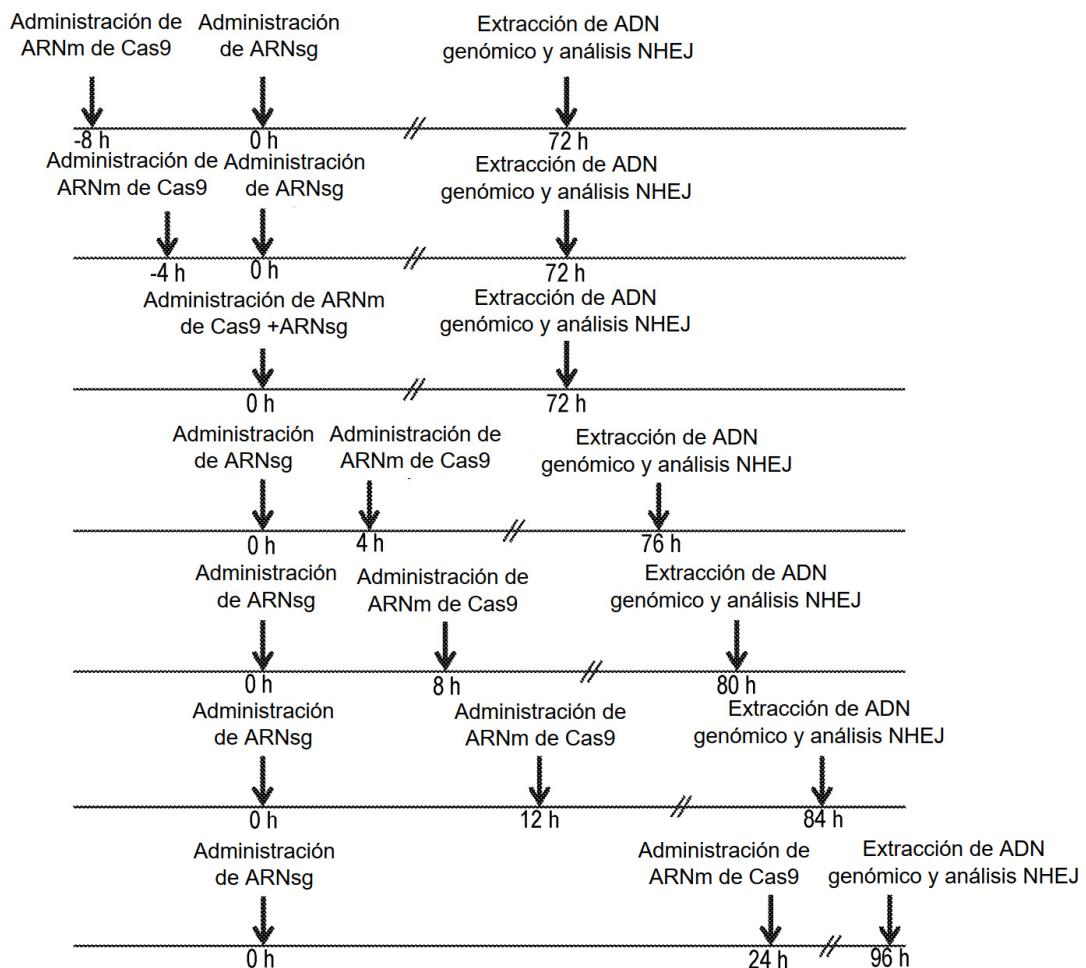
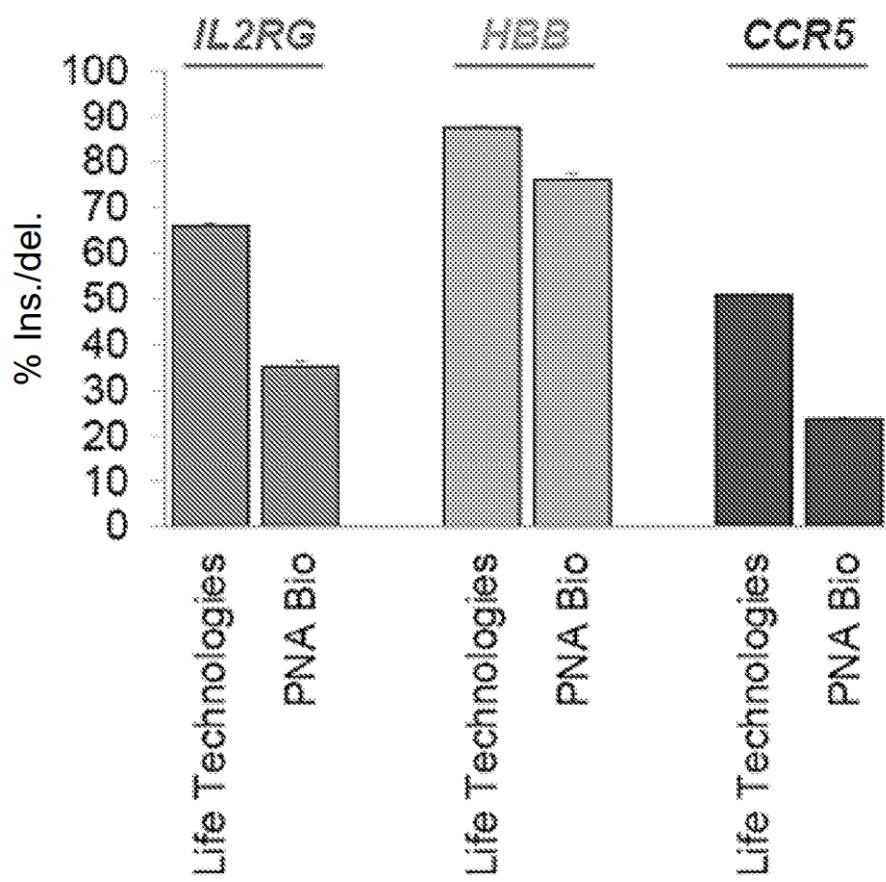
