

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 998 052**

(51) Int. Cl.:

C12N 15/113 (2010.01)
C12N 15/90 (2006.01)
C12N 9/22 (2006.01)
A61P 37/02 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Fecha de presentación y número de la solicitud internacional: **24.10.2018 PCT/US2018/057354**

(87) Fecha y número de publicación internacional: **02.05.2019 WO19084168**

(96) Fecha de presentación y número de la solicitud europea: **24.10.2018 E 18815370 (4)**

(97) Fecha y número de publicación de la concesión europea: **11.09.2024 EP 3701028**

(54) Título: **Sistemas y métodos para el tratamiento del síndrome de hiperigma**

(30) Prioridad:

24.10.2017 US 201762576277 P
30.04.2018 US 201862664800 P
17.05.2018 US 201862673039 P
26.06.2018 US 201862690284 P
17.10.2018 US 201862746900 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
18.02.2025

(73) Titular/es:

EDITAS MEDICINE, INC. (33.33%)
11 Hurley Street
Cambridge, MA 02141, US;
FONDAZIONE TELETHON ETS (33.33%) y
OSPEDALE SAN RAFFAELE S.R.L. (33.33%)

(72) Inventor/es:

COTTA-RAMUSINO, CECILLIA;
MARGULIES, CARRIE, M.;
NALDINI, LUIGI y
GENOVESE, PIETRO

(74) Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 998 052 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

Sistemas y métodos para el tratamiento del síndrome de hiperigma

5 **CAMPO**

[0001] La presente divulgación se refiere a métodos y componentes relacionados con CRISPR/Cas para tratar el Síndrome de Hiper-IgM, métodos de edición de una secuencia de ácido nucleico diana de CD40L y modulación de la expresión de una secuencia de ácido nucleico diana de CD40L, y aplicaciones de los mismos.

10 **ANTECEDENTES**

[0002] Las CRISPR (Repeticiones palindrómicas cortas agrupadas y regularmente espaciadas) evolucionaron en bacterias y arqueas como sistema inmunitario adaptativo para defenderse de ataques víricos. Tras la exposición a un virus, segmentos cortos de ADN viral se integran en el emplazamiento CRISPR. El ARN se transcribe a partir de una porción del emplazamiento CRISPR que incluye la secuencia viral. Este ARN, que contiene una secuencia complementaria al genoma vírico, dirige una proteína Cas9 a una secuencia diana del genoma vírico. La proteína Cas9, a su vez, escinde y silencia la diana vírica.

20 [0003] Recientemente, el sistema CRISPR/Cas se ha adaptado para la edición del genoma en células eucariotas. La introducción de roturas de doble cadena (DSB) en sitios específicos permite alterar la secuencia diana mediante mecanismos endógenos de reparación del ADN, por ejemplo la unión de extremos no homólogos (NHEJ) o la reparación dirigida por homología (HDR).

25 [0004] El síndrome de hiper-IgM (HIGM) es una deficiencia de inmunoglobulinas caracterizada por niveles séricos normales o elevados de IgM, y niveles disminuidos o ausencia de otras inmunoglobulinas séricas, lo que provoca susceptibilidad a las infecciones. La mayoría de los casos de HIGM están ligados al cromosoma X y causados por mutaciones en un gen del cromosoma X que codifica una proteína (ligando CD40, "CD40L") de la superficie de los linfocitos T auxiliares activados. Las mutaciones en el gen CD40L ("CD40LG") en sujetos que tienen el síndrome de hiper IgM incluyen mutaciones sin sentido, mutaciones sin sentido, mutaciones en el sitio de empalme, delecciones e inserciones y son bien conocidas por los expertos en la técnica. Véase, por ejemplo, Seyama et al., (1998) Blood, 92(7):2421-2434. En presencia de citocinas, el ligando CD40 normal interactúa con los linfocitos B y, de este modo, les indica que pasen de producir IgM a producir IgA, IgG y/o IgE. En el síndrome de hiper-IgM ligado al cromosoma X, las células T carecen del ligando CD40 funcional y no pueden indicar a las células B que cambien. Así pues, las células B de los sujetos con HIGM sólo producen IgM, y los niveles de IgM pueden ser normales o elevados.

40 [0005] El tratamiento del síndrome HIGM suele incluir terapia de sustitución con inmunoglobulinas, profilaxis a largo plazo con antibióticos y antifúngicos, y tratamiento con G-CSF para la neutropenia. También es preferible el trasplante de células madre hematopoyéticas si se dispone de un donante hermano HLA-ídntico. Sin embargo, los tratamientos actuales no son eficaces, ya que cerca del 80 % de los pacientes con HIGM mueren antes de los treinta años. Por consiguiente, sigue existiendo una necesidad insatisfecha de desarrollar estrategias y productos terapéuticos novedosos para tratar la HIGM. N. HUBBARD et al. (BLOOD (22 de febrero de 2016), 127 (21): 2513-2522) describe enfoques para corregir el gen CD40L utilizando nucleasas efectoras del tipo activador de la transcripción (TAL).

RESUMEN

50 [0006] Un enfoque de inmunoterapia adoptiva basado en células T autólogas corregidas genéticamente puede ser suficiente para inducir la generación de inmunidad protectora en pacientes receptores de HIGM1 y posiblemente también la producción de algunas células B de memoria de larga vida. Sin embargo, pueden ser necesarias múltiples administraciones de células T editadas para reconstituir un repertorio de células T duradero y amplio que pueda mediar una ayuda eficiente de las células T en respuesta a un amplio intervalo de patógenos. Además, mientras que CD40LG tiene una función crítica en las células T CD4 (Th0, Th1 y Th2), esta molécula también se expresa en varios otros tipos de células hematopoyéticas, tales como células B activadas, plaquetas, células NK, monocitos, basófilos y eosinófilos. Por lo tanto, la presente divulgación amplía sorprendentemente la estrategia de edición genética de la corrección de células T a la corrección de células madre/progenitoras hematopoyéticas autólogas (HSPC), lo que proporciona un beneficio terapéutico mucho más amplio y prolongado.

60 [0007] La presente divulgación se basa, al menos en parte, en el descubrimiento de que las células editadas con ligando CD40 (CD40L) producen sorprendentemente proteína CD40L funcionalmente activa y pueden utilizarse para restaurar la función del cambio de clase de células B para el tratamiento del síndrome de hiper-IgM. En particular, cuando uno o más de los exones 2-5 del gen CD40L, se editan en células, *por ejemplo*, células T o células madre hematopoyéticas (HSC), usando un sistema de edición del genoma, las células editadas demuestran altos niveles de edición del gen CD40L y altos niveles de expresión funcional exógena de CD40L. En concreto, las células T editadas con CD40L restauraron la capacidad de cambio de clase de las células B, aumentando el nivel

de IgG producida y disminuyendo el nivel de IgM. Además, los ratones que recibieron células madre hematopoyéticas que contenían al menos un 10 % de células editadas con CD40L produjeron un mayor nivel de IgG tras la vacunación, lo que demuestra el beneficio terapéutico de las células madre hematopoyéticas editadas con CD40L para su uso en el tratamiento del síndrome de hiper-IgM.

5 [0008] Los sistemas de edición genómica divulgados en el presente documento están diseñados para que no se edite el exón 1 del gen CD40L, lo que evita la posibilidad de que se produzcan eventos de integración aleatoria o eventos fuera de diana. En concreto, la edición del exón 1 puede provocar la pérdida del control del promotor del emplazamiento, afectando a la expresión del gen CD40L y dando lugar a trastornos linfoproliferativos, especialmente en el contexto de los trasplantes de células madre hematopoyéticas.

10 15 [0009] Por consiguiente, en un aspecto, en el presente documento se divulga una plantilla donante de oligonucleótidos aislada que comprende, de 5' a 3', A1-N-UTR-pA-A2, en el que A1 es un brazo de homología que es sustancialmente idéntico a un primer brazo de homología de un ácido nucleico CD40L diana; N es una carga que comprende uno o más de los exones 2-5 de un gen CD40L; UTR es una región no traducida (UTR) 3' de CD40L; pA es una cola poliA; y A2 es un brazo de homología que es sustancialmente idéntico a un segundo brazo de homología del ácido nucleico CD40L diana, y en el que la carga no comprende el exón 1 de CD40L.

20 25 [0010] En un aspecto, se divulga en el presente documento una plantilla donante de oligonucleótidos aislada que comprende, de 5' a 3', A1-S1-N-UTR-pA-S2-A2, en el que A1 es un brazo de homología que es sustancialmente idéntico a un primer brazo de homología de un ácido nucleico CD40L diana; S1 es un primer relleno, N es una carga que comprende uno o más de los exones 2-5 de un gen CD40L; UTR es una región no traducida (UTR) 3' de CD40L; pA es una cola poliA; S2 es un segundo relleno, y A2 es un brazo de homología que es sustancialmente idéntico a un segundo brazo de homología del ácido nucleico CD40L diana, y en el que la carga no comprende el exón 1 de CD40L.

30 35 [0011] En un aspecto, se divulga en el presente documento una plantilla donante de oligonucleótidos aislada que comprende, de 5' a 3', A1-S1-N-UTR-pA-R-S2-A2, en el que A1 es un brazo de homología que es sustancialmente idéntico a un primer brazo de homología de un ácido nucleico CD40L diana; S1 es un primer relleno, N es una carga que comprende uno o más de los exones 2-5 de un gen CD40L; UTR es una región no traducida (UTR) 3' de CD40L; pA es una cola poliA; en el que R es un informador; S2 es un segundo relleno, y A2 es un brazo de homología que es sustancialmente idéntico a un segundo brazo de homología del ácido nucleico CD40L diana, y en el que la carga no comprende el exón 1 de CD40L.

40 45 [0012] En un aspecto, se divulga en el presente documento una plantilla donante de oligonucleótidos aislada que comprende, de 5' a 3', A1-N-UTR-pA-R-A2, en el que A1 es un brazo de homología que es sustancialmente idéntico a un primer brazo de homología de un ácido nucleico CD40L diana; N es una carga que comprende uno o más de los exones 2-5 de un gen CD40L; UTR es una región no traducida (UTR) 3' de CD40L; pA es una cola poliA; R es un informador; y A2 es un brazo de homología sustancialmente idéntico a un segundo brazo de homología del ácido nucleico CD40L diana, y en el que la carga no comprende el exón 1 de CD40L.

50 55 [0013] En cualquier aspecto descrito en el presente documento, la carga puede comprender el exón 2 del gen CD40L; el exón 3 del gen CD40L; el exón 4 del gen CD40L; el exón 5 del gen CD40L; los exones 2-5 del gen CD40L; los exones 2-4 del gen CD40L; los exones 2 y 3 del gen CD40L; los exones 3-5 del gen CD40L; los exones 3 y 4 del gen CD40L; o los exones 4 y 5 del gen CD40L. En una realización, la carga comprende una secuencia CD40L de codón optimizado. Por ejemplo, la carga puede comprender una versión de codón optimizado del exón 2 de CD40L, del exón 3 de CD40L, del exón 4 de CD40L, del exón 5 de CD40L, o una combinación de los mismos, *por ejemplo*, una secuencia de exón 2-5 de codón optimizado, una secuencia de exón 2-4 de codón optimizado; una secuencia de exón 2 y 3 de codón optimizado; una secuencia de exón 3-5 de codón optimizado; una secuencia de exón 3 y 4 de codón optimizado; o una secuencia de exón 4 y 5 de codón optimizado. Una versión ejemplar de codón optimizado de CD40L, exones 2-5, se proporciona como SEQ ID NO:48.

60 65 [0014] En una realización, el gen CD40L es una secuencia CD40L de tipo salvaje. En una realización, el ácido nucleico CD40L diana comprende una mutación. En una realización, la plantilla donante de oligonucleótidos comprende además una secuencia que codifica un informador. En una realización, el reportero es una proteína de fluorescencia verde (GFP), una proteína de fluorescencia amarilla (YFP), DS-Red o luciferasa. En una realización, el reportero puede ser un marcador que puede ser detectado en células vivas usando técnicas estándar, por ejemplo,抗原os de superficie detectables por anticuerpos, péptidos que catalizan o facilitan una reacción química que produce un producto ópticamente detectable, por ejemplo, luciferasa. En una realización, el informador es un marcador seleccionable. En una realización ejemplar, el marcador seleccionable es el receptor del factor de crecimiento nervioso de baja afinidad (NGFR). En otras realizaciones, el marcador seleccionable puede ser una proteína resistente a fármacos (tal como resistencia a neomicina o puromicina, versión mutante del gen MGMT), una versión truncada de una proteína de la superficie celular (por ejemplo CD19, EGFR), un gen que confiere una ventaja selectiva de crecimiento y/o injerto tras el trasplante in vivo de las células editadas (por ejemplo, CXCR4, CD47, receptor de IL2) o una proteína fluorescente informadora (*por ejemplo*, proteína verde fluorescente (GFP), proteína azul fluorescente (BFP), proteína amarilla fluorescente (YFP), etc., o una combinación de las mismas. En

una realización, la plantilla donante de oligonucleótidos aislada es una plantilla donante de oligonucleótidos monocatenaria. En una realización, la plantilla donante de oligonucleótidos aislada es una plantilla donante de oligonucleótidos bicatenarios.

5 [0015] En un aspecto, en el presente documento se divulga una composición que comprende una plantilla donante de oligonucleótidos aislada de la invención y, opcionalmente, un portador farmacéuticamente aceptable.

10 [0016] En un aspecto, en el presente documento se divulga un vector que comprende una plantilla donante de oligonucleótidos aislada de la invención. En una realización, el vector es un vector AAV, un lentivirus, un vector de ADN desnudo o una nanopartícula lipídica. En una realización, el vector AAV es un vector AAV6.

15 [0017] En un aspecto, en el presente documento se divulga una molécula de ARNg aislada que comprende cualquiera de los SEQ ID NOs: 18-25. En una realización, la molécula de ARNg aislada comprende SEQ ID NO: 18. En una realización, la molécula de ARNg aislada comprende SEQ ID NO: 19. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:20. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:21. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:22. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:23. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:24. En una realización, la molécula de ARNg aislada comprende SEQ ID NO:25.

20 [0018] En un aspecto, en el presente documento se divulga un sistema de edición del genoma que comprende una plantilla donante de oligonucleótidos aislada de la invención. En una realización, el sistema de edición genómica comprende además una nucleasa guiada por ARN y al menos una molécula de ARNg. En una realización, el sistema de edición genómica comprende además una molécula de ARNg aislada divulgada en el presente documento.

25 [0019] En un aspecto, en el presente documento se divulga un método in vitro de alteración de una célula que comprende el contacto de la célula con un sistema de edición del genoma de la invención, alterando así la célula.

30 [0020] En un aspecto, en el presente documento se divulga un kit que comprende un sistema de edición genómica de la invención.

35 [0021] En un aspecto, se divulga en el presente documento una plantilla donante de oligonucleótidos aislada de la invención, una composición de la invención, un vector de la invención, una molécula de ARNg aislada de la invención, un sistema de edición del genoma de la invención, una célula preparada por el método de la invención, o un kit de la invención, para su uso en medicina.

40 [0022] En un aspecto, se divulga en el presente documento una plantilla donante de oligonucleótidos para su uso en un método de tratamiento de un sujeto que tiene síndrome de hiper-IgM mediante la alteración de un CD40L diana en una célula, comprendiendo el método las etapas de: formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo homólogo 5' al sitio de escisión y un segundo brazo homólogo 3' al sitio de escisión, y recombinando la plantilla donante de oligonucleótidos con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla donante de oligonucleótidos comprende un primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, una carga que comprende uno o más de los exones 2-5 de un gen CD40L, una región no traducida (UTR) 3' de CD40L, una cola poliA, y un segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L.

50 [0023] En un aspecto, en el presente documento se divulga un método in vitro para alterar una célula, que comprende las etapas de: formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo homólogo 5' al sitio de escisión y un segundo brazo homólogo 3' al sitio de escisión, y recombinando una plantilla exógena de oligonucleótido donante con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla exógena de oligonucleótido donante comprende un primer brazo homológico donante que es sustancialmente idéntico al primer brazo homólogo, una carga que comprende uno o más de los exones 2-5 de un gen CD40L, una región no traducida (UTR) 3' de CD40L, una cola poliA, y un segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L, alterando así la célula.

60 [0024] En un aspecto, en el presente documento se divulga un método in vitro para alterar una célula, que comprende las etapas de: formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo homólogo 5' al sitio de escisión y un segundo brazo homológico 3' al sitio de escisión, y recombinando una plantilla exógena de oligonucleótido donante con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo homológico donante que es sustancialmente idéntico al primer brazo homólogo, un

primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' de CD40L, la cola poliA, un segundo relleno, y el segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L, alterando así la célula.

- 5 [0025] En un aspecto, en el presente documento se divulga un método in vitro para alterar una célula, que comprende las etapas de: formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo homólogo 5' al sitio de escisión y un segundo brazo homólogo 3' al sitio de escisión, y recombinando una plantilla exógena de oligonucleótido donante con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, el primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' de CD40L, la cola poliA, un informador, el segundo relleno, y el segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L, alterando así la célula.

- 10 [0026] En un aspecto, en el presente documento se divulga un método in vitro para alterar una célula, que comprende las etapas de: formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo homólogo 5' al sitio de escisión y un segundo brazo homólogo 3' al sitio de escisión, y recombinando una plantilla exógena de oligonucleótido donante con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' de CD40L, la cola poliA, un reportero, y el segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L, alterando así la célula.

- 15 [0027] En una realización, la carga comprende la totalidad o una parte de: el exón 2 del gen CD40L; el exón 3 del gen CD40L; el exón 4 del gen CD40L; el exón 5 del gen CD40L; los exones 2-5 del gen CD40L; los exones 2-4 del gen CD40L; los exones 2 y 3 del gen CD40L; los exones 3-5 del gen CD40L; los exones 3 y 4 del gen CD40L; o los exones 4 y 5 del gen CD40L.

- 20 [0028] En una realización, el gen CD40L es una secuencia CD40L de tipo salvaje. En una realización, el ácido nucleico CD40L diana comprende una mutación.

- 25 [0029] En una realización, el reportero es una proteína de fluorescencia verde (GFP), una proteína de fluorescencia amarilla (YFP), DS-Red o luciferasa. En una realización, el reportero puede ser un marcador que puede ser detectado en células vivas usando técnicas estándar, por ejemplo, antígenos de superficie detectables por anticuerpos, péptidos que catalizan o facilitan una reacción química que produce un producto ópticamente detectable, por ejemplo, luciferasa. En otra realización, el informador es un marcador seleccionable. En una realización ejemplar, el informador es NGFR.

- 30 [0030] En una realización, la etapa de formación de al menos una rotura de cadena simple o doble comprende el contacto de la célula con una nucleasa guiada por ARN. En una realización, la nucleasa guiada por ARN es una nucleasa asociada a repeticiones regulares intercaladas agrupadas de clase 2 (CRISPR). En una realización, la nucleasa guiada por ARN se selecciona del grupo que consiste en Cas9 de tipo salvaje, una nicasa Cas9, una Cpf1 de tipo salvaje y una nicasa Cpf1.

- 35 [0031] En una realización, el contacto de la nucleasa guiada por ARN con la célula comprende la introducción en la célula de un complejo de ribonucleoproteína (RNP) que comprende la nucleasa guiada por ARN y un ARN guía (ARNg). En una realización, la etapa de recombinación de la plantilla donante de oligonucleótidos exógena en el ácido nucleico CD40L diana mediante recombinación homóloga comprende la introducción de la plantilla donante de oligonucleótidos exógena en la célula. En una realización, la etapa de introducción comprende la electroporación de la célula en presencia del complejo RNP y/o la plantilla exógena donante de oligonucleótidos.

- 40 [0032] En un aspecto, en el presente documento se divulga un método in vitro para alterar un ácido nucleico CD40L diana en una célula, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, el método comprende: Poner en contacto la célula con (a) al menos una molécula de ARNg, (b) una molécula de nucleasa guiada por ARN, y (c) una plantilla exógena de oligonucleótido donante, en el que la plantilla exógena de oligonucleótido donante comprende un primer brazo de homología donante que es sustancialmente idéntico al primer brazo de homología, una carga que comprende uno o más de los exones 2-5 de un gen CD40L, una región no traducida (UTR) 3' CD40L, una cola de poliA, y un segundo brazo homólogo donante que es sustancialmente idéntico al segundo brazo homólogo, y en el que la carga no comprende el exón 1 de CD40L, en el que la molécula de ARNg y la molécula de nucleasa guiada por ARN interactúan con el ácido nucleico CD40L diana, dando lugar a un evento de escisión en o cerca

del sitio de escisión, y en el que el evento de escisión es reparado por al menos una vía de reparación del ADN para producir un ácido nucleico CD40L alterado, alterando así el ácido nucleico CD40L diana en la célula.

[0033] En un aspecto, en el presente documento se divulga un método in vitro para alterar un ácido nucleico CD40L diana en una célula, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, el método comprende: Poner en contacto la célula con (a) al menos una molécula de ARNg, (b) una molécula de nucleasa guiada por ARN, y (c) una plantilla exógena de oligonucleótido donante, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo de homología donante que es sustancialmente idéntico al primer brazo de homología, un primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' CD40L, la cola poliA, un segundo relleno y el segundo brazo homológico donante que es sustancialmente idéntico al segundo brazo homológico, y en el que la carga no comprende el exón 1 de CD40L, en el que la molécula de ARNg y la molécula de nucleasa guiada por ARN interactúan con el ácido nucleico CD40L diana, dando lugar a un evento de escisión en o cerca del sitio de escisión, y en el que el evento de escisión es reparado por al menos una vía de reparación del ADN para producir un ácido nucleico CD40L alterado, alterando así el ácido nucleico CD40L diana en la célula.

[0034] En un aspecto, en el presente documento se divulga un método in vitro para alterar un ácido nucleico CD40L diana en una célula, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, el método comprende: Poner en contacto la célula con (a) al menos una molécula de ARNg, (b) una molécula de nucleasa guiada por ARN, y (c) una plantilla exógena de oligonucleótido donante, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo de homología donante que es sustancialmente idéntico al primer brazo de homología, el primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' CD40L, la cola poliA, un reportero, el segundo relleno, y el segundo brazo homológico donante que es sustancialmente idéntico al segundo brazo homológico, y en el que la carga no comprende el exón 1 de CD40L, en el que la molécula de ARNg y la molécula de nucleasa guiada por ARN interactúan con el ácido nucleico CD40L diana, dando lugar a un evento de escisión en o cerca del sitio de escisión, y en el que el evento de escisión es reparado por al menos una vía de reparación del ADN para producir un ácido nucleico CD40L alterado, alterando así el ácido nucleico CD40L diana en la célula.

[0035] En un aspecto, en el presente documento se divulga un método in vitro para alterar un ácido nucleico CD40L diana en una célula, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, el método comprende: Poner en contacto la célula con (a) al menos una molécula de ARNg, (b) una molécula de nucleasa guiada por ARN, y (c) una plantilla exógena de oligonucleótido donante, en el que la plantilla exógena de oligonucleótido donante comprende el primer brazo homológico donante que es sustancialmente idéntico al primer brazo homológico, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3' CD40L, la cola poliA, un reportero, y el segundo brazo homológico donante que es sustancialmente idéntico al segundo brazo homológico, y en el que la carga no comprende el exón 1 de CD40L, en el que la molécula de ARNg y la molécula de nucleasa guiada por ARN interactúan con el ácido nucleico CD40L diana, dando lugar a un evento de escisión en o cerca del sitio de escisión, y en el que el evento de escisión es reparado por al menos una vía de reparación del ADN para producir un ácido nucleico CD40L alterado, alterando así el ácido nucleico CD40L diana en la célula.

[0036] En una realización, el método comprende además poner en contacto la célula con (d) una segunda molécula de ARNg, en el que la segunda molécula de ARNg y la molécula de nucleasa guiada por ARN interactúan con el ácido nucleico diana, dando lugar a un segundo evento de escisión en o cerca del sitio de escisión, y en el que el segundo evento de escisión es reparado por la al menos una vía de reparación del ADN.

[0037] En una realización, la carga comprende la totalidad o una parte de: el exón 2 del gen CD40L; el exón 3 del gen CD40L; el exón 4 del gen CD40L; el exón 5 del gen CD40L; los exones 2-5 del gen CD40L; los exones 2-4 del gen CD40L; los exones 2 y 3 del gen CD40L; los exones 3-5 del gen CD40L; los exones 3 y 4 del gen CD40L; o los exones 4 y 5 del gen CD40L.

[0038] En una realización, el gen CD40L es una secuencia CD40L de tipo salvaje. En una realización, el ácido nucleico CD40L diana comprende una mutación.

[0039] En una realización, el reportero es una proteína de fluorescencia verde (GFP), una proteína de fluorescencia amarilla (YFB), DS-Red o luciferasa. En una realización, el informador puede ser un marcador que puede detectarse en células vivas utilizando técnicas estándar, por ejemplo, antígenos de superficie detectables mediante anticuerpos, péptidos que catalizan o facilitan de otro modo una reacción química que produce un producto detectable ópticamente, por ejemplo, luciferasa. En otra realización, el informador es un marcador seleccionable. En una realización ejemplar, el informador es NGFR.

[0040] En una realización, el ácido nucleico CD40L alterado comprende, de 5' a 3', el primer brazo de homología donante, la carga, el 3' UTR, la cola poliA y el segundo brazo de homología donante. En una realización, el ácido

nucleico CD40L alterado comprende, de 5' a 3', el primer brazo de homología donante, el primer relleno, la carga, el 3' UTR, la cola poliA, el segundo relleno y el segundo brazo de homología donante. En una realización, el ácido nucleico CD40L alterado comprende, de 5' a 3', el primer brazo homológico donante, el primer relleno, la carga, el 3' UTR, la cola poliA, el informador, el segundo relleno y el segundo brazo homológico donante. En una realización, el ácido nucleico CD40L alterado comprende, de 5' a 3', el primer brazo de homología donante, la carga, el 3' UTR, la cola poliA, el reportero y el segundo brazo de homología donante.

5 [0041] En una realización, la célula se pone en contacto primero con la al menos una molécula de ARNg y la molécula de nucleasa guiada por ARN, seguido del contacto de la célula con la plantilla exógena donante de oligonucleótidos. En una realización, la célula se pone en contacto con la al menos una molécula de ARNg, la molécula de nucleasa guiada por ARN y la plantilla exógena donante de oligonucleótidos al mismo tiempo.

10 15 [0042] En una realización, la plantilla donante de oligonucleótidos exógena es una plantilla donante de oligonucleótidos monocatenarios. En una realización, la plantilla exógena donante de oligonucleótidos es una plantilla donante de oligonucleótidos bicatenarios.

20 [0043] En una realización, la plantilla exógena donante de oligonucleótidos está presente en un vector. En una realización, el vector es un vector viral. En una realización, el vector viral es un vector AAV o un vector lentiviral.

25 [0044] En una realización, la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración dirigida de la plantilla donante de oligonucleótidos exógena. En una realización, el evento de escisión, o tanto el evento de escisión como el segundo evento de escisión, es/son reparado/s por corrección génica.

30 35 [0045] En una realización, el ácido nucleico CD40L alterado comprende una secuencia que es diferente de una secuencia del ácido nucleico CD40L diana.

40 [0046] En una realización, la molécula de ARNg es un ácido nucleico de ARNg, y en el que la molécula de nucleasa guiada por ARN es una proteína nucleasa guiada por ARN. En una realización, la molécula de ARNg es un ácido nucleico de ARNg, y en el que la molécula de nucleasa guiada por ARN es un ácido nucleico de nucleasa guiada por ARN. En una realización, la célula se pone en contacto con la molécula de ARNg y la molécula de nucleasa guiada por ARN como un complejo preformado. En una realización, la nucleasa guiada por ARN se selecciona del grupo que consiste en Cas9 de tipo salvaje, una nicasa Cas9, una Cpfl de tipo salvaje y una nicasa Cpfl.

45 [0047] En una realización, la célula es una célula eucariota. En una realización, la célula eucariota es una célula humana. En una realización, la célula es una célula T. En una realización, la célula es una célula madre hematopoyética (HSC).

50 [0048] En una realización, la célula procede de un sujeto que padece el síndrome de hiper IgM. En una realización, la célula procede de un sujeto que tiene al menos una mutación en el sitio de escisión. En una realización, el método de tratamiento de un sujeto que tiene síndrome de hiper-IgM alterando un CD40L diana en una célula comprende además aislar la célula del sujeto antes de la etapa de formación o la etapa de contacto. En una realización, el método de tratamiento de un sujeto que tiene síndrome de hiper-IgM alterando un CD40L diana en una célula comprende además introducir la célula en un sujeto después de la etapa de recombinación o después de que el evento de escisión sea reparado por la al menos una vía de reparación del ADN.

55 [0049] En una realización, la etapa de formación y la etapa de recombinación, o la etapa de contacto, se realiza *in vitro*. En una realización, la etapa de formación y la etapa de recombinación, o la etapa de contacto, se realiza *ex vivo*. En una realización, la etapa de formación y la etapa de recombinación, o la etapa de contacto, se realiza *in vivo*.

60 [0050] En una realización, el sujeto receptor es acondicionado antes de la administración de las células editadas al sujeto. El condicionamiento puede implicar la linfodepleción del sujeto. Por ejemplo, en una realización, el sujeto se acondiciona con quimioterapia que causa linfodepleción, antes de recibir las células editadas. En una realización, las células T del sujeto se agotan mediante acondicionamiento. Los agentes adecuados para acondicionar al sujeto incluyen agentes que inducen linfodepleción, por ejemplo, depleción de células T. En una realización, el agente acondicionador es la ciclofosfamida. En algunas realizaciones, el método de tratamiento de un sujeto que tiene síndrome de hiper-IgM alterando un CD40L diana en una célula puede comprender opcionalmente una etapa de administración de un agente acondicionador al sujeto, *por ejemplo*, un agente acondicionador que causa linfodepleción. En una realización, el agente acondicionador es un agente quimioterapéutico, *por ejemplo* ciclofosfamida, fludarabina, busulfán, treosulfán, o una combinación de los mismos. En una realización, el agente acondicionador es un anticuerpo depletor policlonal o monoclonal, *por ejemplo* globulina antitimocítica (ATG), anti-CD3, anti-CD4, anti-CD52, anti-CD2, anti-TCR $\alpha\beta$, anti-IL2Ra, o una combinación de los mismos. En otra realización, el sujeto no está condicionado antes de la administración de las células editadas al sujeto.

[0051] Las células pueden estimularse antes de su administración al sujeto. Las células pueden estimularse utilizando citoquinas, *por ejemplo*, IL-7, IL-15, IL-2, o una combinación de las mismas.

5 [0052] En una realización, se restablece el cambio de clase en el sujeto. En una realización, los niveles de IgM disminuyen en un sujeto. En una realización, los niveles de IgG aumentan en un sujeto.

10 [0053] En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 10 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 % o 75 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 15 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 20 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 30 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 40 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 50 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 60 % de las células de la población de células. En una realización, la célula es una población de células, y la vía de reparación del ADN repara el ácido nucleico CD40L diana para dar lugar a la integración diana de la plantilla donante de oligonucleótidos exógena en al menos aproximadamente el 75 % de las células de la población de células.

35 [0054] En una realización, las células son células madre hematopoyéticas (CMH), y en el que al menos aproximadamente el 10 % de las células de la población de CMH comprenden un ácido nucleico CD40L alterado que es un alelo CD40L funcional. En una realización, las células son células T.

40 [0055] En un aspecto, en el presente documento se divulga una población de células, caracterizada porque al menos el 10 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada por ingeniería. En una realización, al menos el 15 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada. En una realización, al menos el 20 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada. En una realización, al menos el 30 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada. En una realización, al menos el 40 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada. En una realización, al menos el 50 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada. En una realización, al menos el 60 % de las células de la población comprenden al menos una copia de una secuencia CD40L modificada.

50 [0056] En una realización, la al menos una copia se inserta en un intrón de un gen CD40L endógeno. En una realización, la población de células, o la progenie de la población de células, son capaces de restaurar el cambio de clase en un sujeto que padece el síndrome de Hiper IgM.

55 [0057] En un aspecto, se divulga en el presente documento una población aislada de células, en la que al menos entre aproximadamente el 5% y aproximadamente el 100 % de las células de la población de células expresan un alelo CD40L funcional, en la que la población de células se ha aislado de un sujeto, y en la que al sujeto se le ha diagnosticado previamente que tiene síndrome de hiper IgM. En una realización, al menos alrededor del 6 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 7 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos aproximadamente el 8 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 9 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 10 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 15 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 20 % de las células de la población de células expresan un alelo CD40L funcional. En una realización, al menos alrededor del 25 % de las células de la población de células expresan un alelo CD40L funcional.

- [0058] Los dibujos adjuntos tienen por objeto proporcionar ejemplos ilustrativos y esquemáticos, más que exhaustivos, de determinados aspectos y realizaciones de la presente divulgación. Los dibujos no se limitan a ninguna teoría o modelo en particular ni son necesariamente a escala. Sin limitar lo anterior, los ácidos nucleicos y los polipéptidos pueden representarse como secuencias lineales, o como estructuras esquemáticas bidimensionales o tridimensionales; estas representaciones pretenden ser ilustrativas más que limitativas o unidas a cualquier modelo o teoría particular con respecto a su estructura.
- 5 Las Figuras. 1A-D muestran la configuración de plantillas donantes CD40L ejemplares para la edición de genes. La Figura. 1A muestra un constructo que comprende brazos de homología, ADNc de CD40L, GFP y secuencias Stuffer. La Figura. 1B muestra un constructo que comprende brazos de homología, ADNc de CD40L y GFP, sin secuencias Stuffer. La Figura. 1C muestra un constructo que comprende brazos de homología, ADNc de CD40L y secuencias Stuffer, sin GFP. La Figura. 1D muestra un constructo que comprende brazos de homología y ADNc de CD40L, sin GFP y sin secuencias Stuffer.
- 10
- 15 Las **Figuras. 2A-2J** representan que los sistemas de edición del genoma descritos en el presente documento dieron lugar a altos niveles de edición del gen CD40LG. En concreto, la Figura. 2A es un esquema de constructos ejemplares para la edición del gen CD40L. La Figura. 2B muestra la posición en el intrón 1 de CD40LG dirigida por 8 S. Pyogenes y 30 ARNg de S. aureus para comprobar la eficacia de corte de CD40LG en células T humanas primarias. La Figura. 2C muestra los porcentajes de edición, incluidas las tasas de inserción/deleción de ARNg ejemplares de la Figura. 2B administrados con nucleasas como un complejo RNP a células T humanas primarias. La Figura. 2D muestra el procedimiento experimental de edición del gen CD40L en linfocitos T. La Figura. 2E muestra el porcentaje de células GFP positivas en poblaciones de células T tras la edición génica. Las células madre T de memoria (TSCM) se definen como CD62L+CD45RA+, las células T de memoria central (TCM) como CD62L+ CD45RA-, las células T efectoras de memoria (TEM) como CD62L- CD45RA- y las células efectoras terminales (TEMRA) como CD62L-CD45RA+. La Figura. 2F muestra gráficos representativos de células T 7 días (panel izquierdo) y 20 días (panel derecho) después del tratamiento que expresan CD45RA y CD62 mediante análisis FACS. La Figura. 2G muestra la composición del cultivo de células T 7 y 19 días (panel izquierdo) y 20 días (panel derecho) después del tratamiento, medida mediante análisis FACS. La Figura. 2H es un protocolo esquemático para la edición génica en células T utilizando una plantilla donante proporcionada por AAV6 o IDLV. La Figura. 2I muestra el resultado de la HDR (evaluado como el porcentaje de células GFP +) en varias subpoblaciones de células T. La Figura. 2J muestra los niveles de edición génica en células de EH y de pacientes.
- 20
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- 65
- Las **figuras 3A-3E** muestran la integración diana y la expresión de CD40L tras la edición con tres ARNg ejemplares (ARNg A, D y G). La Figura. 3A muestra la localización de ARNg ejemplares dentro del intrón 1 de CD40LG. La Figura. 3B muestra los procedimientos experimentales y la cronología de los experimentos presentados en las Figuras. 3D-3E. La Figura. 3C muestra la tasa de integración diana (% de células GFP+) tras el tratamiento con los complejos RNP indicados y la plantilla donante AAV6. La Figura. 3D muestra la expresión de CD40L en la población de células GFP-positivas, determinada por análisis FACS. La Figura. 3E muestra la fracción de células CD40L+ en distintos momentos tras la reestimulación.
- Las **figuras 4A-4F** muestran que el CD40L editado conserva la regulación fisiológica. En concreto, las Figuras. 4A y 4B muestran los constructos y el procedimiento experimental utilizados para la edición y expresión del gen CD40L en linfocitos T. Las Figuras. 4C y 4E representan la cinética de expresión de CD40L en células T editadas tras la inducción de la expresión de superficie de CD40L mediante una estimulación policonal (PMA/Ionomycin). La Figura. 4D muestra un gráfico representativo de la expresión de CD40L en células CD4+ y CD8+ activadas. La Figura. 4F muestra un gráfico representativo de la expresión de CD40L en células CD4+ estimuladas.
- Las **figuras 5A-5E** muestran que el CD40L editado conserva la regulación fisiológica. En concreto, las Figuras. 5A y 5B muestran los constructos y el procedimiento experimental utilizados para la edición y expresión del gen CD40L en linfocitos T. La Figura. 5C muestra la expresión de CD40L y CD8 en células T de tipo salvaje y editadas. La Figura. 5D muestra la cinética de expresión de CD40L en células T editadas. La Figura. 5E muestra el efecto del AAV6 en la eficacia de la edición génica.
- Figuras. 6A y 6B** demuestran que la edición génica produce altos niveles de edición génica de CD40L (Figura. 6A) y altos niveles de expresión exógena de CD40L, utilizando diferentes configuraciones de transgenes (ADNc de tipo salvaje (no recodificado) o secuencia optimizada de codones (recodificada); señal de poliadenilación endógena o señal de poliadenilación viral SV40 (Figura. 6B)).
- Las **Figuras. 7A-7F** representan los resultados de experimentos que evalúan la especificidad de ARNg ejemplares que dirigen CD40LG. La Figura. 7A presenta el análisis Guide-Seq de los ARNg A y ARNg D. Las Figuras. 7B y 7E validan la edición on-target y off-target observada con el ARNg A mediante el análisis

Amplicon-Seq. Las Figuras. 7C-7D presentan los resultados del análisis Digenome para identificar posibles sitios fuera de diana para el ARNg A y el ARNg D. La Figura. 7F muestra una visión general de los resultados del análisis Amplicon-Seq para los sitios de corte fuera de diana identificados para el ARNg A y D.

Figura. 8 representa que las células T editadas con CD40L restauraron la capacidad de cambio de clase de células B en un ensayo de co-cultivo *in vitro*. El histograma muestra el porcentaje de linfocitos B que expresaron IgG en su superficie tras 5 días de cocultivo con linfocitos T de tipo salvaje (No editados), linfocitos T con el gen CD40L editado (GFP+), linfocitos T tratados para la edición génica pero negativos para la integración del transgén corrector (GFP-) o linfocitos T en los que se anuló el gen CD40L endógeno mediante la integración de un casete reportero. Los cultivos con sólo células B y con células B más la adición de proteína CD40L soluble se utilizaron como controles negativo y positivo, respectivamente. Todas las células T se probaron sin ninguna estimulación (en reposo) o tras la estimulación con perlas antiCD3-CD28 (perlas) o PMA/Ionomicina.

Figura. 9 representa un nivel disminuido de IgM y un nivel aumentado de IgG en el medio de células B co-cultivadas con células T que tienen un emplazamiento CD40L editado.

La Fig. 10 representa el nivel de línea de base de la integración diana en cuatro sitios de corte de ARNg utilizando una plantilla donante AAV6 que comprende la siguiente carga: secuencias de relleno, ADNc de CD40LG/3' UTR/polyA, y un casete de expresión de GFP. Cortos brazos homológicos flanqueaban los sitios de corte previstos. Se detectaron diferencias en los niveles de base de integración diana entre los ARNg, con base en la expresión de GFP.

Figura. 11 compara el nivel de integración diana obtenido utilizando plantillas donantes que tienen una longitud de brazo homólogo variada. Las plantillas donantes AAV6 se prepararon como en la Figura. 10, alargando los brazos de homología a 500 pares de bases a cada lado de la secuencia de carga. El aumento de la longitud de los brazos homológicos incrementó el nivel de integración diana en los dos puntos de corte de ARNg probados.

Figuras. 12A y 12B representan el fenotipo de las células T trasplantadas en diferentes momentos tras la inyección. En la Figura. 12A se muestran los porcentajes de células CD8 y CD4 medidos dentro de la población circulante de células T de tipo salvaje. En la Figura. 12B se muestran los porcentajes de células T efectoras (CD44+ CD62L-), de memoria central (MTC, CD44+CD62L+), naive (CD44-CD62L+) y efectoras tempranas (CD44-CD62L-), medidos dentro de las fracciones CD8 y CD4 de células T circulantes de tipo salvaje.

Figura. 13 representa la terapia de células T adoptivas en ratones CD40L KO utilizando células T de tipo salvaje, trasplantadas con y sin un régimen de preacondicionamiento con ciclofosfamida. Los gráficos de la izquierda muestran los porcentajes de células T receptoras y donantes a lo largo del tiempo en los ratones tratados o no con el régimen de preacondicionamiento. El gráfico de puntos de la derecha representa la cantidad de IgG específica del antígeno en el suero de los animales trasplantados antes o después de 14 y 28 días de una inmunización de los ratones trasplantados con el hapteno trinitrofenilo conjugado con el antígeno hemocianina Keyhole Limpet (TNP-KLH). Los ratones HIGM1 se utilizaron como control negativo. Como control positivo se utilizaron ratones trasplantados con una dosis completa de células madre hematopoyéticas de tipo salvaje (HSCT).

Figura. 14A y 14B representan la respuesta a la vacunación de ratones HIGM1 trasplantados con células T de tipo salvaje. La Figura. 14A muestra la cantidad de respuesta IgG específica al antígeno antes o después de 2 refuerzos de inmunización con el antígeno TNP-KLH en el suero de los ratones HIGM1 trasplantados con o sin diferentes dosis de acondicionamiento. La Figura. 14B muestra el gráfico de correlación entre los niveles séricos de IgG específica de TNP/KLH y las cantidades de linfocitos T CD4 de tipo salvaje circulantes (porcentajes en la parte superior y números absolutos en la parte inferior).

Figura. 14C-14E representan la terapia de células T adoptivas en ratones CD40L KO utilizando células T de tipo salvaje, trasplantadas con y sin régimen de preacondicionamiento con antileucocitos (ALS) o anticuerpos anti-CD4. Las Figura. 14C y 14E representan los porcentajes de células T receptoras y donantes a lo largo del tiempo en los ratones tratados o no con el régimen de preacondicionamiento. Los niveles de respuesta IgG específica a la vacunación se correlacionan con la dosis de acondicionamiento administrada (Figura. 14D).

Las **Figuras. 15A-15B** muestran el análisis inmunohistoquímico de células B GC en secciones de bazo de ratones trasplantados adoptivamente de la Figura. 14 tras la inmunización con TNP-KLH. La Figura. 15A muestra el análisis de tinción inmunohistoquímica de células B GC en secciones de bazo con aglutinina de cacahuete en marrón y contratinación de hematoxilina en azul. La Figura. 15B representa un histograma, cuantificando el análisis inmunohistoquímico representado en la Figura. 15A.