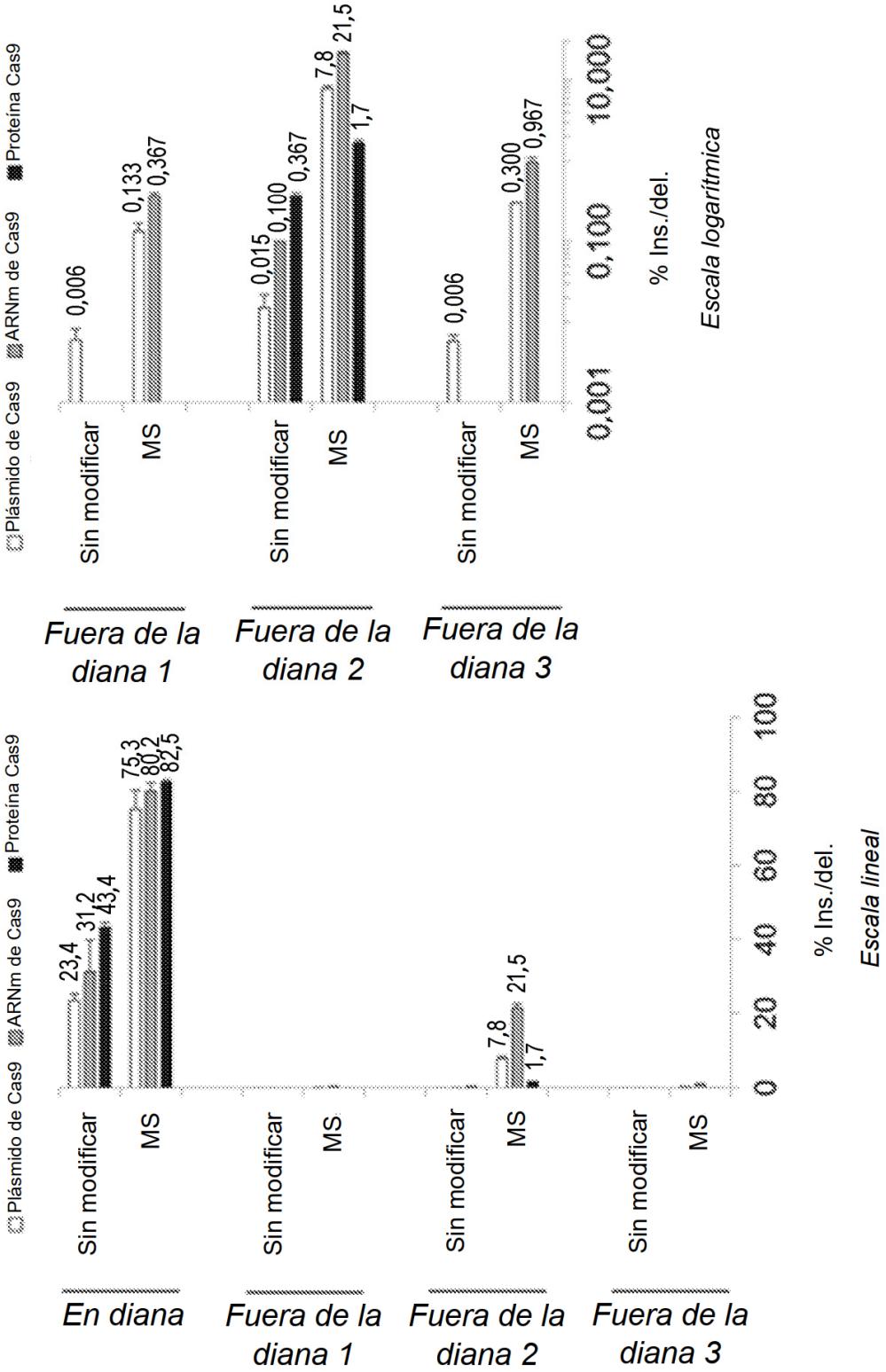