- 5 Las Figuras. 16A-16C representan la terapia con células madre hematopoyéticas en modelos murinos de síndrome de hiper-IgM. Como se muestra en la Figura. 16A, diferentes dosis de células madre hematopoyéticas CD40LG de tipo salvaje, recolectadas de ratones C57BL/6 de tipo salvaje, se trasplantaron en diferentes relaciones junto con HSPC CD40LG negativas, recolectadas de ratones HIGM1, en ratones HIGM1 receptores antes de la inmunización con un antígeno dependiente del timo (TNP-KLH). El gráfico de puntos representa la cantidad de IgG específica de TNP-KLH en el suero de los animales transplantados antes o después de 14 y 28 días de la inmunización de los ratones transplantados con el antígeno TNP-KLH. La Figura. 16B representa un segundo experimento en el que ratones CD40L transplantados con diferentes porcentajes de HSPC de tipo salvaje y HSPC CD40LG negativas como se describe para la Figura. 16A. Los ratones fueron inmunizados con TNP-KLH, y posteriormente fueron inmunizados con ovoalbúmina. El gráfico de puntos representa la cantidad de IgG específica de TNP (panel izquierdo) y la cantidad de IgG específica de OVA (panel derecho) antes y después de la inmunización. La Figura. 16C representa el injerto de células T del donante en sangre periférica, medido en diferentes momentos tras el trasplante.
- 10 20 Las Figuras. 17A-17D muestran una diana CD40L alta y policlonal tanto en células T derivadas de donantes sanos (HD) como de pacientes. Se recogieron células T CD4 de diferentes donantes sanos (HD, n=4) o de un paciente portador de la mutación causante del HIGM1 indicada en el exón 3 (Paciente), se estimularon in vitro y se trataron con 25 pmol de RNP A (S.p.) y AAV6 a MOI: 5*10^4 vg/célula. La Figura. 17A muestra el porcentaje de células editadas con GFP+ medido en diferentes subpoblaciones de células T 17 días después del tratamiento. La Figura. 17B muestra la composición del cultivo medida mediante análisis citofluorimétrico (marcado GFP). La Figura .17C compara la eficiencia de la edición génica medida por análisis molecular ddPCR (barras de la izquierda) y por análisis citofluorimétrico (GFP por FACS, barras de la derecha). La Figura. 17D muestra la diversidad de TCR en células de pacientes no tratadas (paneles grises/derechos) y en células de pacientes editadas (paneles negros/izquierdos), evaluada mediante 5 reacciones PCR multiplex diferentes que discriminan 21 familias de TCR-B.
- 15 25 30 35 Las Figs. 18A y 18B muestran que la integración diana del constructo corrector restauró la expresión de CD40L y su regulación fisiológica. La Figura. 18A muestra gráficos de puntos de citometría de flujo representativos que muestran la expresión de CD40L y GFP 8 horas después de la estimulación con PMA/Ionomicina (P/I) en células de EH y de pacientes de la Figura. 17. La Figura. 18B muestra el curso temporal de la expresión de CD40L en superficie, como Intensidad de Fluorescencia Relativa (RFI, calculada con la relación en T0) y como porcentaje, medido a las 0, 3, 8, 24 y 48 horas tras la activación P/I.
- 40 45 Las Figuras. 19A-19C representan la capacidad de las células T editadas con CD40L para restaurar la proliferación de células B en un ensayo de co-cultivo. La Figura. 19A muestra una representación esquemática del protocolo de ensayos de co-cultivo de células T y B. La Figura. 19B y la Figura. 19C representan gráficos e histogramas de citometría de flujo, respectivamente, que muestran el porcentaje de células B prolíferantes en co-cultivo con células T en reposo, activadas con microesferas (relación 1:1) o estimuladas con PMA/Ionomicina (no tratadas (UT) o editadas (GFP+)). Control negativo: Células B solas; control positivo: Célula B + CD40L soluble.
- 50 55 60 Las Figuras. 20A y 20B representan la capacidad de las células T editadas con CD40L para restaurar el cambio de clase de las células B en un ensayo de co-cultivo. La Figura. 20A y la Figura. 20B muestran imágenes Elispot e histogramas, respectivamente, mostrando el porcentaje de células B con IgG switched en co-cultivo con células T en reposo, activadas con microesferas (relación 1:1) o estimuladas con PMA/Ionomicina (no tratadas (UT) o editadas (GFP+)). Control negativo: Células B solas; control positivo: Célula B + CD40L soluble.
- 65 Las Figuras. 21A-21F representan que los sistemas de edición del genoma descritos en el presente documento dieron lugar a altos niveles de edición del gen CD40LG en células madre/progenitoras hematopoyéticas humanas (HSPCs). La Figura. 21A es un esquema del constructo donante utilizado para editar células CD34+. Las Figura. 21B, Figura. 21C y Figura. 21E muestran el porcentaje de integración diana en células CD34 tratadas a granel y en las subpoblaciones seleccionadas indicadas, respectivamente, medido mediante análisis PCR digital de gotitas realizado en la unión 5' del vector con el genoma. La Figura. 21D representa la composición del cultivo de las células tratadas y no tratadas (UT). La Figura. 21F muestra el porcentaje de células hCD45+ (panel izquierdo) y el porcentaje de células HDR (panel derecho) dentro del injerto, medidos en diferentes momentos tras el trasplante.
- 70 Las Figuras. 22A-22F representan una estrategia de selección desarrollada para enriquecer células T editadas con CD40LG. La Figura. 22A representa el constructo de ADN donante que lleva el gen CD40LG corrector, una secuencia de sitio de entrada de ribosoma interno (IRES) y un gen selector (*por ejemplo* delta LNGFR). La Figura. 22B es una representación esquemática del protocolo de integración diana que indica los diferentes puntos temporales en los que se evaluó la expresión del gen selector mediante

análisis citofluorimétricos. La Figura. 22C representa la expresión del gen informador en los puntos temporales indicados (6 días, 13 días y 20 días) en células T CD4 y CD8 tratadas como se describe en la Figura. 22B. La Figura. 22D representa la cinética de la expresión de CD40L tras la reestimulación de las células T tratadas para la edición génica. La Figura. 22E representa el nivel de expresión (intensidad fluorescente media, MFI) del gen CD40L (izquierda) y el porcentaje de células CD40L+ en la población de células T no tratadas, NGFR+ y NGFR-. La Figura. 22F muestra la capacidad de las células T editadas con CD40L para restaurar el cambio de clase de las células B en un ensayo de co-cultivo, mostrando el porcentaje de células B con intercambio de IgG en co-cultivo con células T en reposo, activadas con microesferas (relación 1:1) o estimuladas con PMA/Ionomicina (no tratadas (UT) o editadas (NGFR+)). Control negativo: Células B solas; control positivo: Célula B + CD40L soluble.

Las **Figuras. 23A-23C** representan la selección de células editadas con CD40LG utilizando un gen de selección NGFR. La Figura. 23A representa el porcentaje de células NGFR+ obtenidas tras la edición génica, medido dentro de diferentes subpoblaciones de células T (de izquierda a derecha: efectores terminales (TEMRA), CD45RA+ CD62L-; memoria de efectores (TEM), CD62L- CD45RA-; memoria central (TCM), CD45RA-CD62L+; y células T madre de memoria (TSCM), CD45RA+CD62L+). La Figura. 23B representa la fracción de subpoblaciones de células T medida antes o después de la selección NGFR. La Figura. 23C representa el porcentaje de edición del gen CD40LG medido por análisis molecular (ddPCR) o por expresión de NGFR, antes o después de la selección de NGFR.

Las **Figuras. 24A-24C** representan el injerto de células editadas con CD40LG tras el xenotrasplante en ratones NSG. La Figura. 24A muestra el porcentaje de células humanas (células hCD45+; paneles superior izquierdo e inferior izquierdo y derecho) y el porcentaje de células NGFR+ (panel superior derecho) dentro del injerto humano, medidos en diferentes momentos tras el trasplante. La Figura. 24B representa la fracción de subpoblaciones de células T (efectores terminales (TEMRA), CD45RA+ CD62L-; memoria de efectores (TEM), CD62L- CD45RA-; memoria central (TCM), CD45RA-CD62L+; y células T madre de memoria (TSCM), CD45RA+CD62L+) de las células T injertadas medidas a las 4, 8 y 16 semanas después del trasplante. La Figura. 24C muestra los niveles de expresión de CD40LG y el porcentaje de células CD40LG+ recuperadas de los ratones NSG.

Las **Figuras. 25A-25D** representan experimentos que demuestran que la edición de HSPC es portátil a la plataforma CRISPR/Cas. La Figura. 25A presenta un protocolo para editar células CD34+ obtenidas de sangre de cordón umbilical. La Figura. 25B muestra la eficiencia de integración diana (expresada como el porcentaje de células GFP+) en varias poblaciones celulares obtenidas mediante edición génica utilizando ZFN o RNP. La Figura. 25C muestra la composición del cultivo tras la edición génica mediante ZFN o RNP, en comparación con células no tratadas (UT). La Figura. 25D muestra el pliegue de inducción de los genes sensibles al interferón (IFN) RIG1, IRF7 y OAS1, tras el tratamiento con RNP, ZFN, Cas9 solo o RNP-mod, en comparación con las células no tratadas (UT).

Las **Figuras. 26A-26D** representan experimentos adicionales que demuestran que la edición de HSPC es portátil a la plataforma CRISPR/Cas. La Figura. 26A presenta un protocolo para editar células CD34+ de sangre de cordón umbilical (CB), médula ósea (BM) o sangre periférica movilizada (MPB). La Figura. 26B muestra la eficiencia de integración diana (expresada como el porcentaje de células GFP+) en varias poblaciones celulares obtenidas utilizando los procedimientos de edición indicados. La Figura. 26C muestra la composición del cultivo después de la edición. La Figura. 26D muestra la expresión de los genes RIG1, IRF7 y OAS1 que responden al IFN en las células editadas.

Figuras. 27A y 27B representan datos que indican que las HSPC humanas de diversas fuentes pueden editarse eficientemente utilizando una variedad de vehículos donantes y configuraciones de nucleasas. La Figura. 27A muestra el porcentaje de subpoblaciones celulares editadas (GFP+) utilizando ZFN o RNP junto con plantillas donantes IDLV o AAV6. La Figura. 27B representa el porcentaje de subpoblaciones celulares editadas (GFP+) utilizando células CD34+ derivadas de sangre de cordón umbilical, médula ósea y sangre periférica movilizada (mPB).

55 DESCRIPCIÓN DETALLADA

Definiciones y abreviaturas

[0059] Salvo que se especifique lo contrario, cada uno de los siguientes términos tiene el significado que se le atribuye en esta sección.

[0060] Los artículos indefinidos "un" y "uno(a)" se refieren a al menos uno de los sustantivos asociados, y se utilizan indistintamente con los términos "al menos uno" y "uno o más". Por ejemplo, "un módulo" significa al menos un módulo, o uno o más módulos.

[0061] Las conjunciones "o" y "y/o" se utilizan indistintamente como disyunciones no excluyentes.

- 5 [0062] "Alt-HDR", "reparación alternativa dirigida por homología" o "HDR alternativa" se utilizan indistintamente para referirse al procedimiento de reparación de daños en el ADN utilizando un ácido nucleico homólogo (*por ejemplo*, una secuencia homóloga endógena, *por ejemplo*, una cromátida hermana, o un ácido nucleico exógeno, *por ejemplo*, un ácido nucleico molde). La Alt-HDR se distingue de la HDR canónica en que el procedimiento utiliza vías diferentes de la HDR canónica, y puede ser inhibida por los mediadores de la HDR canónica, RAD51 y BRCA2. La Alt-HDR también se distingue por la implicación de una plantilla de ácido nucleico homólogo monocatenaria o mellada, mientras que la HDR canónica suele implicar una plantilla homóloga bicatenaria.
- 10 [0063] "HDR canónica", "reparación canónica dirigida por homología" o "HDRc" se refiere al procedimiento de reparación de daños en el ADN utilizando un ácido nucleico homólogo (*por ejemplo*, una secuencia homóloga endógena, *por ejemplo*, una cromátida hermana, o un ácido nucleico exógeno, *por ejemplo*, un ácido nucleico molde). La HDR canónica suele actuar cuando se ha producido una resección significativa en la rotura de la doble cadena, formando al menos una porción de ADN de cadena sencilla. En una célula normal, la HDRc suele implicar una serie de etapas tales como el reconocimiento de la rotura, la estabilización de la rotura, la resección, la estabilización del ADN monocatenario, la formación de un producto intermedio de entrecruzamiento del ADN, la resolución del producto intermedio de entrecruzamiento y la ligadura. El procedimiento requiere RAD51 y BRCA2, y el ácido nucleico homólogo suele ser de doble cadena.
- 15 [0064] A menos que se indique lo contrario, el término "HDR" utilizado en el presente documento engloba tanto el HDR canónico como el alt-HDR.
- 20 [0065] En el presente documento, el término "evento de escisión" se refiere a la rotura de una molécula de ácido nucleico. Un evento de escisión puede ser de una sola hebra o de dos hebras. Un evento de escisión de una sola hebra puede dar lugar a un saliente 5' o a un saliente 3'. Un evento de escisión de doble cadena puede dar lugar a extremos romos, dos salientes 5' o dos salientes 3'.
- 25 [0066] El término "sitio de escisión", tal como se utiliza en el presente documento en referencia a un sitio en una secuencia de ácido nucleico diana, por ejemplo, una secuencia de ácido nucleico CD40L diana, se refiere a una posición diana entre dos residuos de nucleótidos del ácido nucleico diana en la que se produce una rotura de doble cadena, o alternativamente, a una posición diana dentro de un intervalo de varios residuos de nucleótidos del ácido nucleico diana en la que se producen dos roturas de cadena sencilla, según lo mediado por un procedimiento dependiente de nucleasas guiado por ARN. Un sitio de escisión puede ser la posición diana para, *por ejemplo*, una rotura roma de doble cadena. Alternativamente, un sitio de escisión puede ser un sitio diana dentro de un intervalo de varios residuos nucleotídicos del ácido nucleico diana para, *por ejemplo*, dos roturas de cadena sencilla o muescas que forman una rotura de cadena doble y que están separadas por, *por ejemplo*, unos 10 pares de bases. La(s) rotura(s) de doble cadena o la más cercana de las dos melladas de cadena simple en un par estará idealmente dentro de 0-500 pb de una posición diana (*por ejemplo*, a no más de 450, 400, 350, 300, 250, 200, 150, 100, 50, o 25 pb de la posición diana). Cuando se utilizan nicasas duales, los dos nicks de un par se encuentran a una distancia de 25 a 55 pb uno del otro (*por ejemplo*, entre 25 y 50, 25 y 45, 25 y 40, 25 y 35, 25 y 30, 50 y 55, 45 y 55, 40 y 55, 35 y 55, 30 y 55, 30 y 50, 35 y 50, 40 y 50, 45 y 50, 35 y 45, o 40 y 45 pb) y a no más de 100 pb de distancia entre sí (*por ejemplo*, no más de 90, 80, 70, 60, 50, 40, 30, 20, o 10 pb).
- 30 [0067] La expresión "que consiste esencialmente en" significa que las especies mencionadas son las predominantes, pero que otras especies pueden estar presentes en cantidades mínimas o en cantidades que no afecten a la estructura, la función o el comportamiento de la composición en cuestión. Por ejemplo, una composición que consiste esencialmente en una especie particular comprenderá generalmente 90 %, 95 %, 96 %, o más de esa especie.
- 35 [0068] "Dominio" se utiliza para describir un segmento de una proteína o ácido nucleico. A menos que se indique lo contrario, no es necesario que un dominio tenga ninguna propiedad funcional específica.
- 40 [0069] Tal y como se utiliza en el presente documento, el término gen "endógeno", ácido nucleico "endógeno" o región homóloga "endógena" se refiere a un gen, ácido nucleico o región de un gen nativo, que se encuentra en su ubicación natural en el genoma, por ejemplo, cromosoma o plásmido, de una célula. Por el contrario, el término gen "exógeno" o ácido nucleico "exógeno" se refiere a un gen, ácido nucleico o región de un gen que no es nativo dentro de una célula, sino que se introduce en ella durante los métodos de la invención. Un gen exógeno o un ácido nucleico exógeno puede ser homólogo o idéntico a un gen endógeno o a un ácido nucleico endógeno.
- 45 [0070] El término "exón" se refiere a una región codificante de un gen que contiene la información necesaria para codificar una proteína. En los eucariotas, los genes están formados por exones codificantes intercalados con intrones no codificantes. En el empalme del ARN, los intrones se eliminan y los exones se unen covalentemente entre sí para formar un ARN mensajero (ARNm) maduro que pueda traducirse en una proteína. El término exón se refiere tanto a la secuencia de ADN dentro de un gen como a la secuencia correspondiente en los transcritos de ARN.

5 [0071] "Conversión génica" se refiere a la alteración de una secuencia de ADN mediante la incorporación de una secuencia homóloga endógena (*por ejemplo*, una secuencia homóloga dentro de un conjunto de genes). "Corrección génica" se refiere a la alteración de una secuencia de ADN mediante la incorporación de una secuencia homóloga exógena, tal como un ADN molde exógeno donante de cadena sencilla o doble. La conversión génica y la corrección génica son productos de la reparación de roturas de doble cadena del ADN por vías de HDR tales como las que se describen a continuación.

10 15 [0072] Los términos "homología" o "identidad", utilizados indistintamente en el presente documento, se refieren a la identidad de secuencia entre dos secuencias de aminoácidos o dos secuencias de ácidos nucleicos, siendo la identidad una comparación más estricta. Las frases "porcentaje de identidad u homología" y "% de identidad u homología" se refieren al porcentaje de identidad de secuencia encontrado en una comparación de dos o más secuencias de aminoácidos o secuencias de ácidos nucleicos. Dos o más secuencias pueden ser entre 0 y 100 % idénticas, o cualquier valor intermedio. La identidad puede determinarse comparando una posición de cada secuencia que puede alinearse con fines de comparación con una secuencia de referencia. Cuando una posición de la secuencia comparada está ocupada por la misma base nucleotídica o aminoácido, las moléculas son idénticas en esa posición. El grado de identidad de las secuencias de aminoácidos es una función del número de aminoácidos idénticos en posiciones compartidas por las secuencias de aminoácidos. Un grado de identidad entre secuencias de ácidos nucleicos es una función del número de nucleótidos idénticos o coincidentes en posiciones compartidas por las secuencias de ácidos nucleicos. El grado de homología de las secuencias de aminoácidos es una función del número de aminoácidos en posiciones compartidas por las secuencias polipeptídicas.

20 25 30 [0073] Los cálculos de homología o identidad de secuencia entre dos secuencias (los términos se utilizan indistintamente en el presente documento) se realizan del siguiente modo. Las secuencias se alinean con fines de comparación óptima (*por ejemplo*, pueden introducirse brechas en una o ambas de una primera y una segunda secuencia de aminoácidos o ácidos nucleicos para una alineación óptima y pueden despreciarse las secuencias no homólogas con fines de comparación). La alineación óptima se determina como la mejor puntuación utilizando el programa GAP en el paquete de software GCG con una matriz de puntuación Blossum 62 con una penalización de brecha de 12, una penalización de extensión de brecha de 4 y una penalización de brecha de desplazamiento de marco de 5. A continuación, se comparan los residuos de aminoácidos o nucleótidos en las posiciones de aminoácidos o nucleótidos correspondientes. Cuando una posición de la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente de la segunda secuencia, las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias.

35 [0074] El término "brazo homológico", tal como se utiliza en el presente documento, se refiere a una o más regiones que son homólogas a regiones de ADN, *por ejemplo*, un ácido nucleico diana, dentro o cerca (*por ejemplo*, flanqueando o contiguas) de una secuencia diana a escindir, *por ejemplo*, un sitio de escisión.

40 [0075] Un "indel" es una inserción y/o delección en una secuencia de ácido nucleico. Un indel puede ser el producto de la reparación de una rotura de doble cadena de ADN, tal como una rotura de doble cadena formada por un sistema de edición genómica de la presente divulgación. Un indel se forma con mayor frecuencia cuando una rotura es reparada por una vía de reparación "propensa a errores" tal como la vía NHEJ descrita más adelante.

45 50 55 [0076] Los indels, la conversión de genes, la corrección de genes y otros resultados de la edición del genoma se evalúan normalmente mediante secuenciación (más comúnmente por métodos "next-gen" o "secuenciación por síntesis", aunque también puede utilizarse la secuenciación Sanger) y se cuantifican por la frecuencia relativa de cambios numéricos (*por ejemplo*, ±1, ±2 o más bases) en un sitio de interés entre todas las lecturas de secuenciación. Las muestras de ADN para secuenciación pueden prepararse mediante diversos métodos conocidos en la técnica, y pueden implicar la amplificación de sitios de interés mediante reacción en cadena de la polimerasa (PCR), la captura de extremos de ADN generados por roturas de doble cadena, como en el procedimiento GUIDEseq descrito en Tsai et al. (Nat. Biotechnol. 34(5): 483 (2016)) o por otros medios bien conocidos en la técnica. Los resultados de la edición del genoma también pueden evaluarse mediante métodos de hibridación *in situ*, tal como el sistema FiberComb™ comercializado por Genomic Vision (Bagueno, Francia), y mediante cualquier otro método adecuado conocido en la técnica.

60 [0077] El término "intrón" se refiere a un largo tramo de ADN no codificante que se encuentra entre los exones (o regiones codificantes) de un gen. Los intrones suelen comenzar y terminar con una serie específica de nucleótidos y se empalman antes de que la molécula de ARN se traduzca en una proteína. Estas secuencias actúan como límite entre intrones y exones y se conocen como sitios de empalme. El reconocimiento de la frontera entre el ADN codificante y no codificante es crucial para la creación de genes funcionales.

65 [0078] El término "gen aislado" o "ácido nucleico aislado" es un gen o ácido nucleico que, en virtud de su origen o fuente de derivación, no está asociado con componentes naturalmente asociados que lo acompañan en su estado nativo; está sustancialmente libre de otras proteínas de la misma especie; es expresado por una célula de una especie diferente; o no se da en la naturaleza. Así, un ácido nucleico sintetizado químicamente o sintetizado en un sistema celular distinto de la célula de la que procede naturalmente estará "aislado" de sus componentes

naturalmente asociados. Un ácido nucleico también puede quedar sustancialmente libre de componentes naturalmente asociados mediante aislamiento, utilizando técnicas de purificación de proteínas bien conocidas en la técnica.

- 5 [0079] En las secuencias de nucleótidos presentadas en el presente documento se utiliza la notación IUPAC convencional, como se muestra en la **Tabla 1**, que figura a continuación (véase también Cornish-Bowden A, Nucleic Acids Res. 1985 May 10; 13(9):3021-30). Cabe señalar, sin embargo, que "T" denota "Timina o Uracilo" en aquellos casos en los que una secuencia puede estar codificada por ADN o ARN, por ejemplo en dominios dirigidos por ARNg.

10 **Tabla 1: Notación IUPAC de ácidos nucleicos**

| Carácter | Base |
|----------|------------------|
| A | Adenina |
| T | Timina o Uracilo |
| G | Guanina |
| C | Citosina |
| U | Uracilo |
| K | G o T/U |
| M | A o C |
| R | A o G |
| Y | C o T/U |
| S | CorG |
| W | A o T/U |
| B | C, G o T/U |
| V | A, C o G |
| H | A, C o T/U |
| D | A, G o T/U |
| N | A, C, G o T/U |

- 15 [0080] Un "kit" se refiere a cualquier conjunto de dos o más componentes que juntos constituyen una unidad funcional que puede emplearse para un fin específico. A título ilustrativo (y no limitativo), un kit según la presente divulgación puede incluir un ARN guía complejado o capaz de complejarse con una nucleasa guiada por ARN, y acompañado de (*por ejemplo*, suspendido en, o suspendible en) un portador farmacéuticamente aceptable. El kit puede utilizarse para introducir el complejo en, por ejemplo, una célula o un sujeto, con el fin de provocar una alteración genómica deseada en dicha célula o sujeto. Los componentes de un paquete pueden ir empaquetados juntos o por separado. Los kits según la presente divulgación también incluyen opcionalmente instrucciones de uso (DFU) que describen el uso del kit *por ejemplo*, según un método de la presente divulgación. El DFU puede estar físicamente empaquetado con el kit, o puede ponerse a disposición de un usuario del kit, por ejemplo por medios electrónicos.

- 20 [0081] "Unión de extremos no homólogos" o "NHEJ" se refiere a la reparación mediada por ligadura y/o la reparación no mediada por placa, incluyendo la NHEJ canónica (cNHEJ) y la NHEJ alternativa (altNHEJ), que a su vez incluye la unión de extremos mediada por microhomología (MMEJ), el recocido de una sola hebra (SSA) y la unión de extremos mediada por microhomología dependiente de la síntesis (SD-MMEJ).

- 25 [0082] Los términos "poli(A)", "cola de poli(A)" y "poliadenilación" se refieren a la adición de una cola de poli(A) en un ARN mensajero. La cola poli-A es una larga cadena de nucleótidos de adenina que se añade a una molécula de ARN mensajero (ARNm) durante el procesamiento del ARN para aumentar la estabilidad de la molécula. Inmediatamente después de transcribirse un gen en una célula eucariota, la nueva molécula de ARN sufre varias modificaciones conocidas como procesamiento del ARN. Estas modificaciones alteran ambos extremos del transcripto primario de ARN para producir una molécula madura de ARN. El procesamiento del extremo 3' añade una cola poli-A a la molécula de ARN.

- 30 [0083] Los términos "polinucleótido", "secuencia de nucleótidos", "ácido nucleico", "molécula de ácido nucleico", "secuencia de ácido nucleico" y "oligonucleótido" se refieren a una serie de bases nucleotídicas (también

denominadas "nucleótidos") en el ADN y el ARN, y significan cualquier cadena de dos o más nucleótidos. Los polinucleótidos, secuencias de nucleótidos, ácidos nucleicos, etc. pueden ser mezclas químicas o derivados o versiones modificadas de los mismos, monocatenarios o bicatenarios. Pueden modificarse en la fracción de base, la fracción de azúcar o la espina dorsal de fosfato, por ejemplo, para mejorar la estabilidad de la molécula, sus parámetros de hibridación, etc. Una secuencia de nucleótidos suele contener información genética, que incluye, sin limitarse a ella, la información utilizada por la maquinaria celular para fabricar proteínas y enzimas. Estos términos incluyen el ADN genómico de cadena doble o simple, el ARN, cualquier polinucleótido sintético y manipulado genéticamente, y los polinucleótidos en sentido y antisentido. Estos términos también incluyen los ácidos nucleicos que contienen bases modificadas. Un ácido nucleico "CD40L", tal como se utiliza en el presente documento, se refiere a una serie de bases nucleotídicas que codifican una proteína CD40L.

[0084] "Prevenir", "prevenir" y "prevención" se refieren a la prevención de una enfermedad en un mamífero, *por ejemplo*, en un humano, incluyendo (a) evitar o impedir la enfermedad; (b) afectar a la predisposición hacia la enfermedad; o (c) prevenir o retrasar la aparición de al menos un síntoma de la enfermedad.

[0085] Los términos "proteína", "péptido" y "polipéptido" se utilizan indistintamente para referirse a una cadena secuencial de aminoácidos unidos entre sí mediante enlaces peptídicos. Los términos incluyen proteínas individuales, grupos o complejos de proteínas que se asocian entre sí, así como fragmentos o porciones, variantes, derivados y análogos de dichas proteínas. Las secuencias peptídicas se presentan en el presente documento utilizando la notación convencional, comenzando con el amino o N-terminal a la izquierda, y continuando con el carboxilo o C-terminal a la derecha. Pueden utilizarse abreviaturas estándar de una o tres letras.

[0086] "Sustitución" o "reemplazado", cuando se utiliza con referencia a una modificación de una molécula (*por ejemplo*, un ácido nucleico o una proteína), no requiere una limitación del procedimiento, sino que simplemente indica que la entidad de sustitución está presente.

[0087] El término "secuencia de relleno", tal como se utiliza en el presente documento, se refiere a una secuencia de ácido nucleico heteróloga o aleatoria que ha sido seleccionada para facilitar la integración dirigida de una plantilla donante en un sitio diana y la posterior amplificación de un amplicón que comprende la secuencia de relleno. La presencia de la secuencia de relleno también evita la integración de la plantilla donante en otro sitio. Las secuencias de relleno se describen con más detalle en el presente documento.

[0088] "Sujeto" significa un ser humano o un animal no humano. Un sujeto humano puede ser de cualquier edad (*por ejemplo*, un bebé, un niño, un adulto joven o un adulto), y puede padecer una enfermedad o necesitar la alteración de un gen. Alternativamente, el sujeto puede ser un animal, término que incluye, pero no se limita a, mamíferos, aves, peces, reptiles, anfibios y, más particularmente, primates no humanos, roedores (tales como ratones, ratas, hamsters, etc.), conejos, cobayas, perros, gatos, etc. En ciertas realizaciones de esta divulgación, el sujeto es ganado, por ejemplo, una vaca, un caballo, una oveja o una cabra. En determinadas realizaciones, el sujeto es un ave de corral.

[0089] En el presente documento, el término "sustancialmente" se refiere a la condición cualitativa de presentar una extensión o grado total o casi total de una característica o propiedad de interés. Un experto en las técnicas biológicas comprenderá que los fenómenos biológicos y químicos rara vez, si es que alguna vez lo hacen, llegan a completarse y/o proceden a completarse o alcanzan o evitan un resultado absoluto. Por lo tanto, el término "sustancialmente" se utiliza en el presente documento para captar la posible falta de exhaustividad inherente a muchos fenómenos biológicos y químicos.

[0090] El término "sustancialmente idéntico", tal como se utiliza en el presente documento en referencia a dos secuencias de ácido nucleico, se refiere, en algunas realizaciones, a una identidad de secuencia de al menos el 95 % entre al menos 20 nucleótidos contiguos de cada una de las dos secuencias de ácido nucleico. Por ejemplo, una primera secuencia de ácido nucleico es sustancialmente idéntica a una segunda secuencia de ácido nucleico cuando la primera secuencia de ácido nucleico tiene al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad en un tramo contiguo de al menos 20 nucleótidos de la segunda secuencia de ácido nucleico. En algunas realizaciones, una primera secuencia de ácido nucleico es sustancialmente idéntica a una segunda secuencia de ácido nucleico cuando la primera secuencia de ácido nucleico tiene al menos un 95 % de identidad en toda la longitud de la segunda secuencia de ácido nucleico. En algunas realizaciones, el término "sustancialmente idéntico" se utiliza en el contexto de los sitios de cebado, y se refiere a la capacidad de los sitios de cebado para soportar el cebado con el mismo cebador de PCR durante una reacción de amplificación. En algunas realizaciones, el término "sustancialmente idéntico" se utiliza para describir la relación entre el brazo homólogo de una plantilla donante y el brazo homólogo de un ácido nucleico diana, y se refiere a una identidad de ácido nucleico entre los brazos homólogos que permite la recombinación eficiente de la plantilla donante en el ácido nucleico diana con tolerancia a cierto grado de polimorfismo, por ejemplo, para eliminar secuencias PAM o protospacer en el emplazamiento recombinado.

[0091] "Ácido nucleico CD40L diana", tal como se utiliza en el presente documento, se refiere a un ácido nucleico CD40L presente en, *por ejemplo*, el cromosoma, que es diana de modificación mediante un procedimiento

dependiente de nucleasas guiado por ARN utilizando los sistemas de edición del genoma divulgados en el presente documento. Un ácido nucleico CD40L diana puede comprender, por ejemplo, una mutación. En algunas realizaciones, la mutación conduce a la pérdida del cambio de clase de células B debido a una señalización defectuosa de CD40.

[0092] "Posición diana", tal como se utiliza en el presente documento, se refiere a un sitio en un ácido nucleico diana (*por ejemplo*, el cromosoma) que es modificado por un procedimiento dependiente de nucleasas guiado por ARN. En una realización, una posición diana puede ser un sitio entre dos nucleótidos, por ejemplo, nucleótidos adyacentes, en el ácido nucleico diana. La posición diana puede comprender uno o más nucleótidos alterados, por ejemplo, corregidos, por un ácido nucleico molde. En una realización, la posición diana está dentro de una "secuencia diana" (*por ejemplo*, la secuencia a la que se une el ARNg). En una realización, una posición diana está antes o después de una secuencia diana (*por ejemplo*, la secuencia a la que se une el ARNg).

[0093] "Tratar", "tratamiento" y "tratamiento" significan el tratamiento de una enfermedad en un sujeto (*por ejemplo*, un sujeto humano), incluyendo uno o más de inhibir la enfermedad, es decir, detener o evitar su desarrollo o progresión; aliviar la enfermedad, es decir, provocar la regresión del estado de la enfermedad; aliviar uno o más síntomas de la enfermedad; y curar la enfermedad.

[0094] El término "región no traducida" o "UTR" se refiere a una de las dos secciones no codificantes de proteínas del ARNm. La región 5' no traducida (5' UTR), también conocida como secuencia líder o ARN líder, es la región de un ARNm que se encuentra directamente antes del codón de iniciación. La región 3' no traducida (3' UTR) se refiere a la sección del ARNm que sigue inmediatamente al codón de terminación de la traducción.

[0095] El término "variante" se refiere a una entidad tal como un polipéptido, polinucleótido o molécula pequeña que muestra una identidad estructural significativa con una entidad de referencia pero que difiere estructuralmente de la entidad de referencia en la presencia o nivel de una o más fracciones químicas en comparación con la entidad de referencia. En muchas realizaciones, una variante también difiere funcionalmente de su entidad de referencia. En general, el hecho de que una entidad concreta se considere adecuadamente una "variante" de una entidad de referencia se basa en su grado de identidad estructural con la entidad de referencia.

Visión general

[0096] En el presente documento se incluyen sistemas de edición del genoma, plantillas de donantes, moléculas de ARNg y métodos relacionados que permiten el tratamiento del síndrome de hiper IgM, un grupo de trastornos caracterizados por una señalización CD40 defectuosa. Las composiciones y métodos descritos en el presente documento se basan en el uso de plantillas donantes que comprenden uno o más de los exones 2-5 de CD40L para restaurar funcionalmente la señalización CD40 adecuada y la recombinación del interruptor de clase.

Sistemas de edición del genoma

[0097] El término "sistema de edición del genoma" se refiere a cualquier sistema que tiene actividad de edición del ADN guiada por ARN. Los sistemas de edición del genoma de la presente divulgación incluyen al menos dos componentes adaptados de los sistemas CRISPR naturales: un ARN guía (ARNg) y una nucleasa guiada por ARN. Estos dos componentes forman un complejo que es capaz de asociarse con una secuencia específica de ácido nucleico y editar el ADN en o alrededor de esa secuencia de ácido nucleico, por ejemplo haciendo una o más de una rotura de cadena simple (una SSB o nick), una rotura de doble cadena (una DSB) y/o una mutación puntual.

[0098] Los sistemas CRISPR naturales se organizan evolutivamente en dos clases y cinco tipos (Makarova et al. Nat Rev Microbiol. 2011 Jun; 9(6): 467-477 (Makarova)), y aunque los sistemas de edición genómica de la presente divulgación pueden adaptar componentes de cualquier tipo o clase de sistema CRISPR de origen natural, las realizaciones presentadas en el presente documento se adaptan generalmente a partir de sistemas CRISPR de clase 2, y de tipo II o V. Los sistemas de clase 2, que abarcan los tipos II y V, se caracterizan por proteínas nucleasas guiadas por ARN multidominio relativamente grandes (*por ejemplo*, Cas9 o Cpf1) y uno o más ARN guía (*por ejemplo*, un ARNcr y, opcionalmente, un ARNtracr) que forman complejos de ribonucleoproteínas (RNP) que se asocian con (*es decir*, diana) y escinden loci específicos complementarios a una secuencia diana (o espaciadora) del ARNcr. Los sistemas de edición del genoma según la presente divulgación dirigen y editan de forma similar secuencias celulares de ADN, pero difieren significativamente de los sistemas CRISPR que se dan en la naturaleza. Por ejemplo, los ARN guía unimoleculares descritos en el presente documento no se dan en la naturaleza, y tanto los ARN guía como las nucleasas guiadas por ARN según esta divulgación pueden incorporar cualquier número de modificaciones que no se dan en la naturaleza.

[0099] Los sistemas de edición del genoma pueden implementarse (*por ejemplo*, administrarse o suministrarse a una célula o a un sujeto) de diversas maneras, y las diferentes implementaciones pueden ser adecuadas para distintas aplicaciones. Por ejemplo, un sistema de edición del genoma se implementa, en ciertas realizaciones, como un complejo proteína/ARN (una ribonucleoproteína, o RNP), que puede incluirse en una composición farmacéutica que opcionalmente incluye un portador farmacéuticamente aceptable y/o un agente encapsulante, tal

como una micro o nanopartícula lipídica o polimérica, micela, liposoma, etc. En ciertas realizaciones, un sistema de edición de genoma se implementa como uno o más ácidos nucleicos que codifican los componentes de nucleasa guiada por ARN y ARN guía descritos anteriormente (opcionalmente con uno o más componentes adicionales); en ciertas realizaciones, el sistema de edición de genoma se implementa como uno o más vectores que comprenden tales ácidos nucleicos, por ejemplo un vector viral tal como un virus adeno-asociado; y en ciertas realizaciones, el sistema de edición de genoma se implementa como una combinación de cualquiera de los anteriores. Las implementaciones adicionales o modificadas que operan de acuerdo con los principios establecidos en el presente documento serán evidentes para el artesano experto y están dentro del alcance de esta divulgación.

5 [0100] Cabe señalar que los sistemas de edición genómica de la presente divulgación pueden estar dirigidos a una única secuencia nucleotídica específica, o pueden estar dirigidos a -y ser capaces de editar en paralelo- dos o más secuencias nucleotídicas específicas mediante el uso de dos o más ARN guía. El uso de múltiples ARNg se refiere a "multiplexación" a lo largo de esta divulgación, y puede emplearse para dirigir múltiples secuencias diana no relacionadas de interés, o para formar múltiples SSBs o DSBs dentro de un único dominio diana y, en algunos casos, para generar ediciones específicas dentro de dicho dominio diana. Por ejemplo, la Publicación de Patente Internacional Nº WO 2015/138510 de Maeder et al. (Maeder), describe un sistema de edición del genoma para corregir una mutación puntual (C.2991+1655A a G) en el gen CEP290 humano que da lugar a la creación de un sitio de empalme críptico, que a su vez reduce o elimina la función del gen. El sistema de edición genómica de Maeder utiliza dos ARN guía dirigidos a secuencias a ambos lados de (es decir, flanqueando) la mutación puntual, y forma DSBs que flanquean la mutación. Esto, a su vez, promueve la eliminación de la secuencia intermedia, incluida la mutación, eliminando así el sitio de empalme críptico y restaurando la función normal del gen.

10 [0101] Como otro ejemplo, el documento WO 2016/073990 de Cotta-Ramusino, et al. ("Cotta-Ramusino"), describe un sistema de edición del genoma que utiliza dos ARNg en combinación con una mellasa Cas9 (una Cas9 que hace una mella de cadena única tal como *S. pyogenes* D10A), una disposición denominada "sistema de doble mellasa." El sistema de doble nicasa de Cotta-Ramusino está configurado para hacer dos muescas en hebras opuestas de una secuencia de interés que están desplazadas uno o más nucleótidos, las cuales se combinan para crear una rotura de doble hebra que tiene un saliente (5' en el caso de Cotta-Ramusino, aunque también son posibles salientes de 3'). El voladizo, a su vez, puede facilitar los eventos de reparación dirigidos por homología en algunas circunstancias. Y, como otro ejemplo, el documento WO 2015/070083 de Palestant et al. ("Palestant") describe un ARNg dirigido a una secuencia de nucleótidos que codifica Cas9 (al que se refiere como "ARN gobernante"), que puede incluirse en un sistema de edición del genoma que comprenda uno o más ARNg adicionales para permitir la expresión transitoria de un Cas9 que de otro modo podría expresarse constitutivamente, por ejemplo en algunas células transducidas viralmente. Estas aplicaciones de multiplexación pretenden ser ejemplares, más que limitantes, y el artesano experto apreciará que otras aplicaciones de multiplexación son generalmente compatibles con los sistemas de edición de genoma descritos aquí.

25 [0102] Los sistemas de edición del genoma pueden, en algunos casos, formar roturas de doble cadena que son reparadas por mecanismos celulares de rotura de doble cadena del ADN tales como NHEJ o HDR. Estos mecanismos se describen en toda la literatura, por ejemplo por Davis & Maizels, PNAS, 111(10):E924-932, 11 de marzo de 2014 (Davis) (describiendo Alt-HDR); Frit et al. ADN Repair 17(2014) 81-97 (Frit) (describiendo Alt-NHEJ); y Iyama y Wilson III, ADN Repair (Amst.) 2013-Ago; 12(8): 620-636 (Iyama) (describe las vías canónicas HDR y NHEJ en general).

30 [0103] Cuando los sistemas de edición genómica operan mediante la formación de DSB, dichos sistemas incluyen opcionalmente uno o más componentes que promueven o facilitan un modo particular de reparación de roturas de doble cadena o un resultado de reparación particular. Por ejemplo, Cotta-Ramusino también describe sistemas de edición genómica en los que se añade una "plantilla donante" de oligonucleótidos monocatenarios; la plantilla donante se incorpora a una región diana del ADN celular que es escindida por el sistema de edición genómica, y puede dar lugar a un cambio en la secuencia diana.

35 [0104] En ciertas realizaciones, los sistemas de edición genómica modifican una secuencia diana, o modifican la expresión de un gen en o cerca de la secuencia diana, sin causar roturas de cadena simple o doble. Por ejemplo, un sistema de edición genómica puede incluir una nucleasa guiada por ARN fusionada a un dominio funcional que actúa sobre el ADN, modificando así la secuencia diana o su expresión. Por ejemplo, una nucleasa guiada por ARN puede estar conectada a (por ejemplo, fusionada a) un dominio funcional de citidina deaminasa, y puede funcionar generando sustituciones diana de C a A. En Komor et al. Nature 533, 420-424 (19 de mayo de 2016) ("Komor") se describen fusiones nucleasa/deaminasa ejemplares. Alternativamente, un sistema de edición genómica puede utilizar una nucleasa inactivada por corte (es decir, "muerta"), tal como una Cas9 muerta (dCas9), y puede funcionar formando complejos estables en una o más regiones diana del ADN celular, interfiriendo así con funciones que implican la región o regiones diana, incluyendo, sin limitación, la transcripción de ARNm, la remodelación de la cromatina, etc.

Moléculas de ARN guía (ARNg)

- 5 [0105] Los términos "ARN guía" y "ARNg" se refieren a cualquier ácido nucleico que promueva la asociación específica (o "dirigir") de una nucleasa guiada por ARN, tal como una Cas9 o una Cpfl, a una secuencia diana, tal como una secuencia CD40L en una célula. Los ARNg pueden ser unimoleculares (comprenden una única molécula de ARN, y se refieren a ellos alternativamente como químéricos), o modulares (comprenden más de una, y típicamente dos, moléculas de ARN separadas, tales como un ARNcr y un ARNtracr, que normalmente se asocian entre sí, por ejemplo mediante duplexación). Los ARNg y sus componentes se describen en toda la bibliografía, por ejemplo en Briner et al. (Molecular Cell 56(2), 333-339, 23 de octubre de 2014 (Briner)), y en Cotta-Ramusino.
- 10 [0106] En bacterias y arqueas, los sistemas CRISPR de tipo II comprenden generalmente una proteína nucleasa guiada por ARN, tal como Cas9, un ARN CRISPR (ARNcr) que incluye una región 5' complementaria a una secuencia extraña, y un ARNcr transactivador (ARNtrat) que incluye una región 5' complementaria y que forma un dúplex con una región 3' del ARNcr. Sin pretender ceñirnos a ninguna teoría, se cree que este dúplex facilita la formación del complejo Cas9/ARNg y es necesario para su actividad. A medida que los sistemas CRISPR de tipo II se adaptaban para su uso en la edición de genes, se descubrió que el ARNcr y el ARNtracr podían unirse en un único ARN guía unimolecular o químérico, en un ejemplo no limitativo, mediante una secuencia de "tetraloop" o "enlazador" de cuatro nucleótidos (*por ejemplo*, GAAA) que unía regiones complementarias del ARNcr (en su extremo 3') y del ARNtracr (en su extremo 5'). (Mali et al. Science. 2013 Feb 15; 339(6121): 823-826 ("Mali"); Jiang et al. Nat Biotechnol. 2013 Mar; 31(3): 233-239 ("Jiang"); y Jinek et al., 2012 Science Aug. 17; 337(6096): 816-821 ("Jinek").
- 15 [0107] Los ARN guía, ya sean unimoleculares o modulares, incluyen un "dominio diana" que es total o parcialmente complementario a un dominio diana dentro de una secuencia diana, tal como una secuencia CD40L en el genoma de una célula en la que se desea realizar la edición. Los dominios diana se refieren a varios nombres en la literatura, incluyendo sin limitación "secuencias guía" (Hsu et al., Nat Biotechnol. 2013 Sep; 31(9): 827-832, ("Hsu")), "regiones de complementariedad" (Cotta-Ramusino), "espaciadores" (Briner) y genéricamente como "crRNAs" (Jiang). Independientemente de los nombres que reciban, los dominios diana suelen tener entre 10 y 30 nucleótidos de longitud y, en ciertas realizaciones, entre 16 y 24 nucleótidos (*por ejemplo*, 16, 17, 18, 19, 20, 21, 22, 23 o 24 nucleótidos de longitud), y se encuentran en o cerca del extremo 5' en el caso de un ARNg Cas9, y en o cerca del extremo 3' en el caso de un ARNg Cpfl.
- 20 [0108] Además de los dominios diana, los ARNg suelen incluir (aunque no necesariamente, como se explica más adelante) una pluralidad de dominios que pueden influir en la formación o la actividad de los complejos ARNg/Cas9. Por ejemplo, como se ha mencionado anteriormente, la estructura dúplex formada por los dominios de complementariedad primero y secundario de un ARNg (también se refiere a un dúplex de repetición: antirrepetición) interactúa con el lóbulo de reconocimiento (REC) de Cas9 y puede mediar en la formación de complejos Cas9/ARNg. (Nishimasu et al., Cell 156, 935-949, 27 de febrero de 2014 (Nishimasu 2014) y Nishimasu et al., Cell 162, 1113-1126, 27 de agosto de 2015 (Nishimasu 2015)). Cabe señalar que los dominios de complementariedad primero y/o segundo pueden contener uno o más tractos poli-A, que pueden ser reconocidos por las ARN polimerasas como señal de terminación. La secuencia de los dominios de complementariedad primero y segundo se modifican, por tanto, opcionalmente para eliminar estos tractos y promover la transcripción completa *in vitro* de los ARNg, por ejemplo mediante el uso de intercambios A-G como se describe en Briner, o intercambios A-U. Estas y otras modificaciones similares de los dominios de complementariedad primero y segundo están dentro del alcance de la presente divulgación.
- 25 [0109] Junto con los dominios de complementariedad primero y segundo, los ARNg Cas9 suelen incluir dos o más regiones dúplex adicionales que intervienen en la actividad nucleasa *in vivo* pero no necesariamente *in vitro*. (Nishimasu 2015). Un primer bucle de tallo cerca de la porción 3' del segundo dominio de complementariedad se refiere a varios como el "dominio proximal", (Cotta-Ramusino) "bucle de tallo 1" (Nishimasu 2014 y 2015) y el "nexo" (Briner). Por lo general, hay una o más estructuras de bucle madre adicionales cerca del extremo 3' del ARNg, y el número varía según la especie: Los ARNg de *s. pyogenes* suelen incluir dos bucles madre 3' (para un total de cuatro estructuras de bucle madre incluyendo el dúplex repetición:antirrepetición), mientras que *s. aureus* y otras especies tienen sólo una (para un total de tres estructuras de bucle madre). En Briner se ofrece una descripción de las estructuras conservadas del bucle del tallo (y de las estructuras del ARNg en general) organizadas por especies.
- 30 [0110] Aunque la descripción anterior se ha centrado en los ARNg para su uso con Cas9, debe tenerse en cuenta que se han descubierto o inventado (o pueden descubrirse en el futuro) otras nucleasas guiadas por ARN que utilizan ARNg que difieren en algunos aspectos de los descritos hasta ahora. Por ejemplo, Cpfl ("CRISPR de Prevotella y Franciscella 1") es una nucleasa guiada por ARN descubierta recientemente que no necesita un ARN tracr para funcionar. (Zetsche et al., 2015, Cell 163, 759-771 22 de octubre de 2015 (Zetsche I)). Un ARNg para uso en un sistema de edición genómica Cpfl generalmente incluye un dominio diana y un dominio de complementariedad (alternativamente se refiere a un "asa"). También debe tenerse en cuenta que, en los ARNg para su uso con Cpfl, el dominio de diana suele estar presente en o cerca del extremo 3', en lugar del extremo 5' como se ha descrito anteriormente en relación con los ARNg Cas9 (el asa está en o cerca del extremo 5' de un ARNg Cpfl).

[0111] Los expertos en la técnica apreciarán que, aunque pueden existir diferencias estructurales entre los ARNg de diferentes especies procariotas, o entre los ARNg Cpf1 y Cas9, los principios por los que operan los ARNg son generalmente coherentes. Debido a esta coherencia de funcionamiento, los ARNg pueden definirse, en términos generales, por sus secuencias de dominio diana, y los expertos apreciarán que una secuencia de dominio diana dada puede incorporarse en cualquier ARNg adecuado, incluido un ARNg unimolecular o químérico, o un ARNg que incluya una o más modificaciones químicas y/o modificaciones secuenciales (sustituciones, nucleótidos adicionales, truncamientos, etc.). Así, por economía de presentación en esta divulgación, los ARNg pueden describirse únicamente en términos de sus secuencias de dominio diana.

[0112] En términos más generales, los expertos apreciarán que algunos aspectos de la presente divulgación se refieren a sistemas, métodos y composiciones que pueden implementarse utilizando múltiples nucleasas guiadas por ARN. Por este motivo, a menos que se especifique lo contrario, debe entenderse que el término ARNg engloba cualquier ARNg adecuado que pueda utilizarse con cualquier核酸酶 guiada por ARN, y no sólo aquellos ARNg compatibles con una especie concreta de Cas9 o Cpf1. A modo de ilustración, el término ARNg puede, en ciertas realizaciones, incluir un ARNg para su uso con cualquier核酸酶 guiada por ARN que se produzca en un sistema CRISPR de Clase 2, tal como un sistema CRISPR de tipo II o de tipo V o CRISPR, o una核酸酶 guiada por ARN derivada o adaptada del mismo.

Diseño de ARNg

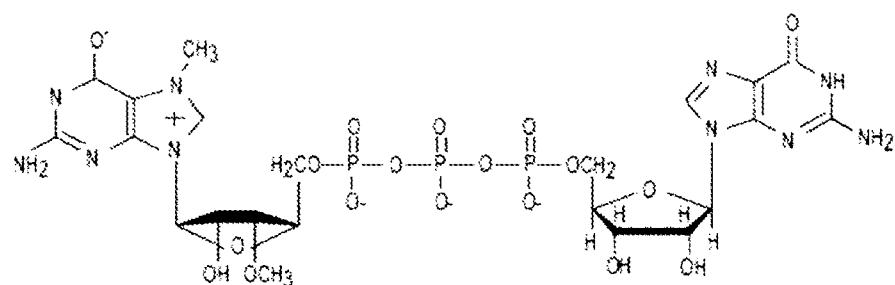
[0113] Los métodos de selección y validación de las secuencias diana, así como los análisis fuera de diana, se han descrito anteriormente, por ejemplo, en Mali; Hsu; Fu et al., 2014 Nat Biotechnol. 32(3): 279-84, Heigwer et al., 2014 Nat methods 11(2): 122-3; Bae et al. (2014) Bioinformatics 30(10): 1473-5; y Xiao A et al. (2014) Bioinformatics 30(8): 1180-1182. Como ejemplo no limitativo, el diseño de ARNg puede implicar el uso de una herramienta de software para optimizar la elección de posibles secuencias diana correspondientes a la secuencia diana de un usuario, por ejemplo, para minimizar la actividad total fuera de diana en todo el genoma. Aunque la actividad fuera de diana no se limita a la escisión, la eficiencia de escisión en cada secuencia fuera de diana puede predecirse, por ejemplo, utilizando un esquema de ponderación derivado experimentalmente. Estos y otros métodos de selección de guías se describen detalladamente en Maeder y Cotta-Ramusino.

Modificaciones del ARNg

[0114] La actividad, estabilidad u otras características de los ARNg pueden alterarse mediante la incorporación de determinadas modificaciones. Por ejemplo, los ácidos nucleicos expresados o administrados transitoriamente pueden ser propensos a la degradación por, por ejemplo, nucleasas celulares. En consecuencia, los ARNg descritos en el presente documento pueden contener uno o más nucleósidos o nucleótidos modificados que introducen estabilidad frente a las nucleasas. Aunque no se desea estar unido por la teoría, también se cree que ciertos ARNg modificados descritos en el presente documento pueden mostrar una respuesta inmunitaria innata reducida cuando se introducen en las células. Los expertos en la técnica conocerán ciertas respuestas celulares comúnmente observadas en células, por ejemplo, células de mamífero, en respuesta a ácidos nucleicos exógenos, en particular los de origen vírico o bacteriano. Tales respuestas, que pueden incluir la inducción de la expresión y liberación de citoquinas y la muerte celular, pueden reducirse o eliminarse por completo mediante las modificaciones presentadas en el presente documento.

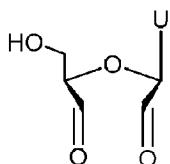
[0115] Ciertas modificaciones ejemplares discutidas en esta sección pueden incluirse en cualquier posición dentro de una secuencia de ARNg incluyendo, sin limitación en o cerca del extremo 5' (por ejemplo, dentro de 1-10, 1-5, o 1-2 nucleótidos del extremo 5') y/o en o cerca del extremo 3' (por ejemplo, entre 1-10, 1-5, o 1-2 nucleótidos del extremo 3'). En algunos casos, las modificaciones se sitúan dentro de motivos funcionales, tales como el dúplex repetición-anti-repetición de un ARNg Cas9, una estructura de bucle madre de un ARNg Cas9 o Cpf1, y/o un dominio diana de un ARNg.

[0116] Como un ejemplo, el extremo 5' de un ARNg puede incluir una estructura de casquete de ARNm eucariota o un análogo de casquete (por ejemplo, un G(5')*ppp*(5')G análogo de caperuza, un *m7G*(5')*ppp*(5')G análogo de caperuza, o un 3'-O-Mem7G(5')*ppp*(5')G análogo de caperuza antirreverso (ARCA)), como se muestra a continuación:



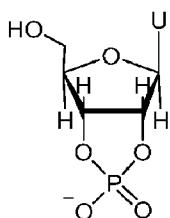
La caperuza o el análogo de la caperuza pueden incluirse durante la síntesis química o la transcripción *in vitro* del ARNg.

- 5 [0117] De forma similar, el extremo 5' del ARNg puede carecer de un grupo trifosfato 5'. Por ejemplo, los ARNg transcritos *in vitro* pueden tratarse con fosfatasa (*por ejemplo*, usando fosfatasa alcalina intestinal de ternera) para eliminar un grupo 5' trifosfato.
- 10 [0118] Otra modificación común implica la adición, en el extremo 3' de un ARNg, de una pluralidad (*por ejemplo*, 1-10, 10-20, o 25-200) de residuos de adenina (A) que se refiere a un tracto poliA. El tracto poliA puede añadirse a un ARNg durante la síntesis química, tras la transcripción *in vitro* utilizando una poliadenosina polimerasa (*por ejemplo*, *E. coli* Poli(A)Polimerasa), o *in vivo* mediante una secuencia de poliadenilación, como se describe en Maeder.
- 15 [0119] Debe tenerse en cuenta que las modificaciones descritas en el presente documento pueden combinarse de cualquier manera adecuada, *por ejemplo*, un ARNg, ya sea transcrita *in vivo* a partir de un vector de ADN, o *ARNg transcrita in vitro*, puede incluir una estructura de caperuza 5' o un análogo de caperuza y un tracto poliA 3'.
- 20 [0120] Los ARN guía pueden modificarse en una U ribosa terminal 3'. Por ejemplo, los dos grupos hidroxilo terminales de la ribosa U pueden oxidarse a grupos aldehído y una apertura concomitante del anillo de ribosa para dar lugar a un nucleósido modificado como se muestra a continuación:



- 25 en el que "U" puede ser una uridina no modificada o modificada.

[0121] La U ribosa terminal 3' puede modificarse con un fosfato cíclico 2'3' como se muestra a continuación:



- 30 en el que "U" puede ser una uridina no modificada o modificada.

[0122] Los ARN guía pueden contener nucleótidos 3' que pueden estabilizarse contra la degradación, *por ejemplo*, incorporando uno o más de los nucleótidos modificados descritos en el presente documento. En ciertas realizaciones, las uridinas pueden sustituirse con uridinas modificadas, *por ejemplo*, 5-(2-amino)propil uridina, y 5-bromo uridina, o con cualquiera de las uridinas modificadas descritas en el presente documento; las adenosinas y guanosinas pueden sustituirse con adenosinas y guanosinas modificadas, *por ejemplo*, con modificaciones en la posición 8, *por ejemplo*, 8-bromo guanosina, o con cualquiera de las adenosinas o guanosinas modificadas descritas en el presente documento.

40 [0123] En ciertas realizaciones, pueden incorporarse ribonucleótidos modificados con azúcar en el ARNg, *por ejemplo*, en el que el grupo 2' OH se sustituye por un grupo seleccionado de H, -OR, -R (en el que R puede ser, *por ejemplo*, alquilo, cicloalquilo, arilo, aralquilo, heteroarilo o azúcar), halo, -SH, -SR (en el que R puede ser, *por ejemplo*, alquilo, cicloalquilo, arilo, aralquilo, heteroarilo o azúcar), amino (en el que amino puede ser, *por ejemplo*, NH₂; alquilamino, dialquilamino, heterociclico, arilamino, diarylamino, heteroarilamino, diheteroarilamino, o aminoácido); o ciano (-CN). En ciertas realizaciones, la espina dorsal de fosfato puede modificarse como se describe en el presente documento, *por ejemplo*, con un grupo fosftoato (PhTx). En ciertas realizaciones, uno o más de los nucleótidos del ARNg pueden ser cada uno independientemente un nucleótido modificado o no modificado incluyendo, pero no se limitan a 2'-azúcar modificado, tal como, 2'-O-metil, 2'-O-metoxietil, o 2'-Fluoro modificado incluyendo, *por ejemplo*, 2'-F o 2'-O-metil, adenosina (A), 2'-F o 2'-O-metil, citidina (C), 2'-F o 2'-O-metil, uridina (U), 2'-F o 2'-O-metil, timidina (T), 2'-F o 2'-O-metil, guanosina (G), 2'-O-metoxietil-5-metiluridina (Teo), 2'-O-metoxietiladenosina (Aeo), 2'-O-metoxietil-5-metilcitidina (m5Ceo), y cualquier combinación de los mismos.

[0124] Los ARN guía también pueden incluir ácidos nucleicos "bloqueados" (LNA) en los que el grupo OH 2' puede estar conectado, por ejemplo, mediante un puente de alquíleno C1-6 o heteroalquíleno C1-6, al carbono 4' del mismo azúcar ribosa. Puede usarse cualquier fracción adecuada para proporcionar tales puentes, incluyendo sin limitación puentes de metileno, propileno, éter o amino; O-amino (en el que amino puede ser, *por ejemplo*, NH₂; alquilamino, dialquilamino, heterociclico, arilamino, diarilamino, heteroarilamino o diheteroarilamino, etilendiamina o poliamino) y aminoalcoxi u O(CH₂)_n-amino (en el que amino puede ser, *por ejemplo*, NH₂; alquilamino, dialquilamino, heterociclico, arilamino, diarilamino, heteroarilamino o diheteroarilamino, etilendiamina o poliamino).

[0125] En ciertas realizaciones, un ARNg puede incluir un nucleótido modificado que es multicíclico (*por ejemplo*, triciclo; y formas "desbloqueadas", tales como ácido nucleico glicólico (GNA) (*por ejemplo*, R-GNA o S-GNA, donde la ribosa se sustituye por unidades de glicol acopladas a enlaces fosfodiéster), o ácido nucleico de treosa (TNA, donde la ribosa se sustituye por α-L-treofuranosil-(3'→2')).

[0126] Generalmente, los ARNg incluyen el grupo de azúcares ribosa, que es un anillo de 5 miembros que tiene un oxígeno. Los ARNg modificados ejemplares pueden incluir, sin limitación, la sustitución del oxígeno en la ribosa (*por ejemplo*, con azufre (S), selenio (Se), o alquíleno, tal como, *por ejemplo*, metileno o etileno); adición de un doble enlace (*por ejemplo*, para sustituir la ribosa por ciclopentenilo o ciclohexenilo); contracción del anillo de ribosa (*por ejemplo*, para formar un anillo de 4 miembros de ciclobutano u oxetano); expansión del anillo de ribosa (*por ejemplo*, para formar un anillo de 6 o 7 miembros que tiene un carbono o heteroátomo adicional, tal como, *por ejemplo*, anhidrohexitol, alritol, manitol, ciclohexanilo, ciclohexenilo y morfolino que también tiene una espina dorsal de fosforamidato). Aunque la mayoría de las alteraciones de los análogos del azúcar se localizan en la posición 2', otros sitios son susceptibles de modificación, incluida la posición 4'. En ciertas realizaciones, un ARNg comprende una modificación 4'-S, 4'-Se o 4'-C-aminometil-2'-O-Me.

[0127] En ciertas realizaciones, pueden incorporarse al ARNg nucleótidos deaza, *por ejemplo*, 7-deaza-adenosina. En ciertas realizaciones, pueden incorporarse al ARNg nucleótidos O- y N-alquilados, *por ejemplo*, N6-metil adenosina. En ciertas realizaciones, uno o más o todos los nucleótidos de un ARNg son desoxinucleótidos.

CD40L Plantillas de donantes

[0128] Las plantillas donantes, *por ejemplo*, plantillas donantes de oligonucleótidos, según la presente divulgación pueden implementarse de cualquier forma adecuada, incluyendo sin limitación ADN monocatenario o bicatenario, lineal o circular, desnudo o comprendido dentro de un vector, y/o asociado, covalentemente o no covalentemente (*por ejemplo*, mediante hibridación directa o hibridación por empalme) con un ARN guía. Algunos diseños preferentes de plantillas donantes se representan esquemáticamente en la Figura. 1.

[0129] En algunas realizaciones, la plantilla donante es un ssODN. Cuando se utiliza un ssODN lineal, puede configurarse para (i) recocido a una cadena mellada del ácido nucleico diana, (ii) recocido a la cadena intacta del ácido nucleico diana, (iii) recocido a la cadena positiva del ácido nucleico diana, y/o (iv) recocido a la cadena negativa del ácido nucleico diana. Un ssODN puede tener cualquier longitud adecuada, *por ejemplo*, aproximadamente, o no más de 150-200 nucleótidos (*por ejemplo*, 150, 160, 170, 180, 190, o 200 nucleótidos). En otras realizaciones, la ssODN tiene una longitud de unos 100-3000 nucleótidos, por ejemplo, unos 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 nucleótidos. En otras realizaciones, la ssODN tiene una longitud de al menos 100-3000 nucleótidos, por ejemplo, al menos 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 nucleótidos. En otras realizaciones, la ssODN tiene una longitud de no más de 100-3000 nucleótidos, por ejemplo, no más de 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 nucleótidos.

[0130] En otras realizaciones, la plantilla donante es un dsODN. En una realización, la plantilla donante comprende una primera cadena. En otra realización, una plantilla donante comprende una primera hebra y una segunda hebra. Un dsODN puede tener cualquier longitud adecuada, *por ejemplo*, aproximadamente, o no más de 150-200 pares de bases (*por ejemplo*, 150, 160, 170, 180, 190 o 200 pares de bases). En otras realizaciones, el dsODN tiene una longitud de unos 100-3000 pares de bases, por ejemplo, unos 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 pares de bases. En otras realizaciones, el dsODN tiene una longitud de al menos 100-3000 pares de bases, por ejemplo, al menos 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 pares de bases. En otras realizaciones, el dsODN tiene una longitud de no más de 100-3000 pares de bases, por ejemplo, no más de 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, o 3000 pares de bases.

[0131] En algunas realizaciones, una plantilla donante es un oligonucleótido exógeno, *por ejemplo*, un oligonucleótido que no está presente de forma natural en una célula.

[0132] Cabe señalar que una plantilla donante también puede estar comprendida dentro de un vector de ácido nucleico, tal como un genoma viral o ADN circular de doble cadena, *por ejemplo*, un plásmido. En algunas realizaciones, la plantilla donante puede ser un ADN con forma de hueso de perro (véase, *por ejemplo*, Patente estadounidense nº 9499847). Los vectores de ácido nucleico que comprenden plantillas donantes pueden incluir otros elementos codificantes o no codificantes. Por ejemplo, un ácido nucleico molde donante puede entregarse como parte de un genoma viral (*por ejemplo*, en un genoma AAV o lentiviral) que incluye ciertos elementos de la espina dorsal genómica (*por ejemplo*, repeticiones terminales invertidas, en el caso de un genoma AAV) y opcionalmente incluye secuencias adicionales que codifican un ARNg y/o una nucleasa guiada por ARN. En ciertas realizaciones, la plantilla donante puede ser adyacente a, o estar flanqueada por, sitios diana reconocidos por uno o más ARNg, para facilitar la formación de DSB libres en uno o ambos extremos de la plantilla donante que pueden participar en la reparación de SSB o DSB correspondientes formados en el ADN celular utilizando los mismos ARNg. En Cotta-Ramusino se describen ejemplos de vectores de ácidos nucleicos adecuados para su uso como plantillas donantes. En una realización, la plantilla donante está presente en un vector AAV6. En otra realización, la plantilla donante está presente en un vector IDLV.

A. Brazos de homología

[0133] Ya sean monocatenarias o bicatenarias, las plantillas donantes generalmente incluyen una o más regiones homólogas a regiones de ADN, *por ejemplo*, un ácido nucleico diana, secuencia CD40L, dentro o cerca (*por ejemplo*, flanqueando o contiguas) de una secuencia diana que se va a escindir, *por ejemplo*, el sitio de escisión. En el presente documento se hace referencia a estas regiones homólogas como "brazos de homología", y se ilustran esquemáticamente en la Figura. 1.

[0134] Los brazos homológicos de las plantillas donantes descritas en el presente documento pueden tener cualquier longitud adecuada, con la condición de que dicha longitud sea suficiente para permitir la resolución eficaz de un sitio de corte en un ácido nucleico diana mediante un procedimiento de reparación del ADN que requiera una plantilla donante. En algunas realizaciones, en las que se desea la amplificación mediante, *por ejemplo*, PCR, del brazo de homología, el brazo de homología tiene una longitud tal como para que pueda realizarse la amplificación. En algunas realizaciones, cuando se desea secuenciar el brazo homológico, éste tiene una longitud tal que permite realizar la secuenciación. En algunas realizaciones, en las que se desea una evaluación cuantitativa de los amplicones, los brazos de homología tienen una longitud tal como para conseguir un número similar de amplificaciones de cada amplicón, *por ejemplo*, teniendo un contenido G/C, temperaturas de amplificación, etc. similares. En algunas realizaciones, el brazo de homología es de doble cadena. En algunas realizaciones, el brazo homológico bicatenario es monocatenario.

[0135] En algunas realizaciones, el brazo homológico 5' tiene una longitud de entre 150 y 250 nucleótidos. En algunas realizaciones, el brazo homológico 5' tiene una longitud de 700 nucleótidos o menos. En algunas realizaciones, el brazo homológico 5' tiene una longitud de 650 nucleótidos o menos. En algunas realizaciones, el brazo homológico 5' tiene una longitud de 600 nucleótidos o menos. En algunas realizaciones, el brazo homológico 5' tiene una longitud de 550 nucleótidos o menos. En algunas realizaciones, el brazo homológico 5' tiene una longitud de 500 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 400 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 300 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 250 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 200 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 150 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 100 nucleótidos o menos. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 50 nucleótidos o menos. En algunas realizaciones, el brazo de homología 5' tiene una longitud de 250, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, o 20 nucleótidos. En algunas realizaciones, el brazo homólogo 5' tiene 40 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene una longitud de 250 nucleótidos o menos. En otras realizaciones, el brazo homólogo 5' tiene al menos 50 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 100 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 150 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 200 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 250 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 300 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 350 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 400 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 450 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 500 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 550 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 600 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 650 nucleótidos de longitud. En algunas realizaciones, el brazo homólogo 5' tiene al menos 700 nucleótidos de longitud. En una realización, el brazo de homología 5' tiene una longitud de entre 200 y 700

nucleótidos. En una realización, el brazo de homología 5' tiene una longitud de entre 200 y 300 nucleótidos. En una realización, el brazo de homología 5' tiene una longitud de entre 200 y 400 nucleótidos. En una realización, el brazo de homología 5' tiene una longitud de entre 200 y 500 nucleótidos. En una realización, el brazo de homología 5' tiene una longitud de entre 400 y 700 nucleótidos. En una realización, el brazo de homología 5' tiene una longitud de entre 500 y 700 nucleótidos. En una realización, el brazo de homología 5' tiene unos 500 nucleótidos de longitud.

homológico 5' tiene entre 200-700 pares de bases de longitud. En una realización, el brazo homológico 5' tiene entre 200-300 pares de bases de longitud. En una realización, el brazo homológico 5' tiene entre 200-400 pares de bases de longitud. En una realización, el brazo homológico 5' tiene entre 200-500 pares de bases de longitud. En una realización, el brazo homológico 5' tiene entre 400-700 pares de bases de longitud. En una realización, el brazo homológico 5' tiene entre 500-700 pares de bases de longitud. En una realización, el brazo homológico 5' tiene unos 500 pares de bases de longitud.

[0138] En algunas realizaciones, el brazo homólogo 3' tiene entre 150 pares de bases y 250 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 700 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 650 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 600 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 550 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 500 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 400 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 300 pares de bases o menos de longitud. En algunas realizaciones, el brazo homólogo 3' tiene una longitud de 250 pares de bases o menos. En algunas realizaciones, el brazo homólogo 3' tiene una longitud de 200 pares de bases o menos. En algunas realizaciones, el brazo homólogo 3' tiene una longitud de 150 pares de bases o menos. En algunas realizaciones, el brazo homólogo 3' tiene una longitud de 100 pares de bases o menos. En algunas realizaciones, el brazo homólogo 3' tiene una longitud de 50 pares de bases o menos. En algunas realizaciones, el brazo homólogo 3' tiene 250, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, o 20 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene 40 pares de bases de longitud. En otras realizaciones, el brazo homólogo 3' tiene al menos 50 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 100 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 200 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 250 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 300 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 350 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 400 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 450 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 500 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 550 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 600 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 650 pares de bases de longitud. En algunas realizaciones, el brazo homólogo 3' tiene al menos 700 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 200-700 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 200-300 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 200-400 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 200-500 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 400-700 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene entre 500-700 pares de bases de longitud. En una realización, el brazo homólogo 3' tiene unos 500 pares de bases de longitud.

[0139] Los brazos homológicos 5' y 3' pueden ser de la misma longitud o pueden diferir en longitud. En algunas realizaciones, los brazos de homología 5' y 3' se amplifican para permitir la evaluación cuantitativa de los eventos de edición génica, tal como la integración dirigida, en un ácido nucleico diana, *por ejemplo*, CD40L. En algunas realizaciones, la evaluación cuantitativa de los eventos de edición génica puede basarse en la amplificación tanto de la unión 5' como de la unión 3' en el sitio de integración diana mediante la amplificación de la totalidad o de una parte del brazo de homología utilizando un único par de cebadores de PCR en una única reacción de amplificación. En consecuencia, aunque la longitud de los brazos homológicos 5' y 3' puede diferir, la longitud de cada brazo homológico debe poder amplificarse (*por ejemplo*, mediante PCR), según se deseé. Además, cuando se desea la amplificación tanto del brazo 5' como de la diferencia de longitudes de los brazos homológicos 5' y 3' en una única reacción PCR, la diferencia de longitud entre los brazos homológicos 5' y 3' debe permitir la amplificación PCR utilizando un único par de cebadores PCR.

[0140] En algunas realizaciones, la longitud de los brazos homológicos 5' y 3' no difiere en más de 75 nucleótidos. Así, en algunas realizaciones, cuando los brazos homológicos 5' y 3' difieren en longitud, la diferencia de longitud entre los brazos homológicos es inferior a 70, 60, 50, 40, 30, 20, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 nucleótidos o pares de bases. En algunas realizaciones, los brazos homológicos 5' y 3' difieren en longitud en al menos 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, 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, o 75 nucleótidos. En algunas realizaciones, la diferencia de longitud entre los brazos homológicos 5' y 3' es inferior a 70, 60, 50, 40, 30, 20, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 pares de bases. En algunas realizaciones, los brazos homológicos 5' y 3' difieren en longitud en al menos 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, 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, o 75 pares de bases.

[0141] Sea cual sea el formato utilizado, se puede diseñar una plantilla de donante para evitar secuencias indeseables. En determinadas realizaciones, uno o ambos brazos homólogos pueden acortarse para evitar el solapamiento con determinados elementos de repetición de secuencias, por ejemplo, repeticiones Alu, elementos LINE, etc.

5 **B. Carga - Exones CD40L**

[0142] La plantilla donante de los sistemas de edición génica descritos en el presente documento comprende una carga (*por ejemplo*, carga de corrección del gen CD40L). La carga divulgada en el presente documento puede comprender uno o más de los exones 2-5 de un ácido nucleico CD40L. La carga puede ser de cualquier longitud necesaria para lograr el resultado deseado. Por ejemplo, una secuencia de carga puede tener menos de 2500 pares de bases o menos de 2500 nucleótidos de longitud. Los expertos en la técnica comprobarán fácilmente que cuando la plantilla donante se administra utilizando un vehículo de administración (*por ejemplo*, un vehículo de administración viral tal como un vehículo de administración de virus adenoasociado (AAV) o virus del herpes simple (HSV)) con limitaciones de tamaño, el tamaño de la plantilla donante, incluida la carga, no debe exceder la limitación de tamaño del sistema de administración.

[0143] En algunas realizaciones, la carga comprende una secuencia de sustitución. En una realización, la secuencia de sustitución es una secuencia CD40L de tipo salvaje. En una realización, la secuencia de sustitución es una secuencia CD40L de codón optimizado, que aumentaría la traducción de la proteína CD40L. En una realización, la secuencia CD40L de tipo salvaje no comprende una mutación que haría CD40L no funcional o afectaría a su expresión y/o empalme.

[0144] En algunas realizaciones, la carga comprende todo o una porción de un exón de una secuencia genética, CD40L, *por ejemplo*, todo o una porción de uno o más de los exones 2, 3, 4 y 5 de CD40L. En algunas realizaciones, la carga comprende todo o parte del exón 2 de CD40L. En algunas realizaciones, la carga comprende todo o parte del exón 3 de CD40L. En algunas realizaciones, la carga comprende todo o parte del exón 4 de CD40L. En algunas realizaciones, la carga comprende todo o parte del exón 5 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 2-5 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 2 y 3 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 2-4 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 3 y 4 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 3-5 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 3 y 5 de CD40L. En algunas realizaciones, la carga comprende la totalidad o una parte de los exones 4 y 5 de CD40L. La carga no comprende el exón 1 de CD40L. En una realización preferente, la carga comprende una señal de aceptor de empalme (SA) antes del CD40L.

[0145] La secuencia del gen CD40L, así como los límites de los exones e intrones, son bien conocidos por los expertos en la técnica. En una realización, la secuencia de ARNm de CD40L se describe al menos en NM_000074. En una realización, la secuencia del gen CD40L RefSeq del cromosoma X se describe al menos en NG_007280.1. Una versión optimizada de codones de uno o más exones de CD40L puede utilizarse en la plantilla donante, en algunas realizaciones, como se describe en el presente documento. Una versión ejemplar de codón optimizado de CD40L, exones 2-5, se proporciona como SEQ ID NO:48.

ES 2 998 052 T3

ES 2 998 052 T3

| | |
|---|--|
| CD40L RefSeqGene en el cromosoma (NG_007280) (SEQ ID NO: 47) | <p>X</p> <p>ACGATTCAACTTAAACACAGCATGATCGAAACATAACAAACAAACTCTCCCGAATGGGGCAC TGCAGTCGCCATCAGCAAGAAATTTCCTGTATTTCTACTCTTTTCATACAGATGAT TGGGTCAGCACTTTTGCTGTCTATCTCTAGAGGTGACAAGCTAACATGATGAGACAGCT TTAACTAAATTGGGGCTCTACTAAATTCATAGCTGCTCTACCCAAATGATGGATGAAGCT AGAAACCAAATAGAACAAATGGCTTGTGGCAAAATGTTGCTTGCCTAGAACGCTCAATT CTTCTCTGCTTAGGATCTTGGATCTGGACTCAGACTCTGGGTTCAAATCTGCTCTGCC ATACCATCTCTGTTATCTGGGCAAGTCGCTCACTTCCACACCTGAGAACAGGCTGCTAATG CTTCATTCATAGATTAATGCTGCTTAAAGCTTACGGAGGGCTCTGATTCATAG TAAGCACTATGTAACGCACTACCTATAATTGCTAAAGTTCTGCTTAAAGCTTATTTT TTATGGAGACAGCTTTCTTAAATTCCAGUTAGGCAAGAAGGGCTCAATTGATCTAAA TTTCATTAATGCTTCAGATTAACATAGACATGCTAAAGCTTACGGAGCTTCTGCTTAAAG TAGTATTCTGCTTAATGCTCACAAACACTCATGCTTGCACGATGCCAGTCTGCTCTGCA GACPTTACATTTAAATGCTTAGRCCTSCAGGAACAAAGAGGAGCTAACCTTACCTTACG TTTGCTTCTAGGAAAGAAAAACCATCAGGACTTATTGTTCTGTTCTGTTCTGCTTAC TGAGGAGTAAATTGGCTGGCTGACAAACATACCATCTAACATGCTAACGACAAAGCTGATAG TTTCTGGCTGTTCTAAATTCTGGATGAGAACCTGTTGGCTAGAACGCTCTGGCTGATTGAC TTTCACGGCATACTTCACTTCTCTTACLAATCTGGACTACAAATTAAATCTGATGCTT CTAAACGATTAAGCTGAGCTGAGCTGAGGGCTCAAGTCATGCTACCTGGCTTATCCAATTTAGGCTC ACTGCTCACAAAAGACGGGCTTCTGGCTGGCTTATTCTCTTGGCTCTGCTCTGCTC AACAAATGATGGACATAGGCTAGCTTCAAGGGGKAACTGACCTCAGTGAGCTGATAATT ACCCTACTATTACATGTGCTTAAATACACCAGAACCTGCCACCTAACCTAACCTGTTGAA TCTAAAAAACTACTCTTAAACCAAGAAATGGCTTAAATTAGTGTATAATGCAAGCTGATCTG</p> |
|---|--|

ES 2 998 052 T3

ES 2 998 052 T3

AAGTGAAGAAGAACCTTTAAGAAGTGTGCGACTTGCCTTGAACATCAACACCATTGGGAGATGG
CCTGTGTTCCGAAATGCTGATATTCTAAGTAATAACAGTGCAACTATCAATAAGAGAAATCTCTT
CAGCCCATTGAAAGGGATAGCAAATTAAAATGTCTGAGGGCTTTTCATAGTCGGCATTTCTC
CCCAAGCTCAAACCTACTATTATCTTTCCACAGGATTCAAGACCAAATTATTCTAATAGATAC
ACACCATGCTTATGTTAATAATATTCCATATACCAGTCCCAGGGTAGAATCATCTCCCCATTC
GGCATTATTGTCAATATCTGCAAAGCCAAGGAGGTGAGGTCAAGGAAGGGTCAGGATCACAG
CCTCTGGTCTGGAGAGAGCACTGGAATGGAGATAATAAGGCTGGATTTTACTTCAGATTCTCCC
CTGGCTTCTGGTTCTGGCTCATCTGTCAGATCCATGGACTCCAAATTGGCATATGCAATT
ATGACAGGATCTGAGTCTATATGATAATCTCAGGAAACAGACAACAGAGTAATGACAGATGCA
AAACGAATGATAATTAAACACCCACAGCAGAGGCCCTGTCAAAATGACCTCTTGCATGCTTCT
TATTAGGATATAATGTTAAACACAGGAGCAGAAGAAGAAACAGCTTTGAAATGCAAAAAAG
GTAGGTTGCTATTGCTAATTCTATGAAATGCCCTAAACATAAAGAAGCTTTAGGCTGATCAT
ATTGACAACCCAGTCTTGTGATCAGGAACTTGTGAACTTGGAAATAAAACAGGAAACACAATT
GTCAAATTGACACCTCTCTGGTCCCTGTGAAATTGGAAAGACTTGTACATATAATTGAA
AAGGATGTGTTCTTAAATGCCATGATACAAATCTGAAGAAATCCATTGTTCAATACCTTA
ATACAACCAACCATAACGCTGATACCACCTACAGTGGATAACAAGACAGGAAAGTCATCATTG
GTAACAGTGGCATCATCACTCATTGATAACAGTTTCTATGGGCACTGAGTGGCCGGTGGAGCCTC
TGGGATCAAGGAGTGACAATGTCACAGTGTCTTATTGCCCCGGTCTTAAAGTGAGAGCATCC
TGAACATCTCAGGGTTGGAAAGGAACACTTGAGGTTCTCAATCCAGCACCCTCCACAAACAAAA
TCTCCTTCACAATAACACTCACCTCCACCTCTGATCAACATGTCAGGGATGAGGACCTTCC
ACCTCATAAGGCAGCCTGATCCCTTTGAATGGCTTAATACCAACATTACTACTACTCC
AGAGAAGTCTTCCCTCAAGTCAAACCTTGTCTTATATCTCCACTATTGGCCAGTTCTG
CTCTTGAGGCCCTACTAAACAAAGTATAATTGCTCTCCTACCCAGCAGCTGTCCAGATATGGAAG
ACAGCAATCATGCTGGCCAACCTTGACTGAGCTTTCTCTCCAGGCTAAAGATCCCTGATGTC
TTCCACTGTTCTCTATGACCCTTCCAGGACCTTCTCTGCCACTCACCTCCTTCTGGAC
ACACTAACGTTTCTCAGGTTAGAATGTTGCAATCGCAAAACAAATACAATAAATGCGGAGAAGTGA
CTTCAGCACCACATTATGCCAAAGACCCCTGTTGTTACACACAATTATATACAAACTTGGT
AACATGATGTTATGGCTCTGCAAATTAAACACACATACATACACAGGGATAGAGA
TACTATTCTGAGGAAACAGAGTACTCAGACCTTGCCTTAACGTTGATGTTCTGGATACTAAATGG
TCATCCCACITCCATCAAGCTTATCTTCAAGATGACTGCAACATAATGTTGAGTAATACTACAC
GCTGCTGTTAATTACAGAAATCTGAGGAAACACTTATGAGATGCTGCTAGAAGTCTTCA
GGAAATGATAATTAAACAAACTGGTCAATTAGGTCATGCAATTAAACTCAACATTATAGGCA
CTTACAAAGTGCCAAATATCAGGCTCATACGGACAAAAGAAACTTCCACACAGTCTGCGCT
TAGAAGATTGACACATCTCATAGGGAGCAGGGCTTAACACAAGAAATAAAGAGACAGATA
ATAGTCAGCACTTCTGACAATTCAAGAACATAAGAATCTACTAAGTGTGAGACTTTGG
AGCCCACAAATCCCCAGTGTATGGAGTTCTCTAAAGCAAGATCAGGTATGTTAATGAA
GACAGTGTGTTAGCCTGTCATATGTTACAGGAAACACTTATGAGTGTGCTAGTGTAGGAG
AATGAGACCGATGAAACACCAATTACTCAATTACTCAATTAGAACAAACTTTCTCAGTGG
CACACATCTCAGACCAATTATCTTCTTTATATATCTTATTTCCACACTCTTATTACACA
GTTGAGAACTGGACCAGATAGTTGTCAGGCTTACATGTCATAAGTGAGGCCAGTGG
TTCCATTCTCAGGTGATCAGAATCTCAAATTGCGGCACATGTCATAAGTGAGGCCAGTAAAC
AACATCTGTAAGTCACACAGCATCTGAGCGGTAGCCACCCAAGGGGAAAGGCTGGGATGCCAG
TCATGTTACCTAATGGTAAACCTCTTCTCCCTGGACCAATTACAAACCTACCCCTACAC
TTCTCTATTCCCTCTTGTCTCAAAGTGAGTTCAAATGCAAGATGGGACTTAGAGGGACAAA
AGGAGGTGGAATGCAATCTGGATGTTCTCATTATGTTCTGCTCAATGGCTGATTCTAAATGATGA
ATTACTGGGTGGAGGGACATTGTTCTGACACATAGAAGAAATGGCATGAGTGCACCTCTGACT
GGGAGCATCCCTCTCTAACCCATCTTCACTGAGTGGAAATGGGCTCATGGGTATTTCTGCC
ATCTGCAATCCCTGATCATTAAGCTCACCTCACTGAGGCCAACCTCAGGGAAAGTAAAGTAA
AATCATTCTGTAAGATCAATAGGTCCAAGACGTTACATTTCACATGAAAGTAACACAGAC
ATATTGTGATCTTCAACTCTGAACGATTATTTCCATATACGTTCTGCCACATTCTAGCCTT
TAGATATTCTGAAATGTCATCTGCGATAACTGGTGCACAAAGAAATATGTCATCTGATAA
ATGGATGGAAACATGCACTAACATAAGTCTCCCATCACACATAAAGGCAAGAGCGTCAGAGGAG
TCTTGAAGGAAATTCTACACACTGCTCCGGAAATGGACTTCAAGCAGTGCACTGTTGTCATATG
TGTATGTTGTCAGGGAGAAAGAGAGATGGCAGAGAGAGAGAAGAAAGACACTGCTTCTAC
TGAAGTGGCTTGGCTCTCAGTAGGCGTAACACATGGACAGTTATCATTATCATGGATCATGGTA
CCAAAGTAAGAGCACTGAATAGGGAGTTTGAACACTGGGATTCAAGGACCATGACCACTGCTG
CTGGGTGACCTTGAGCAAGACCTTACCTATGCCACAGTTCTACTTCACCTACTTACAGGGT
GGCTTGAGCATCAAATCAGCTAATGTCAGGAAACTGATGCTGTCAGTGCTGACAACCGTAAG
GTGACACTACTTAGTTACTTCAACATGGCTTAGATGTCACAAAGGGTAGACATAAAGCCCTCACTA
ATACCACTTAGTTACACATATTAAATAATTGTCAGTACCCCTCTCTGGATCAGATG
ACAACAAACAGAGAAATCTCCTAGAAGAATAGCTCCACTGGTCITTCTTGCCTGATCTAAACC

- [0146] En algunas realizaciones, la carga comprende un intrón de una secuencia genética. En algunas realizaciones, la carga comprende una secuencia de ADNc. En algunas realizaciones, la carga comprende un elemento regulador transcripcional. En algunas realizaciones, la carga comprende un complemento inverso de una secuencia de sustitución, un exón de una secuencia génica, un intrón de una secuencia génica, una secuencia de ADNc o un elemento regulador transcripcional. En algunas realizaciones, la carga comprende una porción de una secuencia de sustitución, un exón de una secuencia génica, un intrón de una secuencia génica, una secuencia de ADNc o un elemento regulador transcripcional.

[0147] Las secuencias de reemplazo en plantillas donantes se han descrito en otros lugares, incluyendo en Cotta-Ramusino *et al.* Una secuencia de reemplazo puede ser de cualquier longitud adecuada (incluyendo cero nucleótidos, cuando el resultado de reparación deseado es una delección), y típicamente incluye una, dos, tres o más modificaciones de secuencia relativas a la secuencia de origen natural dentro de una célula en la que se desea la edición. Una modificación común de la secuencia implica la alteración de la secuencia natural para reparar

una mutación que se refiere a una enfermedad o condición cuyo tratamiento se desea. Otra modificación común de la secuencia implica la alteración de una o más secuencias que son complementarias a, o codifican para, la secuencia PAM de la nucleasa guiada por ARN o el dominio diana del ARNg(s) que se utiliza para generar un SSB o DSB, para reducir o eliminar la escisión repetida del sitio diana después de que la secuencia de sustitución se haya incorporado al sitio diana.

[0148] En algunas realizaciones, la plantilla donante de oligonucleótidos puede comprender además una secuencia que codifica un informador. En una realización, el reportero es una proteína de fluorescencia verde (GFP), una proteína de fluorescencia amarilla (YFP), DS-Red o luciferasa. En una realización, el informador puede ser un marcador que puede detectarse en células vivas utilizando técnicas estándar, *por ejemplo*, antígenos de superficie detectables mediante anticuerpos, péptidos que catalizan o facilitan de otro modo una reacción química que produce un producto detectable ópticamente, *por ejemplo*, luciferasa. En una realización, el informador es un marcador seleccionable. En una realización ejemplar, el marcador seleccionable es el receptor del factor de crecimiento nervioso de baja afinidad (NGFR). En otras realizaciones, el marcador seleccionable puede ser una proteína resistente a fármacos (tal como resistencia a neomicina o puromicina, versión mutante del gen MGMT), una versión truncada de una proteína de la superficie celular (*por ejemplo*, CD 19, EGFR), un gen que confiere una ventaja selectiva de crecimiento y/o injerto tras el trasplante *in vivo* de las células editadas (*por ejemplo* CXCR4, CD47, receptor de IL2) o una proteína reportera fluorescente (*por ejemplo*, proteína verde fluorescente (GFP), proteína azul fluorescente (BFP), proteína amarilla fluorescente (YFP), etc., o una combinación de las mismas.

C. Rellenos

[0149] En algunas realizaciones, la plantilla donante puede comprender opcionalmente una o más secuencias de relleno. Generalmente, una secuencia de relleno es una secuencia de ácido nucleico heteróloga o aleatoria que se ha seleccionado para (a) facilitar (o no inhibir) la integración dirigida de una plantilla donante de la presente divulgación en un sitio diana y la amplificación posterior de un amplicón que comprende la secuencia de relleno según ciertos métodos de esta divulgación, pero (b) para evitar dirigir la integración de la plantilla donante en otro sitio. La secuencia de relleno puede colocarse, por ejemplo, entre un brazo de homología y una secuencia de exón CD40L para ajustar el tamaño del amplicón que se generará cuando la secuencia de plantilla donante se integre en el sitio diana. Tales ajustes de tamaño pueden emplearse, por ejemplo, para equilibrar el tamaño de los amplicones producidos por sitios diana integrados y no integrados y, en consecuencia, para equilibrar las eficiencias con las que se produce cada amplicón en una única reacción de PCR; esto, a su vez, puede facilitar la evaluación cuantitativa de la tasa de integración diana basada en la abundancia relativa de los dos amplicones en una mezcla de reacción.

[0150] Para facilitar la integración y amplificación diana, la secuencia de relleno puede seleccionarse para minimizar la formación de estructuras secundarias que puedan interferir con la resolución del sitio de corte por la maquinaria de reparación del ADN (*por ejemplo*, mediante recombinación homóloga) o que puedan interferir con la amplificación.

[0151] En una realización, el primer relleno tiene una secuencia que comprende al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 55, al menos 60, al menos 65, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 105, al menos 110, al menos 115, al menos 120, al menos 125, al menos 130, al menos 135, al menos 140, al menos 145, al menos 150, al menos 155, al menos 160, al menos 165, al menos 170, al menos 175, al menos 180, al menos 185, al menos 190, al menos 195, al menos 200, al menos 205, al menos 210, al menos 215, al menos 220, al menos 225, al menos 230, al menos 235, al menos 240, al menos 245, al menos 250, al menos 275, al menos 300, al menos 325, al menos 350, al menos 375, al menos 400, al menos 425, al menos 450, al menos 475, o al menos 500 nucleótidos. En otra realización, el segundo relleno tiene una secuencia que comprende al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 55, al menos 60, al menos 65, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 105, al menos 110, al menos 115, al menos 120, al menos 125, al menos 130, al menos 135, al menos 140, al menos 145, al menos 150, al menos 155, al menos 160, al menos 165, al menos 170, al menos 175, al menos 180, al menos 185, al menos 190, al menos 195, al menos 200, al menos 205, al menos 210, al menos 215, al menos 220, al menos 225, al menos 230, al menos 235, al menos 240, al menos 245, al menos 250, al menos 275, al menos 300, al menos 325, al menos 350, al menos 375, al menos 400, al menos 425, al menos 450, al menos 475, o al menos 500 nucleótidos.

[0152] Es preferible que la secuencia de relleno no interfiera con la resolución del sitio de escisión en el ácido nucleico diana. Así, la secuencia de relleno debe tener una identidad de secuencia mínima con la secuencia de ácido nucleico en el sitio de corte del ácido nucleico diana. En algunas realizaciones, la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 %, o 10 % idéntica a cualquier secuencia de ácido nucleico dentro de 500, 450, 400, 350, 300, 250, 200, 150, 100, 50 nucleótidos del sitio de escisión del ácido nucleico diana. En algunas realizaciones, la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 %, o 10 % idéntica a cualquier secuencia de ácido nucleico

dentro de 500, 450, 400, 350, 300, 250, 200, 150, 100, 50 pares de bases desde el sitio de escisión del ácido nucleico diana.

[0153] Para evitar eventos de recombinación molecular fuera del objetivo, es preferible que la secuencia de relleno tenga una homología mínima con una secuencia de ácido nucleico en el genoma de la célula diana. En algunas realizaciones, la secuencia de relleno tiene una identidad de secuencia mínima con un ácido nucleico del genoma de la célula diana. En algunas realizaciones, la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntica a cualquier secuencia de ácido nucleico de la misma longitud (medida en pares de bases o nucleótidos) en el genoma de la célula diana. En algunas realizaciones, un tramo de 20 pares de bases de la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntico a cualquier tramo de al menos 20 pares de bases de ácido nucleico del genoma de la célula diana. En algunas realizaciones, un tramo de 20 nucleótidos de la secuencia de relleno es menos del 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntico a cualquier tramo de al menos 20 nucleótidos de ácido nucleico del genoma de la célula diana.