FIG. 7

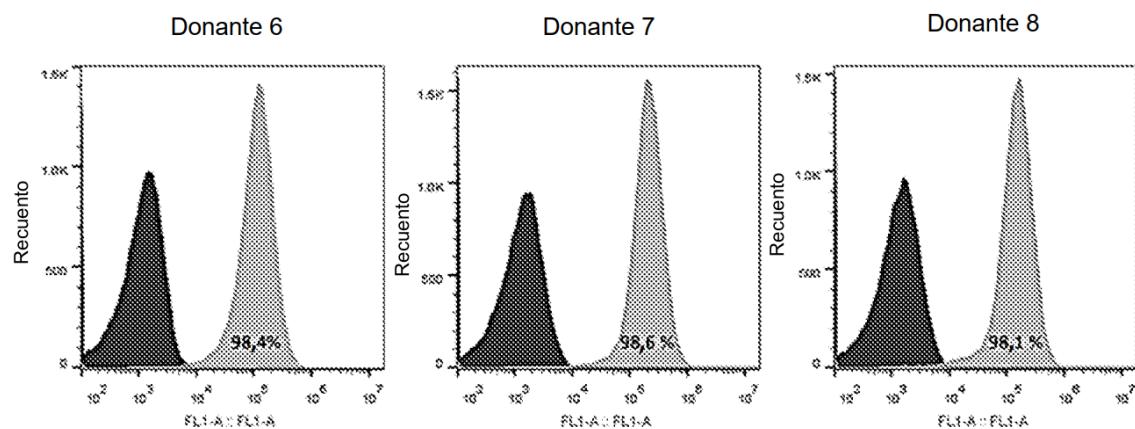


**FIG. 8**

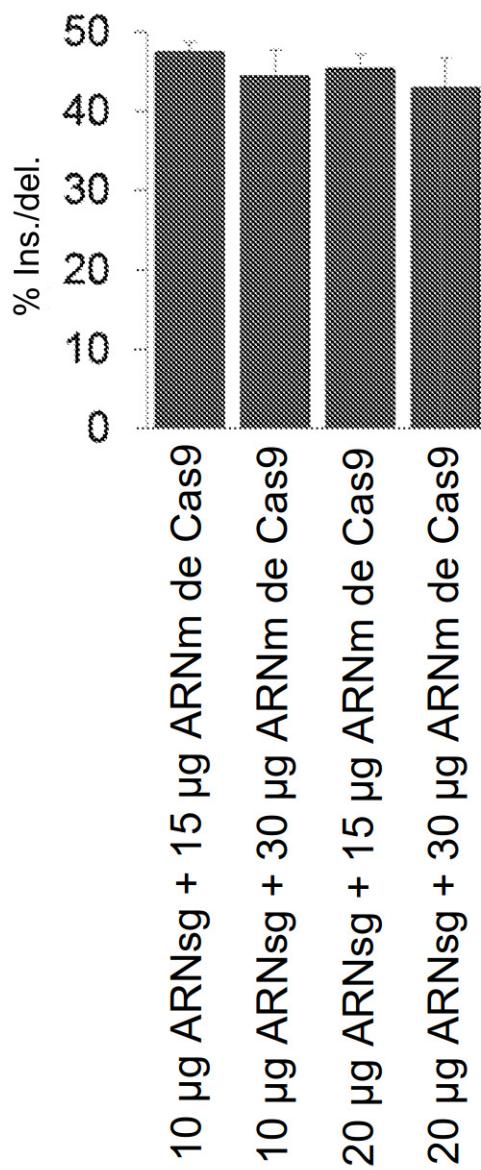
FIG. 9A



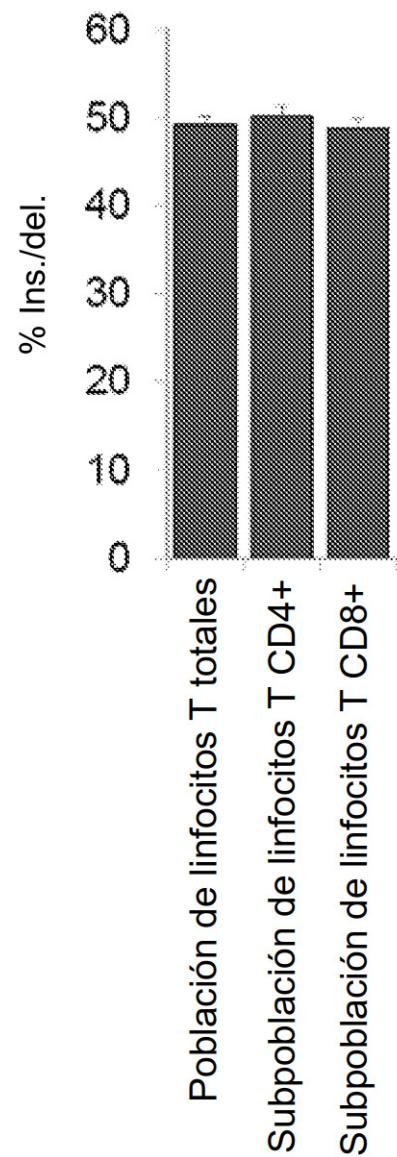