[0154] En algunas realizaciones, la secuencia de relleno tiene una identidad de secuencia mínima con una secuencia de ácido nucleico en la plantilla donante (*por ejemplo*, la secuencia de ácido nucleico de la carga, o la secuencia de ácido nucleico de un sitio de cebado presente en la plantilla donante). En algunas realizaciones, la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntica a cualquier secuencia de ácido nucleico de la misma longitud (medida en pares de bases o nucleótidos) en la plantilla donante. En algunas realizaciones, un tramo de 20 pares de bases de la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntico a cualquier tramo de 20 pares de bases de ácido nucleico de la plantilla donante. En algunas realizaciones, un tramo de 20 nucleótidos de la secuencia de relleno es menos del 80 %, 70 %, 60 %, 55 %, 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 % o 10 % idéntico a cualquier tramo de 20 nucleótidos de ácido nucleico de la plantilla donante.

[0155] Las secuencias de relleno se describen con más detalle en, la solicitud provisional estadounidense n.º 62/532509, presentada el 14 de julio de 2017.

30 Nucleasas guiadas por ARN

[0156] Las nucleasas guiadas por ARN según la presente divulgación incluyen, pero no se limitan a, nucleasas CRISPR de clase 2 de origen natural tales como Cas9 y Cpf1, así como otras nucleasas derivadas u obtenidas a partir de las mismas. En términos funcionales, las nucleasas guiadas por ARN se definen como aquellas nucleasas que: (a) interactuar con (*por ejemplo*, complejarse con) un ARNg; y (b) junto con el ARNg, asociarse con, y opcionalmente escindir o modificar, una región diana de un ADN que incluye (i) una secuencia complementaria al dominio que dirige el ARNg y, opcionalmente, (ii) una secuencia adicional a la que se refiere como "motivo adyacente al protoespaciador", o "PAM", que se describe con más detalle a continuación. Como se ilustrará en los siguientes ejemplos, las nucleasas guiadas por ARN pueden definirse, en términos generales, por su especificidad PAM y su actividad de corte, aunque puedan existir variaciones entre nucleasas guiadas por ARN individuales que compartan la misma especificidad PAM o actividad de corte. Los expertos apreciarán que algunos aspectos de la presente divulgación se refieren a sistemas, métodos y composiciones que pueden implementarse utilizando cualquier nucleasa guiada por ARN adecuada que tenga una cierta especificidad de PAM y/o actividad de corte. Por esta razón, a menos que se especifique lo contrario, el término nucleasa guiada por ARN debe entenderse como un término genérico, y no se limita a ningún tipo particular (*por ejemplo*, Cas9 frente a Cpf1), especie (*por ejemplo*, *S. pyogenes* frente a *S. aureus*) o variación (*por ejemplo*, de longitud completa frente a truncada o dividida; especificidad PAM de origen natural frente a especificidad PAM de ingeniería, etc.) de la nucleasa guiada por ARN.

[0157] La secuencia PAM toma su nombre de su relación secuencial con la secuencia "protospacer" que es complementaria a los dominios (o "espaciadores") que dirigen los ARNg. Junto con las secuencias de protoespaciadores, las secuencias PAM definen regiones o secuencias diana para combinaciones específicas de ARN guiado por nucleasas / ARNg.

[0158] Diversas nucleasas guiadas por ARN pueden requerir diferentes relaciones secuenciales entre las PAM y los protoespaciadores. En general, las Cas9 reconocen las secuencias PAM que se encuentran a 3' del protoespaciador como se visualiza en la hebra inferior o no complementaria:

| | |
|-----------------------------|-------------------|
| 5'-----[protospacer]-----3' | complementario |
| 3'-----[PAM] 5' | no complementario |

[0159] Cpf1, por otra parte, reconoce generalmente las secuencias PAM que están a 5' del protospacer como se visualiza en la hebra inferior o no complementaria:

| | |
|-----------------------------|-------------------|
| 5'-----[protospacer]-----3' | complementario |
| 3'-----[PAM]-----5' | no complementario |

[0160] Además de reconocer orientaciones secuenciales específicas de PAM y protoespaciadores, las nucleasas guiadas por ARN también pueden reconocer secuencias específicas de PAM. *S. aureus* Cas9, por ejemplo, reconoce una secuencia PAM de NNGRRT o NNGRRV, en el que los residuos N están inmediatamente 3' de la

5 región reconocida por el dominio de diana del ARNg. *S. pyogenes* Cas9 reconoce las secuencias NGG PAM. Y *F. novicida* Cpf1 reconoce una secuencia TTN PAM. Se han identificado secuencias PAM para una variedad de nucleasas guiadas por ARN, y una estrategia para identificar nuevas secuencias PAM ha sido descrita por Shmakov et al., 2015, Molecular Cell 60, 385-397, 5 de noviembre de 2015. También debe tenerse en cuenta que las nucleasas guiadas por ARN manipuladas pueden tener especificidades de PAM diferentes de las especificidades de PAM de las moléculas de referencia (por ejemplo, en el caso de una nucleasa guiada por ARN manipulada, la molécula de referencia puede ser la variante natural de la que se deriva la nucleasa guiada por ARN, o la variante natural que tenga la mayor homología de secuencia de aminoácidos con la nucleasa guiada por ARN manipulada).

10 [0161] Además de por su especificidad PAM, las nucleasas guiadas por ARN pueden caracterizarse por su actividad de corte del ADN: las nucleasas guiadas por ARN que se encuentran en la naturaleza suelen formar DSB en los ácidos nucleicos diana, pero se han producido variantes de ingeniería que generan solo SSB (ya comentadas) Ran & Hsu, et al., Cell 154(6), 1380-1389, 12 de septiembre de 2013 (Ran)), o que no cortan en absoluto.

20 Cas9

[0162] Se han determinado estructuras cristalinas para *S. pyogenes* Cas9 (Jinek 2014), y para *S. aureus* Cas9 en complejo con un ARN guía unimolecular y un ADN diana (Nishimasu 2014; Anders 2014; y Nishimasu 2015).

25 [0163] Una proteína Cas9 natural consta de dos lóbulos: un lóbulo de reconocimiento (REC) y un lóbulo de nucleasa (NUC), cada uno de los cuales comprende determinados dominios estructurales y/o funcionales. El lóbulo REC comprende un dominio de hélice puente (BH) rico en arginina, y al menos un dominio REC (por ejemplo, un dominio REC1 y, opcionalmente, un dominio REC2). El lóbulo REC no comparte similitud estructural con otras proteínas conocidas, lo que indica que se trata de un dominio funcional único. Sin querer ceñirnos a ninguna teoría, 30 los análisis mutacionales sugieren papeles funcionales específicos para los dominios BH y REC: el dominio BH parece desempeñar un papel en el reconocimiento ARNg:ADN, mientras que se cree que el dominio REC interactúa con el dúplex repetición:antirrepetición del ARNg y media en la formación del complejo Cas9/ARNg.

35 [0164] El lóbulo NUC comprende un dominio RuvC, un dominio HNH y un dominio de interacción PAM (PI). El dominio RuvC comparte similitud estructural con los miembros de la superfamilia de integrasas retrovirales y escinde la cadena no complementaria (es decir, inferior) del ácido nucleico diana. Puede estar formado por dos o más motivos RuvC divididos (tales como RuvC I, RuvCII, y RuvCIII en *s. pyogenes* y *s. aureus*). El dominio HNH, por su parte, es estructuralmente similar a los motivos de endonucleasas HNN, y escinde la cadena complementaria (es decir, superior) del ácido nucleico diana. El dominio PI, como su nombre indica, contribuye a 40 la especificidad de las PAM.

45 [0165] Si bien ciertas funciones de Cas9 están vinculadas a los dominios específicos antes mencionados (aunque no necesariamente determinadas por ellos), éstas y otras funciones pueden estar mediadas o influenciadas por otros dominios de Cas9, o por múltiples dominios en cualquiera de los lóbulos. Por ejemplo, en *S. pyogenes* Cas9, como se describe en Nishimasu 2014, el dúplex repetición: antirrepetición del ARNg cae en un surco entre los lóbulos REC y NUC, y los nucleótidos del dúplex interactúan con aminoácidos en los dominios BH, PI y REC. Algunos nucleótidos en la estructura del primer bucle del tallo también interactúan con aminoácidos en múltiples dominios (PI, BH y REC1), al igual que algunos nucleótidos en el segundo y tercer bucles del tallo (dominios RuvC y PI).

50 Cpf1

55 [0166] La estructura cristalina de *Acidaminococcus sp.* Cpf1 en complejo con ARNcr y una diana de ADN de doble cadena (ds) que incluye una secuencia TTTN PAM ha sido resuelta por Yamano et al. (Cell. 2016 May 5; 165(4): 949-962 (Yamano)). Cpf1, al igual que Cas9, tiene dos lóbulos: un lóbulo REC (reconocimiento) y un lóbulo NUC (nucleasa). El lóbulo REC incluye los dominios REC1 y REC2, que carecen de similitud con cualquier estructura proteica conocida. El lóbulo NUC, por su parte, incluye tres dominios RuvC (RuvC-I, -II y -III) y un dominio BH. Sin embargo, a diferencia de Cas9, el lóbulo REC de Cpf1 carece de un dominio HNH, e incluye otros dominios que también carecen de similitud con estructuras proteicas conocidas: un dominio PI estructuralmente único, tres dominios Wedge (WED) (WED-I, -II y -III), y un dominio nucleasa (Nuc).

5 [0167] Aunque Cas9 y Cpf1 comparten similitudes en estructura y función, debe apreciarse que ciertas actividades de Cpf1 están mediadas por dominios estructurales que no son análogos a ninguno de los dominios de Cas9. Por ejemplo, el corte de la cadena complementaria del ADN diana parece estar mediado por el dominio Nuc, que difiere secuencial y espacialmente del dominio HNH de Cas9. Además, la porción no dirigida del ARNg Cpf1 (el asa) adopta una estructura pseudoknot, en lugar de una estructura de bucle de tallo formada por el dúplex repetición:antirrepetición en los ARNg Cas9.

Modificaciones de las nucleasas guiadas por ARN

10 [0168] Las nucleasas guiadas por ARN descritas anteriormente tienen actividades y propiedades que pueden ser útiles en una variedad de aplicaciones, pero el artesano experto apreciará que las nucleasas guiadas por ARN también pueden modificarse en ciertos casos, para alterar la actividad de corte, la especificidad PAM u otras características estructurales o funcionales.

15 [0169] Pasando primero a las modificaciones que alteran la actividad de escisión, se han descrito anteriormente mutaciones que reducen o eliminan la actividad de dominios dentro del lóbulo NUC. Mutaciones ejemplares que pueden realizarse en los dominios RuvC, en el dominio HNH de Cas9 o en el dominio Nuc de Cpf1 se describen en Ran y Yamano, así como en Cotta-Ramusino. En general, las mutaciones que reducen o eliminan la actividad en uno de los dos dominios nucleasa dan lugar a nucleasas guiadas por ARN con actividad nicasa, pero debe tenerse en cuenta que el tipo de actividad nicasa varía en función de qué dominio se inactiva. Por ejemplo, la inactivación de un dominio RuvC de una Cas9 dará lugar a una mellasa que corta la cadena complementaria o superior. Por otra parte, la inactivación de un dominio HNH de Cas9 da lugar a una mellasa que corta la hebra inferior o no complementaria.

25 [0170] Kleinstiver et al. han descrito modificaciones de la especificidad PAM en relación con moléculas de referencia Cas9 naturales tanto para *S. pyogenes* (Kleinstiver et al., Nature. 2015 Jul 23;523(7561):481-5 (Kleinstiver I) y *S. aureus* (Kleinstiver et al., Nat Biotechnol. 2015 Dic; 33(12): 1293-1298 (Kleinstiver II)). Kleinstiver et al. también han descrito modificaciones que mejoran la fidelidad diana de Cas9 (Nature, 2016 January 28; 529, 490-495 (Kleinstiver III)).

30 [0171] Las nucleasas guiadas por ARN se han dividido en dos o más partes, como describen Zetsche et al. (Nat Biotechnol. 2015 Feb; 33(2):139-42 (Zetsche II)), y por Fine et al. (Sci. Rep. 2015 Jul 1;5:10777 (Fine)).

35 [0172] Las nucleasas guiadas por ARN pueden, en ciertas realizaciones, optimizarse en tamaño o truncarse, por ejemplo mediante una o más delecciones que reduzcan el tamaño de la nucleasa al tiempo que conservan la asociación con el ARNg, el reconocimiento de dianas y PAM, y las actividades de corte. En ciertas realizaciones, las nucleasas guiadas por ARN están unidas, de forma covalente o no covalente, a otro polipéptido, nucleótido u otra estructura, opcionalmente mediante un enlazador. Algunos ejemplos de nucleasas y enlazadores unidos se describen en Guilinger et al., Nature Biotechnology 32, 577-582 (2014).

40 [0173] Las nucleasas guiadas por ARN también incluyen opcionalmente una etiqueta, tal como, pero no se limitan a, una señal de localización nuclear para facilitar el movimiento de la proteína nucleasa guiada por ARN al núcleo. En ciertas realizaciones, la nucleasa guiada por ARN puede incorporar señales de localización nuclear C- y/o N-terminal. Las secuencias de localización nuclear son conocidas en la técnica y se describen en Maeder y en otros lugares.

45 [0174] La anterior lista de modificaciones pretende ser de naturaleza ejemplar, y el artesano experto apreciará, a la vista de la presente divulgación, que otras modificaciones pueden ser posibles o deseables en ciertas aplicaciones. Por lo tanto, en aras de la brevedad, los sistemas, métodos y composiciones ejemplares de la presente divulgación se presentan con referencia a nucleasas guiadas por ARN concretas, pero debe entenderse que las nucleasas guiadas por ARN utilizadas pueden modificarse de formas que no alteren sus principios de funcionamiento. Tales modificaciones están dentro del alcance de la presente divulgación.

Ácidos nucleicos que codifican nucleasas guiadas por ARN

50 [0175] En el presente documento se proporcionan ácidos nucleicos que codifican nucleasas guiadas por ARN, por ejemplo, Cas9, Cpf1 o fragmentos funcionales de las mismas. Anteriormente se han descrito ácidos nucleicos ejemplares que codifican nucleasas guiadas por ARN (véase, por ejemplo, Cong 2013; Wang 2013; Mali 2013; Jinek 2012).

55 [0176] En algunos casos, un ácido nucleico que codifica una nucleasa guiada por ARN puede ser una secuencia de ácido nucleico sintético. Por ejemplo, la molécula sintética de ácido nucleico puede modificarse químicamente. En ciertas realizaciones, un ARNm que codifica una nucleasa guiada por ARN tendrá una o más (por ejemplo, todas) de las siguientes propiedades: puede estar capado; poliadenilado; y sustituido con 5-metilcitolina y/o pseudouridina.

5 [0177] Las secuencias sintéticas de ácidos nucleicos también pueden estar optimizadas en cuanto a codones, por ejemplo, al menos un codón no común o un codón menos común ha sido sustituido por un codón común. Por ejemplo, el ácido nucleico sintético puede dirigir la síntesis de un ARNm mensajero optimizado, *por ejemplo*, optimizado para la expresión en un sistema de expresión de mamíferos, *por ejemplo*, descrito en el presente documento. En Cotta-Ramusino se presentan ejemplos de secuencias codificantes de Cas9 optimizadas mediante codones.

10 [0178] Además, o alternativamente, un ácido nucleico que codifica una nucleasa guiada por ARN puede comprender una secuencia de localización nuclear (NLS). Las secuencias de localización nuclear son conocidas en la técnica.

Análisis funcional de las moléculas candidatas

15 [0179] Las nucleasas candidatas guiadas por ARN, los ARNg y sus complejos pueden evaluarse mediante métodos estándar conocidos en la técnica. Véase, por ejemplo, Cotta-Ramusino. La estabilidad de los complejos RNP puede evaluarse mediante fluorimetría diferencial de barrido, como se describe a continuación.

Fluorimetría diferencial de barrido (DSF)

20 [0180] La termoestabilidad de los complejos de ribonucleoproteínas (RNP) que comprenden ARNg y nucleasas guiadas por ARN puede medirse mediante DSF. La técnica DSF mide la termoestabilidad de una proteína, que puede aumentar en condiciones favorables tales como la adición de una molécula de ARN de unión, *por ejemplo*, un ARNg.

25 [0181] Un ensayo DSF puede realizarse de acuerdo con cualquier protocolo adecuado, y puede emplearse en cualquier entorno adecuado, incluyendo sin limitación (a) la prueba de diferentes condiciones (*por ejemplo*, diferentes relaciones estequiométricas de ARNg: proteína nucleasa guiada por ARN, diferentes soluciones tampón, etc.) para identificar las condiciones óptimas para la formación de la RNP; y (b) probar modificaciones (*por ejemplo*, modificaciones químicas, alteraciones de la secuencia, etc.) de una nucleasa guiada por ARN y/o un ARNg para identificar aquellas modificaciones que mejoran la formación o estabilidad de la RNP. Una lectura de un ensayo DSF es un cambio en la temperatura de fusión del complejo RNP; un cambio relativamente alto sugiere que el complejo RNP es más estable (y por tanto puede tener una mayor actividad o una cinética de formación más favorable, una cinética de degradación u otra característica funcional) en relación con un complejo RNP de referencia caracterizado por un cambio menor. Cuando el ensayo DSF se utiliza como herramienta de cribado, puede especificarse un umbral de cambio de temperatura de fusión, de modo que la salida sea una o más RNP que tengan un cambio de temperatura de fusión igual o superior al umbral. Por ejemplo, el umbral puede ser de 5-10 °C (*por ejemplo*, 5°, 6°, 7°, 8°, 9°, 10°) o más, y la salida puede ser una o más RNP caracterizadas por un desplazamiento de la temperatura de fusión mayor o igual que el umbral.

40 [0182] A continuación se exponen dos ejemplos no limitativos de condiciones de ensayo de DSF: Para determinar la mejor solución para formar complejos de RNP, se dispensa en una placa de 384 pocillos una concentración fija (*por ejemplo*, 2 µM) de Cas9 en agua+10x SYPRO Orange® (Life Technologies cat#S-6650). A continuación, se añade una cantidad equimolar de ARNg diluido en soluciones con pH y sal variables. Tras incubar a temperatura ambiente durante 10 minutos y centrifugar brevemente para eliminar cualquier burbuja, se utiliza un Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler con el software Bio-Rad CFX Manager para ejecutar un gradiente de 20 °C a 90 °C con un aumento de temperatura de 1 °C cada 10 segundos.

50 [0183] El segundo ensayo consiste en mezclar varias concentraciones de ARNg con una concentración fija (*por ejemplo*, 2 µM) de Cas9 en el tampón óptimo del ensayo 1 anterior e incubar (*por ejemplo*, a RT durante 10') en una placa de 384 pocillos. Se añade un volumen igual de tampón óptimo + 10x SYPRO Orange® (Life Technologies cat#S-6650) y se sella la placa con adhesivo Microseal® B (MSB-1001). Tras una breve centrifugación para eliminar cualquier burbuja, se utiliza un Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler con el software Bio-Rad CFX Manager para ejecutar un gradiente de 20 °C a 90 °C con un aumento de temperatura de 1 °C cada 10 segundos.

Estrategias de edición del genoma

55 [0184] Los sistemas de edición genómica descritos anteriormente se utilizan, en diversas realizaciones de la presente divulgación, para generar ediciones en (es decir, para alterar) regiones diana de ADN dentro de u obtenidas de una célula. En el presente documento se describen varias estrategias para generar ediciones particulares, y estas estrategias se describen generalmente en términos del resultado de reparación deseado, el número y la posición de las ediciones individuales (*por ejemplo*, SSB o DSB), y los sitios diana de dichas ediciones.

60 [0185] Las estrategias de edición del genoma que implican la formación de SSB o DSB se caracterizan por resultados de reparación que incluyen: (a) supresión total o parcial de una región diana; (b) inserción o sustitución total o parcial de una región diana; o (c) interrupción total o parcial de una región diana. Esta agrupación no

pretende ser limitativa ni vinculante para ninguna teoría o modelo en particular, y se ofrece únicamente por economía de presentación. Los artesanos expertos comprenderán que los resultados enumerados no se excluyen mutuamente y que algunas reparaciones pueden dar lugar a otros resultados. No debe entenderse que la descripción de una estrategia o método de edición concreto requiera un resultado de reparación determinado, a menos que se especifique lo contrario.

[0186] La sustitución de una región diana suele implicar la sustitución total o parcial de la secuencia existente dentro de la región diana por una secuencia homóloga, por ejemplo mediante la corrección o la conversión génica, dos resultados de reparación mediados por las vías de HDR. La HDR se promueve mediante el uso de una plantilla donante, que puede ser monocatenaria o bicatenaria, como se describe con más detalle a continuación. Las plantillas de cadena simple o doble pueden ser exógenas, en cuyo caso promoverán la corrección génica, o pueden ser endógenas (*por ejemplo*, una secuencia homóloga dentro del genoma celular), para promover la conversión génica. Las plantillas exógenas pueden tener salientes asimétricos (*es decir*, la porción de la plantilla que es complementaria al sitio del DSB puede estar desplazada en dirección 3' o 5', en lugar de estar centrada dentro de la plantilla donante), por ejemplo, como describen Richardson et al. (Nature Biotechnology 34, 339-344 (2016), (Richardson)). En los casos en que la plantilla es monocatenaria, puede corresponder a la cadena complementaria (superior) o no complementaria (inferior) de la región diana.

[0187] La conversión y la corrección génicas se ven facilitadas, en algunos casos, por la formación de una o más muescas en la región diana o alrededor de ella, como se describe en Ran y Cotta-Ramusino. En algunos casos, se utiliza una estrategia de nicasa dual para formar dos SSB desplazados que, a su vez, forman un único DSB que tiene un saliente (*por ejemplo*, un saliente 5').

[0188] La interrupción y/o delección de toda o parte de una secuencia diana puede conseguirse mediante diversos resultados de reparación. Por ejemplo, una secuencia puede eliminarse generando simultáneamente dos o más DSB que flanqueen una región diana, que se eliminará cuando se reparen los DSB, como se describe en Maeder para la mutación LCA10. Como otro ejemplo, una secuencia puede interrumpirse por una delección generada por la formación de una rotura de doble cadena con salientes monocatenarios, seguida de un procesamiento exonucleolítico de los salientes antes de la reparación.

[0189] Un subconjunto específico de interrupciones de la secuencia diana está mediado por la formación de una indel dentro de la secuencia diana, donde el resultado de la reparación está mediado típicamente por vías NHEJ (incluyendo Alt-NHEJ). La NHEJ se refiere a una vía de reparación "propensa a errores" debido a su asociación con las mutaciones indel. En algunos casos, sin embargo, un DSB es reparado por NHEJ sin alteración de la secuencia que lo rodea (una reparación llamada "perfecta" o "sin cicatriz"); esto generalmente requiere que los dos extremos del DSB estén perfectamente ligados. Por su parte, se cree que los indels surgen del procesamiento enzimático de los extremos libres del ADN antes de su ligación, que añade y/o elimina nucleótidos de una o ambas cadenas de uno o ambos extremos libres.

[0190] Dado que el procesamiento enzimático de los extremos libres de DSB puede ser de naturaleza estocástica, las mutaciones indel tienden a ser variables, se producen a lo largo de una distribución y pueden verse influídas por diversos factores, como el sitio diana específico, el tipo de célula utilizada, la estrategia de edición del genoma empleada, etc. Aun así, es posible extraer generalizaciones limitadas sobre la formación de indels: las delecciones formadas por la reparación de un único DSB suelen estar en el intervalo de 1-50 pb, pero pueden alcanzar más de 100-200 pb. Las inserciones formadas por la reparación de un único DSB tienden a ser más cortas y a menudo incluyen duplicaciones cortas de la secuencia que rodea inmediatamente el lugar de la rotura. Sin embargo, es posible obtener inserciones de gran tamaño y, en estos casos, la secuencia insertada se ha rastreado a menudo hasta otras regiones del genoma o hasta ADN plasmídico presente en las células.

[0191] Las mutaciones indel -y los sistemas de edición genómica configurados para producir indels- son útiles para interrumpir secuencias diana, por ejemplo, cuando no se requiere la generación de una secuencia final específica y/o cuando se toleraría una mutación frameshift. También pueden ser útiles en entornos en los que se prefieren determinadas secuencias, en la medida en que las determinadas secuencias deseadas tienden a producirse preferentemente a partir de la reparación de un SSB o DSB en un sitio determinado. Las mutaciones indel también son una herramienta útil para evaluar o cribar la actividad de determinados sistemas de edición genómica y sus componentes. En estos y otros contextos, los indels pueden caracterizarse por (a) sus frecuencias relativas y absolutas en los genomas de células contactadas con sistemas de edición genómica y (b) la distribución de las diferencias numéricas en relación con la secuencia no editada, *por ejemplo*, ± 1 , ± 2 , ± 3 , etc. Como un ejemplo, en un entorno de búsqueda de pistas, se pueden cribar múltiples ARNg para identificar aquellos ARNg que dirigen más eficientemente el corte en un sitio diana basándose en una lectura de indel en condiciones controladas. Las guías que producen indels con una frecuencia igual o superior a un umbral, o que producen una distribución particular de indels, pueden ser seleccionadas para su posterior estudio y desarrollo. La frecuencia y distribución de indel también puede ser útil como indicador para evaluar diferentes implementaciones o formulaciones de sistemas de edición genómica y métodos de administración, por ejemplo, manteniendo el ARNg constante y variando otras condiciones de reacción o métodos de administración.

Estrategias multiplex

[0192] Mientras que las estrategias ejemplares discutidas anteriormente se han centrado en los resultados de reparación mediados por DSB únicos, los sistemas de edición del genoma de acuerdo con esta divulgación también pueden ser empleados para generar dos o más DSB, ya sea en el mismo emplazamiento o en diferentes loci. Las estrategias de edición que implican la formación de múltiples DSB, o SSB, se describen, por ejemplo, en Cotta-Ramusino.

Diseño de plantillas para donantes

[0193] El diseño de la plantilla de donantes se describe detalladamente en la bibliografía, por ejemplo en Cotta-Ramusino. Las plantillas donantes de oligómeros de ADN (oligodesoxinucleótidos u ODN), que pueden ser monocatenarios (ssODN) o bicatenarios (dsODN), pueden utilizarse para facilitar la reparación de DSB basada en HDR, y son particularmente útiles para introducir alteraciones en una secuencia diana de ADN, insertar una nueva secuencia en la secuencia diana o sustituir la secuencia diana por completo.

[0194] Ya sean monocatenarias o bicatenarias, las plantillas donantes suelen incluir regiones homólogas a regiones de ADN dentro o cerca (*por ejemplo*, flanqueando o contiguas) de una secuencia diana que se va a escindir. En el presente documento se hace referencia a estas regiones homólogas como "brazos de homología", y se ilustran esquemáticamente a continuación:
 [brazo homológico 5'] - [secuencia de sustitución] – [brazo homológico 3'].

[0195] Los brazos homológicos pueden tener cualquier longitud adecuada (incluyendo 0 nucleótidos si sólo se utiliza un brazo homológico), y los brazos homológicos 3' y 5' pueden tener la misma longitud, o pueden diferir en longitud. La selección de las longitudes adecuadas de los brazos homológicos puede verse influida por diversos factores, como el deseo de evitar homologías o microhomologías con determinadas secuencias, tales como las repeticiones Alu u otros elementos muy comunes. Por ejemplo, un brazo homológico 5' puede acortarse para evitar un elemento de repetición de secuencia. En otras realizaciones, un brazo homológico 3' puede acortarse para evitar un elemento de repetición de secuencia. En algunas realizaciones, tanto el brazo homológico 5' como el 3' pueden acortarse para evitar incluir ciertos elementos de repetición de secuencias. Además, algunos diseños de brazos homológicos pueden mejorar la eficacia de la edición o aumentar la frecuencia de un resultado de reparación deseado. Por ejemplo, Richardson et al. Nature Biotechnology 34, 339-344 (2016) (Richardson), descubrieron que la asimetría relativa de los brazos de homología 3' y 5' de las plantillas donantes monocatenarias influía en las tasas de reparación y/o en los resultados.

[0196] Las secuencias de sustitución en las plantillas donantes se han descrito en otros lugares, como en Cotta-Ramusino et al. Una secuencia de sustitución puede tener cualquier longitud adecuada (incluso cero nucleótidos, cuando el resultado de la reparación deseada es una delección), y normalmente incluye una, dos, tres o más modificaciones de la secuencia en relación con la secuencia natural dentro de una célula en la que se desea la edición. Una modificación común de la secuencia implica la alteración de la secuencia natural para reparar una mutación que se refiere a una enfermedad o condición cuyo tratamiento se desea. Otra modificación común de la secuencia implica la alteración de una o más secuencias que son complementarias a, o codifican para, la secuencia PAM de la nucleasa guiada por ARN o el dominio diana del ARNg(s) que se utiliza para generar un SSB o DSB, para reducir o eliminar la escisión repetida del sitio diana después de que la secuencia de sustitución se haya incorporado al sitio diana.

[0197] Cuando se utiliza un ssODN lineal, puede configurarse para (i) recocido a la cadena mellada del ácido nucleico diana, (ii) recocido a la cadena intacta del ácido nucleico diana, (iii) recocido a la cadena positiva del ácido nucleico diana, y/o (iv) recocido a la cadena negativa del ácido nucleico diana. Un ssODN puede tener cualquier longitud adecuada, *por ejemplo*, aproximadamente, al menos, o no más de 150-200 nucleótidos (*por ejemplo*, 150, 160, 170, 180, 190, o 200 nucleótidos).

[0198] Cabe señalar que un ácido nucleico molde también puede ser un vector de ácido nucleico, tal como un genoma viral o ADN circular de doble cadena, *por ejemplo*, un plásmido. Los vectores de ácido nucleico que comprenden plantillas donantes pueden incluir otros elementos codificantes o no codificantes. Por ejemplo, un ácido nucleico molde puede entregarse como parte de un genoma viral (*por ejemplo*, en un genoma AAV o lentiviral) que incluye ciertos elementos de espina dorsal genómica (*por ejemplo*, repeticiones terminales invertidas, en el caso de un genoma AAV) y opcionalmente incluye secuencias adicionales que codifican un ARNg y/o una nucleasa guiada por ARN. En ciertas realizaciones, la plantilla donante puede ser adyacente a, o estar flanqueada por, sitios diana reconocidos por uno o más ARNg, para facilitar la formación de DSB libres en uno o ambos extremos de la plantilla donante que pueden participar en la reparación de SSB o DSB correspondientes formados en el ADN celular utilizando los mismos ARNg. En Cotta-Ramusino se describen ejemplos de vectores de ácidos nucleicos adecuados para su uso como plantillas donantes.

[0199] Sea cual sea el formato utilizado, un ácido nucleico molde puede diseñarse para evitar secuencias indeseables. En ciertas realizaciones, uno o ambos brazos homológicos pueden acortarse para evitar el

solapamiento con ciertos elementos de repetición de secuencias, *por ejemplo*, repeticiones Alu, elementos LINE, etc.

[0200] Además de los brazos de homología, las plantillas donantes pueden comprender elementos adicionales tales como secuencias de relleno y/o secuencias informadoras, *por ejemplo*, un informador GFP (véase la Figura. 5 1), o una secuencia NGFR (véase la Figura. 22). En algunas realizaciones, una plantilla donante comprende un primer brazo homológico, una primera secuencia de relleno, una secuencia de carga, *por ejemplo*, uno o más de los exones 2-5 de CD40L, una secuencia 3' UTR, una secuencia poliA, una secuencia informadora, *por ejemplo*, una secuencia informadora de GFP o una secuencia NGFR, una segunda secuencia de relleno y un segundo brazo 10 homológico. En otra realización, una plantilla donante comprende un primer brazo de homología, una secuencia de carga, *por ejemplo*, una o más de los exones 2-5 de CD40L, una secuencia 3' UTR, una secuencia poliA, una secuencia informadora, *por ejemplo*, una secuencia informadora GFP o una secuencia NGFR, y un segundo brazo 15 de homología. En otra realización, una plantilla donante comprende un primer brazo de homología, una primera secuencia de relleno, una secuencia de carga, *por ejemplo*, uno o más de los exones 2-5 de CD40L, una secuencia 3' UTR, una secuencia poliA, una segunda secuencia de relleno y un segundo brazo de homología. En otra 20 realización, una plantilla donante comprende un primer brazo de homología, una secuencia de carga, *por ejemplo*, uno o más de los exones 2-5 de CD40L, una secuencia 3' UTR, una secuencia poliA, y un segundo brazo de homología.

20 Células diana

[0201] Los sistemas de edición del genoma según la presente divulgación pueden utilizarse para manipular o alterar una célula, *por ejemplo*, para editar o alterar un ácido nucleico diana. La manipulación puede tener lugar, 25 en diversas realizaciones, *in vivo* o *ex vivo*.

[0202] Pueden manipularse o alterarse diversos tipos celulares de acuerdo con las realizaciones de la presente divulgación, y en algunos casos, tal como *aplicaciones in vivo*, se alteran o manipulan una pluralidad de tipos celulares, por ejemplo mediante la administración de sistemas de edición del genoma de acuerdo con la presente divulgación a una pluralidad de tipos celulares. En otros casos, sin embargo, puede ser deseable limitar la 30 manipulación o alteración a un tipo o tipos celulares concretos. Por ejemplo, en algunos casos puede ser deseable editar una célula con un potencial de diferenciación limitado o una célula diferenciada terminalmente, tal como una célula fotorreceptora en el caso de Maeder, en la que se espera que la modificación de un genotipo dé lugar a un cambio en el fenotipo celular. En otros casos, sin embargo, puede ser deseable editar una célula madre o progenitora menos diferenciada, multipotente o pluripotente. A modo de ejemplo, la célula puede ser una célula 35 madre embrionaria, una célula madre pluripotente inducida (iPSC), una célula madre/progenitora hematopoyética (HSPC) u otro tipo de célula madre o progenitora que se diferencie en un tipo celular de relevancia para una aplicación o indicación determinada.

[0203] Para el tratamiento del síndrome de hiper IgM, una célula diana puede ser una célula madre hematopoyética (HSC) o una población de HSC. Alternativamente, una célula diana puede ser una célula T, o una población de 40 células T.

[0204] Como corolario, la célula que se altera o manipula puede ser una célula en división o una célula que no se divide, dependiendo del tipo de célula que se dirija y/o del resultado de edición deseado.

[0205] Cuando las células se manipulan o alteran *ex vivo*, las células pueden utilizarse (*por ejemplo*, administrarse a un sujeto) inmediatamente, o pueden mantenerse o almacenarse para su uso posterior. Los expertos en la técnica apreciarán que las células pueden mantenerse en cultivo o almacenarse (*por ejemplo*, congeladas en nitrógeno líquido) utilizando cualquier método adecuado conocido en la técnica.

50 Método de tratamiento

[0206] Los sistemas de edición del genoma según la presente divulgación pueden utilizarse para tratar el síndrome de Hiper-IgM (HIGM) y trastornos asociados. La HIGM es una inmunodeficiencia caracterizada por niveles séricos normales o elevados de IgM y niveles disminuidos o ausencia de otras inmunoglobulinas séricas, lo que provoca susceptibilidad a las infecciones bacterianas. Los pacientes con síndrome HIGM tienen una incapacidad para pasar de la producción de anticuerpos de tipo IgM a anticuerpos de tipo IgG, IgA y/o IgE. En consecuencia, los sujetos que padecen esta enfermedad presentan niveles reducidos de IgG e IgA, pero niveles normales o elevados de IgM en la sangre. Estos distintos tipos de anticuerpos desempeñan funciones diferentes y todos son importantes para combatir las infecciones. Normalmente, los linfocitos B pueden producir anticuerpos IgM por sí solos, pero necesitan la ayuda interactiva de los linfocitos T para pasar de IgM a IgG, IgA o IgE. La HIGM es el resultado de diversos defectos genéticos que afectan a esta interacción entre los linfocitos T y los linfocitos B.

[0207] La forma más común del síndrome HIGM es el resultado de un defecto o deficiencia de una proteína que se encuentra en la superficie de los linfocitos T activados. La proteína afectada se denomina ligando CD40 (CD40L) porque se une, o liga, a una proteína de los linfocitos B llamada CD40. El ligando CD40 es producido por un gen

del cromosoma X. Por lo tanto, esta enfermedad de inmunodeficiencia primaria se hereda como un rasgo recesivo ligado al cromosoma X.

[0208] Como consecuencia de la deficiencia del ligando CD40, los linfocitos T de los pacientes con hiper IgM ligada al cromosoma X (XHIGM) son incapaces de ordenar a los linfocitos B que cambien su producción de inmunoglobulinas de IgM a IgG, IgA y/o IgE. El ligando CD40 también es importante para otras funciones que desempeñan los linfocitos T, por lo que los pacientes con síndrome de hiper IgM ligado al cromosoma X (XHIM) tienen una inmunidad celular defectuosa y también son susceptibles a muchos tipos de infecciones, sobre todo oportunistas, y a algunos tipos de cáncer. Los pacientes con esta forma también pueden tener neutropenia grave y a menudo presentan durante la infancia neumonía por *Pneumocystis jirovecii*. El tejido linfoide es muy pequeño porque la señalización deficiente del ligando CD40 no activa las células B.

[0209] El diagnóstico del síndrome de hiper-IgM se sospecha con base en criterios clínicos. Se miden los niveles séricos de Ig; los niveles séricos normales o elevados de IgM y los niveles bajos o la ausencia de otras inmunoglobulinas apoyan el diagnóstico. Debe realizarse una prueba de citometría de flujo de la expresión del ligando CD40 en la superficie de las células T. Cuando es posible, el diagnóstico se confirma mediante pruebas genéticas. Se pueden ofrecer pruebas genéticas prenatales a las mujeres que estén pensando en quedarse embarazadas si tienen antecedentes familiares de deficiencia de ligando CD40. Las pruebas genéticas de otros familiares no se realizan de forma rutinaria. Otros hallazgos de laboratorio incluyen un número reducido de células B de memoria (CD27) y la ausencia de células B de memoria de clase comutada (IgD-CD27).

[0210] Los tratamientos que han demostrado ayudar a evitar las infecciones recurrentes asociadas al síndrome de Hiper-IgM incluyen la administración profiláctica de medicación antibiótica y/o infusiones con anticuerpos (inmunoglobulinas) obtenidos del plasma. Además, aunque la terapia con esteroides suele ser eficaz en el tratamiento de la neutropenia, los trastornos autoinmunitarios en niños con síndrome de hiper-IgM pueden presentar un difícil dilema terapéutico, ya que el uso de medicamentos esteroideos suele suprimir un sistema inmunitario ya de por sí débil. En algunos casos, los fármacos antiinflamatorios no esteroideos pueden ser útiles para controlar los síntomas de tipo autoinmunitario evitando el uso de corticosteroides.

[0211] Los datos presentados en el presente documento proporcionan pruebas de que, cuando las células T y las células madre hematopoyéticas se editan para corregir la expresión y la función de CD40L utilizando los sistemas de edición del genoma, como se describe en el presente documento, las células T editadas restauraron la capacidad de cambio de clase de células B. Del mismo modo, los ratones que recibieron células madre hematopoyéticas con un 10 % de células CD40L editadas presentaron un nivel mucho mayor de IgG, lo que sugiere que las células madre hematopoyéticas con al menos un 10 % de CD40L editadas demuestran un gran potencial terapéutico para tratar el síndrome de hiper-IgM. En consecuencia, los sistemas de edición del genoma, tal como se describen en el presente documento, proporcionan un nuevo tratamiento importante para los sujetos con síndrome de hiper-IgM.

[0212] En una realización, al menos el 5 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 6 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 7 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 8 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 9 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 10 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 15 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase. En otra realización, al menos el 20 % de las células editadas en una población de células son suficientes para restaurar funcionalmente el cambio de células de clase.

[0213] En algunas realizaciones, puede ser deseable que los sujetos con síndrome de hiper-IgM estén linfodeplecionados antes de la administración de las células editadas. Por consiguiente, en una realización, el sujeto es acondicionado antes de la administración de las células editadas. El condicionamiento puede implicar la linfodeplección del sujeto. Por ejemplo, en una realización, el sujeto se acondiciona con quimioterapia que causa linfodeplección, antes de recibir las células editadas. En una realización, las células T del sujeto se agotan mediante acondicionamiento. Los agentes adecuados para acondicionar al sujeto incluyen agentes que inducen linfodeplección, por ejemplo, depleción de células T. Los agentes, es decir, acondicionadores, que inducen linfodeplección son conocidos en la técnica. En una realización, el agente acondicionador es un agente quimioterapéutico. En una realización, el agente acondicionador es la ciclofosfamida. En algunas realizaciones, el método puede comprender opcionalmente una etapa de administración de un agente acondicionador al sujeto, por ejemplo, un agente acondicionador que cause linfodeplección. En una realización, el agente acondicionador es un agente quimioterapéutico. En una realización, el agente acondicionador es la ciclofosfamida.

[0214] En otra realización, las células editadas se administran al sujeto sin linfodepleción. En esta realización, el sujeto no está condicionado antes de la administración de las células editadas.

5 **[0215]** En una realización, las células se estimulan antes de la administración al sujeto. En una realización, las células se estimulan utilizando citoquinas, *por ejemplo*, citoquinas estimuladoras. Por ejemplo, las células pueden estimularse con IL-7, IL-15, IL-2 o una combinación de las mismas.

Aplicación de sistemas de edición genómica: suministro, formulaciones y vías de administración

10 **[0216]** Como se ha comentado anteriormente, los sistemas de edición del genoma de la presente divulgación pueden implementarse de cualquier forma adecuada, lo que significa que los componentes de dichos sistemas, incluyendo sin limitación la nucleasa guiada por ARN, el ARNg y el ácido nucleico de plantilla donante opcional, pueden entregarse, formularse o administrarse en cualquier forma adecuada o combinación de formas que dé lugar a la transducción, expresión o introducción de un sistema de edición del genoma y/o provoque un resultado de reparación deseado en una célula, tejido o sujeto. **Las Tablas 5 y 6** exponen varios ejemplos, no limitantes, de implementaciones de sistemas de edición del genoma. Sin embargo, los expertos en la técnica comprenderán que estos listados no son exhaustivos y que son posibles otras implementaciones. Con referencia a la **Tabla 5** en particular, la Tabla enumera varias implementaciones ejemplares de un sistema de edición genómica que comprende un único ARNg y una plantilla donante opcional. Sin embargo, los sistemas de edición genómica según esta divulgación pueden incorporar múltiples ARNg, múltiples nucleasas guiadas por ARN, y otros componentes tales como proteínas, y una variedad de implementaciones serán evidentes para el artesano experto con base en los principios ilustrados en la Tabla. En la Tabla, [N/A] indica que el sistema de edición genómica no incluye el componente indicado.

25 **Tabla 5**

| Componentes del sistema de edición del genoma | | | |
|---|-----|-----------------------|--|
| Nucleasa guiada por ARN | ARN | Plantilla de donantes | Comentarios |
| Proteína | ARN | [N/A] | Una proteína nucleasa guiada por ARN en complejo con una molécula de ARNg (un complejo RNP). |
| Proteína | ARN | ADN | Un complejo RNP como el descrito anteriormente más una plantilla donante monocatenaria o bicatenaria. |
| Proteína | ADN | [N/A] | Una proteína nucleasa guiada por ARN más ARNg transcripto a partir del ADN. |
| Proteína | ADN | ADN | Una proteína nucleasa guiada por ARN más ADN que codifica el ARNg y una plantilla donante de ADN. |
| Proteína | ADN | | Una proteína nucleasa guiada por ARN y un ADN único que codifica tanto un ARNg como una plantilla donante. |
| ADN | | | Un vector de ADN o ADN que codifica una nucleasa guiada por ARN, un ARNg y una plantilla donante. |
| ADN | ADN | [N/A] | Dos ADN separados, o dos vectores de ADN separados, que codifican la nucleasa guiada por ARN y el ARNg, respectivamente. |
| ADN | ADN | ADN | Tres ADN separados, o tres vectores de ADN separados, que codifican la nucleasa guiada por ARN, el ARNg y la plantilla donante, respectivamente. |
| ADN | | [N/A] | Un ADN o vector de ADN que codifica una nucleasa guiada por ARN y un ARNg. |
| ADN | | ADN | Un primer vector de ADN o ADN que codifica una nucleasa guiada por ARN y un ARNg, y un segundo vector de ADN o ADN que codifica una plantilla donante. |
| ADN | ADN | | Un primer vector de ADN o ADN que codifica una nucleasa guiada por ARN y un segundo vector de ADN o ADN que codifica un ARNg y una plantilla donante. |
| ADN | | | Un primer vector de ADN o ADN que codifica una nucleasa guiada por ARN y una plantilla donante, y un segundo vector de ADN o ADN que codifica un ARNg. |

| Componentes del sistema de edición del genoma | | | |
|---|-------|-----------------------|---|
| Nucleasa guiada por ARNg | ARN | Plantilla de donantes | Comentarios |
| ADN | | | Un ADN o vector de ADN que codifica una nucleasa guiada por ARN y una plantilla donante, y un ARNg |
| ARN | [N/A] | | Un ARN o ARN vector que codifica una nucleasa guiada por ARN y que comprende un ARNg |
| ARN | ADN | | Un vector de ARN o ARN que codifica una nucleasa guiada por ARN y que comprende un ARNg, y un vector de ADN o ADN que codifica una plantilla donante. |

[0217] La Tabla 6 resume diversos métodos de administración para los componentes de los sistemas de edición del genoma, tal como se describen en el presente documento. Una vez más, la enumeración pretende ser ejemplar y no limitativa.

5

Tabla 6

| Vector/modo de entrega | Liberación en células que no se dividen | Duración de la expresión | Integración del genoma | Tipo molécula suministrada | |
|--|---|--------------------------|------------------------|------------------------------|------------------------------|
| Físicas (por ejemplo, electroporación, pistola de partículas, transfección con fosfato cálcico, compresión o estrujamiento celular). | Sí | Transitorios | NO | Ácidos nucleicos y proteínas | |
| <i>Viral</i> | Retrovirus | NO | Estable | Sí | ARN |
| | Lentivirus | Sí | Estable | Sí/NO con modificaciones | ARN |
| | Adenovirus | Sí | Transitorios | NO | ADN |
| | Virus adenoasociado (AAV) | Sí | Estable | NO | ADN |
| | Virus Vaccinia | Sí | Muy transitorio | NO | ADN |
| | Virus del herpes simple | Sí | Estable | NO | ADN |
| <i>No viral</i> | Liposomas catiónicos | Sí | Transitorios | Depende de lo que entregue | Ácidos nucleicos y proteínas |
| | Nanopartículas poliméricas | Sí | Transitorios | Depende de lo que entregue | Ácidos nucleicos y proteínas |
| <i>Vehículos biológicos no virales</i> | Bacterias atenuadas | Sí | Transitorios | NO | Ácidos nucleicos |
| | Bacteriófagos de ingeniería | Sí | Transitorios | NO | Ácidos nucleicos |
| | Partículas similares a virus de mamíferos | Sí | Transitorios | NO | Ácidos nucleicos |

| Vector/modo de entrega | Liberación en células que no se dividen | Duración de la expresión | Integración del genoma | Tipo de molécula suministrada | |
|------------------------|---|--------------------------|------------------------|-------------------------------|------------------|
| | Liposomas biológicos: Fantasmas exosomas eritrocitarios | sí y | Transitorios | NO | Ácidos nucleicos |

Sistemas de edición genómica basados en ácidos nucleicos

[0218] Los ácidos nucleicos que codifican los diversos elementos de un sistema de edición del genoma según la presente divulgación pueden administrarse a sujetos o introducirse en células mediante métodos conocidos en la técnica o como se describe en el presente documento. Por ejemplo, el ADN codificante de nucleasas guiadas por ARN y/o el ADN codificante de ARNg, así como los ácidos nucleicos molde donantes pueden administrarse mediante, *por ejemplo*, vectores (*por ejemplo*, vectores virales o no virales), métodos no basados en vectores (*por ejemplo*, utilizando ADN desnudo o complejos de ADN), o una combinación de los mismos.

5 [0219] Los ácidos nucleicos que codifican los sistemas de edición del genoma o sus componentes pueden administrarse directamente a las células como ADN o ARN desnudos, por ejemplo mediante transfección o electroporación, o pueden conjugarse con moléculas (*por ejemplo*, N-acetilgalactosamina) que favorecen la captación por las células diana (*por ejemplo*, eritrocitos, HSC). También pueden usarse vectores de ácido nucleico, tales como los vectores resumidos en la **Tabla 6**.

10 [0220] Los vectores de ácido nucleico pueden comprender una o más secuencias que codifican componentes del sistema de edición del genoma, tales como una nucleasa guiada por ARN, un ARNg y/o una plantilla donante. Un vector también puede comprender una secuencia que codifica un péptido señal (*por ejemplo*, para localización nuclear, localización nucleolar o localización mitocondrial), asociado con (*por ejemplo*, insertado en o fusionado con) una secuencia que codifica una proteína. Como ejemplo, un vector de ácido nucleico puede incluir una secuencia codificante de Cas9 que incluya una o más secuencias de localización nuclear (*por ejemplo*, una secuencia de localización nuclear de SV40).

15 [0221] El vector de ácido nucleico también puede incluir cualquier número adecuado de elementos reguladores/controladores, *por ejemplo*, promotores, potenciadores, intrones, señales de poliadenilación, 20 secuencias consenso de Kozak o sitios internos de entrada de ribosomas (IRES). Estos elementos son bien conocidos en la técnica y se describen en Cotta-Ramusino.

25 [0222] Los vectores de ácido nucleico según la presente divulgación incluyen vectores virales recombinantes. Vectores virales ejemplares se exponen en la **Tabla 6**, y vectores virales adecuados adicionales y su uso y producción se describen en Cotta-Ramusino. También pueden utilizarse otros vectores virales conocidos en la técnica. Además, las partículas virales pueden utilizarse para suministrar componentes del sistema de edición del genoma en forma de ácido nucleico y/o péptido. Por ejemplo, las partículas virales "vacías" pueden ensamblarse para contener cualquier carga adecuada. Los vectores virales y las partículas virales también pueden modificarse para incorporar ligandos que dirijan la diana y alteren la especificidad del tejido.

30 [0223] Además de los vectores virales, pueden utilizarse vectores no virales para suministrar ácidos nucleicos que codifiquen sistemas de edición del genoma según la presente divulgación. Una categoría importante de vectores de ácidos nucleicos no virales son las nanopartículas, que pueden ser orgánicas o inorgánicas. Las nanopartículas son bien conocidas en la técnica, y se resumen en Cotta-Ramusino. Se puede utilizar cualquier diseño de nanopartícula adecuado para suministrar componentes del sistema de edición genómica o ácidos nucleicos que codifiquen dichos componentes. Por ejemplo, las nanopartículas orgánicas (*por ejemplo*, lipídicas y/o poliméricas) 35 pueden ser adecuadas para su uso como vehículos de administración en ciertas realizaciones de la presente divulgación. En la **Tabla 7**, se muestran lípidos ejemplares para su uso en formulaciones de nanopartículas, y/o transferencia de genes y en la **Tabla 8** se enumeran polímeros ejemplares para su uso en formulaciones de transferencia de genes y/o nanopartículas.

Tabla 7: Lípidos utilizados para la transferencia de genes

| Lípidos | Abreviatura | Característica |
|---|-------------|----------------|
| 1,2-Dioleoil-sn-glicero-3-fosfatidilcolina | DOPC | Ayudante |
| 1,2-Dioleoil-sn-glicero-3-fosfatidiletanolamina | DOPE | Ayudante |
| Colesterol | | Ayudante |

| Lípidos | Abreviatura | Característica |
|--|----------------|----------------|
| Cloruro de <i>N</i> -[1-(2,3-Dioleiloxi)propil] <i>N,N,N</i> -trimetilamonio | DOTMA | Catiónico |
| 1,2-Dioleoiloxy-3-trimetilamonio-propano | DOTAP | Catiónico |
| Diocatadecilamidoglicilspermina | DOGS | Catiónico |
| Bromuro de <i>N</i> -(3-Aminopropil)- <i>N,N</i> -dimetil-2,3-bis(dodeciloxi)-1-propanamino | GAP-DLRIE | Catiónico |
| Bromuro de cetiltrimetilamonio | CTAB | Catiónico |
| 6-Lauroxihexil omitinato | LHON | Catiónico |
| 1-(2,3-Dioleiloxipropil)-2,4,6-trimetilpiridinio | 2Oc | Catiónico |
| Trifluoroacetato de 2,3-Dioleiloxi- <i>N</i> -[2(espermínecarboxamido-etyl]- <i>N,N</i> -dimetil-1-propanamino | DOSPA | Catiónico |
| 1,2-Dioleil-3-trimetilamonio-propano | DOPA | Catiónico |
| <i>N</i> -(2-hidroxietil)- <i>N,N</i> -bromuro de dimetil-2,3-bis(tetradeciloxi)-1-propanamonio | MDRIE | Catiónico |
| Bromuro de dimeristooxipropil dimetil hidroxietil amonio | DMRI | Catiónico |
| 3β-[<i>N</i> -(<i>N',N'</i> -Dimetilaminoetano)-carbamoi]colesterol | DC-Chol | Catiónico |
| Bis-guanidio-tren-colesterol | BGTC | Catiónico |
| 1,3-Diodeoxi-2-(6-carboxi-espermil)-propilamida | DOSPER | Catiónico |
| Bromuro de dimetiloctadecilamonio | DDAB | Catiónico |
| Diocatadecilamidoglicilspermidina | DSL | Catiónico |
| cloruro de rac-[(2,3-Dioctadeciloxipropil)(2-hidroxietil)]-dimetilamonio | CLIP-1 | Catiónico |
| bromuro de rac-[2(2,3-Dihexadeciloxipropil-oximetiloxi)etyl]trimetilamonio | CLIP-6 | Catiónico |
| Etildimiristoilfosfatidilcolina | EDMPC | Catiónico |
| 1,2-Disteariloxi- <i>N,N</i> -dimetil-3-aminopropano | DSDMA | Catiónico |
| 1,2-Dimiristoil-trimetilamonio propano | DMTAP | Catiónico |
| <i>O,O'</i> -Dimiristil- <i>N</i> -lisl aspartato | DMKE | Catiónico |
| 1,2-Distearoil-sn-glicero-3-ethylfosfocolina | DSEPC | Catiónico |
| <i>N</i> -Palmitoil D-eritro-sfingosil carbamoi-spermina | CCS | Catiónico |
| <i>N-t</i> -Butil- <i>N</i> 0-tetradecil-3-tetradecilaminopropionamidina | diC 14-amidina | Catiónico |
| Cloruro de octadecenoloxi[etyl-2-heptadecenil-3-hidroxietil] imidazolinio | DOTIM | Catiónico |
| <i>N</i> 1-colesteriloxicarbonil-3,7-diazanonano-1,9-diamina | CDAN | Catiónico |
| 2-(3-[Bis(3-amino-propil)-amino]propilamino)- <i>N</i> -ditetradecilcarbamoiile-etyl-acetamida | RRP209120 | Catiónico |
| 1,2-dilinoiloxy-3- dimetilaminopropano | DLinDMA | Catiónico |
| 2,2-dilinoleil-4-dimetilaminoethyl-[1,3]-dioxolano | DLin-KC2-DMA | Catiónico |
| dilinoleil-metil-4-dimetilaminobutirato | DLin-MC3-DMA | Catiónico |

Tabla 8: Polímeros utilizados para la transferencia de genes

| Polímero | Abreviatura |
|----------------------------------|-------------|
| Poli(etileno)glicol | PEG |
| Polietilenimina | PEI |
| Ditiobis(succinimidilpropionato) | DSP |

| Polímero | Abreviatura |
|---|-------------|
| Dimetil-3,3'-ditiobispropionimidato | DTBP |
| Poli(étileno imina) biscarbamato | PEIC |
| Poli(L-lisina) | PLL |
| PLL modificado con histidina | |
| Poli(N-vinilpirrolidona) | PVP |
| Poli(propilenimina) | PPI |
| Poli(amidoamina) | PAMAM |
| Poli(amidoetilenimina) | SS-PAEI |
| Trietilentetramina | TETA |
| Poli(β -aminoéster) | |
| Poli(éster 4-hidroxi-L-prolina) | PHP |
| Poli(alilamina) | |
| Poli(α -[4-aminobutil]-ácido L-glicólico) | PAGA |
| Poli(ácido D,L-láctico-co-glicólico) | PLGA |
| Bromuro de poli(<i>N</i> -etil-4-vinilpiridinio) | |
| Poli(fosfazeno)s | PPZ |
| Poli(fosfoésteres) | EPI |
| Poli(fosforamidato)s | CCE |
| Poli(<i>N</i> -2-hidroxipropilmetacrilamida) | pHPMA |
| Poli(2-(dimetilamino)metacrilato de etilo) | pDMAEMA |
| Poli(2-aminoethyl propileno fosfato) | PPE-EA |
| Chitosán | |
| Quitosano galactosilado | |
| <i>N</i> -Citosán dodacilado | |
| Histona | |
| Colágeno | |
| Dextrano-espermina | D-SPM |

[0224] Los vectores no virales incluyen opcionalmente modificaciones diana para mejorar la captación y/o dirigirse selectivamente a determinados tipos celulares. Estas modificaciones diana pueden incluir *por ejemplo*, antígenos específicos de células, anticuerpos monoclonales, anticuerpos de cadena única, aptámeros, polímeros, azúcares (por ejemplo, N-acetilgalactosamina (GalNAc)) y péptidos penetrantes de células. Tales vectores también utilizan opcionalmente péptidos/polímeros fusogénicos y desestabilizadores del endosoma, sufren cambios conformacionales desencadenados por ácidos (*por ejemplo*, para acelerar el escape endosomal de la carga), y/o incorporan un polímero escindible por estímulos, *por ejemplo*, para su liberación en un compartimento celular. Por ejemplo, pueden utilizarse polímeros catiónicos basados en disulfuro que se escinden en el entorno celular reductor.

5
10
15
20

[0225] En ciertas realizaciones, se suministran una o más moléculas de ácido nucleico (*por ejemplo*, moléculas de ADN) distintas de los componentes de un sistema de edición del genoma, *por ejemplo*, el componente nucleasa guiado por ARN y/o el componente ARNg descritos en el presente documento. En ciertas realizaciones, la molécula de ácido nucleico se suministra al mismo tiempo que uno o más de los componentes del sistema de edición del genoma. En ciertas realizaciones, la molécula de ácido nucleico se suministra antes o después (*por ejemplo*, menos de unos 30 minutos, 1 hora, 2 horas, 3 horas, 6 horas, 9 horas, 12 horas, 1 día, 2 días, 3 días, 1 semana, 2 semanas o 4 semanas) de que se suministren uno o más de los componentes del sistema de edición del genoma. En ciertas realizaciones, la molécula de ácido nucleico se administra por un medio diferente al que se administran uno o más de los componentes del sistema de edición del genoma, *por ejemplo*, el componente nucleasa guiado por ARN y/o el componente ARNg. La molécula de ácido nucleico puede administrarse mediante cualquiera de los

5 métodos de administración descritos en el presente documento. Por ejemplo, la molécula de ácido nucleico puede entregarse mediante un vector viral, *por ejemplo*, un lentivirus de integración deficiente, y el componente de molécula nucleasa guiada por ARN y/o el componente de ARNg puede entregarse mediante electroporación, *por ejemplo*, tal que pueda reducirse la toxicidad causada por ácidos nucleicos (*por ejemplo*, ADN). En ciertas realizaciones, la molécula de ácido nucleico codifica una proteína terapéutica, *por ejemplo*, una proteína descrita en el presente documento. En ciertas realizaciones, la molécula de ácido nucleico codifica una molécula de ARN, *por ejemplo*, una molécula de ARN descrita en el presente documento.

Entrega de RNPs y/o ARN que codifican componentes del sistema de edición del genoma.

10 [0226] Las RNP (complejos de ARNg y nucleasas guiadas por ARN) y/o los ARN que codifican nucleasas guiadas por ARN y/o ARNg, pueden administrarse a las células o a los sujetos mediante métodos conocidos en la técnica, algunos de los cuales se describen en Cotta-Ramusino. *In vitro*, el ARN guiado por nucleasas y/o el ARN codificante de ARNg puede administrarse, por ejemplo, mediante microinyección, electroporación, compresión celular transitoria o estrujamiento (véase, *por ejemplo*, Lee 2012). Para la administración *in vitro* e *in vivo* también pueden utilizarse la transfección mediada por lípidos, la administración mediada por péptidos, la administración mediada por GalNAc u otros conjugados, y combinaciones de los mismos.

15 [0227] *In vitro*, la administración mediante electroporación comprende mezclar las células con el ARN que codifica nucleasas guiadas por ARN y/o ARNg, con o sin moléculas de ácido nucleico molde donante, en un cartucho, cámara o cubeta y aplicar uno o más impulsos eléctricos de duración y amplitud definidas. Los sistemas y protocolos de electroporación son conocidos en la técnica, y puede utilizarse cualquier herramienta y/o protocolo de electroporación adecuado en relación con las diversas realizaciones de la presente divulgación.

20 **Vía de administración**

25 [0228] Los sistemas de edición del genoma, o las células alteradas o manipuladas mediante dichos sistemas, pueden administrarse a los sujetos por cualquier modo o vía adecuados, ya sea local o sistémica. Los modos de administración sistémica incluyen las vías oral y parenteral. Las vías parenterales incluyen, a modo de ejemplo, las vías intravenosa, intramédula, intrarterial, intramuscular, intradérmica, subcutánea, intranasal e intraperitoneal. Los componentes administrados sistémicamente pueden modificarse o formularse para dirigir, *por ejemplo*, HSC, células madre/progenitoras hematopoyéticas, o progenitores o células precursoras eritroides.

30 [0229] Los modos de administración local incluyen, a modo de ejemplo, la inyección intramuscular en el hueso trabecular o la inyección intrafemoral en el espacio medular, y la infusión en la vena porta. En ciertas realizaciones, cantidades significativamente menores de los componentes (en comparación con los enfoques sistémicos) pueden ejercer un efecto cuando se administran localmente (por ejemplo, directamente en la médula ósea) en comparación con cuando se administran sistémicamente (por ejemplo, por vía intravenosa). Los modos de administración local pueden reducir o eliminar la incidencia de efectos secundarios potencialmente tóxicos que pueden producirse cuando se administran sistémicamente cantidades terapéuticamente eficaces de un componente.

35 [0230] La administración puede proporcionarse como un bolo periódico (por ejemplo, por vía intravenosa) o como infusión continua desde un depósito interno o desde un depósito externo (por ejemplo, desde una bolsa intravenosa o una bomba implantable). Los componentes pueden administrarse localmente, por ejemplo, mediante liberación continua a partir de un dispositivo de liberación sostenida de fármacos.

40 [0231] Además, los componentes pueden formularse para permitir su liberación durante un periodo de tiempo prolongado. Un sistema de liberación puede incluir una matriz de un material biodegradable o un material que libere los componentes incorporados por difusión. Los componentes pueden estar distribuidos homogénea o heterogéneamente en el sistema de liberación. Sin embargo, la elección del sistema apropiado dependerá de la tasa de liberación que requiera una aplicación concreta. Pueden utilizarse tanto sistemas de liberación no degradables como degradables. Los sistemas de liberación adecuados incluyen polímeros y matrices poliméricas, matrices no poliméricas o excipientes y diluyentes inorgánicos y orgánicos tales como, pero no se limitan a, carbonato cálcico y azúcar (por ejemplo, trehalosa). Los sistemas de liberación pueden ser naturales o sintéticos. 45 Sin embargo, se prefieren los sistemas de liberación sintética porque generalmente son más fiables, más reproducibles y producen perfiles de liberación más definidos. El material del sistema de liberación puede seleccionarse de modo que los componentes que tienen diferentes pesos moleculares se liberen por difusión a través del material o por degradación del mismo.

50 [0232] Los polímeros sintéticos biodegradables representativos incluyen, por ejemplo: poliamidas tales como polí(aminoácidos) y polí(peptidos); poliésteres tales como polí(ácido láctico), polí(ácido glicólico), polí(ácido láctico-glicólico), y polí(caprolactona); poli(anhídridos); poliortoésteres; policarbonatos; y derivados químicos de los mismos (sustituciones, adiciones de grupos químicos, por ejemplo, alquilo, alquileno, hidroxilaciones, oxidaciones, y otras modificaciones realizadas rutinariamente por los expertos en la técnica), copolímeros y mezclas de los mismos. Los polímeros sintéticos no degradables representativos incluyen, por ejemplo: poliéteres tales como polí(óxido de etileno), polí(glicol de etileno) y polí(óxido de tetrametileno); polímeros de vinilo-poliacrilatos y

5 polimetacrilatos tales como metil, etil, otros alquilos, metacrilato de hidroxietilo, ácidos acrílico y metacrílico, y otros tales como poli(alcohol vinílico), polí(pirolidona vinílica) y poli(acetato de vinilo); polí(uretanos); celulosa y sus derivados, tales como alquilos, hidroxialquilos, éteres, ésteres, nitrocelulosa y diversos acetatos de celulosa; polisiloxanos; y cualesquiera derivados químicos de los mismos (sustituciones, adiciones de grupos químicos, por ejemplo, alquilos, alquilenos, hidroxilaciones, oxidaciones y otras modificaciones efectuadas habitualmente por los expertos en la técnica), copolímeros y mezclas de los mismos.

10 [0233] También pueden utilizarse microesferas de poli(lactida-co-glicolida). Normalmente, las microesferas están compuestas por un polímero de ácido láctico y ácido glicólico, que se estructuran para formar esferas huecas. Las esferas pueden tener un diámetro aproximado de 15-30 micras y pueden cargarse con los componentes descritos en el presente documento.

Entrega multimodal o diferencial de componentes

15 [0234] Los artesanos expertos apreciarán, en vista de la presente documento, que los diferentes componentes de los sistemas de edición genómica divulgados en el presente documento pueden suministrarse juntos o por separado y simultáneamente o no simultáneamente. La entrega separada y/o asíncrona de los componentes del sistema de edición del genoma puede ser particularmente deseable para proporcionar un control temporal o espacial sobre la función de los sistemas de edición del genoma y limitar ciertos efectos causados por su actividad.

20 [0235] Los modos diferentes o diferenciales, tal como se utilizan en el presente documento, se refieren a modos de administración que confieren propiedades farmacodinámicas o farmacocinéticas diferentes a la molécula componente en cuestión, *por ejemplo*, una molécula nucleasa guiada por ARN, ARNg, ácido nucleico molde o carga útil. Por ejemplo, los modos de administración pueden dar lugar a una distribución tisular diferente, una semivida diferente o una distribución temporal diferente, *por ejemplo*, en un compartimento, tejido u órgano seleccionado de.

25 [0236] Algunos modos de administración, *por ejemplo*, la administración mediante un vector de ácido nucleico que persiste en una célula, o en la progenie de una célula, *por ejemplo*, por replicación autónoma o inserción en ácido nucleico celular, dan lugar a una expresión más persistente de y presencia de un componente. Algunos ejemplos son la administración viral, *por ejemplo*, AAV o lentivirus.

30 [0237] A modo de ejemplo, los componentes de un sistema de edición del genoma, *por ejemplo*, una nucleasa guiada por ARN y un ARNg, pueden administrarse por modos que difieren en términos de semivida resultante o persistencia del componente administrado el cuerpo, o en un compartimento, tejido u órgano particular. En ciertas realizaciones, un ARNg puede administrarse por tales modos. El componente de la molécula nucleasa guiada por ARN puede administrarse de un modo que resulte en una menor persistencia o menor exposición al cuerpo o a un compartimento o tejido u órgano en particular.

35 [0238] Más generalmente, en ciertas realizaciones, se utiliza un primer modo de suministro para suministrar un primer componente y un segundo modo de suministro para suministrar un segundo componente. El primer modo de administración confiere una primera propiedad farmacodinámica o farmacocinética. La primera propiedad farmacodinámica puede ser, por ejemplo, la distribución, persistencia o exposición del componente, o de un ácido nucleico que codifica el componente, en el cuerpo, un compartimento, tejido u órgano. El segundo modo de administración confiere una segunda propiedad farmacodinámica o farmacocinética. La segunda propiedad farmacodinámica puede ser, *por ejemplo*, la distribución, persistencia o exposición, del componente, o de un ácido nucleico que codifica el componente, en el cuerpo, un compartimento, tejido u órgano.

40 [0239] En ciertas realizaciones, la primera propiedad farmacodinámica o farmacocinética, *por ejemplo*, distribución, persistencia o exposición, está más limitada que la segunda propiedad farmacodinámica o farmacocinética.

45 [0240] En ciertas realizaciones, el primer modo de administración se selecciona para optimizar, *por ejemplo*, minimizar, una propiedad farmacodinámica o farmacocinética, *por ejemplo*, la distribución, la persistencia o la exposición.

50 [0241] En ciertas realizaciones, el segundo modo de administración se selecciona para optimizar, *por ejemplo*, maximizar, una propiedad farmacodinámica o farmacocinética, *por ejemplo*, la distribución, la persistencia o la exposición.

55 [0242] En ciertas realizaciones, el primer modo de administración comprende el uso de un elemento relativamente persistente, *por ejemplo*, un ácido nucleico, *por ejemplo*, un plásmido o vector viral, *por ejemplo*, un AAV o lentivirus. Como tales vectores son relativamente persistentes, el producto transcrita a partir de ellos sería relativamente persistente.

60 [0243] En ciertas realizaciones, el segundo modo de administración comprende un elemento relativamente transitorio, *por ejemplo*, un ARN o una proteína.

5 [0244] En ciertas realizaciones, el primer componente comprende ARNg, y el modo de administración es relativamente persistente, *por ejemplo*, el ARNg se transcribe a partir de un plásmido o vector viral, *por ejemplo*, un AAV o lentivirus. La transcripción de estos genes tendría pocas consecuencias fisiológicas porque los genes no codifican para un producto proteico, y los ARNg son incapaces de actuar de forma aislada. El segundo componente, una molécula nucleasa guiada por ARN, se suministra de forma transitoria, *por ejemplo* como ARNm o como proteína, lo que garantiza que el complejo completo molécula nucleasa/ARN guiada por ARN sólo esté presente y activo durante un breve periodo de tiempo.

10 [0245] Además, los componentes pueden administrarse en diferentes formas moleculares o con diferentes vectores de administración que se complementan entre sí para mejorar la seguridad y la especificidad tisular.

15 [0246] El uso de modos de administración diferenciales puede mejorar el rendimiento, la seguridad y/o la eficacia, *por ejemplo*, puede reducirse la probabilidad de una eventual modificación fuera de la diana. La administración de componentes inmunogénicos, *por ejemplo*, moléculas Cas9, por modos menos persistentes puede reducir la inmunogenicidad, ya que los péptidos de la enzima Cas de origen bacteriano son visualizados en la superficie de la célula por moléculas MHC. Un sistema de suministro en dos partes puede paliar estos inconvenientes.

20 [0247] Se pueden utilizar modos de administración diferenciales para administrar componentes a regiones diana diferentes pero solapadas. El complejo activo de formación se minimiza fuera de la superposición de las regiones diana. Así, en ciertas realizaciones, un primer componente, *por ejemplo*, un ARNg se administra mediante un primer modo de administración que da lugar a una primera distribución espacial, *por ejemplo*, tisular. Un segundo componente, *por ejemplo*, una molécula nucleasa guiada por ARN se administra mediante un segundo modo de administración que da lugar a una segunda distribución espacial, *por ejemplo*, tisular. En ciertas realizaciones, la primera modalidad comprende un primer elemento seleccionado de entre un liposoma, nanopartícula, *por ejemplo*, nanopartícula polimérica, y un ácido nucleico, *por ejemplo*, vector viral. El segundo modo comprende un segundo elemento seleccionado de entre el grupo. En ciertas realizaciones, el primer modo de administración comprende un primer elemento diana, *por ejemplo*, un receptor específico celular o un anticuerpo, y el segundo modo de administración no incluye dicho elemento. En ciertas realizaciones, el segundo modo de administración comprende un segundo elemento diana, *por ejemplo*, un segundo receptor celular específico o un segundo anticuerpo.

35 [0248] Cuando la molécula nucleasa guiada por ARN se administra en un vector de administración de virus, un liposoma o una nanopartícula polimérica, existe la posibilidad de administración y actividad terapéutica en múltiples tejidos, cuando puede ser deseable dirigirla a un solo tejido. Un sistema de administración en dos partes puede resolver este problema y mejorar la especificidad tisular. Si el ARNg y la molécula de nucleasa guiada por ARN se empaquetan en vehículos de administración separados con un tropismo tisular distinto pero superpuesto, el complejo plenamente funcional sólo se formará en el tejido diana de ambos vectores.

40 [0249] La presente invención se ilustra además mediante los siguientes ejemplos, que no pretenden ser limitativos en modo alguno.

Ejemplos

45 [0250] Los siguientes Ejemplos son meramente ilustrativos y no pretenden limitar en modo alguno el alcance o contenido de la invención.

Ejemplo 1: Edición genética de CD40L en linfocitos T

50 [0251] Se realizaron experimentos de integración diana en células T con *S. pyogenes* de tipo salvaje Cas9 ribonucleoproteína (RNP) dirigida al locus CD40L. Las secuencias de ARNg se muestran en la **Tabla 9**. Las secuencias de los donantes de CD40L se muestran en la **Tabla 10**. La Figura. 2A muestra una representación esquemática del locus CD40L tras la integración diana. 8 *S. Pyogenes* y 30 *S. Aureus* Los ARNg dirigidos al exón 1 de CD40LG se examinaron en células T humanas primarias para comprobar la eficacia de la edición junto con *S. Pyogenes* o *S. Aureus* Cas9 de tipo salvaje. La posición de los ARNg en el exón 1 de CD40LG se representa en la Figura. 2B. Aproximadamente el 90 % - 95 % de las mutaciones CD40LG se corrigieron para restaurar tanto la expresión como la función del gen, en *S. Pyogenes* y *S. Aureus*. La Figura. 2C muestra las tasas de inserción/deleción para cada ARNg medidas en células T. Para determinar el nivel de edición y los tipos de edición en múltiples sitios de ARNg identificados en el intrón 1 de CD40LG , se introdujo ribonucleoproteína (RNP) en células T CD4⁺ masculinas humanas primarias. Las células T se descongelaron y se activaron durante 2 días con microesferas magnéticas anti-CD3/CD28, tras lo cual se retiraron las microesferas y se dejó que las células se expandieran durante 2 días en medio X-Vivo 15 antes de realizar las nucleofeciones. Los medios durante todas las etapas contenían citocinas humanas (IL-2, IL-7 e IL-15). Las células T se nucleofectaron con 2.27 µM de RNP en una relación 2:1 de ARNg:Cas9 utilizando el sistema de nucleofeción de Lonza. A continuación, se dejó que las células se expandieran durante 4 días en medio de células T antes de aislar el ADN genómico. A continuación, el ADN genómico (ADNg) se amplificó por PCR utilizando cebadores que rodean el sitio de corte de interés, se secuenció y se analizó en busca de inserciones y delecciones en el sitio de corte. Se logró una edición global

superior al 70 % para cada ARNg presentado en la Figura. 2C, aunque la relación entre inserciones y delecciones varió. Los resultados representados en la Figura. 2C representan al menos 2 muestras independientes. El porcentaje de edición conseguido con cada RNP se presenta también en la Tabla 9.

- 5 [0252] En los experimentos de integración dirigida se utilizaron ARNg con una alta eficacia de edición junto con un vector AAV6 que contenía la plantilla donante CD40LG representada en la Figura. 2A. AAV6 se añadió después de la electroporación de RNP en el día 3 y la fluorescencia de GFP se midió por FACS 20 días después del experimento, como se muestra en la Figura. 2D. El porcentaje de células GFP positivas en poblaciones de células T tras la edición génica se muestra en la Figura. 2E. Las células T madre de memoria (TSCM) se definen como CD62L+CD45RA+, las T de memoria central (TCM) como CD62L+ CD45RA-, las T efectoras de memoria (TEM) como CD62L- CD45RA- y las efectoras terminales (TEMRA) como CD62L-CD45RA+. La combinación CRISPR/Cas9 y AAV6 permitió altos niveles de edición del gen CD40L en múltiples poblaciones de células T. En la Figura. 2F se presentan gráficos representativos de las células T que expresan CD45RA y CD62L 7 días (panel izquierdo) y 20 días (panel derecho) después del tratamiento. La composición del cultivo de células T 7 y 19 días (panel izquierdo) y 20 días (panel derecho) después del tratamiento se muestra en la Figura. 2G. Estos datos indican que no hay desviación de las subpoblaciones de células T tras la edición génica en el emplazamiento CD40L, en relación con el control sin editar ("sólo células"). También se realizaron experimentos de integración diana junto con un vector IDLV que contenía la plantilla donante CD40LG, como se muestra en la Figura. 2H. El nivel de HDR (evaluado mediante la determinación del porcentaje de células GFP +) dentro de varios linajes de células T se muestra en la Figura. 2I. La Figura. 2J muestra que se observan niveles similares y elevados de edición génica tanto en las células de la HD como en las del paciente.

Tabla 9. ARN guía, protoespaciador y secuencias PAM

| Nombre del ARNg | Secuencia Protospacer (5' a 3') | PAM | | |
|-----------------|--------------------------------------|------------------------|--|--|
| A | TGGATGATTGCACTTTATCA (SEQ ID NO: 2) | GGG (SEQ ID NO: 10) | | |
| B | TTTTCTAACAGGATAAGGTG (SEQ ID NO: 3) | AGG (SEQ ID NO: 11) | | |
| C | CGGTAAATATCAGTCCACTG (SEQ ID NO: 4) | AGG (SEQ ID NO: 12) | | |
| D | AGTGAGGGCTGAAGTCATCCA (SEQ ID NO: 5) | CTGGGT (SEQ ID NO: 13) | | |
| E | ACCTAACATTTGGATAACCCA (SEQ ID NO: 6) | GTGGAT (SEQ ID NO: 14) | | |
| F | CAATGAGAAAATGTGACAATT (SEQ ID NO: 7) | CAGAAT (SEQ ID NO: 15) | | |
| G | AGAATAGCTCTGATTCTACC (SEQ ID NO: 8) | TGGAGT (SEQ ID NO: 16) | | |
| H | GAGGACTTCAGGCATAAATG (SEQ ID NO: 9) | GAGAAT (SEQ ID NO: 17) | | |

| Nombre del ARNg | Cas9 | Secuencia del ARNg | PAM | % Edición (secuenciación) |
|-----------------|--------------------|---------------------------------------|------------------------|---------------------------|
| A | <i>S. pyogenes</i> | UGGAUGAUUGCACUUUAUCA (SEQ ID NO: 18) | GGG (SEQ ID NO: 26) | 90% |
| B | <i>S. pyogenes</i> | UUUCUACAGGUAAGGUG (SEQ ID NO: 19) | AGG (SEQ ID NO: 27) | 70% |
| C | <i>S. pyogenes</i> | CGGUAAAUAUCAGUCCACUG (SEQ ID NO: 20) | AGG (SEQ ID NO: 28) | 85% |
| D | <i>S. aureus</i> | AGUGAGGGCUGAAGUCAUCCA (SEQ ID NO: 21) | CTGGGT (SEQ ID NO: 29) | 90% |
| E | <i>S. aureus</i> | ACCUAAUAUUUGGAUACCCA (SEQ ID NO: 22) | GTGGAT (SEQ ID NO: 30) | 90% |
| F | <i>S. aureus</i> | CAAUGAGAAAUGUGACAUUA (SEQ ID NO: 23) | CAGAAT (SEQ ID NO: 31) | 80% |

| Nombre del ARNg | Cas9 | Secuencia del ARNg | PAM | % Edición (secuenciación) |
|-----------------|-----------|--|------------------------|---------------------------|
| G | S. aureus | AGAAUAGCUCUGAUUUCUACC (SEQ ID NO: 24) | TGGAGT (SEQ ID NO: 32) | 80% |
| H | S. aureus | GAGGACTTCAGGCATAAATG (SEQ ID NO: 25) | GAGAAT (SEQ ID NO: 33) | 15-20% |

Tabla 10. Secuencias de donantes de CD40L

| Nombre del donante | del | Descripción del donante | Secuencia de donantes |
|--------------------|-----|-------------------------|--|
| 1 | | gARN-A_HA | (SEQ ID NO: 34) |
| | | + ADNc | TGTTTGCAATTCTTAGGAAAAGAAAACCATCAGGACTTATTTGTTTCATGTA |
| | | + GFP | TTTTTCACCTCCACTGAGGAGTATAATTGGCTGGTGTGACAAAATACCAATC |
| | | + Relleno | ATAGATGTAAGGAGAAAGTTGATTAGTTCTGGCTGTTCTAAATTCTGG |
| | | | ATGCAGGAACTGTGGCTAGAAAGCATCTGGATGATTGCACTTATACTCTAA |
| | | | TTCATTACATATTGCGGTGCAATTCAAGGAGGCCATAATGCGGTTACAATA |
| | | | ATTCCCTATACTAAATATAAAAGATTAAATTCAAAAAATGGTACCAAGCA |
| | | | TCGTTAGTGCCTATACATCAAGAGGCCACGTCCCCGGAGACAGCAAGTAAGC |
| | | | TCTTAAACACCTGATGTGCTGTCAATCAAATGTAAGCTTCTTAGGTTA |
| | | | CATGTGCTCTTAATTACAGCAGAACGGCTCTGACCTCTCTCCCTCCACAG |
| | | | ATCGAGGACGAGAGAAACCTGCACGAGGACTCTGTGTTCATGAAGACCATCC |
| | | | AGCGGTGCAACACCGCGAGAGAAGTCTGAGCCTGCTGAACCTGCGAGGAAA |
| | | | TCAAGAGCCAGTCGAGGGCTCGTGAAGGACATCATGCTGAACAAAGAGGA |
| | | | AACGAAGAAAGAAAACCTTCGAGATGCAAGAGGGCGACCAGAATCCTCAG |
| | | | ATCGCCGTCACGTGATCAGCGAGGCCAGCAGCAAGACAACAAGCGTGTG |
| | | | AGTGGGCGAGAAGGGCTACTACACCATGAGCAACAACCTGTCACCTGGA |
| | | | AAACGGCAAGCAGCTGACAGTGAAGCGCAGGGCTGTACTACATCTAGCC |
| | | | CAAGTGCACCTCTGCAAGCACAGAGAGGCCAGCTCTCAGGCCCTTATCGC |
| | | | CAGCCTGTGCCCTGAAGTCCCCTGGCAGATTGAGCGGATTCTGCTGAGAGCC |
| | | | GCCAACACACACAGCAGGCCAACCTTGTGCCAGCAGTCTATTCACCTCGG |
| | | | CGGAGTGTGAGCTGCAGCCTGGCAGCGTGTGTAATGTGACAGAC |
| | | | CCTAGCCAGGTGTCCTCACGGCACCGGCTTACATCTTGGACTGCTGAAGCT |
| | | | GTGAACAGTGCACCTGCAGGCTGTTGGAGCTGACGCTGGAGTCTTA |
| | | | TAATACAGCACAGCGGTTAGCCCACCCCTGTTAAGCTGCTTATAACCT |
| | | | AGGATCTCTTATGGAGAACTATTATACACTCCAAGGCATGTAGAACT |
| | | | GAGCTTATATCTGAAGCAGCAACCCACTGATGCAAGACATCCAGAGAGTCC |
| | | | TATGAAAAGACAAGGCCATTATGCACAGGTTGAATTCTGAGTAAACAGCAGA |
| | | | TAACCTGCCAAGTTCACTGTTGTTCTCGTGCACTGCTTCCATGGATAA |
| | | | TGCATTGATTTCAGTGAAGATGCAAGAGGGAAATGGGGAGCCTCAGCTC |
| | | | ACATTCACTTATGGTTGACTCTGGGTTCTATGGCCTGTTGGAGGGGCCAG |
| | | | GCTCTAGAACGCTAACACAGTGGAGAACCGAAACCCCCCCCCCGCCACCC |
| | | | TCTCGGACAGTTATTCACTCTCTCAACCTCTTCCAACTCTCTTCTCAATCTCTGTTCCCT |
| | | | TTGTCAGTCTCTCCCTCCCCAGTCTCTCTCAATCCCCCTTCAACACACA |
| | | | CACACACACACACACACACACACACACACACACACAGAGTCAG |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------|--------------------------------|--|
| | | | <p>GGCGTTGCTAGTCAGTCTTCTTCCACCCGTCCCTATCTCTACCACTATAG ATGAGGGTGAGGAGTAGGGAGTGCAGCCCTGAGCCTGCCACTCCTCATTAC GAAATGACTGTATTAAGGAATCTATTGTATCTACCTGCAGTCTCCATTGTT TCCAGAGTGAACCTGTAACTTATCTTGTATTATTTTGAGATAATAAGACCTC TTAACATTACCGCTTAACATTATCGTTGTTGAGTACCTAAAGCTCCAG CCAGGTTGGGAAAGAGGAAGCATTGGAGGGAAATTCCCACCTTGTGA TGTTTCTATAAACTTGTCTCAAGCTACTTACATTACCGTACTAGTTGGCTCC GGTGCCGTCAGTGGCAGAGCGCACATGCCACAGTCCCCGAGAAGTTGG GGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTA AACTGGGAAAGTGTGTCGTACTGGCTCCGCCCTTCCCAGGGTGGGG GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTCGACACGGG TTGCGGCCAGAACACAGGGTGTGACCGGATCGTACCCGGGTTAGGGC GAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCACCCTGCCACCATG GTGAGCAAGGGCGAGGAGCTGTCACCGGGGTGGTGCCTACCTGGTCAG CTGGACGGCAGCTAAACGGCACAAGTTCAGCGTGTCCGGCGAGGGCGAG GGCATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGACCCACGGCA AGCTGCCGTGCCCTGGCCACCCCTGTGACCCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCATGAAAGCAGCACGACTTCTCAAGTCCGC CATGCCCGAAGGCTACGTCAGGAGCCACATCTTCTCAAGGACGACGGC AACTACAAGACCCGCGCCAGGTGAAGTTGAGGGCGACACCCCTGGTGAACC GCATCGAGCTGAAGGGCATCGACTTAAGGAGGACGGCAACATCTGGGGC ACAAGCTGGAGTACAACACTACAACAGCCACAACGCTTATATCATGGCGACAA GCAGAAGAACGGCATCAAGGTGAACCTCAAGATCGCCACAACATCGAGGAC GGCAGCGTGCAGCTGCCGACCACTACCAGCAGAACACCCCCATGGCGACG GCCCGCGTGTGCTGCCGACAACCAACTACCTGAGCACCCAGTCCGCCGTGAGC AAAGACCCCAACGAGAAGCGCGATCACATGGCTCTGCTGGAGTTCGTGACCG CCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGT CGAGTCTAGAGGGCCGTTAAACCCGCTGATCAGCCTGACTGTGCCTTCA GTTGCAGCCATCTGTTGTTGCCCTCCCCGTGCCCTCTGACCCCTGGAAAG GTGCCACTCCACTGCCCTTCTTAATAAAATGAGGAAATTGACATCGCATTGTC TGAGTAGGTGTATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGG GGGAGGATTGGAAGACAATAGCAGGCATGCTGGGATGCCGTGGCTCTA TGGTACAGCTAAAGAGCAGAGTAACCTAGTAGGCTGCTTGACATACGATT TTTAATAAAACATGAGCATTGAATAAAAACGACTCCTCATACTGTAACACATC ACGCATGCACATTAGACAATAATCCAGTAACGAAACGGCTCAGTCGTAATCG CCCATATAGTGGCTACAGAATGTTGGATAGAGAACTTAAGTACGCTAAGGC GGCATTTCTTAATATTAGGGTATTCAAGGATACTTGAGTGTCCCTCTCT TAGGATCTGGACCTAGAATTAATGTCATGAGATTCTAACAGGATAAGGTG AGGTAGTGAAGGGCTGAAGTCATCCACTGGGTATCAAATATTAGGTTACT GCTGACAAAAGAGGGGGCTCTGGTCTGGTTGGTATTGTGTTGG</p> |
| 2 | gARN-A_HA | (SEQ ID NO: 35) | |
| | + ADNc | | TGTTTGCAATTCTAGGAAAAGAAAACCATCAGGACTTATTGTTGTTCTGTA |
| | + GFP | | TTTTTCACTTCACTGAGGAGTATAATTGGCTGGTGTGACAAAATACCAATC |
| | - Relleno | | ATAGATGTAAAGGAGAAAGTTGATTAGTGTCTGGCTGTTCTAAATTCTGG ATGCAGGAACTGTGGCTAGAAAGCATCTGGATGATTGACTTACCTGATGTG CTCTGTCATCAAATGTAAGCTCCTAGGTTACATGTGCTCTTAATTACAG CAGAACCGGCTGACCTCTCTCCACAGATCGAGGACGAGAGAAC <u>CTGCACGAGGACTTCGTGTTCAAGGACATCCAGCGGTGCAACACGGCG</u> <u>AGAGAAGTCTGAGCCTGTAACCTGCGAGGAATCAAGAGCCAGTCGAGG</u> <u>GCTTCGTGAAGGACATCATGCTGAACAAAGAGGAAACGAAGAAAGAAAAC</u> <u>CCTCGAGATGCCAGAAGGGCGACCAGAACCTCAGATGCCGTCACGTGAT</u> |