# ES 2 884 838 T3



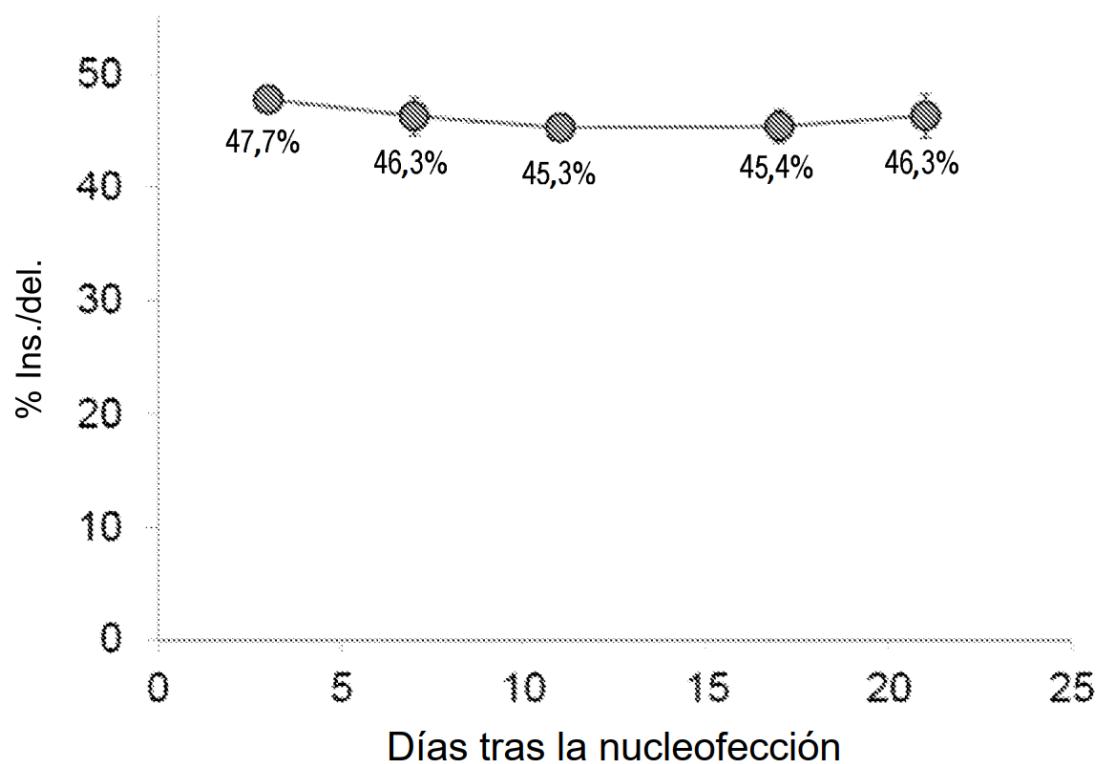
**FIG. 10**



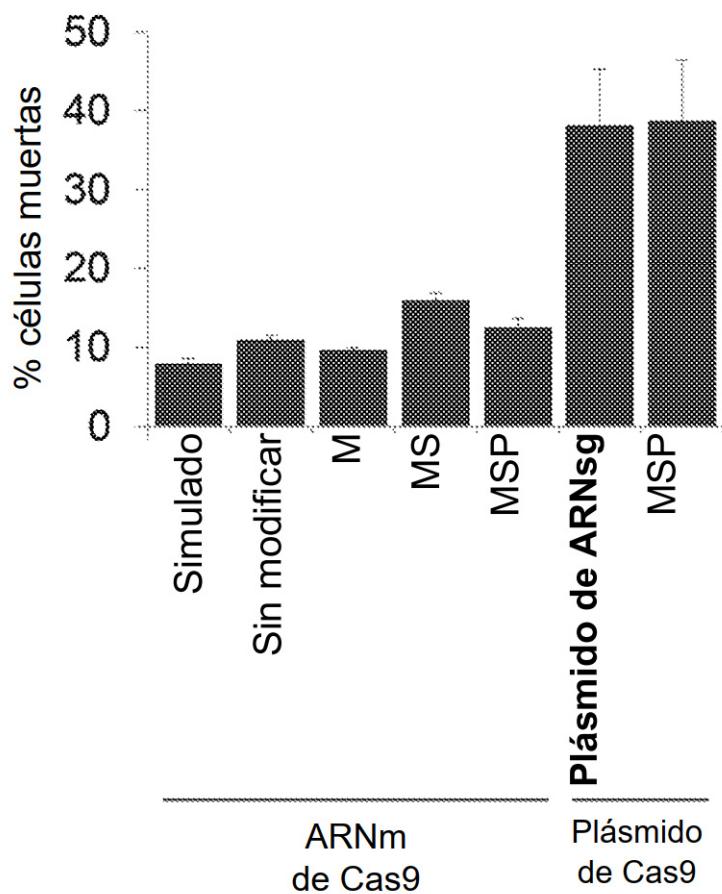