ES 2 998 052 T3

| Nombre del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------------------|---|
| | | <p>CAGCGAGGCCAGCAGCAAGACAACAAGCGTGTGCAGTGGGCCGAGAAGGG <u>CTACTACACCATGAGCAACAACCTGGTACCCCTGGAAAACCGCAAGCAGCTG</u> <u>ACAGTGAAGCGGCAGGGCTGTACTACATCTACGCCAACGTGACCTCTGCA</u> <u>GCAACAGAGAGGCCAGCTCTAGGCCCTTTATGCCAGGCCGTGCCCTGAAG</u> <u>TCCCCCTGGCAGATTGAGCGGATTCTGCTGAGAGGCCAACACACACAGCA</u> <u>GCGCAAACCTTGTCGGCCAGCAGTCTATTACCTCGGGGAGTGTTGAGCTG</u> <u>CAGCCTGGCGAAGCGTGTCTGTGAATGTGACAGACCCAGCCAGGTGTCCC</u> <u>ACGGCACCGGCTTACATCTTGGACTGCTGAAGCTGTGAACAGTGTACCT</u> <u>TGCAGGCTGTGGTGGAGCTGACGCTGGGAGTCTTCATAATACAGCACAGCG</u> <u>TTAAGCCCACCCCTGTTAACCTGCTATTITATAACCCCTAGGATCTCTTATGG</u> <u>AGAACTATTATTATAACTCCAAGGCATGAGAACTGTAATAAGTGAATTAC</u> <u>AGGTACATGAAACCAAAACGGGCCCTGCTCCATAAGAGCTTATATATCTGAA</u> <u>GCAGCAACCCCACTGATGCAGACATCCAGAGAGTCTATGAAAAGACAAGGC</u> <u>CATTATGCACAGGTTGAATTCTGAGTAAACAGCAGATAACTGCAAGTTCA</u> <u>TTTGTCTTGTGCTGAGTGTCTTCCATGGATAATGCAATTGATTATTCAGT</u> <u>GAAGATGCAGAAGGAAATGGGAGCCTCAGCTCACATTCACTTATGGTTGA</u> <u>CTCTGGGTCTATGGCTTGGAGGGGGCCAGGCTCTAGAACGTCTAACAC</u> <u>CAGTGGAGAACCGAAACCCCCCCCCCCCCGCAACCTCTCGGACAGTTATTCA</u> <u>TCTCTTCAATCTCTCTCCATCTCTCTTCACTTCACTCTCTCTCTCAACCTCT</u> <u>TTCTTCAATCTCTCTCAATCTCTGTTCCCTTGTCACTTCACTCTCTCTCT</u> <u>CCCAGTCTCTCTCAATCCCCCTTCTAACACACACACACACACACACACA</u> <u>CACACACACACACACACACACACAGAGTCAGGCCGTTGCTAGTCAGTTC</u> <u>TCTCTTCCACCCCTGCTCCCTATCTTCACTTCACTTCACTTCACTTCACTT</u> <u>GGAGTGAGGCCCTGAGCCTGCCCACTCTCATTACGAAATGACTGTATTAA</u> <u>GGAAATCTATTGTATCTACCTGCACTCTCATTGTTCAAGGTGAACCTGAA</u> <u>TTATCTTGTATTATTATTTGAATAATAAGACCTCTAACATTACGCGCTAA</u> <u>CATTATCGTTGTGTTGAGTACCTAAAGCTCCAGGCCAGGTTGGGAAAGAG</u> <u>GAAGCATTGGAGGGAATTTCACCTTGTGATGTTTCATAAACTTGT</u> <u>CTCAAGCTACTTACATTACGCGTACTAGTTGGCTCCGGTCCCCGTCACTGGC</u> <u>AGAGCGCACATGCCACAGTCCCCGAGAAGTGGGGGGAGGGGTGGCAA</u> <u>TTGAACCGGTGCTAGAGAAGGTGGCGGGGTAAACTGGAAAGTGTATGT</u> <u>CGTGTACTGGCTCGCTTTCCGAGGGTGGGGGAGAACCGTATATAAGT</u> <u>GCAGTAGTCGCCGTGAACGTTCTTTCGCAACGGGTTGCCGCCAGAACACA</u> <u>GGTGTGTGACGCCGATGTCACCCGGTTAAGGGCAATTCCAGCACACTGG</u> <u>CGGGCGTTACTAGTGGATCCACCGGTGCCACCATGGTGAAGCAAGGGCGAGG</u> <u>AGCTGTTACCGGGGGTGGGCCATCTGGTCAAGCTGGACGGCAGCTAACGTA</u> <u>CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGGATGCCACCTACGG</u> <u>CAAGCTGACCTGAAAGTTCTGACCACCGGCAAGCTGCCGTGCCCTGGC</u> <u>CCACCCCTGTGACCACCCCTGACCTACGGCGTGCAGTGTCTAGCCGTACCC</u> <u>GACCACATGAAGCAGCACGACTCTCAAGTCCGCCATGCCGAAGGCTACGT</u> <u>CCAGGAGCGCACCATCTCTCAAGGACGACGGCAACTACAAGACCCGCC</u> <u>GAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATGAGCTGAAGGGC</u> <u>ATCGACTTCAAGGAGGACGCCAACATCTGGGGCACAAGCTGGAGTACAAC</u> <u>ACAACAGCCACAACGTCTATATCATGGGCCACAAGCAGAAGAACGGCATCAA</u> <u>GGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCC</u> <u>GACCACTACCAAGCAGAACACCCCCATCGGCACGGCCCCGTGCTGTGCCG</u> <u>ACAACCAACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAA</u> <u>GCGCGATCACATGGCTCTGCTGGAGTTCTGACCGCCGCCGGGATCACTCTC</u> <u>GGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGTCAAGTCTAGAGGGCCC</u> <u>GTTAAACCCGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTT</u> <u>GTTGCCCTCCCCGTGCCCTTGAACCTGGAAAGGTGCCACTCCACTGCC</u> </p> |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------|--------------------------------|--|
| | | | CTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCAGTAGGTGTCATTCT ATTCTGGGGGGTGGGGTGGGCAGGACAGCAAGGGGGAGGATTGGGAAGA CAATAGCAGGCATGCTGGGATCGGGCTATGGTACCAAGCTAAAGA GCAGAGTAACTAGTAGGGCTAGGGATACTTGAGTGTCTCTAGGATCTG GACCTAGAATTAATGTCATGAGATTCTAACAGGATAAGGTGAGGTAGTG AGGGCTGAAGTCATCCACTGGGTATCCAATATTAGGTTACTGCTGACAA AAGAGGGGGCTCTGGCTGGTTATTGTGTTGG |
| 3 | gARN-A_HA | (SEQ ID NO: 36) | |
| | | + <u>ADNc</u> | TGTTTGCACTCTAGGAAAAGAAAACCACAGGACTTATTTGTTTATGTA TTTTTCACTCCACTGAGGAGTATAATTGGCTGGTGTGACAAAATACCAATC ATAGATGTAAGGAGAAAGTTGATTAGTTCTGGCTGTTCTAAATTCTGG ATGCAGGAACGTGGCTAGAAAGCATCTGGATGATTGCACTTATACTCTAA TTCATTACATATTGTGCGGTGCAATTAGGGAGGCCATAATGGGTTACAATA ATTCCCTATACTAAATATACAAAGATTAAATTCAAAAAATGGTACACAGCA TCGTTAGTGCATACATCAAGAGGACACGTGCCCCGGAGACAGCAAGTAAGC TCTTAAACACCTGATGTGCTCTGTCATCAAATGTAAGCTTCTTAGGTTA CATGTCCTTAATTACAGCAGAACCGGCTGACCTCTCTCCCTCCACAG ATCGAGGACGAGAGAAACCTGCACGAGGACTTCGTGTTCATGAAGACCATCC <u>AGCGGTGCAACACCGCGAGAGAAGTCTGAGCCTGCTGAACCTGAGGAAA</u> <u>TCAAGAGCCAGTCGAGGCTCGTGAAGGACATCATGTCGAACAAAGAGGA</u> <u>AACGAAGAAAGAAAATCCTCGAGATGCAAGGGCACCAGAAATCTCAG</u> <u>ATGCCGCTACGTACAGCGAGGCCAGCAGCAAGACAACAAGCGTGC</u> <u>AGTGGGCCAGAAGGGCTACTACACCATGAGCAACACCTGGTCACCCCTGG</u> <u>AAACGGCAAGCAGCTGACAGTGAAGCGCAGGGCTGTACTACATCTACGCC</u> <u>CAAGTGCACCTCTGCAGCACAGAGAGGCCAGCTCTCAGGCCCTTATCGC</u> <u>CAGCCTGTGCCAGTCCACGGCACGGCTTACATCTCGGACTGCTGAAGCT</u> <u>GCCAAACACACACAGCAGGCCAACCTTGTGGCCAGCAGTCTATTACACCTCGG</u> <u>CGGAGTGTGAGCTGCAGCCTGGCGAAGCGTGTGTAATGTGACAGAC</u> <u>CCTAGCCAGGTGCTCCACGGCACGGCTTACATCTCGGACTGCTGAAGCT</u> <u>GTGAACAGTGTACCTTGCAAGGCTGGTGGAGCTGACGCTGGAGTCTTA</u> TAATACAGCACAGCGGTTAAGCCCACCCCTGTTAATGCTATTATAACCT AGGATCCTCTTATGGAGAACTATTATTACACTCCAAGGCATGTAGAACT GTAAATAAGTGAATTACAGGTCACATGAAACAAACGGGCCCTGCTCCATAA GAGCTTATATCTGAAGCAGCAACCCACTGATGCAAGACATCCAGAGAGTC TATGAAAAGACAAGGCCATTATGCACAGGTTGAATTCTGAGTAAACAGCAGA TAACTTGCAAGTTCACTGTTCTTGTGCTGAGTGTCTTCCATGGATAA TGCATTGATTATCAGTGAAGATGCAAGGGAAATGGGGAGCCTCAGCTC ACATTCACTGGTTGACTCTGGGTTCTATGGCCTTGTGGAGGGGCCAG GCTCTAGAACGCTAACACAGTGGAGAACGGAAACCCCCCCCCCGCCACCC TCTCGGACAGTTATTCACTCTCTCACTCTCTCCATCTCTCTTCTCAG TCTCTCTCTCAACCTCTTCCAATCTCTCTTCTCAATCTCTGTTCCCT TTGTCAGTCTCTCCCTCCCCAGTCTCTCTCAATCCCCCTTCTAACACACA CACACACACACACACACACACACACACACACACAGAGTCAG GCCGTTGCTAGTCAGTTCTCTTCACTCTGAGCTGCCCCACTCTCATTAC ATGAGGGTGAGGAGTAGGGAGTGCAGCCCTGAGCTGCCCCACTCTCATTAC GAAATGACTGTATTAAAGGAAATCTATTGTATCTACCTGCACTGCTCATTGTT TCCAGAGTGAACTGTAATTATCTGTTATTATTTGAATAATAAGACCTC TTAACATTACCGCTTAACATTATCGTTGTTGAGTACCTAAAGCTCCAG CCAGGTTGGGAAAGAGGAAGCATTGGAGGGAAATTCCACCTTGTGA TGTTTCTAAACTTTGTTCAAGCTACTACATTACGCGTACTAGTACCAAGC TTAAAGAGCAGAGTAACCTAGTAGGCTGCTTGCACATACGATTAAATAAAA |
| | | - GFP | |
| | | + relleno | |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|--|
| | | | CATGAGCATTGAATAAAAACGACTCCTCATACTGTAAACATCACGCATGCA CATTAGACAATAATCCAGTAACGAAACGGCTTCAGTCGAATGCCCATATAG TTGGCTACAGAATGTTGGATAGAGAACCTAAGTACGCTAAGGCAGCGTATTT CTTAATATTAGGGGTATTCAGGGACTTGAAGTGCTCTCTAGGATCTGG ACCTAGAATTAAATGTCTAGAGATTCTAACAGGATAAGGTGAGGTAGTGA GGGCTGAAGTCATCCACTGGGTTATCCAATATTAGGTTACTGCTGACAAA AGAGGGGGCTCTGGTCTGGTGGTTATTGTGTTGG |
| 4 | gARN-A_HA | (SEQ ID NO: 37) | |
| | + ADNc | | TGTTTGCAATTCTAGGAAAAGAAAACCACAGGACTTATTTGTTTCACTGTA TTTTTCACTCCACTGAGGAGTATAATTGGCTGGTGTGACAAAATACCAATC ATAGATGTAAAGGAGAAAGTTGATTAGTTCTGGCTGTTCTAAATTCTGG ATGCAGGAACTGTGGCTAGAAAGCATTGGATGATTGACTTACCTGATGTG CTCTGTCAATCAAATGTAAGCTCTTAGGTTACATGTGCTCTAAATTACAG CAGAACCGGTCTGACCTCTCTCCCTCCCACAGATCGAGGACGAGAGAAC CTGCACGAGGACTTCGTGTTCATGAAGACCATCCAGCGGTGCAACACCGCG AGAGAAGTCTAGCCTGCTGAACCTGAGGAAATCAAGAGCCAGTCGAGG GCTCGTGAAGGACATCATGCTGAACAAAGAGGAAACGAAGAAAAGAAA CCTCGAGATGCAGAAGGGCGACCAGAACCTCAGATGCCGCTACGTGAT CAGCGAGGCCAGCAGCAAGACAACAAGCGTGCAGTGGGCCAGAAGGG CTACTACACCAGAGCAACACCTGGTCACCTGGAAAACGGCAAGCAGCTG ACAGTGAAGCGGCAGGGCCTGACTACATCTACGCCAAGTGACCTCTGCA GCAACAGAGAGGCCAGCTCAGGCCCTTTATGCCAGCCTGTGCCAG TCCCTGGCAGATTGAGCGGATTCTGCTGAGAGGCCAAACACACAGCA GCGCCAAACCTTGTGGCCAGCAGTCTATTCACTCGGGAGTGTGTTGAGCTG CAGCCTGGCGCAAGCGTGTGTAATGTGACAGACCTAGCCAGGTGTC ACGGCACGGCTTACATCTTGGACTGCTGAAGCTGTGAACAGTGTACCT TGCAGGCTGTGGAGCTGACGCTGGAGTCTCATAATACAGCACAGCG TTAAGCCCACCCCTGTTAATGCTTATAACCCCTAGGATCTCTTATGG AGAACTATTATACACTCCAAGGCATGTAGAACTGTAATAAGTGAATTAC AGGTACATGAAACAAAACGGGCCCTGCTCCATAAGAGCTTATATCTGAA GCAGCAACCCACTGATGCAGACATCCAGAGAGTCTATGAAAAGACAAGGC CATTATGCACAGGTTGAATTCTGAGTAAACAGCAGATAACTGCCAAGTTCAG TTTTGTTCTTGTGCTGAGTGTCTTCCATGGATAATGCATTGATTATCAGT GAAGATGCAGAAGGGAAATGGGGAGCCTCAGCTCACATTAGCTTATGGTGA CTCTGGGTTCTATGGCCTTGTGGAGGGGGCCAGGCTCTAGAACGTCTAAC CAGTGGAGAACCGAAACCCCCCCCCCCCCGCCACCCCTCTGGACAGTTATT TCTCTTCAATCTCTCTCCATCTCTTCAATCTCTGTCTTCCCTTGCA CCAGTCTCTCTCAATCCCCCTTCTAACACACACACACACACACACA CACACACACACACACACACACAGAGTCAGGCCGTTGCTAGTCAGTTC TCTCTTCCACCCCTGCTCCATCTTCAACTAGTGTCTCTCTCTCTCA GGAGTGCAGCCCTGAGCCTGCCACTCTCATTAGCAAATGACTGTATTAAA GGAAATCTATTGTATCTACCTGCACTGCTCCATTGTTCCAGAGTGAAC TTATCTGTTATTATTTGAAATAATAAGACCTTAACATTACGCGCTAA CATTATCGTTGTTGAGTACCTAAAGCTCCAGGCCAGGTTGGGGAAAGAG GAAGCATTGGAGGGAAATTCTCAACCTTGTGATGTTCTATAAAACTTGT CTCAAGCTACTTACATTACCGTACTAGTACCGCTAAAGAGCAGAGTAAC TTAGTAGGCTCAGGGACTTGAGTGCTCTCTTAGGATCTGGACCTAGAAT TAATGTCATGAGATTCTAACAGGATAAGGTGAGGTAGTGAGGGCTGAAG TCATCCACTGGGTTATCCAATATTAGGTTACTGCTGACAAAAGAGGGGC TTCTGGTCTGGTGGTTATTGTGTTGG |
| | - GFP | | |
| | - relleno | | |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|--|
| 5 | gARN-G_HA | (SEQ ID NO: 38) | |
| | + ADNc | | TTCTTTTATAAAGCACTGCATCACAAACACTAAAATGAAGTGGGCAAATTAG |
| | + GFP | | CTCTGCAGAAAACATTTCTAGGTGATGTTATAATGACCAATCATTACTGA |
| | + Relleno | | AGCAATGAGAAATGTGACAATTACAGAAATTGCTGCTAGTAGTATGTTGAAAA AATATGCATTTGAGTGAACATTAGTAGAATAGCTGATTTCTTACTCTTA ATTCAATTACATATTGTGCGGTGAATTAGGGAGCCATAATGCGGTTACAAT AATTCCTATACTAAATATAACAAAGATTAAATTCAAAAAATGGTTACCAGC ATCGTTAGTGCGTATACATCAAGAGGGACGTGCCCGGAGACAGCAAGTAAG CTCTTAAACATGCTTGAACATACGATTTAATAAAACATGAGCATTTGAATA AAAACGACTTCTCTACTGTAAACATCACGCGATGCACATTAGACAATAGATA CAGAGATGCAACACAGGAGAAGCTTCTTAGGTTACATGTGCTCTAAC AGCAGAACGGGCTGACCTCTCTTCTCCCACAGATCGAGGAGCAGAGAGAA <u>ACCTGCACGAGGACTTCGTGTCATGAAGACCATCCAGCGGTGCAACACCGG</u> <u>CGAGAGAAGTCGAGGCTGTCGAACGAGGAAATCAAGAGCCAGTCGA</u> <u>GGGCTCGTGAAGGACATCATGCTGAACAAAGAGGAAACAGAAAGAAAAA</u> <u>CTCCCTCGAGATGCAAGGGCGACCAATCTCAGATGCCGCTCACGTG</u> <u>ATCAGCGAGGCCAGCAGCAAGACAACAAGCGTGTGAGTGGGGAGAAG</u> <u>GGCTACTACACCATGAGCAACACCTGGTACCCCTGAAAAACGGCAAGCAGC</u> <u>TGACAGTGAAGGGCAGGGCTGTACTACATCTACGCCAAGTGACCTCTGC</u> <u>AGCAACAGAGAGGCCAGCTCTAGGCCCTTTATGCCAGCTGTGCTGAA</u> <u>GTCCTCTGGCAGATTGAGGGATTCTGCTGAGAGCCCAACACACAGC</u> <u>AGGCCAAACCTTGCGCAGCACTTACATCTTCGGACTGTGACAGACCC</u> <u>GCAGCCTGGCGCAAGCGTGTGTAATGTGACAGACCCAGGTGTCC</u> <u>CACGGCACCGGCTTACATCTTCGGACTGTGAGCTGTGAAACAGTGTACC</u> <u>TGCAAGGCTGTGGAGCTGACGCTGGAGTCTTCTATAATACAGCACAGC</u> <u>GTTAAGCCCACCCCTGTTAAGTGCCTTATAACCTAGGATCTCTTATG</u> <u>GAGAACTATTATTATAACTCCAAGGCATGTAGAACTGTAATAAGTGAATT</u> <u>CAGGTACATGAAACCAACAGGGCCCTGCTCCATAAGAGCTTATATCTGA</u> <u>AGCAGCAACCCACTGATGCAAGACATCCAGAGAGTCTATGAAAAGACAAG</u> <u>CCATTATGCAAGGTTGAATTCTGAGTAAACAGCAGATAACTGCAAGTTCA</u> <u>GTTTGTCTTCTGCGTCAGTGTCTTCCATGGATAATGCTTGTGATTATCA</u> <u>GTGAAGATGCAAGGGAAATGGGGAGCCTCAGCTCACATTGAGTATGGTT</u> <u>GACTCTGGTTCTATGGCTTGTGGAGGGGCCAGGCTAGAACAGTCTA</u> <u>ACACAGTGGAGAACCGAAACCCCCCCCCCGCACCCCTCTGGACAGTTATT</u> <u>CATTCTTTCAATCTCTCTCCATCTCTCTCTTCTGAGTCTCTCTCTCAACC</u> <u>TCTTCTTCCAACTCTCTTCTCAATCTCTCTGTTCCCTTGTGAGTCTCTCCC</u> <u>TCCCCCAGTCTCTCTCAATCCCCCTTCTAAACACACACACACACACAC</u> <u>ACACACACACACACACACACACACACACAGAGTCAGGCCGTTGCTAGTC</u> <u>GTTCTTCTTCTCCACCTGTCCTATCTTCTACACTATAGATGAGGGTGAGGA</u> <u>GTAGGGAGTGCAGCCCTGAGCTGCCACTCTCATTACGAAATGACTGTATT</u> <u>TAAAGGAAATCTATTGTATCTACCTGCACTGCTCCATTGTTCCAGAGTGAAC</u> <u>TTGTTCTCAAGCTACTACATACGCGTACTAGTTGGCTCCGTGCCGTCACT</u> <u>GGGCAGAGGCCACATGCCACAGTCCCCGAGAAGTGGGGGGAGGGTCTG</u> <u>GCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGTAAACTGGGAAAGTG</u> <u>ATGTCGTGACTGGCTCCGCTTTTCCCGAGGGTGGGGGAGAACCGTATATA</u> <u>AGTGCAGTAGTCGCGTGAACGTTCTTCTGCAACGGGTTGCCGCCAGAAC</u> <u>ACAGGTGTCGTGACCGGGATCGTACCCGGGTTAAGGGCGAATTCCAGCACAC</u> <u>TGGCGGCCGTTACTAGTGGATCCACCGGTGCCACCATGGTGAGCAAGGGCG</u> |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------|--------------------------------|--|
| | | | AGGAGCTTCAACGGGGTGGTGCCCACCTGGTCAGCTGGACGGCAGCT AAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTA CGGCAAGCTGACCTGAAGTTCTGCACCAACGGCAAGCTGCCGTGCCCT GGCCACCCCTGTGACCACCTGACCTACGGCGTGCAGTGCTTCAGCCCTAC CCCGACCACATGAAGCAGCACGACTTCAAGTCCGCCATGCCGAAGGCTA CGTCCAGGAGCGCACCATCTCTCAAGGACGACGGCAACTACAAGACCCGC GCCGAGGTGAAGTTCAGGGCGACACCTGGTAACCGCATCGAGCTGAAG GGCATCGACTCAAGGAGGACGGCAACATCCTGGGGACAAGCTGGAGTAC AACTACAACAGCCACAACGTCTATATCATGGCCACAAGCAGAAGAACGGCA TCAAGGTGAACTCAAGATCCGCCAACACATCGAGGACGGCAGCGTGCAGCT CGCCGACCACCTACAGCAGAACACCCCCATCGGCAGGGCCCGTGTGCTG CCGACAACCAACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCAACGA GAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT CTCGGCATGGACGAGCTGACAAGTAAGCGGCCGCGTGAAGTCTAGAGGG CCCGTTAAACCCGCTGATCAGCCTCGACTGTGCCCTTAGTTGCCAGCCATCT GTTTTGCCCTCCCCGTGCCCTTGACCCCTGGAAGGTGCCACTCCACT GCCCTTCTAATAAAATGAGGAATTGATCGCATGCTGTAGTAGGTGTCA TTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGG AAGACAATAGCAGGCATGCTGGGATGCGGTGGCTATGGTACAGCTTA ACTCGTTAATGACACATTGCAAATTAAACATCCAGTAACGAAACGGCTTCAGT CGTAATGCCCATATAGTGGCTACAGAATGTTGGATAGAGAACTTAAGTAG CTAAGGCGCGTATTCTTAATATTAGGGTATTGCCAGTCATTACAGA TAACCGCCTATGCCCATGCCAGGATTAGATAACTTTAACATTAGCCG CAGAGGTGGACTAGCACGTAATACCTGGAGTTCTGATAACATGACATCTTA ATTGCTGTCTTTATAGTTTAAACTGCAAATACAAATAGCAATCAGCAA TATAATAACTTATTCTCATTATGCCCTGAAAGTCCCTCTGTTGATGCC GTGGAAATGAATGTAGAGGCAGATATCATTAGCTGATTCTCCCTGAATGA CATTATCATATCCTGTTATTCCAAAATAGATAGAAGATGAAAGGAATCTCA TGAAGATTTGTATTGATGAAAAC |
| 6 | gARN-G_HA | (SEQ ID NO: 39) | |
| | + ADNc | | TTCTTTTATAAAGCACTGCATCACAAACACTAAATGAAGTGGGCAAATTAG |
| | + GFP | | CTCTGCAGAAAACATTCTAGGCTGATGTTATAATGACCAATCATTACTGA |
| | - Relleno | | AGCAATGAGAAATGTGACAATTACAGAAATTGCTGCTATAGTATGTTAAAA AATATGCATTTGTAGTGAACATTAGAATAGCTGTGATTCTGATACAGA GATGCAACACAGGAGAACGCTCCTAGGTTACATGTGCTCTAATTACAGCA GAACCGGTCTGACCTCTCTCCACAGATCGAGGACGAGAGAACCT <u>GCACGAGGACTCGTGTTCATGAAGACCATCCAGCGGTCAACACCGGCAG</u> <u>AGAAGTCTGAGCCTGCTGAACTCGCAGGAAATCAAGAGCCAGTCAGGGCT</u> <u>TCTGTAAGGACATCATGCTGAACAAAGAGGAAACGAAGAAAGAAAACCTCTT</u> <u>CGAGATGCGAGAAGGGCGACCGAGAACCTCAGATCGCGCTCACGTGATCAGC</u> <u>GAGGCCAGCAGCAAGACAACAAGCGTGCTGAGCTGGGCCGAGAAGGGCTAC</u> <u>TACACCATGAGCAACAACCTGGTCACCCCTGGAAAACGGCAAGCAGCTGACAG</u> <u>TGAAGCGGCAGGGCTGTACTACATCTACGCCAAGTGACCTCTGAGCAAC</u> <u>AGAGAGGCCAGCTCTAGGCCAGCTATTACACCTGGCGAGTGTGAGCTGAGCC</u> <u>AAACCTTGCGCCAGCAGTCTATTACACCTGGCGAGTGTGAGCTGAGCC</u> <u>TGGCGCAAGCGTGTGTAATGTGACAGACCCCTAGCCAGGTGTCCACGGC</u> <u>ACCGGCTTACATCTTGTGACTGCTGAAGCTGTGAACAGTGTACCTTGAG</u> <u>GCTGTGGTGGAGCTGACGCTGGAGTCTCATAATACAGCACAGCGGTTAAG</u> <u>CCCACCCCTGTTAATGCTTATTAAACCTAGGATCCTCCTTATGGAGAAC</u> <u>TATTATATACACTCCAAGGCATGTAGAACTGTAATAAGTGAATTACAGGTC</u> |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------|--------------------------------|---|
| | | | ACATGAAACCAAACGGGCCCTGCTCCATAAGAGCTTATATCTGAAGCAGC AACCCCCACTGATGCAGACATCCAGAGAGTCTATGAAAAGACAAGGCCATT TGCACAGGTTGAATTCTGAGTAAACAGCAGATAACTGCCAAGTTCAGTTTG TTCTTCTCGTGCAGTGTCTTCCATGGATAATGCATTGATTTACAGTGAAG ATGCAGAAGGGAAATGGGGAGCCTCAGCTCACATTAGTTATGGTTGACTCT GGGTTCCATGCCCTGTTGGAGGGGGCAGGCTCTAGAACGCTAACACAG TGGAGAACCGAAACCCCCCCCCCCCACCCCTCTGGAGCAGTTATCATTCTC TTCAATCTCTCTCTCCATCTCTCTCTCAGTCTCTCTCAACCTCTTCT TCCAATCTCTCTCTCAATCTCTCTGTCTCCCTTGTCACTCTCTCCCTCCCC AGTCTCTCTCTCAATCCCCCTTCTAACACACACACACACACACACACA CACACACACACACACACACAGAGTCAGGCCGTGCTAGTCAGTTCTC TTCTTCCACCCCTGCTCCATCTTACCACTATAGATGAGGGTGAGGAGTAGG GAGTGCAGCCCTGAGCCTGCCACTCTCATTACGAAATGACTGTATTAAAG GAAATCTATTGTATCTACCTGCAGTCTCATTGTTCAAGAGTGAACCTGTAAT TATCTTGTATTATTTTGAATAATAAAGACCTTAAACATTACCGCCTTAAC ATTATCGTTGTTGTTGAGTACCTAAAGCTCCAGCCAGGTGGGGAAAGAG GAAGCATTGGAGGAAATTCTCCAACCTTGTGATGTTTICATAAAACTTGT CTCAAGCTACITACATTACCGTACTAGITGGCTCCGTGCCCCGTAGTGGGC AGAGCGCACATGCCACAGTCCCCAGAAGTGGGGGGAGGGGTGGCAA TTGAACCGGTGCTAGAGAAGTGGCGGGTAAACTGGAAAGTGTATGATG CGTGTACTGGCTCCGCTTTTCCGAGGGTGGGGAGAACCGTATATAAGT GCAGTAGTCGCGTGAACGTTCTTCTCGAACGGGTTGCCGCAGAACACA GGTGTGTGACCGGGATCGTACCCGGGTTAAGGGCGAATTCCAGCACACTGG CGGGCGTTACTAGTGGATCACCCTGCGCCACCATGGTGAGCAAGGGCGAGG AGCTGTTACCCGGGTGGTGCCATCTGGTGAGCTGGACGGCGACGTAAA CGGCCACAGTTCAAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGG CAAGCTGACCCCTGAAGTTCTGACCTACGGCGTGCAGTGTCTAGCCGCTACCC CCACCCCTGTGACCACCCCTGACCTACGGCGTGCAGTGTCTAGCCGCTACCC GACCACATGAAGCAGCACGACTTCTCAAGTCCGCATGCCGAAGGCTACGT CCAGGAGCGCACCATTCTCAAGGAGCAGCGCAACTACAAGACCCCGGCC GAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGC ATCGACTTCAAGGAGGACGCGAACATCTGGGGACAAGCTGGAGTACAAC ACAACAGCCACAACGTCTATATCATGGCGACAAGCAGAAGAACGGCATCAA GGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC GACCACTACCGCAGAACACCCCCATGGCGACGGGGCGTGTGCTGCCCG ACAACCACTACCTGAGCACCCAGTCGCCCTGAGCAAGACCCACGAGAA GCGCGATCACATGGCTCTGTTGAGTGTGACCGCCGGGATCACTCTC GGCATGGACGAGCTGTACAAGTAAAGGGCGCGTGCAGTGTAGAGGGCCC GTTAAACCCGCTGATCAGCCTCGACTGTGCTTCTAGTGTGCAAGCCATCTGTT GTTGCCCTCCCCGTGCCCTCTGACCTGGAAAGGTGCCACTCCACTGCC CTTCTCTAATAAAATGAGGAAATTGATCGATTGTCAGTAGGTGTATTCT ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGA CAATAGCAGGCATGCTGGGATGCGGTGGCTCTATGGTACAGCTTAAC GTAACTGACACATTGCAAATTACACCTGGAGTTCTGATAACATGACATCT AATTGCTGTCTTATAGATTAAACTGCAAATACAAATAGCAATCAGCCA ATATAATAACTTATTCTCCATTATGCCGAAAGTCTCCTTGTGATGC CGTGGAAATGAATGTAGAGGGCAGATATCATTAGCTGATTCTCCTCCGAATG ACATTATCATATCCTTGTATTCCAAAATAGATAGAAGATGAAAGGAATCTC ATGAAGATTTGTATTGATGAAAAC |
| 7 | | gARN-G_HA | (SEQ ID NO: 40) |
| | | + ADNc | |
| | | - GFP | |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|--|
| | + Relleno | | TTCTTTTATAAAGCACTGCATCACAAACACTAAAATGAAGTGGGCAAATTAG CTCTGCAGAAAACATTTCAGGCTGATGTTATAATGACCAATCATTACTGA AGCAATGAGAAATGTGACAATTACAGAATATTGCTGCTATAGTAGTGTGAAAA AATATGCACTTGAGTGAACATTTAGTAGAATAGCTGATTTCTTACTCTTA ATTCAATTACATATTGTGGTGAATTCAAGGGAGCCATAATGCGGTTACAAT AATTCTTACATCTAAATACAAAGATTTAAATTCAAAATGGTTACCAGC ATCGTTAGTGCGTATACATCAAGAGGCACGTGCCCGGAGACAGCAAGTAAG CTCTTAAACATGCTTGTGACATACGATTTAAATAAAACATGAGCATTGAATA AAAAGCACTTCTCATACTGTAAACATCAAGCATGCGACATTAGACAATAGATA CAGAGATGCAACACAGGAGAAGCTTCTTAGGTTTACATGTGCTCTTAATTAC AGCAGAACCGGGTCTGACCTCTCTTCCACAGATGAGGACGAGAGAA <u>ACCTGCACGAGGAACTCGTGTCTATGAAAGACCATCAAGGGTCAACACCGG</u> <u>CGAGAGAAGTCAGGCTGTAACCTGCAGGAAATCAAGAGCCAGTTCGA</u> <u>GGGCTTCGTGAAGGACATCATGCTGAACAAAGAGGAAACCAAGAAAGAAAA</u> <u>CTCTTCGAGATGCAAGAGGGCAACAGAATCTCAGATGCCGCTCACGTG</u> <u>ATCAGCGAGGCCAGCAAGACAACAAGCGTGTGCAAGTGGGGGAGAAG</u> <u>GGCTACTACACCATGAGCAACAAACCTGGTCACTGGAAACGGCAAGCAGC</u> <u>TGACAGTGAAGCGGCAGGGCTGTACTACATCTACGCCAAGTGACCTCTGC</u> <u>AGCAACAGAGAGGCCAGCTCAGGCCCTTATGCCAGCCTGTGCTGAA</u> <u>GTCCCCCTGGCAGATTGAGGGGATTCTGCTGAGAGCCGCAACACACAGC</u> <u>AGCGCCAACACCTGTGGGCCAGCAGTCTATTACCTGGCGGAGTGTGAGCT</u> <u>GCAGCCTGGCGCAAGCGTGTCTGTGAATGTGACAGACCCTAGCCAGGTGTCC</u> <u>CACGGCACCGGCTTACATCTTCGACTGCTGAAGCTGTGAACAGTGTCA</u> <u>TTGCAAGCTGTGGTGGAGCTGACGCTGGAGTCTCTATAATACAGCACAGC</u> <u>GTAAAGCCACCCCTGTTAACTGCTATTATAACCTAGGATCTCTTATG</u> <u>GAGAACTATTATTATAACTCCAAGGATGTAAGAACTGTAATAAGTGAATT</u> <u>CAGGTCAATGAAACCAACGGGCCCTGCTCCATAAGAGCTTATATCTGA</u> <u>AGCAGCAACCCACTGATGCAAGACATCCAGAGAGTCTATGAAAAGACAAGG</u> <u>CCATTATGCACAGGTTGAATTCTGAGTAAACAGCAGATAACTGCAAGTTCA</u> <u>GTTTTGTCTTGCGTCAGTGTCTTCCATGGATAATGCAATTGATTATCA</u> <u>GTGAAGATGCAAGGGAAATGGGGAGCCTCAGCTCACATTCAAGTTATGGTT</u> <u>GAECTGGGTTCTATGGCTTGTGGAGGGGGCCAGGCTCTAGAACGTCTA</u> <u>ACACAGTGGAGAACCGAAACCCCCCCCCCGCCACCCCTCGGACAGTTATT</u> <u>CATTCTTCAATCTCTCTCCATCTCTCTTCACTGCTCTCTCTCAACC</u> <u>TCTTCTTCAATCTCTCTCAATCTCTCTCCATCTCTCTTGTCACTCTCTCCC</u> <u>TCCCCCAGTCTCTCTCAATCCCCCTTCTAACACACACACACACACACAC</u> <u>AC</u> <u>GTTCTCTCTTCCACCTGTCCCTATCTCTACCACTATAGATGAGGGTGGAGGA</u> <u>GTAGGGAGTGCAGCCCTGAGCCTGCCACTCTCATTACGAAATGACTGTATT</u> <u>TAAGGAAATCTATTGTATCTACCTGCACTGCTCCATTGTTCCAGAGTGAACCT</u> <u>GTAATTATCTGTTATTATTTGATAATAAAGACCTTAAACATTACGCC</u> <u>TTAACATTATGTTGTTGAGTACCTAAAGCTCCAGCCAGGTTGGGAA</u> <u>AGAGGAAGCATTGGAGGAAATTCCAAACCTTGTGATGTTCTATAAAACT</u> <u>TTGTTCTCAAGCTACTTACATTACGCTACTAGTTACAGCTTAACCTGTTAAC</u> <u>TGACACATTGCAAATTAAACATCCAGTAACGAAACGGCTCAGTCGAATCGCC</u> <u>CATATAGTGGTACAGAATGTTGAGTACGAACTTAAGTACGCTAACGGGG</u> <u>CGTATTCTTAAATATTAGGGGATTGCGCAGTCATTACAGATAACCGCCTA</u> <u>TGCGGCCATGCCAGGATTAGATAACTTTAACATTAGCCGAGAGGTGG</u> <u>GAECTAGCAGTAATACCTGGAGTTCTGATAACATGACATCTTAAATTGCTGTCT</u> <u>TTTATGATTAAACTGCAAATACAAAGATGCAATCAGCCAATAATAACT</u> |
| | | | TATTATTCTCCATTATGCTGAAAGTCTCTCTTGTGATGCCGTGGAAATG AATGTAGAGGCAGATATCATTAGCTGATTCTCTCCGAATGACATTATCAT ATCCTTGTATTCCAAAATAGATAGAAGATGAAAGGAATCTCATGAAGATT TGTATTGAAACAC |
| 8 | gARN-G_HA | (SEQ ID NO: 41) | |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|---|
| | | + ADNc | TTCTTTTATAAAGCACTGCATCACAAACACTAAAATGAAGTGGGCAAATTAG |
| | | - GFP | CTCTGCAGAAAACATTCTAGGCTGATGTTATAATGACCAATCATTACTGA |
| | | - Relleno | AGCAATGAGAAATGTGACAATTACAGAAATTGCTGCTATAGTATGTTAAAAA AATATGCATTTGAGTGAACATTAGTAGAATAGCTCTGATTCTGATAACAGA GATGCAACACAGGAGAACGCTCTTAGGTTACATGTGCTCTAATTACAGCA GAACCGGTCTGACCTCTCTCCCTCCCACAGATCGAGGACGAGAGAAACCT <u>GCACGAGGACTTCGTGTTATGAAGACCATCCAGCGGTCAACACCGGCAG</u> <u>AGAAGTCTGAGCCTGCTGAACTCGCAGGAAATCAAGAGGCCAGTCAGGGCT</u> <u>TCGTGAAGGACATCATGCTGAACAAAGAGGAAACGAAGAAAAGAAAACCTCTT</u> <u>CGAGATGCAGAAGGGCGACCAGAACCTCAGATGCCGCTCACGTGATCAGC</u> <u>GAGGCCAGCAGCAAGACAACAAGCGTGCTCAGTGGGCCAGAAGGGCTAC</u> <u>TACACCATGAGCAACAACCTGGTACCCCTGGAAAACGGCAAGCAGCTGACAG</u> <u>TGAAGCGGCAGGGCCTGTAATACATCTACGCCAAGTGACCTCTGAGCAAC</u> <u>AGAGAGGCCAGCTCTCAGGCCCCCTTTATCGCCAGCCTGCTGAAGTCCCC</u> <u>TGGCAGATTGAGCGGATTCTGCTGAGAGGCCAACACACAGCAGCGCC</u> <u>AAACCTTGCCAGCAGCTTACCTCACCTCGCGAGTGTGAGCTGAGCC</u> <u>TGGCGCAAGCGTGTGTAATGTGACAGACCTAGCCAGGTGTCCCACGGC</u> <u>ACCGGCTTACATCTCGACTGCTGAAGCTGTGAACAGTGTACCTTGAG</u> GCTGTTGGAGCTGACGCTGGAGTCTCATAATACAGCACAGCGGTTAAG CCCACCCCTGTTAATGCTTATTAACCTAGGATCCTCCTTATGGAGAAC TATTATTATACACTCCAAGGCATGTAGAACTGTAATAAGTGAATTACAGGTC ACATGAAACCAAAACGGGCCCTGCTCCATAAGAGCTTATATCTGAAGCAGC AACCCACTGATGCAGACATCCAGAGACTCTATGAAAGACAAGGCCATTA TGCACAGGTTGAATTCTGAGTAAACAGCAGATAACTGCCAAGTTCACTTGT TTCTTGCCTGAGTGTCTTCCATGGATAATGCAATTGATTATCAGTGAAG ATGCAGAAGGAAATGGGGAGCCTCAGCTCACATTCTAGTTATGGTTGACTCT GGGTTCTATGCCCTGTTGGAGGGGCCAGGCTCTAGAACGCTAACACAG TGGAGAACGAAACCCCCCCCCCCCCGCCACCCCTCTGGACAGTTATTCTC TTCAATCTCTCTCCATCTCTTCACTGCTCTTCCATGGATAATGCAATTGATT TCCAACTCTCTTCTCAATCCCCCTTCAACACACACACACACACACACA AGTCTCTTCTCAATCCCCCTTCAACACACACACAGAGTCAGGGCTGCTAGT TTCTTCCACCTGCTCTTCTCTTCACTGCTCTTCCATGGATAATGCAATTG GAGTCAGCCCTGAGCCTGCCACTCTCATTAGCAGGAAATGACTGTATTAAAG GAAATCTATTGATCTACCTGCAGTCTCATTGTTCCAGAGTGAACCTGTAAT TATCTGTTATTATTTGAATAATAAGACCTCTAACATTACGGCTAAC ATTATCGTTGTTGAGTACCTAAAGCTCCAGCCAGGTTGGGGAAAGAG GAAGCATTGGAGGGAAATTCCAACCTTGTGATGTTTCTAAACACTTGT CTCAAGCTACTTACATTACCGTACTAGTTACCAAGCTTAACCTGTTA CATTGCAAATTAAACACCTGGAGTTCTGATAACATGACATCTTAATTGCT TTTATAGATTTAAACTGCAAATACAAAATAGCAATCAGCCAATATAAACT |
| | | | TATTATTCTCCATTATGCCGAAAGCTCTCTTGTGATGCCGTGGAAATG AATGTAGAGGCAGATATCATTAGCTGATTCTCCTCCGAATGACATTATCAT ATCCTTGTATTCCAAAATAGATAGAAGATGAAAGGAATCTCATGAAGATT TGTATTGAAAC |
| 9 | gARN-C_HA | (SEQ ID NO: 42) | |
| | + ADNc | | |
| | + GFP | | |

| Nombre del donante | del | Descripción del donante | Secuencia de donantes |
|--------------------|-----|-------------------------|---|
| | | + Relleno | GATACTTGAGTGTCCCTCTTAGGATCTGGACCTAGAATTAAATGTCATGAGAT TTTCTAACAGGATAAGGTGAGGTAGTGAGGGCTGAAGTCATCCACTGGTT ATCCAATATTAGGTTCACTGCTGACAAAAGAGGGGCTCTGGTCTGGTTG GTTATTGTGTTGGCCTGATGTGCTCTGCAATCAAATGTATGGACATAGGC CTAGCTCTAAAGGGCAATAGTGACCTCAGTACTCTAATTCTTACATATTG TGGGTCGAATTCAAGGGAGCCGATAATGCGGTTACAATAATTCTTACATATTAA ATATACAAAGATTAAAATTCAAAAATGGTACAGCATCGTTAGTGCCTA TAGCAAGTGAATTCTGTACATTAAATTCTAAAGCTCTTAGGTTACAT GTGCTCTAATTACAGCAGAACCGGCTGACCTCTCTCCCTCCCACAGATC GAGGACGAGAGAACCTGACGAGGACTTCGTGTTCATGAAGACCATCCAGC <u>GGTCAACACCCGGCAGAGAGAAGTCTGAGCCTGCTGAAC</u> TGCGAGGAATCA AGAGCCAGTCGAGGGCTCGTGAAGGACATCATGCTGAACAAAGAGGAAA CGAAGAAAAGAAAACTCCTCGAGATGCAGAAGGGCAGCAGAATCCTCAGAT CGCCGCTCACGTGATCAGCGAGGCCAGCAGCAAGACAACAAGCGTGTGAG TGGCCGAGAAGGGCTACTACACCATGAGCAACACCTGGTACCCCTGGAAA ACGGCAAGCAGCTGACAGTGAAGCGGCAGGGCTGTAACATCTACGCCA AGTACCTTCTGCAGCAACAGAGAGGCCAGCTCTCAGGCCCCTTATGCCA GCCTGTGCTGAAGTCCCCGGCAGATTGAGCGGATTCTGCTGAGAGCCG CAACACACACAGCAGCGCAAACCTTGTGGCCAGCAGTCTATTCACCTCGCG GAGTGTTGAGCTGCAGCTGGCGCAAGCGTGTGTAATGTGACAGACCC TAGCCAGGTGTCACCGGACCGGCTTACATCTTCCGACTGCTGAAGCTGT <u>GAACAGTGTACCTTGAGGCTGTGGAGCTGACGCTGGAGTCTTACAT</u> ATACAGCACAGCGGTTAACCCCACCCCTGTTAACCTGCTATTATAACCTAG GATCCTCCTATGGAGAACTATTATACACTCCAAGGCATGTAGAACTGTAA ATAAGTGAATTACAGGTACATGAAACAAAACGGGCCCTGCTCCATAAGAG CTTATATATCTGAAGCAGCAACCCCACGTGACAGGTTGAATTCTGAGTAACAGCAGATAA CTTGCCAAGTTCAGTTGTTCTTGCCTGAGTGTCTTCCATGGATAATGC ATTTGATTATCAGTGAAGATGCAGAAGGGAAATGGGAGGCTCAGCTCACA TTCAGTTATGGTGAACCTGGGTTCTATGGCTTGTGGAGGGGCCAGGCT CTAGAACGTCTAACACAGTGGAGAACGAAACCCCCCCCCCGCCACCCCT CGGACAGTTATTCTTCAATCTCTCTCCATCTCTCTTCAACACACACA CAGAGTCAGGC CGTTGCTAGTCAGTTCTCTTCCACCCCTGTCCTATCTTCAACTCTCT TCAGTCTCTCCCTCCCCAGTCTCTCTCAATCCCCCTTCAACACACACA CAGAGTCAGGC GAGGGTGAGGAGTAGGGAGTGCAGCCCTGAGCCCTGCCCCACTCTCATTACGA AATGACTGTATTAAAGGAAATCTATTGTATCTACCTGCAGTCTCCATTGTT CAGAGTGAACCTGAAATTCTTGTATTATTTTGAATAATAAGACCTT AACATTACCGCCTAACATTATCGTTGTGTTGAGTACCTAACGCTCCAGCC AGGTTGGGAAAGAGGAAGCATTGGAGGGAAATTCCAACCTTGTGATG |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|--|
| | | | TTTCATAAACTTGTCTCAAGCTACTTACATTACCGGTACTAGTTGGCTCCG GTGCCCGTCAGTGGCAGAGCGCACATGCCAACAGTCCCAGAAGTGGG GGGAGGGGTCGGCAATTGAACCGGTGCTAGAGAAGGTGGCGGGGTAAG ACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTCCCAGGGTGGGG AGAACCGTATAAAGTCAGTAGTCGCCGTGAACGTTCTTCGCAACGGG TTGCCGCCAGAACACAGGTGTCGTGACGCCGATCGTACCCGGGTTAAGGG AATTCAGCACACTGGCGGCCGTTACTAGTGGATCCACCAGGTCGCCACCATGG TGAGCAAGGGCAGGAGCTGTTACCGGGGTTGGTGGCCATCTGGTCGAGCT GGACGGCGACGTAACGCCACAAGTCAGCGTCCGGCGAGGGCGAGGG CGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAACGGCAAGC TGCCCCTGCCCCACCCCTCGTGACCAACCCCTGACCTACGGCGTGCAGTGC TTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCAAGTCCGCAT GCCCGAAGGCTACGTCAGGAGCGCACCACATCTTCAGGACGACGGCAAC TACAAGACCCGCCGGAGGTGAAGTTGAGGGCGACACCCGGTGAACCGC ATCGAGCTGAAGGGCATGACTTCAAGGAGGACGGCAACATCTGGGGCAC AAGCTGGAGTACAACACAGCCACAAGTCTATATCATGCCGACAAGC AGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGG CAGCGTGCAGCTCGCGACCACCTACCGCAGAACACCCCCATGGCGACGGC CCCGTGTGCTGCCGACAACCAACTACCTGAGCACCCAGTCCGCCGTGAGCAA AGACCCCAACGAGAACGGCATCACATGGCTCTGGAGTTGACCG GCCGGGATCACTCTGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCTCG AGTCTAGAGGGCCGTTAAACCGCTGATCAGCCTGACTGTGCTCTAGT TGCCAGCCATCTGTTGTTGCCCTCCCCGTGCTTCTGACCTGGAAAGGT GCCACTCCCCTGCTTCTTAATAAAATGAGGAATTGCATCGCATTGCTG AGTAGGTGTCATTCTATTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GAGGATTGGGAAGACAATAGCAGGCTGCTGGGGATGCGGTGGGCTATG GTACCAGCTTCACTGGATGATTGACTTATCAGGCATCAAGAGGACAGTGC CCCGGAGACAGCAAGTAAGCTTTAACATGCTTACAGATACGATTAAAT AAAACATGAGCATTGAATAAAACGACTCCTCATACTGTAACACATCACGCA TGCACATTAGACAATAATCCAGTAACGAAACGGCTCAGTCGTAATGCCCAT ATAGTTGGCTACAGAATGTTGAGAGAACTTAAGTACGCTAAGGCGGCGT ATTITCTTAATATTGGACTGATATTACCGTACTATTACATGTGCTTAATT ACAGCAGAAGCTGCCAGCTAACTGAATCTGTTGAATCTAAAAACTACT CTTAAAGCAAGAAATGGTATAAAATTAGTTGATAAT |
| 10 | gARN-D_HA | (SEQ ID NO: 43) | |
| | + ADNc | | AGATCTTGATTAGTTCTGGCTGTTCTAAATTCTGGATGCAGGAACGTG GCTAGAAAGCATCTGGATGATTGCACTTATCAGGGATACTTGAGTGTCTCT |
| | + GFP | | CTTAGGATCTGGACCTAGAATTATGTCATGAGATTCTAACAGGATAAGG TGAGGTAGTGGGGCTGAAGTCATTACTCTTAATTACATATTGTGCGGT CGAATTCAAGGAGCCGATAATGCCGTACAATAATTCTATACTTAAATATAC AAAGATTAAATTCAAAAAATGGTACAGCATCGTTAGTGCCTATACATC AAGAGGACCGTCCCCGGAGACAGCAAGTAAGCTTAAACATGCTTGAC ATACGATTAAATAAAACATGAGCATTGAATAAAACGACTCCTCATACTG TAAACATCACGCATGCACATTAGACAATAATCCAGTAACGAAAGAATTCTGTA CATTTAATTCTAAGACATTGGAGCTTCTAGGTTACATGTGCTTTAA TTACAGCAGAACCGGCTGACCTCTTCTTCTCCACAGATCGAGGACGAG AGAAACCTGCACTGGGGACTTCGTGTTCTGAGACCATCCAGGGTGCAACA CCGGCGAGAGAAAGTCTGAGGCTGTAAGGACATCATGCTGAACAAAGAGGAAACGAAG TCGAGGGCTCGTGAAGGACATCATGCTGAACAAAGAGGAAACGAAGAAG AAAACCTCTCGAGATGCAAGAAGGGCGACCAGAATCCTCAGATGCCGCTCA CGTATCAGCGAGGCCAGCAGCAAGACAACAAGCGTGTGCAAGTGGGCCGA |
| | + Relleno | | |