**FIG. 11**



**FIG. 12**



**FIG. 13**



**FIG. 14**

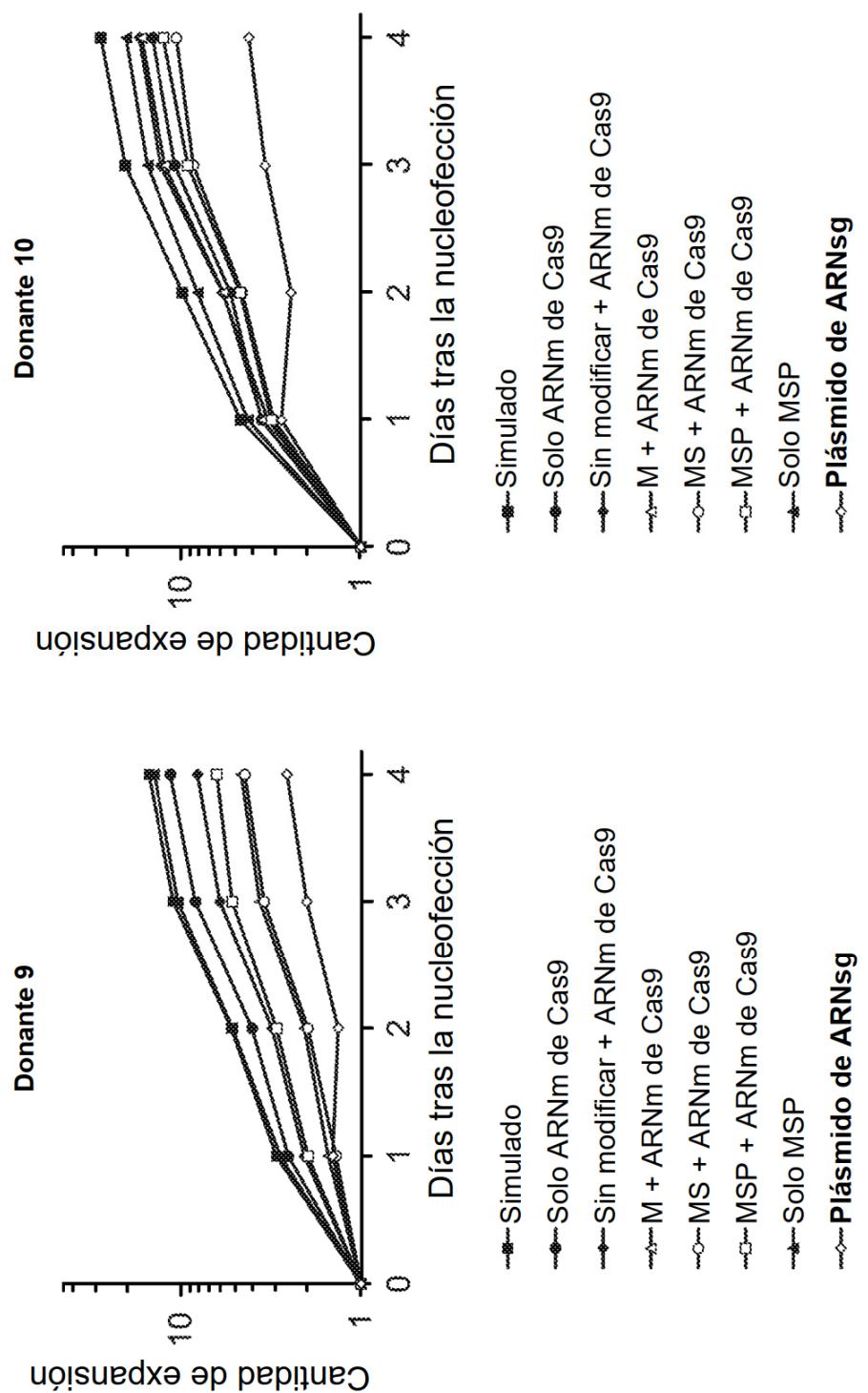
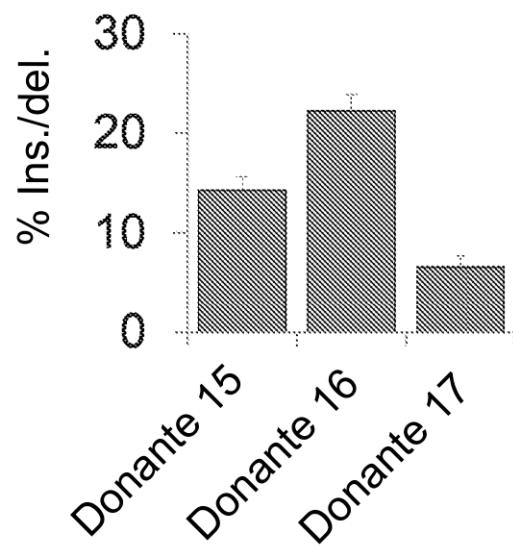


FIG. 15



**FIG. 16**

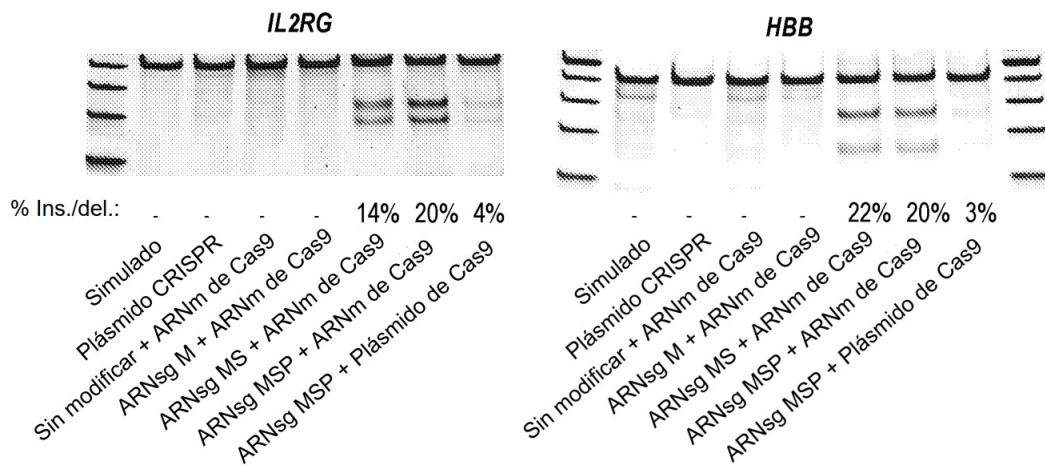
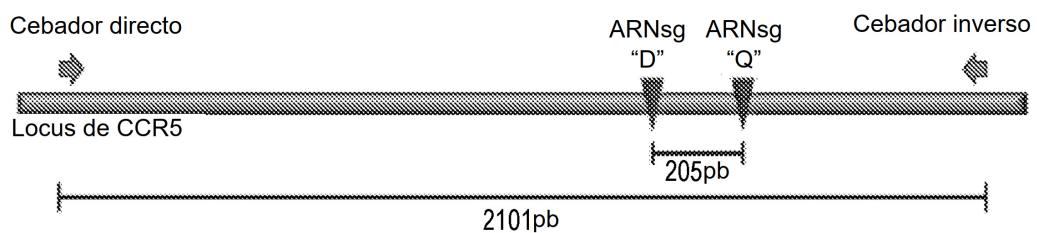
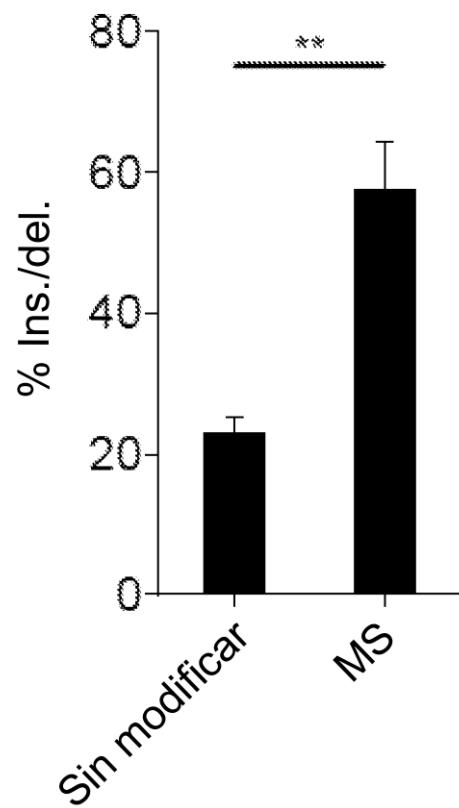