ES 2 998 052 T3

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|---|
| | | | <p>GAAGGGCTACTACACCATGAGCAACAACCTGGTCACCCGGAAAACGGCAAG <u>CAGCTGACAGTGAA</u>GCGCAGGGCCTGTACTACATCTACGCCAAGTGACCT <u>TCTG</u>CAGCAACAGAGAGGGCCAGCTCTCAGGCCCTTTATGCCAGCCTGTGC <u>CTGAAGTCCCCTGG</u>CAGATTGAGCGGATCTGCTGAGAGGCCAACACAC <u>ACAGCAGCGCAA</u>ACCTTGTGGCCAGCAGTCTTACACCTGGCGAGTGTT <u>GAGCTGCAGCCTGGC</u>GAAGCGTGTGAAATGTGACAGACCCTAGCCAGG <u>TG</u>TCCCACGGCACCGGTTACATCTTCCGACTGTGAAAGCTGTGAAACAGTG <u>TCAC</u>CTTGAGGTGTGGTGAGCTGACGCTGGGAGTCTCATAATAACGCA <u>CAGCGGTTAAGCCCACCCCTG</u>TAACTGCCATTATAACCCTAGGATCCTCC <u>TTATGAGAACT</u>ATTATTATACACTCCAAGGCATGAGAACTGTATAAGTG <u>AATTACAGGT</u>CACATGAAACAAAACGGCCCTGCTCCATAAGAGCTTATATA <u>TCTGAAGCAGCAACCC</u>ACTGATGAGACATCCAGAGAGTCCTATGAAAAGA <u>CAAGGCCATTATG</u>CACAGGTGAATTCTGAGTAAACAGCAGATAACCTTGCAA <u>GTTAGT</u>TTTGTCTTGCCTGAGTGTCTTCCATGGATAATGCATTGATT <u>ATCAGTGAAGATG</u>CAGAAGGGAAATGGGAGGCTCAGCTCACATTCAAGTT <u>GGTTGACTCTGGG</u>TCCTATGGCTGTTGGAGGGGGCAGGCTCTAGAACG <u>TCTAACACAGTGGAGAACC</u>GAACCCCCCCCCCCCCGCACCCCTCAGGACAGT <u>TATT</u>CATTCTCTTCAATCTCTCTCCATTCTCTCTCAGTCTCTCTCTC <u>AACCTCTTCTTCA</u>ATCTCTCTCAATCCCCCTTCAACACACACACACACA <u>CACACACACACACACACACACACACACACACAGAGTCAGGCCGTTGCTA</u> <u>GTCAGTTCTCTTCTTCCACCC</u>GTCTTCTCTCTCTCTCTCTCTCTCTCT <u>AGGAGTAGGGAGTGCAGC</u>CCCTGAGCTCCCCTCCTCATTACGAAATGACT <u>GTATTTAAAGGAAATCTATTG</u>TATCTACCTGAGTCTCATTGTTCCAGAGTG <u>AACTTGT</u>TAATTATCTGTTATTATTGAAATAATAAAGACCTTAAACATTA <u>CGCGCTTAACATTATCG</u>TTGTTGAGTACCTAAAGCTCCAGCCAGGTTG <u>GGGAAAGAGGAAGCAT</u>TTGGAGGGAAATTTCCTAACCTTGTGATGTTTCA <u>AAACTTTGTCTCAAGT</u>ACTTACATTACCGCTACTAGTTGGCTCCCGTGC <u>TCAGTGGGCAGAGCGCACATCG</u>CCCACAGTCCCGAGAAGTTGGGGGAGG <u>GGTCGGCA</u>ATGAACCGGTGCCTAGAGAAGGTGGCGGGGTAACTGAAA <u>AAAGTGA</u>GTGCTGACTGGCTCCGCTTCCGAGGGTGGGGAGAACCG <u>TATATAAGTGCAGT</u>AGTCGCCGTGAACGTTCTTTCGCAACGGGTTGCCGC <u>CAGAACACAGGTG</u>CGTGACGCCATCGTACCCGGGTTAAGGGCGAACCTCA <u>GCACACTGGCGCC</u>TTACTAGTGGATCCACCGGTGCCACCATGGTAGCA <u>AGGGCGAGGAGCTG</u>TTCACCGGGGTGGGCCATCTGGTGAGCTGGACG <u>GCGACGTAACCGGCC</u>ACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG <u>CCACCTACGGCAAGT</u>GACCTGAAGTTCATCTGCACCCACGGCAAGCTGCC <u>GTGCCCTGGCC</u>ACCCCTCGTGACCAACCTGACCTACGGCGTGAAGTCTCAG <u>CCGCTACCCGACCACATG</u>AAGCAGCAGCACTTCTCAAGTCCGCACTGCCCG <u>AAGGCTACGT</u>CCAGGAGCGCACCATCTTCAAGGACGACGGCAACTACAA <u>SACCCGCGCC</u>GAGGTGAAGTCAGGGGGACACCCCTGGTGACCGCATCGA <u>GCTGAAGGGCATG</u>ACTTCAAGGAGGAGGGCAACATCTGGGGCACAAGCT <u>GGAGTACA</u>ACTACAACAGCACAACGTCTATATCATGCCGACAAGCAGAAG <u>AA</u>CGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGAGGGCAGCG <u>TGCAGCTGCCG</u>ACCAACTACAGCAGAACACCCCATGGCGACGGCCCGT <u>GCTGCTGCCG</u>ACAACCAACTACCTGAGCAGGCCAGTCCGCCCTGAGCAAAGACC <u>CCAACGAGAAGCGCG</u>ATCACATGGCTCTGCTGAGTTGAGTCCGACCGCCCG <u>GATCACTCTGCCG</u>CATGGACGAGCTGACAAGTAAAGCGGCCGCTGAGTCT <u>AGAGGGCCG</u>TTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCA <u>GCCATCTGTTG</u>TTGCCCTCCCCGTGCTTCTTGACCCCTGGAAAGGTGCCAC <u>TCCCACTGCC</u>TTCTCTAATAAAATGAGGAATTGCATCGCATTGCTGAGTA </p> |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|---|
| | | | GGTGTCAATTCTATTCGGGGGGTGGGGTGGGCAGGACAGCAAGGGGGAG GATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGCTATGGTA CCAGCTTAAATACCAATCATAGATGTAAGGAGAAAGCGGCCAGTCAGTCAA TCGCCATATAGTTGGCTACAGAATGTTGGATAGAGAACTTAAGTACGCTAAC GCGCGTATTTCTTAATATTAGGGTATTGCCGCAGTCAGATAACC GCCTATGCAGGCCATGCCAGGATTAGATAACTTTAACATTAGCCGCAGAG GTGGCCACTGGGTTACCAAATTAGGTTCACTGCTGACAAAAGAGGGGG CTTCTGGTCTGGTGGTTATTGTGTTGGCTGATGTGCTGTCAATCAAAT GTATGGACATAGGCCTAGCTTCTAAAGGGGCAATAGTGACCTCAGTGGACTG ATATTACCGTACTATTACATGTGCTTAATTACAGCAGAAGCTGCCAGCTA ACTGAATCTTGAATCTAAAGGAACTACTCTAAAGCAAGAAAATGGTAT AAAATTAGTTGATAATGCAAGTAGATCT |
| 11 | gARN-D_HA | (SEQ ID NO: 44) | |
| | + ADNc | | AGATCTTGATTAGTTCTGGCTGTTCTAAAATTCTGGATGCAGGAACGTG GCTAGAAAGCATCTGGATGATTGCACTTTATCAGGGATACTTGAGTGTCT CTTAGGATCTGGACCTAGAATTAATGTCATGAGATTCTAACAGGATAAGG TGAGGTAGTGAGGGCTGAAGTCATGAATTCTGTACATTAAATTCTAACAGAC ATTGGAAGCTCCTTAGGTTACATGTGCTCTTAATTACAGCAGAACGGCT GACCTCTCTCCCTCCACAGATCGAGGACGAGAGAAACCTGCACGAGGAC <u>TTCGTGTTCATGAAGACCATCCAGCGGTCAACACCGCGAGAGAAAGTCTGA</u> <u>GCCTGCTGAACTGCGAGGAAATCAAGAGCCAGTCGAGGGCTCGTGAAGG</u> <u>ACATCATGCTGAACAAAGAGGAAACGAAGAAAGAAAACCTTCGAGATGCA</u> <u>GAAGGGCGACCAGAACATCTCAGATCGCCGTCACGTGATCAGCGAGGCCAGC</u> <u>AGCAAGACAACAAGCGTGTGAGTGGGCGAGAAGGGCTACTACACCATG</u> <u>AGCAACAACCTGGTCACCTGGAAACCGCAAGCAGCTGACAGTGAAGCGG</u> <u>CAGGGCTGTACTACATCTACGCCAAGTGACCTCTGCAGCACAGAGAGG</u> <u>CCAGCTCTCAGGCCCTTATGCCAGCCTGTGCTGAAGTCCCCTGGCAGA</u> <u>TCAGGAGCGATTCTGCTGAGAGCCGCAACACACAGCAGCGCCAACCTT</u> <u>GTGGCCAGCAGTCTATTACCTCGGCGGAGTGTGAGCTGCAGCCTGGCG</u> <u>AAGCGTGTGTAATGTGACAGACCTAGCCAGGTGTCACCGCACCGC</u> TITACATCTTGGACTGCTGAAGCTGTGAAACAGTGTACCTTGAGGCTGTG GTGGAGCTGACGCTGGAGTCTCATAATACAGCACAGCGGTTAAGCCCACC CCCTGTTAATGCCTATTATAACCTAGGATCCTCTTATGGAGAAACTATTAA TTATACACTCCAAGGCATGTAGAACTGTAATAAGTAATTACAGGTACATGA AACCAAAACGGGCCCTGCTCCATAAGAGCTTATATCTGAAGCAGCAACCCC ACTGATGCGACACATCCAGAGAGTCCTATGAAAAGACAAGGCATTATGACA GGTGAATTCTGAGTAAACAGCAGATAACTTGCAAGTTCAAGTTGTTCTT GCGTGCAGTGTCTTCCATGGATAATGCAATTGATTATCAGTGAAGATGCG AAGGGAAATGGGAGCCTCAGCTCACATTCACTGTTATGGTTGACTCTGGGTC CTATGGCCTTGTGGAGGGGCCAGGCTCTAGAACGCTAACACAGTGGAGA ACCGAAACCCCCCCCCCCCCGCCACCTCTCGGACAGTTATTCTCTTCAA TCTCTCTCTCCATCTCTCTTCACTGCTCTCTCAACCTCTTCTCCTCAA CTCTCTCTCAATCTCTGTTCCCTTGTGAGTCACTGCTCTTCTCCTCCTC TCTCTCAATCCCCCTTCTAACACACACACACACACACACACACACA CACACACACACACACACAGAGTCAGGCCGTTGCTAGTCAGTTCTCTTCT CACCTGTCCTATCTCTACCACTATAGATGAGGGTGAGGAGTAGGGAGTC AGCCCTGAGCCTGCCACTCTCATTACGAAATGACTGTATTAAAGGAAC TATTGTATCTACCTGCAGTCTCCATTGTTCCAGAGTGAACTTGAAATTATCTG TTATTTATTGAAATAATAAGACCTCTAACATTACGCGCTAACATTATCG TTGTTGTTGAGTACCTAAAGCTCCAGCCAGGTTGGGAAAGAGGAAGCAT TTGGAGGGAATTTCACCTTGTGATGTTCTAACACTTGTCAAG |
| | + GFP | | |
| | - Relleno | | |

| Nombre del donante | del donante | Descripción del donante | Secuencia de donantes |
|--------------------|-------------|-------------------------|--|
| | | | TACCTACATTACCGTACTAGTGGCTCCGTGCCGTCAGTGGGCAGAGCGC ACATCGCCCACAGTCCCCGAGAAGTGGGGGAGGGTCGGAATTGAACC GGTGCCTAGAGAAGGTGGCGGGTAAACTGGAAAGTGTGCGTAG TGGCTCCGCCTTTTCCGAGGGTGGGGAGAACCGTATAAGTGAGTAG TCGCCGTGAACGTTCTTTCGCAACGGTTGCCAGAACACAGGTGCG TGACGCGGATCGTACCCGGTTAAGGGCGAATTCAAGCACACTGGCGGCCGT TACTAGTGGATCCACCGGTGCCACATGGTGAGCAAGGGCGAGGAGCTT CACCGGGGTGGTGCCTACCTGGTCGAGCTGGACGGCGACGTAACGGCAC AAGTTCAAGCGTGTCCGGCAGGGCGAGGGCGATGCCACCTACGGCAAGCTG ACCTGAAGTTCATCTGACCCACCGGAAGCTGCCGTGCCCTGGCCAC CGTGACCAACCTGACCTACGGCGTCAGTGTCTCAGCGCTACCCGAC TGAAGCAGCACGACTCTCAAGTCCGCGATGCCAAGGCTACGTCCAGGA GCGCACCATCTTCAAGGACGACGGCAACTACAAGACCCGCGCCAGGTG AAGTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCG TCAAGGAGGACGGCAACATCTGGGCAAAAGCTGGAGTACAACATAACA GCCACAACGTCTATATCATGGCGACAAGCAGAAGAACGGCATCAAGGTGAA CTTCAAGATCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCGAC TACAGCAGAACACCCCCATCGGCACGGCCCCGTGCTGCCGACAACC ACTACCTGAGCACCCAGTCCGCGCTGAGCAAAGACCCAACGAGAAC TCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGATCACTCTGGCATGG ACGAGCTGTACAAGTAAAGCGGCCGCGTCAGTGTAGAGGGCCGTTAAC CCGCTGATCAGCCTGACTGTGCTCTAGTTGCCAGCCATCTGTTGCC CTCCCCCGTGCCTCTGACCTGGAGGTGCCACTCCACTGCCCTTCTA ATAAAATGAGGAAATTGCATCGATTGCTGAGTAGGTGTCAATTCTGG GGGGTGGGGTGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC AGGCATGCTGGGATGCGGTGGCTATGGTACAGCTTAAATACCAATC ATAGATGTAAGGAGAAAGCCACTGGTTATCCAATATTAGTTCACTGCT GACAAAAGAGGGGCTCTGGCTGGTGGTTATTGTGTTGGCCTGATGT GCTCTGCAATCAAATGTATGGACATAGGCCAGCTTCTAAAGGGCAATAGT GACCTCAGTGGACTGATATTACCGTACTATTACATGTGCTCTAAATTACAGC AGAAGCTGCCAGCTAAGTGAATTGTTGAATCTAAAAACTACTCTTAA GCAAGAAAATGGTATAAAATTAGTTGATAATGCAAGTAGATCT |
| 12 | gARN-D_HA | (SEQ ID NO: 45) | |
| | + ADNc | | AGATCTTGATTAGTTCTGGCTGTTCTAAATTCTGGATGAGGAACGTG GCTAGAAAGCATCTGGATGTTGACTTATCAGGGATACTTGAGTGTCT CTTAGGATCTGGACCTAGAATTATGTATGAGATTCTAACAGGATAAGG TGAGGTAGTGGGGCTGAAGTCATTACTCTTAATTCTAACATATTGCG CGAATTCAAGGAGCCGATAATGCGGTTACAATAATTCTATACTAAATATAC AAAGATTAAATTCAAAAATGGTACAGCATCGTAGTGTGCTATACATC AAGAGGACGTGCCGGAGACAGCAAGTAAGCTTAAACATGCTTGA ATACGATTAAATAAAATGAGCATTGAATAAAACGACTCTCTCATACTG TAAACATCACGCGATGCACATTAGACAATAATCCAGTAACGAAAGAATTCTGTA CATTTAATTATTCTAAGACATTGGAGCTCTTCTAGGTTACATGTGCTTAA TTACAGCAGAACGGCTGACCTCTTCTCCACAGATGAGGAC AGAAAACCTGCACTGGGACTTCGTGTTCATGAAGACCATCCAGCGGTG CCGGCGAGAGAAGTCTGAGCCTGCTGAACGTGCGAGGAATCAAGAGCCAGT TCGAGGGCTCGTGAAGGACATCATGCTGAACAAAGAGGAACGAAGAAAG AAAACCTCGAGATGCAGAAGGGCGACCAGAACCTCAGATGCCGCTCA CGTATCAGCGAGGCCAGCAGCAAGACAACAAGCGTGTGCTGAGTGGGCCGA GAAGGGCTACTACACCATGAGCAACAACCTGGTACCCCTGGAAAACGGCAAG CAGTGAAGCGGGAGGGCTGACTACATCTACGCCAAGTGA |
| | - GFP | | |
| | + Relleno | | |

ES 2 998 052 T3

| Nombre del donante | Descripción del donante | Secuencia de donantes |
|---------------------------|--------------------------------|---|
| | | <p>TCTGCAGCAACAGAGAGGCCAGCTCTAGGCCCCTTTATGCCAGCCTGTGC <u>CTGAAGTCCCTGGCAGATTGAGCGGATTCTGCTGAGAGGCCAACACAC</u> <u>ACAGCAGGCCAACCTTGTGGCAGCAGTCTATTCACTCGCGGAGTGT</u> <u>GAGCTGCAGCTGGCGAAGCGTGTCTGAATGTGACAGACCCTAGCCAGG</u> <u>TGTCCCACGGCACCGGCTTACATCTTCGGACTGCTGAAGCTGTGAACAGTG</u> <u>TCACCTTGAGGCTGTGGTGGAGCTGACGCTGGGAGTCTTCATAATACAGCA</u> <u>CAGCGGTTAAGCCCACCCCTGTTAACCTGCCTATTATAACCCTAGGATCCTCC</u> <u>TTATGGAGAACTATTTATTATAACTCCAAGGCATGTAGAACTGTAATAAGTG</u> <u>AATTACAGGTACATGAAACACAAACGGGCCCTGCTCCATAAGAGCTTATATA</u> <u>TCTGAAGCAGCAACCCCCTGATGCAGACATCCAGAGAGTCCTATGAAAAGA</u> <u>CAAGGCCATTATGCACAGGTTGAATTCTGAGTAACAGCAGATAACTTGCAA</u> <u>GTTAGTTTGTCTTGCCTGCAGTGTCTTCATGGATAATGCATTGATT</u> <u>ATCAGTGAAGATGCAGAAGGGAAATGGGAGCCTCAGCTCACATTAGTT</u> <u>GGTTGACTCTGGGTTCTATGGCCTGTTGGAGGGGGCCAGGCTCTAGAACG</u> <u>TCTAACACAGTGGAGAACCGAACCCCCCCCCCCCCGCCACCCCTCGGACAGT</u> <u>TATTCACTCTCTTCAATCTCTCTCCATCTCTCTTCAGTCTCTCTCTC</u> <u>AAACCTTCTTCAATCTCTCTCAATCTCTGTTCCCTTGTAGTCT</u> <u>TTCCCTCCCCAGTCTCTTCAATCCCCCTTCTAACACACACACACACA</u> <u>CACACACACACACACACACACACACACAGAGTCAGGCCGTTGCTA</u> <u>GTCAGTTCTCTTCCACCCCTGTCCTATCTTACCACTATAGATGAGGGTG</u> <u>AGGAGTAGGGAGTGCAGCCCTGAGCCTGCCACTCTCATTACGAAATGACT</u> <u>GTATTTAAAGGAAATCTATTGTATCTACCTGCAGTCTCATTGTTCCAGAGTG</u> <u>AACTTGTAAATTATCTGTTATTATTTGAATAATAAGACCTTAACATTA</u> <u>CGCGCTTAACATTATCGTTGTTGAGTACCTAAAGCTCCAGCCAGGTTG</u> <u>GGGAAAGAGGAAGCATTTGGAGGAAATTCCAACCTTGTGATGTTTCA</u> <u>AAACTTGTTCTCAAGCTACTTACATTACCGCTACTAGTAATACCAATCATAG</u> <u>ATGAAAGGAGAAAGCGGCTCAGTCGAATGCCATATAGTTGGCTACAG</u> <u>AATGTTGGATAGAGAACTTAAGTACGCTAAGGCAGCGTATTTCTTAATATT</u> <u>AGGGGTATTGCCGCAGTCATTACAGATAACCGCCTATGCCAGGA</u> <u>TTATAGATAACTTTAACATTAGCCGAGAGGTGCCACTGGGTATCCAAA</u> <u>TATTAGGTTCACTGCTGACAAAAGAGGGGGCTCTGGTCTGGTTGGTTATT</u> <u>GTGTTGGCCTGATGTGCTGTCATCAAATGTATGGACATAGGCCAGCT</u> <u>CTAAAGGGGCAATAGTGACCTCAGTGGACTGATATTACCGTACTATTACAT</u> <u>GTGCTCTTAATTACAGCAGAAGCTGCCAGCTAAGTGAATCTGTTTGAATCTA</u> <u>AAAAATCTACTCTAAAGCAAGAAAATGGTATAAAATTAGTTGATAATGCAAG</u> <u>TAGATCT</u> </p> |
| 13 | gARN-D_HA | (SEQ ID NO: 46) |
| | + ADNc | |
| | - GFP | |

| Nombre donante | del | Descripción donante | del | Secuencia de donantes |
|-------------------|-----|------------------------|-----|---|
| | | - Relleno | | <p>AGATCTTGATTAGTTCTGGCTGTCCTAAATTCTGGATGCAGGAACGTG GCTAGAAAAGCATCTGGATGATTGCACTTATCAGGGATACTTGAGTGTCTCT CTTAGGATCTGGACCTAGAATTAATGTCATGAGATTCTAACAGGATAAGG TGAGGTAGTGAGGGCTGAAGTCATGAATTCTGTACATTAAATTCTAACAG ATTGGAAGCTTCCTAGGTTACATGTGCTTAAATTACAGCAGAACCCGGTCT GACCTCTTCTTCCTCCCACAGATCGAGGACGAGAGAACCTGCACGAGGAC <u>TTCGTGTTCATGAAGACCATCCAGGGTCAACACCGGCCAGAGAACGAGTCTGA</u> <u>GCCTGCTGAACTCGAGGAAATCAAGAGCCAGTCAGGGCTCGTAAGG</u> <u>ACATCATGTCAGAACAAAGAGGAACGAAGAAAGAAAACCTTCGAGATGCA</u> <u>GAAGGGCGACCAGAACCTCTAGATGCCGTCACGTGATCAGCGAGGCCAGC</u> <u>AGCAAGACAACAAGCGTGTGCACTGGGGAGAACGGCTACTACACCATG</u> <u>AGCAACAACTGGTACCCCTGAAACAGGCAAGCAGCTGACAGTGAAGCGG</u> <u>CAGGGCCTGACTACATCTACGCCAAGTGACCTCTGCAGCACAGAGAGG</u> <u>CCAGCTCTAGGCCCTTTATGCCAGCCTGCTGAAGTCCCCTGGCAGA</u></p> <p><u>TTCGAGCGGATTCTGCTGAGAGGCCAACACACACAGCAGGCCAACCTT</u> <u>GTGGCCAGCAGTCTATTCACCTCGCGGAGTGTGAGCTGCAGCCTGGCGC</u> <u>AAGCGTGTGTAATGTGACAGACCCCTAGCCAGGTGTCACGGCACCGGC</u> <u>TTTACATCTTCGGACTGCTGAAGCTGTGACAGCTGACCTTGAGGCTG</u> <u>GTGGAGCTACGCTGGAGTCTCATAATACAGCACAGCGGTTAACGCCACC</u> <u>CCCTGTTAACTGCCTATTATAACCCCTAGGATCCTCTTATGGAGAACTATT</u> <u>TTATACACTCCAAGGCATGAGAACGTTAATAAGTGAATTACAGGTACATGA</u> <u>AACCAAAACGGGCCCTGCTCCATAAGAGCTTATATATCTGAAGCAGCACCCCC</u> <u>ACTGATGCAGACATCCAGAGAGTCTATGAAAAGACAAGGCCATTATGCACA</u> <u>GGTTGAATTCTGAGTAACAGCAGATAACTGCCAAGTTCAGTTGTTCTT</u> <u>GCGTGCAGTGTCTTCATGGATAATGCAATTGATTATCAGTGAAGATGCAG</u> <u>AAGGGAAATGGGGAGCCTCAGCTCACATTAGCTTATGGTTGACTCTGGTC</u> <u>CTATGGCCTTGTGGAGGGGGCCAGGCTCTAGAACGTCACACAGTGGAGA</u> <u>ACCGAAACCCCCCCCCCCCCGCCACCTCTCGGACAGTTACCTCTTCAA</u> <u>TCTCTCTCTCCATCTCTCTTCACTGCTCTCTCAACCTCTTCTCAA</u> <u>CTCTCTTCTCAATCTCTGTTCCTTGTCACTGCTCTCTCAACCTCTTCTCAA</u> <u>TCTCTCAATCCCCCTTCTAACACACACACACACACACACACACACA</u> <u>CACACACACACACACAGAGTCAGGCCGTTCTAGTCAGTCTCTTCTCAA</u> <u>CACCTGCCCCATCTTCACTGCTTCACTGAGTGAACAGTGTATTAAAGGAAATC</u> <u>TATTGATCTACCTGCAGTCTCATTGTTCCAGAGTGAACCTGAAATTCTTG</u> <u>TTATTATTTTGAATAATAAAGACCTCTAACATTACCGCCTAACATTATCG</u> <u>TTGTTGTTGAGTACCTAAAGCTCCAGGCCAGGTTGGGGAAAGAGGAACAT</u> <u>TTGGAGGAAATTCCAAACCTTGTGATGTTCTAAACTTGTCTCAAGC</u> <u>TACTTACATTACCGTACTAGTAAATACCAATCATAGATGAAAGGAGAAAGC</u> <u>CACTGGGTTATCAAATATTAGGTTCACTGCTGACAAAGAGGGGGCTCTG</u> <u>GTCTGGTTGGTATTGTGTTGGCCTGATGTCGTCATCAATCAAATGATGG</u> <u>ACATAGGCCCTAGCTCTAAAGGGGAATAGTGACCTCAGTGGACTGATATT</u> <u>CCGTACTATTACATGTGCTCTAACAGCAGAACGCTGCCAGCTAACATGAAT</u> <u>CTTGTTTGAATCTAAAAAATCTACTCTAAAGCAAGAAAATGGTATAAAATT</u> <u>GTTGATAATGCAAGTAGATCT</u></p> |

[0253] Se determinó la integración diana y la expresión de CD40L tras la edición con tres complejos de RNP que contenían ARNg ejemplares (ARNg A, D y G). La localización de CD40LG, intrón 1, a la que dirigen estos ARNg se muestra en la Figura. 3A. El procedimiento experimental se describe en la Figura. 3B. La tasa de integración

5 se muestra en la Figura. 3A. El procedimiento experimental se describe en la Figura. 3B. La tasa de integración diana (presentada como el porcentaje de células GFP⁺) tras el tratamiento con complejos RNP y la plantilla donante AAV6 se proporciona en la Figura. 3C. El nivel de integración diana obtenido con el ARNg A fue comparable al nivel de integración diana obtenido con el ARNg D, y fue superior al obtenido con el ARNg G. La Figura. 3D muestra la expresión de CD40L en la población de células positivas para GFP, determinada mediante análisis FACS. La Figura. 3E muestra la fracción de células CD40L⁺ en distintos momentos tras la reestimulación. La integración de la plantilla de reparación correctiva en el extremo 3' del intrón 1 utilizando el ARNg G no cambió la expresión de

CD40L, en relación con la integración de la plantilla de reparación correctiva en el centro del intrón 1 utilizando el ARNg A o el ARNg D.

[0254] Para probar si la edición génica de CD40L preserva la regulación fisiológica, se determinó la cinética de expresión de CD40L. Como se muestra en la Figura. 4B, tras la electroporación de RNP y la adición de ADN donante, se monitorizó la cinética de la expresión de CD40L durante un periodo de 48 horas a partir del día 15, tras la inducción de la expresión de CD40L con PMA/Ionomicina. Las Figuras. 4C y 4E muestran la cinética de expresión de CD40L en células T CD4⁺ editadas tras el tratamiento con PMA/Ionomicina. La expresión de CD40L en células CD4⁺ y CD8⁺ activadas se muestra en las Figuras. 4D y 4F. Estos datos indican que la expresión normal de CD40L se restableció tras la edición génica.

[0255] Para confirmar estos resultados, se realizó otro experimento. Se electroporaron células T humanas con RNP y, 15 minutos después, se transdijeron con un vector AAV6 portador del donante corrector. El análisis FACS se realizó 6 días después de la electroporación RNP, y la cinética de expresión de CD40L se determinó el día 15 (Figuras. 5B). La edición de CD40LG permite la expresión normal de la proteína CD40L y preserva la regulación fisiológica de CD40L, como se muestra en las Figuras. 5C y 5D. Las Figuras. 5C y 5D muestran la cinética de expresión de CD40L en las células T CD4⁺, editadas con un donante CD40LG de tipo salvaje o de codón optimizado ("co") tras el tratamiento con PMA/Ionomicina. La expresión de CD40L en células CD4⁺ y CD8⁺ activadas se muestra en la Figura. 5C. Los datos muestran que el ADNc CD40LG con codón optimizado permite una mejor restauración de la expresión de CD40L. Además, la eficacia de la edición génica mejoró cuando se añadieron mayores concentraciones del donante AAV6 a las células (Figura. 5E). Mientras que la expresión de CD40L corrector alcanza niveles más bajos que la de CD40L endógeno, ambas muestran un fuerte descenso de la expresión después de aproximadamente 6 horas (*datos no mostrados*). Del mismo modo, se mejoró la eficacia de la edición génica optimizando las dosis de RNP y AAV6. Las Figuras. 6A y 6B demuestran que los métodos descritos en el presente documento producen altos niveles de edición génica de CD40L (Figura. 6A) y altos niveles de expresión exógena de CD40L (Figura. 6B). Se observó que el AAV inhibía la expansión de células T tras la electroporación en función de su MOI (*datos no mostrados*).

[0256] La especificidad de los gRNA A y gRNA D se evaluó mediante análisis Guide-Seq y Amplicon-Seq (Amp-Seq). Se realizó un análisis de Guide-Seq para evaluar el número de eventos de edición dentro y fuera de la diana que se produjeron tras la edición con RNPs que contenían ARNg A o ARNg D, y los resultados se muestran en la Figura. 7A. Guide-Seq se realizó en células T masculinas tras la activación y expansión de las células. Las células se nucleofectaron con RNPs que contenían gRNA A o D y un oligo corto de doble cadena (Nat. Biotech. 2015, 33: 187-197). Se aisló el ADNg, se esquiló y se añadieron adaptadores para la amplificación PCR antes de la amplificación PCR. Las secuencias adyacentes al oligo Guide-Seq se alinearon con el genoma para encontrar la ubicación de la inserción del oligo de doble cadena. Los resultados del análisis Guide-Seq se resumen en la Tabla 11.

Tabla 11: Análisis Guide-Seq

| ARNg | Cas9 | Edición media dianas por | Promedio de lecturas guía-seq diana | # Dianas detectadas | Edición media fuera de diana | Promedio de lecturas de Guide-Seq no dirigidas |
|--------|------|--------------------------|-------------------------------------|---------------------|------------------------------|--|
| ARNg A | S.p. | 90% | 9173 | 1 | 0,3% | 116 |
| ARNg D | S.a. | 90% | 9287 | 0 | No detectado | N/A |

[0257] Se realizó un análisis Amp-Seq para validar el evento de edición fuera de diana observado con el ARNg A, como se muestra en la Figura. 7B. Amp-Seq se llevó a cabo mediante amplificación por PCR alrededor del sitio de corte fuera de diana hallado a partir del análisis Guide-Seq y, a continuación, utilizando secuenciación de alto rendimiento y alineación computacional para determinar si había eventos de edición en el sitio de corte. Guide-Seq identificó un sitio de corte no diana para el ARNg A. Este sitio de corte estaba en el cromosoma 8, y a ~300 kb de otro gen. El corte fuera de diana se redujo a niveles de antecedente utilizando una Cas9 de mayor fidelidad (eCas9), como se muestra en la Figura. 7B. Los eventos de edición en diana para el ARNg A se muestran en las Figuras. 7B y 7E. El análisis del digenoma reveló un número significativamente mayor de posibles sitios de corte *in vitro* para el ARNg A (205), en comparación con el ARNg D (10) (Figuras. 7C-7D). El ARNg A y el ARNg D mostraron tasas muy eficientes de edición e integración diana, como se muestra arriba, en combinación con un alto nivel de especificidad. La expresión de CD40L a partir de la carga integrada fue comparable a los niveles endógenos de CD40L. La Figura. 7F muestra una visión general de los resultados del análisis Amplicon-Seq para los sitios de corte fuera de diana identificados para los ARNg A y D.

Ejemplo 2: Las células CD40L editadas son funcionalmente activas y restauran la recombinación del interruptor de clase de células B

5 [0258] Para determinar si se restaura la función en las células editadas con CD40L, se co-cultivaron células B y células T editadas, y se determinó la expresión superficial de IgG en las células B. La edición génica de CD40L en células T se realizó como se ha descrito anteriormente. Las células B se cultivaron conjuntamente con células T de tipo salvaje (no editadas), células T editadas y seleccionadas para la integración diana del gen CD40L corrector (GFP+), células T tratadas para la edición génica pero negativas para la integración del gen CD40L corrector (GFP-), y células T en las que se eliminó el gen CD40L endógeno mediante la integración de un casete informador (CD40L KO). Los cultivos que contenían (i) células B solas, y (ii) células B y proteína CD40L soluble (sCD40L), se utilizaron como controles negativos y positivos, respectivamente. Las células T se probaron sin estimulación (en reposo), tras estimulación con microesferas anti-CD3/CD28 (microesferas) o tras estimulación con PMA/Ionomicina (PMA/Iono). Tras 5 días de co-cultivo, se realizaron análisis FACS. Como se muestra en la Figura. 8, las células T editadas fueron capaces de inducir el cambio de clase de las células B en el ensayo de cocultivo. Además, los niveles de IgM se redujeron significativamente y los de IgG aumentaron significativamente en el medio de las células B co-cultivadas con células T tras la edición génica (Figura.9). En conjunto, los datos comunicados en la Figura. 8 y la Figura. 9 demuestran que las células T editadas expresaron un CD40L funcional, que es capaz de proporcionar señales dependientes del contacto a las células B que indujeron el cambio de clase *in vitro*.

10

15

Ejemplo 3: Integración dirigida en el emplazamiento de CD40L

20 [0259] Se determinó el nivel de línea de base de la integración diana en cuatro sitios de corte de ARNg del ejemplo 1: ARNg A (*S. pyogenes*), ARNg C (*S. pyogenes*), gRNA D (*S. aureus*) y gRNA G (*S. aureus*). Las plantillas de donantes AAV6 se diseñaron siguiendo la representación de la Figura. 1A, de modo que todos los donantes contuvieran la misma carga (secuencias de relleno, CD40LG cDNA/3' UTR/polyA y casete de expresión de GFP) dentro de brazos cortos de homología diseñados para flanquear los sitios de corte previstos. En la Tabla 12 se indica la longitud de los brazos homológicos de cada constructo.

25

Tabla 12: Homología Longitud del brazo

| | 5' HA | 3' HA |
|--------|--------|--------|
| gRNA-A | 205 pb | 178 pb |
| gRNA-C | 242 pb | 131 pb |
| gRNA-D | 178 pb | 284 pb |
| gRNA-G | 206 pb | 270 pb |

30 [0260] Se cultivaron células T CD4⁺ masculinas humanas primarias y se nucleofectaron con el ARNg indicado como se ha descrito anteriormente. Las células se trasladaron a un nuevo medio de cultivo y, 15 minutos después, se añadieron partículas AAV6 al cultivo con una multiplicidad de infección (MOI) de 5e4 o 1e5. Las células se expandieron durante 7 días y, a continuación, se evaluó la integración diana en función de la expresión de GFP mediante citometría de flujo. En este punto temporal tras la nucleofeccción, se detecta muy poca expresión de GFP en el donante AAV no integrado (datos no mostrados), lo que indica que las células con altos niveles de expresión de GFP se sometieron a integración diana. Se encontraron diferencias en los niveles de base de la integración dirigida entre los ARNg, con base en la expresión de GFP, como se muestra en la Figura 10. (gRNA A ~35 %, gRNA C ~15 %, gRNA D ~28 %, y gRNA G ~17 %). No se observó una reducción dependiente de la dosis en la integración diana al reducir la cantidad de donante AAV6 administrada a las células. Los resultados representan al menos 2 muestras independientes.

35

40

[0261] Para aumentar el nivel de base de la integración diana, se examinó el efecto de alargar los brazos de homología. Se diseñaron plantillas de donantes AAV6 con brazos homológicos largos (500 pb a cada lado de la carga) para los principales sitios de ARNg basados en la integración diana determinada en la Figura. 10 (ARNg A y ARNg D). Los experimentos de integración diana se llevaron a cabo en linfocitos T CD4⁺ masculinos primarios, tal y como se ha descrito anteriormente. Las células recibieron una plantilla donante AAV6 con HA corta (los mismos donantes AAV6 utilizados para la Figura. 10), o HA larga. La expresión de GFP se determinó 7 días después de la nucleofeción mediante citometría de flujo y se representa en la Figura. 11. Los HA largos mejoraron los niveles de integración diana para ambos sitios de corte del ARNg. La integración dirigida en el sitio A del ARNg aumentó de ~35 % a -44 %, y la integración diana en el sitio D del ARNg aumentó del ~28 % al ~42 %.

45

50

Ejemplo 4: Terapia con células T adoptivas en modelos murinos

55 [0262] Los ratones fueron pretratados o acondicionados con ciclofosfamida el día 0 para agotar las células receptoras y crear espacio en el nicho linfocitario, permitiendo un mejor injerto de las células trasplantadas. Se inyectaron células T de tipo salvaje el día 2 y se midió longitudinalmente su injerto a lo largo del tiempo. La relación CD8/CD4 permaneció estable en el grupo no condicionado, mientras que aumentó significativamente en los ratones condicionados (véase la Figura. 12A). Durante ese tiempo, la gran mayoría de las células T CD8 y CD4

circulantes del donante mostraron un fenotipo de memoria central T y de memoria efector, respectivamente (Figura. 12B). La vacunación con TNP/KLH se introdujo el día 21, seguida de ensayos serológicos los días 35 y 50. Como se muestra en la Figura. 13, el acondicionamiento con ciclofosfamida dio lugar a un mayor nivel de IgG específica de TNP/KLH en ratones. Los niveles de respuesta IgG específica a la vacunación se correlacionan con la dosis de acondicionamiento administrada (Figura. 14A) y con la cantidad de linfocitos T CD4 de tipo salvaje circulantes (porcentajes y números absolutos, Figura. 14B). Sin embargo, la etapa de acondicionamiento no es absolutamente necesaria, ya que los ratones sin acondicionamiento también mostraron un aumento de los niveles de IgG tras la vacunación (Figura. 13 y Figura. 14A). En un segundo experimento, los ratones fueron pretratados o acondicionados con anticuerpos antileucocitos (ALS) o anti-CD4 el Día 0. Las Figuras. 14C y 14E representan los porcentajes de células T receptoras y donantes a lo largo del tiempo en los ratones tratados o no con el régimen de preacondicionamiento. Los niveles de respuesta IgG específica a la vacunación se correlacionan con la dosis de acondicionamiento administrada (Figura. 14D). Estos datos demuestran que el acondicionamiento linfodepletor también podría realizarse mediante estrategias inmunobiológicas. El trasplante de células T adoptivas restauró también la presencia de centros germinales en el bazo de los ratones trasplantados (Figura. 15A y Figura. 15B).

Ejemplo 5: Terapia con células madre hematopoyéticas en modelos murinos

[0263] Los ratones CD40L negativos fueron tratados con HSPC CD40L de tipo salvaje, recolectadas de ratones C57B/6 de tipo salvaje de 6-8 semanas de edad, mezcladas en diferentes porcentajes (100 %, 10 %, 1 % y 0 %) con células CD40LG negativas, recolectadas de ratones HIGM1. A continuación, se inmunió a los ratones receptores con TNP-KLH y se midieron los niveles de IgG específica de TNP-KLH. Como se muestra en la Figura. 16A, los ratones que recibieron células madre hematopoyéticas con un 10 % y un 100 % de células CD40L de tipo salvaje tenían un nivel mucho mayor de IgG, lo que sugiere que las células madre hematopoyéticas con al menos un 10 % de células CD40L editadas demuestran un efecto terapéutico para tratar enfermedades, tales como el síndrome de hiper-IgM.

[0264] En un segundo experimento, los ratones fueron desafiados con dos抗ígenos, TNP-KLH y ovoalbúmina (OVA), como se muestra en la Figura. 16B. Los ratones CD40LG-negativos fueron trasplantados con diferentes porcentajes de HSPC de tipo salvaje en el día 0. Los ratones fueron inmunizados con TNP-KLH a las 12 y 15 semanas tras el trasplante, y posteriormente fueron inmunizados con OVA a las 20 y 23 semanas tras el trasplante. Se evaluó la respuesta IgG específica de TNP (i) antes de la inmunización con TNP-KLH (pre-desafío), (ii) 14 días después de la inmunización inicial con TNP-KLH (día 14), y (iii) 28 días después de la inmunización inicial con TNP-KLH (post-desafío; día 28). Se evaluó la respuesta IgG específica a OVA (i) antes de la inmunización con OVA (pre-desafío), (ii) 14 días después de la inmunización inicial con OVA (día 14), y (iii) 28 días después de la inmunización inicial con OVA (post-desafío; día 28). Los resultados indican que la restauración de al menos un 10 % de HSPC CD40LG positivas permite la restauración parcial de la respuesta inmunitaria a diferentes vacunas dependientes de células T. La Figura. 16C muestra el injerto de células T del donante en sangre periférica.

Ejemplo 6: Prueba de concepto de la corrección de CD40LG en células T derivadas de un paciente con HIGM1

[0265] Para confirmar la corrección de las mutaciones causantes de HIGM1 mediante la integración diana de la plantilla donante correctora, se realizó la edición génica en células T obtenidas de un sujeto varón con una mutación inactivadora genotipada en el exón 3 del gen CD40L (c.334 G>T). Las células T CD4⁺ derivadas del paciente y de donantes sanos (HD) como control (n=4) se cultivaron y nucleofectaron con RNP (ARNg A) y, 15 minutos después, se transdujeron con un vector AAV6 portador de la plantilla donante correctora representada en la Figura. 2A, a una MOI de 5e4. La integración diana se midió mediante un análisis citofluorimétrico que puntuaba el porcentaje de células GFP+ dentro de las siguientes subpoblaciones de células T CD4: la memoria de células madre T (TSCM, CD45RA+ CD62L+), la memoria central (CM, CD45RA-CD62L+), la memoria efectora (EM, CD45RA-CD62L-) y la memoria efectora RA (TEMRA, CD45RA+CD62L-; Figura. 17A, 17B). La inserción diana se verificó mediante análisis de PCR digital de gotitas cuantificando la unión 5' del vector con el genoma, lo que confirmó cantidades similares de reparación dirigida por homología a las medidas por la expresión de GFP (Figura. 17C). La diana de células T policlonales se confirmó mediante la cuantificación de la diversidad de receptores de células T a través de la espectrotipificación TCR-beta (Figura. 17D). La integración del constructo corrector restauró la expresión y la regulación fisiológica de CD40L en las células T editadas del paciente tras la estimulación con PMA/Ionomicina (Figura. 18). Para evaluar la funcionalidad de las células editadas, se probaron células GFP+ seleccionadas en un ensayo de co-cultivo de células T-B (Figura. 19A), que confirmó la restauración de una señal dependiente de células T a células B, induciendo su proliferación (Figuras. 19B,C) y la producción de IgG emparejadas por clase (Figura. 20).