FIG. 17

**FIG. 18A****FIG. 18B**

	NIñOS	HSFC CDR4		
Total de fragmentos de PCR clonados analizados	181	186	222	216
% modificación de alelos $\pm$ SEM	73,9 ( $\pm$ 1,8)	93,1 ( $\pm$ 1,1)	37,8 ( $\pm$ 0,5)	43,0 ( $\pm$ 4,3)
% alelos SV $\pm$ SEM	26,1 ( $\pm$ 1,8)	6,9 ( $\pm$ 1,1)	62,2 ( $\pm$ 0,5)	57,0 ( $\pm$ 4,3)
% alelos con delección entre los sitios de corte $\pm$ SEM	53,4 ( $\pm$ 4,2)	69,7 ( $\pm$ 2,1)	21,1 ( $\pm$ 1,9)	23,1 ( $\pm$ 3,1)
% alelos con secuencia invertida entre los sitios de corte $\pm$ SEM	5,3 ( $\pm$ 2,2)	9,6 ( $\pm$ 2,2)	5,9 ( $\pm$ 1,6)	6,9 ( $\pm$ 1,1)
% ins/del en ambos sitios $\pm$ SEM	12,7 ( $\pm$ 2,8)	12,6 ( $\pm$ 1,6)	3,6 ( $\pm$ 1,2)	7,9 ( $\pm$ 1,9)
% ins/del solo en el sitio de la derecha $\pm$ SEM	2,4 ( $\pm$ 1,7)	1,1 ( $\pm$ 0,8)	3,6 ( $\pm$ 0,9)	2,3 ( $\pm$ 1,2)
% ins/del solo en el sitio de la izquierda $\pm$ SEM	0	0	3,6 ( $\pm$ 0,5)	2,8 ( $\pm$ 0,8)



**FIG. 19**

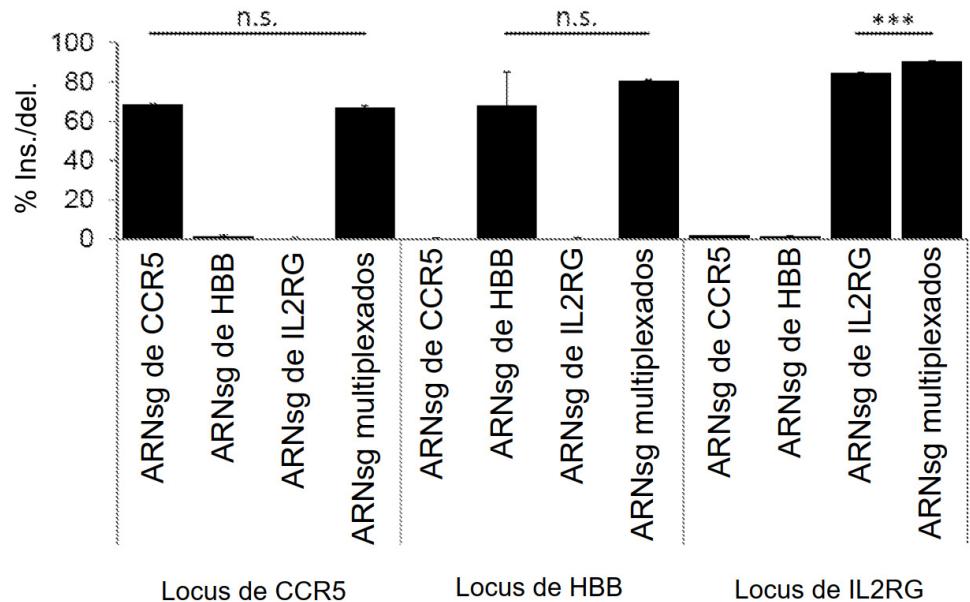


FIG. 20

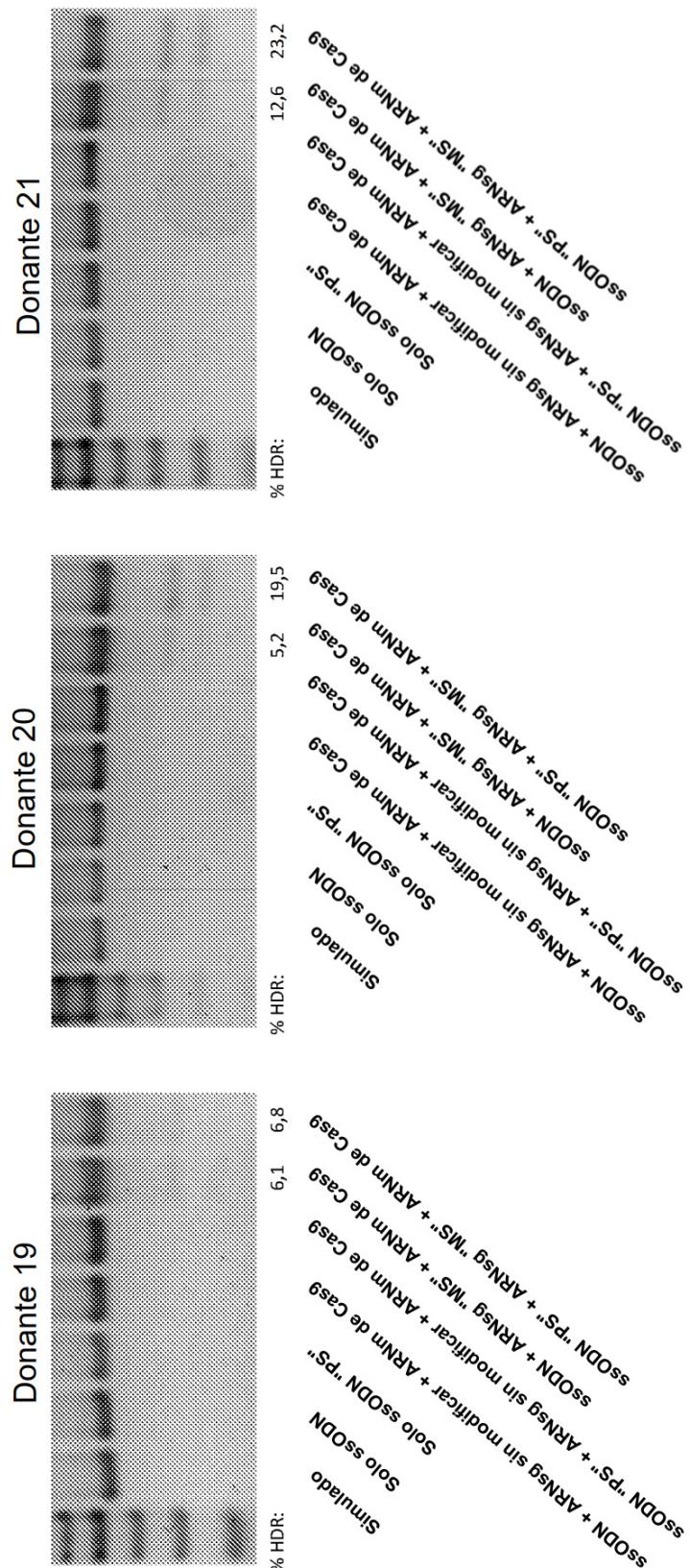


FIG. 21