Ejemplo 7: Selección de las células editadas CD40LG antes del trasplante

[0266] Además de los métodos descritos anteriormente, el constructo corrector utilizado para fijar las mutaciones CD40LG puede equiparse con un gen de selección con el fin de enriquecer las células editadas antes del trasplante. Mediante el acoplamiento del ADNc CD40LG corrector con un gen de selección (*por ejemplo*, receptor del factor de crecimiento nervioso (NGFR)) precedido por una secuencia de sitio de entrada del ribosoma interno

(IRES) (Figura. 22A), es posible expresar el gen de selección sólo en células en las que se ha producido la integración diana que expresan concomitantemente el gen CD40LG. Mientras que CD40LG necesita una fuerte estimulación para exponerse en la superficie de los linfocitos T, los niveles basales de expresión del promotor de CD40LG son suficientes para permitir la expresión en superficie del gen de selección NGFR en las células editadas a partir de 6 días después de la edición génica, seguida de la selección de anticuerpos en el día 13 (Figura. 22B y Figura. 22C). Tras la estimulación con PMA/ionomicina de las células tratadas, el promotor CD40LG potenció la expresión de las proteínas CD40L y NGFR (Figura. 22D). Sin embargo, la cinética de la expresión superficial de CD40L es comparable a la medida en las células T simuladas (no editadas, UT) (Figura. 22E). El casete de expresión NGFR se utilizó para obtener la expresión fisiológica del gen CD40LG en las células editadas, tanto en términos del nivel de expresión, como del porcentaje de células CD40L+ (Figura. 22E). La Figura. 22F muestra la capacidad de las células T editadas con CD40L para restaurar el cambio de clase de las células B en un ensayo de co-cultivo, mostrando el porcentaje de células B con intercambio de IgG en co-cultivo con células T en reposo, activadas con microesferas (relación 1:1) o estimuladas con PMA/ionomicina (no tratadas (UT) o editadas (NGFR+)). La eficacia de integración dirigida del donante corrector con el casete de selección (Figura. 23A) fue similar a la medida con la plantilla original del donante corrector (Figura. 6A). Las células diana que expresan CD40LG pueden enriquecerse mediante la selección de NGFR sin ninguna desviación en la composición del cultivo de células T (Figura. 23B), y el análisis molecular confirmó la presencia de integración diana en las células NGFR+ (Figura. 23C). Los ratones NSG fueron trasplantados con células T CD4⁺ editadas, que fueron o no clasificadas para la expresión de NGFR. Los niveles de injerto fueron comparables a los observados en ratones trasplantados con células T no editadas, y la fracción de células editadas se mantuvo estable a lo largo del tiempo (Figura. 24A). Estos datos demuestran que el procedimiento de edición de CD40LG y la selección NGFR no perjudican la capacidad de repoblación de las células T tratadas, que normalmente son capaces de injertarse a largo plazo y de diferenciarse fisiológicamente en células efectoras tras el trasplante (Figura. 24B). La Figura. 24C muestra que el CD40LG editado es expresado por las células T editadas recuperadas de los ratones NSG.

[0267] La estrategia de selección descrita en el presente documento puede aprovecharse para aumentar la fracción y el número absoluto de células T corregidas terapéuticamente relevantes que pueden trasplantarse adoptivamente en un paciente con HIGM, al tiempo que se minimiza el número de células T no corregidas, *cultivadas in vitro*, que no proporcionan ningún efecto terapéutico.

Ejemplo 8: Edición genética de CD40L en células progenitoras hematopoyéticas humanas

[0268] Un enfoque de inmunoterapia adoptiva basado en células T autólogas corregidas genéticamente puede ser suficiente para inducir la generación de inmunidad protectora en pacientes receptores de HIGM1 y posiblemente también la producción de algunas células B de memoria de larga vida. Sin embargo, pueden ser necesarias múltiples administraciones de células T editadas para reconstituir un repertorio de células T duradero y amplio que pueda mediar una ayuda eficiente de las células T en respuesta a un amplio intervalo de patógenos. Además, mientras que CD40LG tiene una función crítica en las células T CD4 (Th0, Th1 y Th2), esta molécula también se expresa en varios otros tipos de células hematopoyéticas, tales como células B activadas, plaquetas, células NK, monocitos, basófilos y eosinófilos. Por lo tanto, será útil ampliar la estrategia de edición genética de la corrección de las células T a la corrección de las células madre/progenitoras hematopoyéticas autólogas (HSPC), lo que puede proporcionar un beneficio terapéutico mucho más amplio y prolongado.

[0269] Para demostrar la portabilidad de la estrategia de corrección genética descrita en HSPC humanas, primero adaptamos el protocolo de edición génica basado en ZFN previamente descrito al sistema CRISPR/Cas9. Brevemente, las células CD34+ fueron aisladas de la sangre del cordón umbilical y preestimuladas durante dos días utilizando SR1+dMPGE2, como se muestra en la Figura. 25A. Las células se infectaron con el donante IDLV en el día 2, y se electroporaron en el día 3 con un complejo de ribonucleoproteína (RNP) compuesto por la proteína Cas9 ensamblada con un ARNg transcrita *in vitro* o con una versión modificada ARNg transcrita *in vitro* (ARNgmod: Purificación HPLC, tapado ARCA, 5' Fosfatasa). A continuación, las células se analizaron mediante FACS para determinar la expresión de GFP y de marcadores de superficie celular el día 7. En la Figura. 25B se muestra la eficacia de integración diana conseguida con las dos RNP diferentes y un ZFN de referencia en poblaciones celulares específicas. La composición del cultivo tras la edición con ZFN y las RNP, en comparación con las células no tratadas, se muestra en la Figura. 25C. En la Figura. 25D se muestra el pliegue de inducción de los genes RIG1, IRF7 y OAS1 que responden al interferón (IFN) en las células editadas. Estos resultados demuestran que la administración de una RNP modificada permite niveles de edición mediada por homología en las HSPC humanas comparables a nuestro protocolo histórico de ZFN estándar de oro, sin inducir una respuesta IFN perjudicial y sin afectar al fenotipo de las células tratadas.

[0270] Se realizaron experimentos adicionales con células CD34+ obtenidas de sangre de cordón umbilical (CB), médula ósea (BM) y sangre periférica movilizada (MPB). Brevemente, las células fueron preestimuladas con SR1, dMPGE2 y UM171 durante tres días, momento en el cual las células fueron electroporadas con un complejo RNP compuesto por la proteína Cas9 ensamblada con un ARNg sintético de una o dos partes e infectadas con el donante AAV6, como se muestra en la Figura. 26A. A continuación, las células se clasificaron FACS para la expresión de GFP y marcadores de superficie celular el día 6. La eficacia de integración diana (expresada como porcentaje de células GFP+) se muestra en la Figura. 26B. La composición del cultivo celular después de la edición

se muestra en la Figura. 26C. La expresión de los genes que responden al IFN se muestra en la Figura. 26D. Estos resultados confirmaron que la edición mediada por homología en HSPC humanas puede lograrse, sin inducir una respuesta IFN perjudicial y sin afectar al fenotipo de las células tratadas, también con el uso de AAV6 como vehículos de entrega para el ADN donante o con RNP ensambladas con ARNg sintéticos.

5 [0271] Las HSPC humanas de diversas fuentes pueden editarse eficientemente utilizando una variedad de vehículos donantes y configuraciones de nucleasas, como se muestra en la Figura. 27. Los tipos celulares indicados se editaron utilizando ZFN o RNP, con una plantilla donante IDLV o AAV6. El porcentaje de células editadas (GFP+) en cada subpoblación se muestra en la Figura. 27A. El porcentaje de poblaciones celulares editadas (GFP+) derivadas de sangre de cordón umbilical, médula ósea y sangre periférica movilizada (mPB) se muestra en la Figura. 27B.

10 [0272] Para confirmar que la estrategia de corrección del gen CD40L es factible en HSPC humanas, se realizó la integración diana del donante corrector en células CD34+ de sangre de cordón umbilical obtenidas de donantes varones sanos. Las HSPC se cultivaron y nucleofectaron con diferentes dosis de RNP (ARNg A) y, 15 minutos después, se transdujeron con diferentes dosis de un vector AAV6 (MOI: 1e4, Se4 y 1e5 vg/célula) portadoras del donante corrector (Figura. 21A). La integración diana se midió mediante análisis de PCR digital en gotitas cuantificando la unión 5' del vector con el genoma. Para confirmar la edición en todas las diferentes subpoblaciones de células HSPC, las células tratadas se clasificaron aprovechando los marcadores de superficie que identifican prospectivamente los progenitores más primitivos (CD34+CD133+CD90+), tempranos (CD34+CD133+CD90-) y comprometidos (CD34+CD133-), y las células diferenciadas (CD34-). Se midieron altos niveles de integración diana tanto en el grueso como en las diferentes poblaciones clasificadas (Figuras. 21B, C y E), sin alterar la composición general del cultivo (Figura. 21D). Los niveles de injerto fueron comparables a los observados en ratones trasplantados con células T no editadas, y la fracción de células editadas se mantuvo estable a lo largo del tiempo (Figura. 21F).

REIVINDICACIONES

1. Una plantilla donante de oligonucleótidos aislada que comprende, de 5' a 3', A1-N-UTR-pA-A2, en el que
 5 A1 es un brazo homológico sustancialmente idéntico a un primer brazo homólogo de un ácido nucleico CD40L diana;
 N es una carga que comprende uno o más de los exones 2-5 de un gen CD40L;
 UTR es una región 3' no traducida (UTR) de CD40L;
 pA es una cola de poliA; y
 10 A2 es un brazo homólogo sustancialmente idéntico a un segundo brazo homológico del ácido nucleico CD40L diana,
 y en el que la carga no comprende el exón 1 de CD40L.
2. La plantilla donante de oligonucleótidos aislada de la reivindicación 1, en el que la plantilla donante de oligonucleótidos aislada comprende
 15 (a) de 5' a 3', A1-S1-N-UTR-pA-S2-A2,
 en el que S1 es un primer relleno, en el que S2 es un segundo relleno;
 (b) de 5' a 3', A1-S1-N-UTR-pA-R-S2-A2, en el que S1 es un primer relleno, en el que S2 es un segundo relleno, en el que R es un informador;
 20 (c) de 5' a 3', A1-N-UTR-pA-R-A2, en el que R es un informador; o bien
 (d) de 5' a 3', A1-N-IRES-Sel-UTR-pA-A2, en el que IRES es un sitio interno de entrada al ribosoma, y en el que Sel es un gen selector.
3. La plantilla donante de oligonucleótidos aislada de cualquiera de las reivindicaciones 1-2, en el que la carga comprende: el exón 2 del gen CD40L; el exón 3 del gen CD40L; el exón 4 del gen CD40L; el exón 5 del gen CD40L; los exones 2-5 del gen CD40L; los exones 2-4 del gen CD40L; los exones 2 y 3 del gen CD40L; los exones 3-5 del gen CD40L; los exones 3 y 4 del gen CD40L; o los exones 4 y 5 del gen CD40L; opcionalmente en el que la carga comprende una secuencia CD40L optimizada en codones.
 25
4. La plantilla donante de oligonucleótidos aislada de cualquiera de las reivindicaciones 2-3, en el que el reportero es una proteína de fluorescencia verde (GFP), una proteína de fluorescencia amarilla (YFP), DS-Red, o luciferasa; un receptor del factor de crecimiento nervioso de baja afinidad (NGFR); CD 19 truncado; EGFR truncado; proteína de resistencia a la neomicina; proteína de resistencia a la puromicina.
 30
5. Una composición que comprende la plantilla donante de oligonucleótidos aislada de cualquiera de las reivindicaciones 1-4 y, opcionalmente, un portador farmacéuticamente aceptable.
 35
6. Un vector que comprende la plantilla donante de oligonucleótidos aislada de cualquiera de las reivindicaciones 1-4, opcionalmente en el que el vector es un vector AAV, un lentivirus, un vector de ADN desnudo o una nanopartícula lipídica.
 40
7. Una molécula de ARNg aislada que comprende cualquiera de los SEQ ID NOs: 18-25.
8. Un sistema de edición genómica que comprende la plantilla donante de oligonucleótidos aislada de cualquiera de las reivindicaciones 1-4, que opcionalmente comprende además:
 45 (a) una nucleasa guiada por ARN y al menos una molécula de ARNg; o bien
 (b) la molécula de ARNg aislada de la reivindicación 7.
9. Un método in vitro de alteración de una célula que comprende el contacto de la célula con el sistema de edición del genoma de la reivindicación 8, alterando así la célula.
 50
10. Un kit que comprende el sistema de edición genómica de la reivindicación 8.
11. La plantilla donante de oligonucleótidos aislada según cualquiera de las reivindicaciones 1-4, la composición según la reivindicación 5, el vector según la reivindicación 6, la molécula de ARNg aislada según la reivindicación 7, el sistema de edición génica según la reivindicación 8, una célula preparada por el método según la reivindicación 9, o el kit según la reivindicación 10, para su uso en medicina.
 55
12. Una plantilla donante de oligonucleótidos para su uso en un método de tratamiento de un sujeto que tiene síndrome de hiper-IgM mediante la alteración de una diana CD40L en una célula, el método comprende las etapas de:
 60 formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, y
 65 recombinación de la plantilla donante de oligonucleótidos con el ácido nucleico CD40L diana mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla donante de oligonucleótidos comprende un primer brazo homológico donante que es sustancialmente idéntico al

primer brazo homológico, una carga que comprende uno o más de los exones 2-5 de un gen CD40L, una región no traducida (UTR) 3' de CD40L, una cola poliA, y un segundo brazo homológico donante que es sustancialmente idéntico al segundo brazo homológico, y en el que la carga no comprende el exón 1 de CD40L.

- 5 13. Un método in vitro de alteración de una célula, que comprende las etapas de:
 formar, en un ácido nucleico CD40L diana de la célula, al menos una rotura de cadena simple o doble en
 un sitio de escisión, en el que el ácido nucleico CD40L diana comprende: un primer brazo de homología
 5' al sitio de escisión y un segundo brazo de homología 3' al sitio de escisión, y
 10 recombinación de una plantilla donante de oligonucleótidos exógena con el ácido nucleico CD40L diana
 mediante recombinación homóloga para producir un ácido nucleico CD40L alterado, en el que la plantilla
 donante de oligonucleótidos exógena comprende un primer brazo de homología donante que es
 sustancialmente idéntico al primer brazo de homología, una carga que comprende uno o más de los
 15 exones 2-5 de un gen CD40L, una región no traducida (UTR) 3' de CD40L, una cola poliA, y un segundo
 brazo homológico donante que es sustancialmente idéntico al segundo brazo homológico, y en el que la
 carga no comprende el exón 1 de CD40L,
 alterando así la célula.
- 20 14. El método de la reivindicación 13, en el que la plantilla exógena donante de oligonucleótidos comprende:
 (a) el primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, un
 primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida
 (UTR) 3' CD40L, la cola poliA, un segundo relleno, y el segundo brazo homólogo donante que es
 sustancialmente idéntico al segundo brazo homólogo;
 25 (b) el primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, un
 primer relleno, la carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida
 (UTR) 3' CD40L, la cola poliA, un informador, un segundo relleno, y el segundo brazo homólogo donante
 que es sustancialmente idéntico al segundo brazo homólogo; o bien
 30 (c) el primer brazo homólogo donante que es sustancialmente idéntico al primer brazo homólogo, la
 carga que comprende uno o más de los exones 2-5 del gen CD40L, la región no traducida (UTR) 3'
 CD40L, la cola poliA, un reportero, y el segundo brazo homólogo donante que es sustancialmente
 idéntico al segundo brazo homólogo.
- 35 15. El método de cualquiera de las reivindicaciones 13-14, en el que la carga comprende el exón 2 del gen CD40L;
 el exón 3 del gen CD40L; el exón 4 del gen CD40L; el exón 5 del gen CD40L; los exones 2-5 del gen CD40L; los
 exones 2-4 del gen CD40L; los exones 2 y 3 del gen CD40L; los exones 3-5 del gen CD40L; los exones 3 y 4 del
 gen CD40L; o los exones 4 y 5 del gen CD40L.
- 40 16. El método de cualquiera de las reivindicaciones 13-15, en el que la etapa de formación de la al menos una
 rotura de cadena simple o doble comprende el contacto de la célula con una nucleasa guiada por ARN.

FIGURA 1

Plantillas de donantes

DIBUJOS

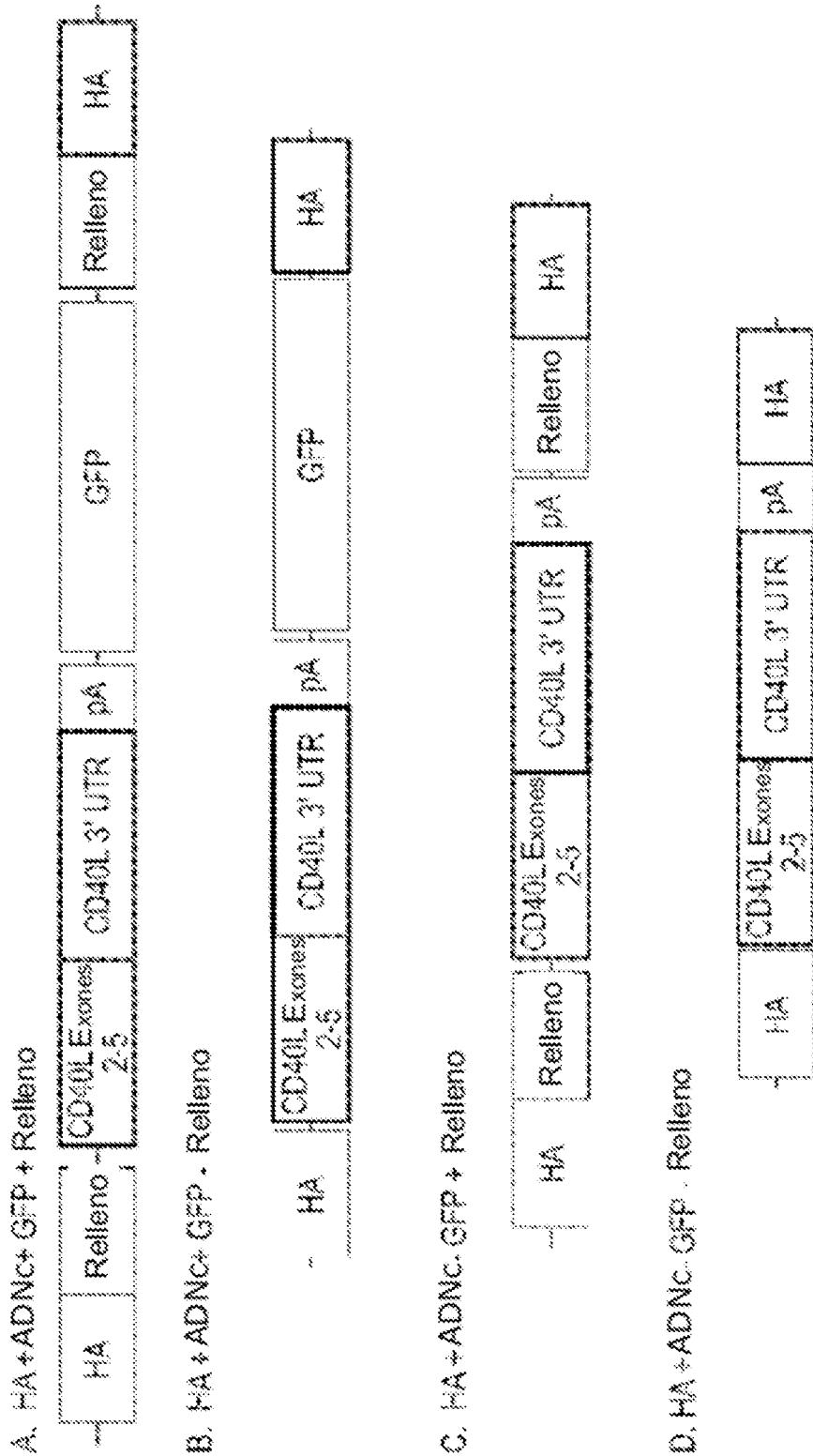
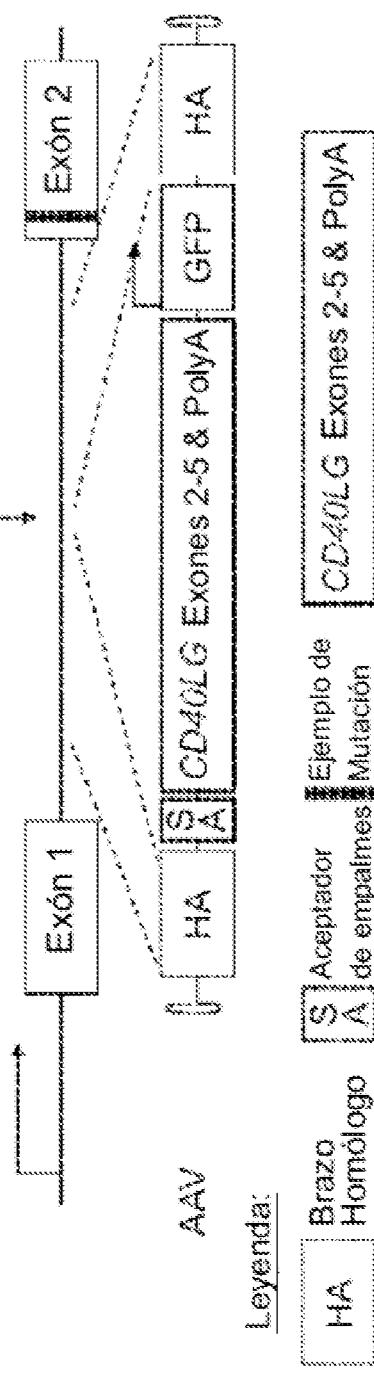
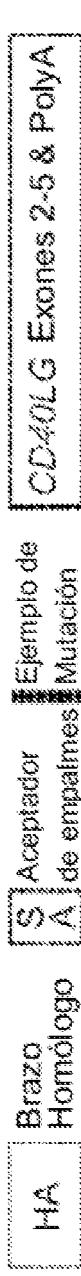


FIGURA 2A**Paciente CD40LG**

ARNm mutado

CRISPR/Cas9
Intrón 1**Leyenda:****CD40LG Corregido**

ARNm Corregido

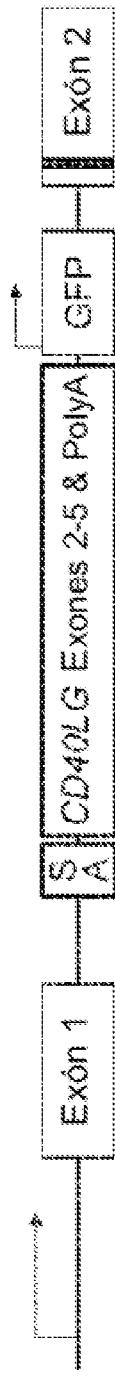


FIGURA 2B

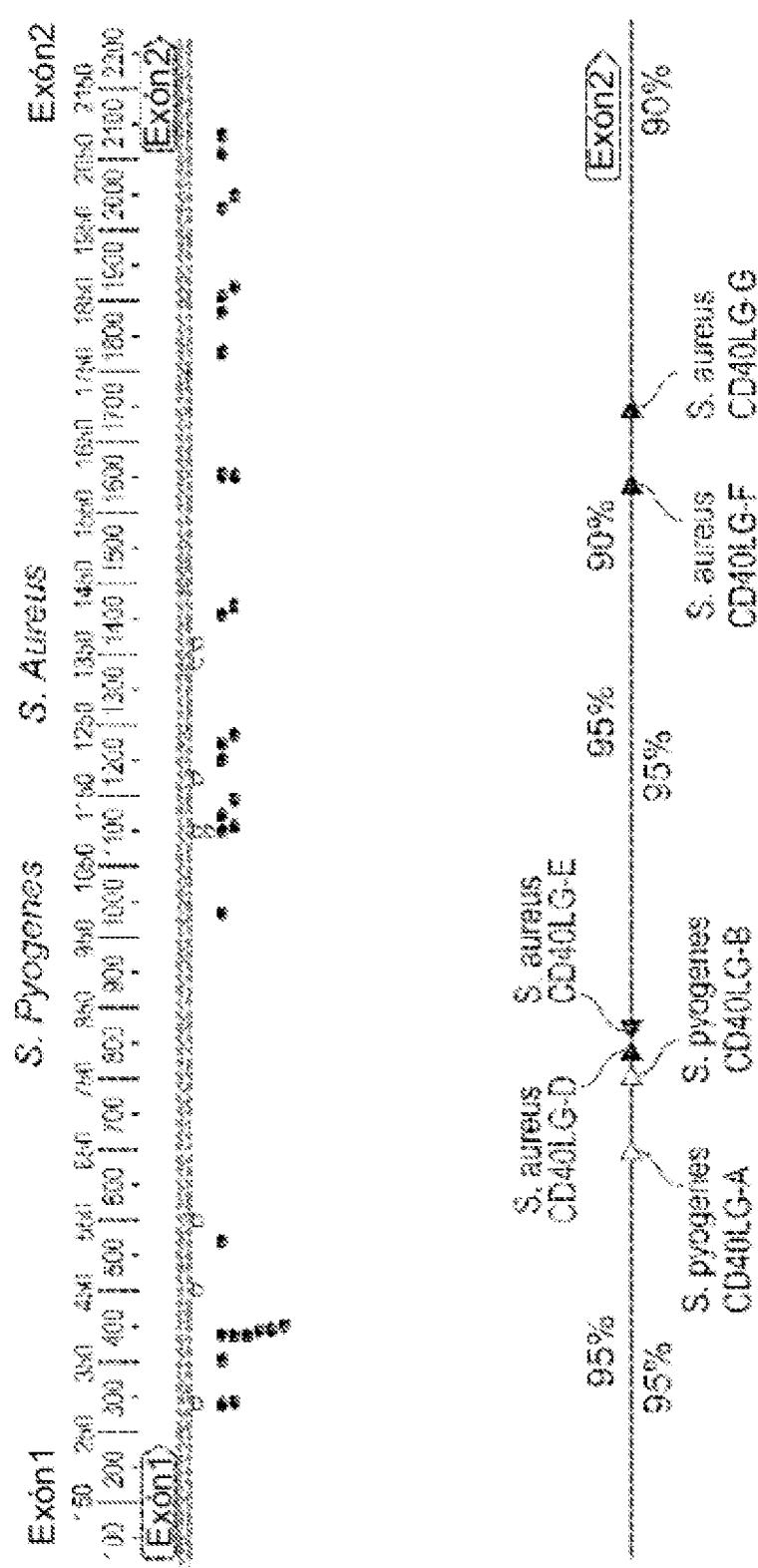


FIGURA 2C



FIGURA 2D

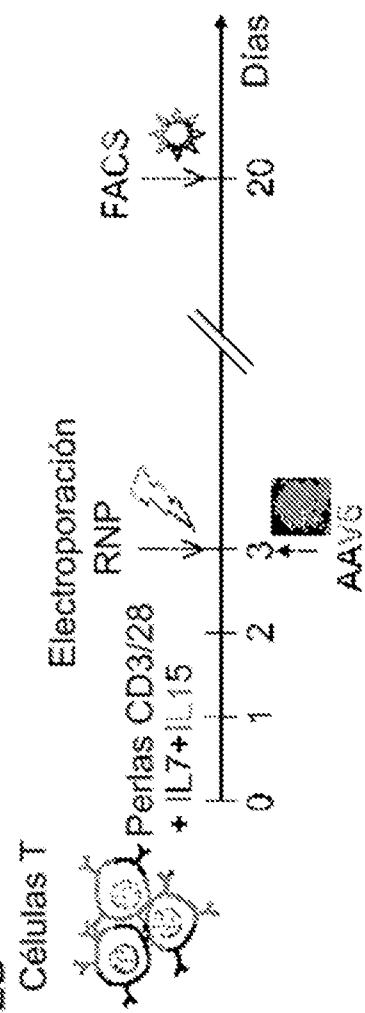


FIGURA 2E
Eficacia de la HDR en poblaciones de células T

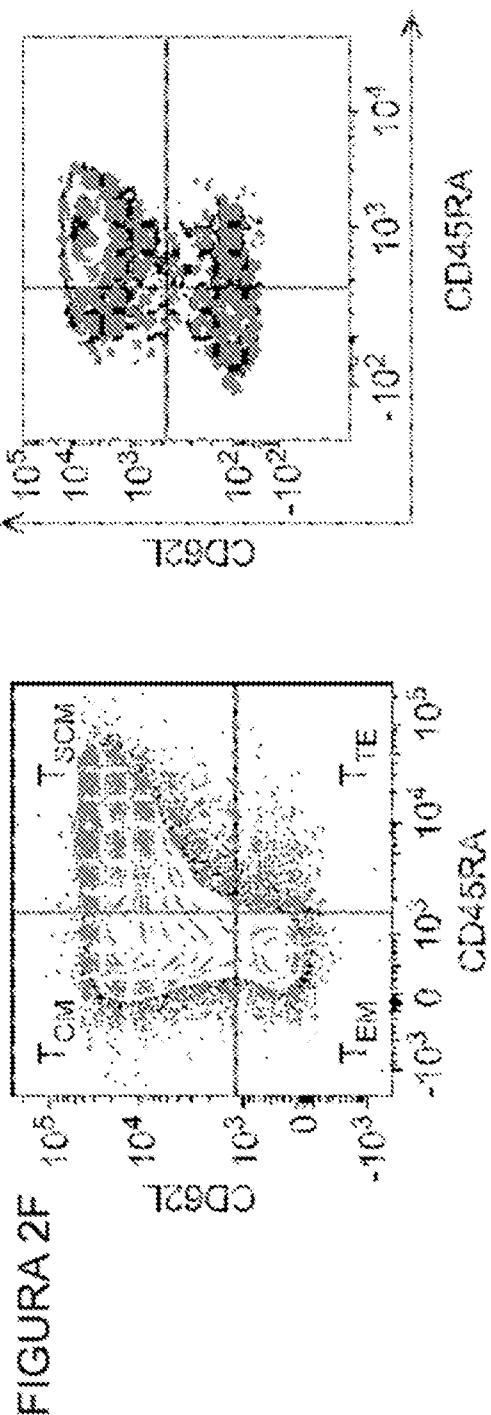
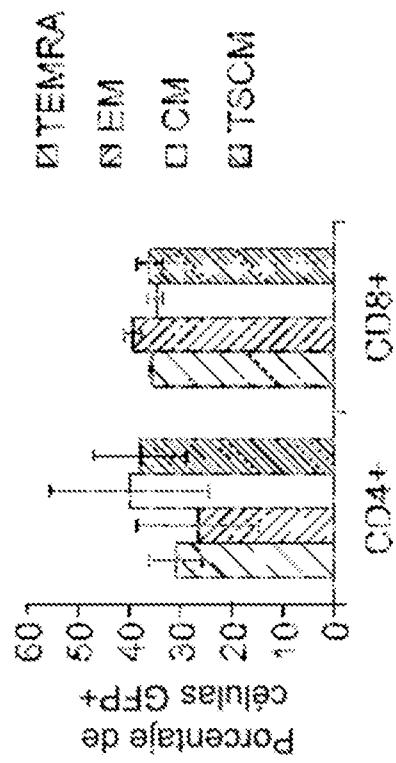


FIGURA 2G

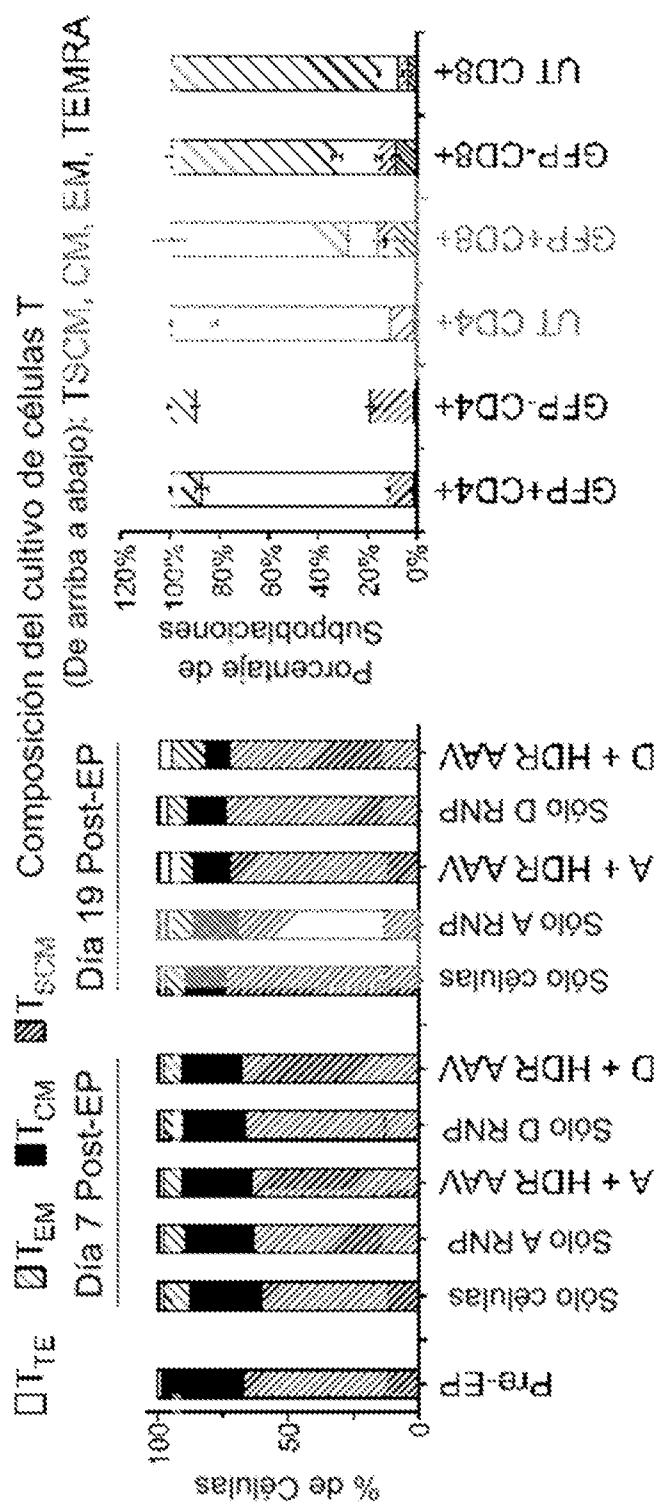


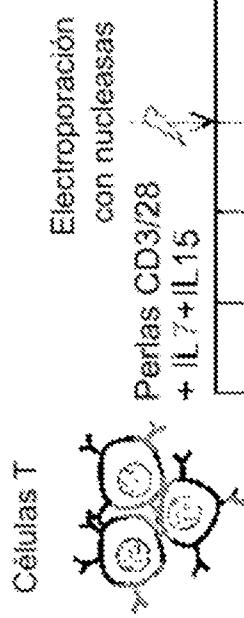
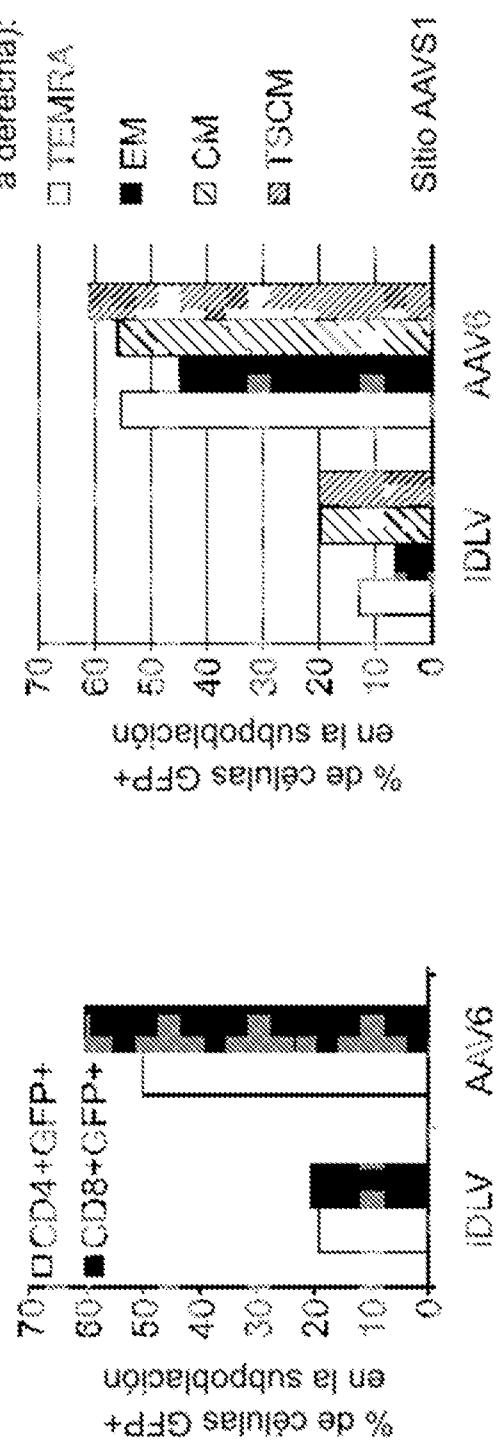
FIGURA 2H**FIGURA 2I**

FIGURA 2J

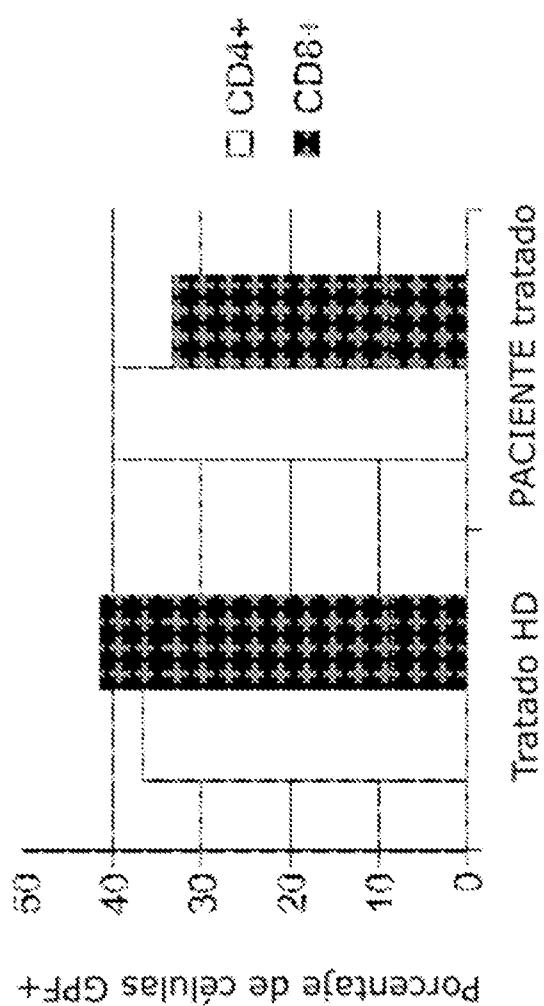


FIGURA 3A



FIGURA 3B

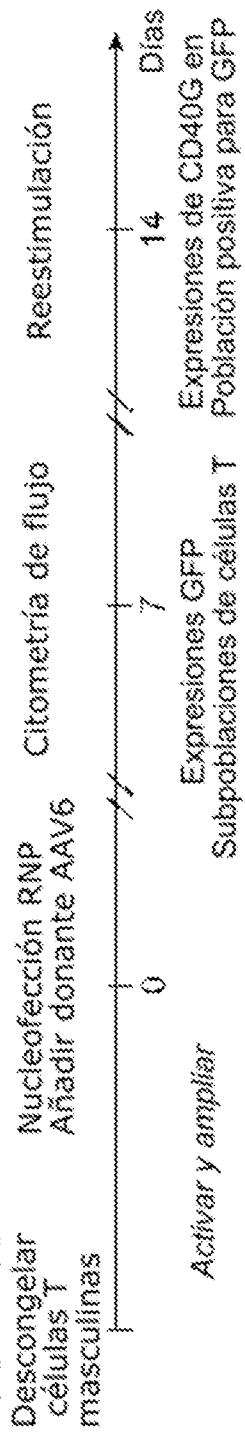


FIGURA 3C

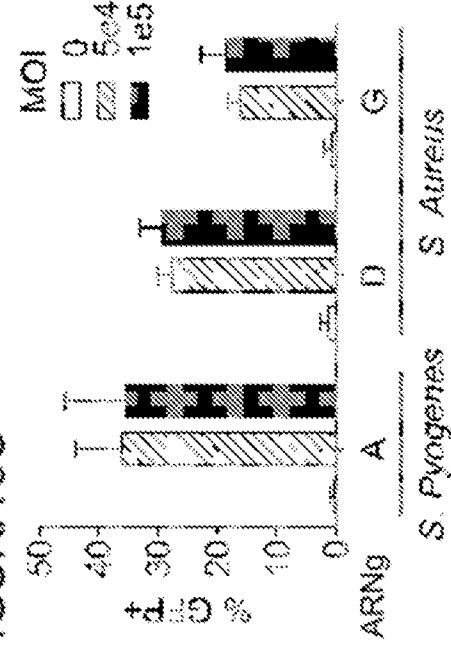


FIGURA 3D

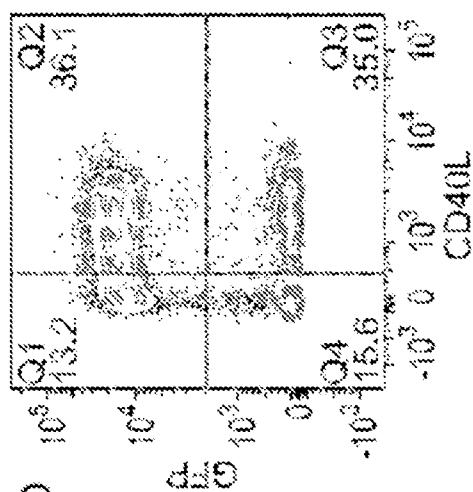
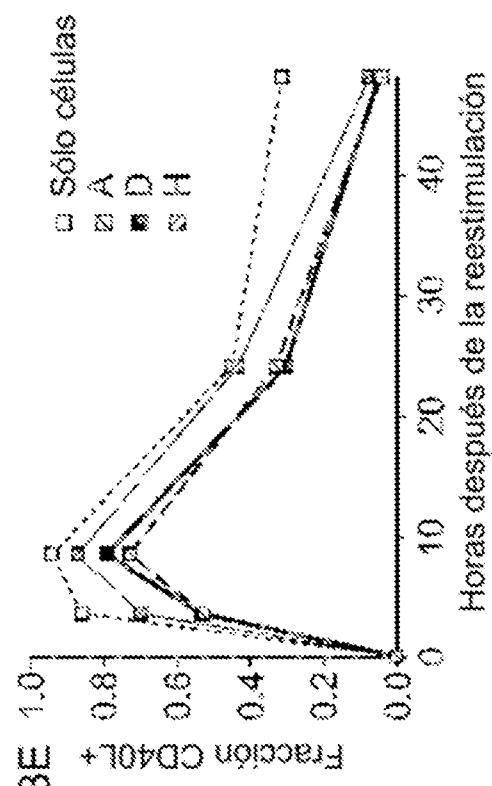


FIGURA 3E



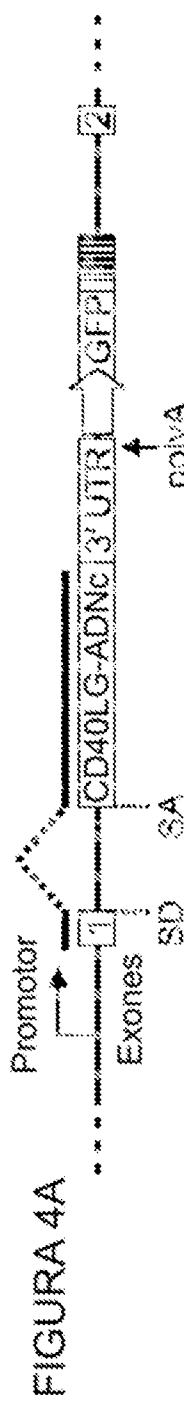
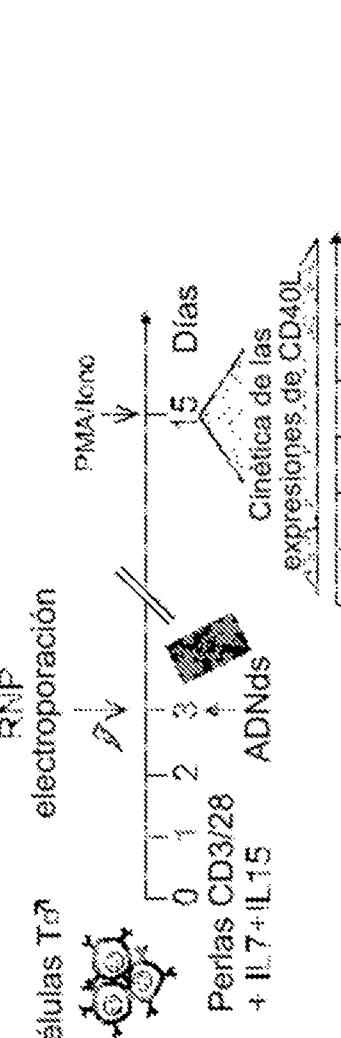
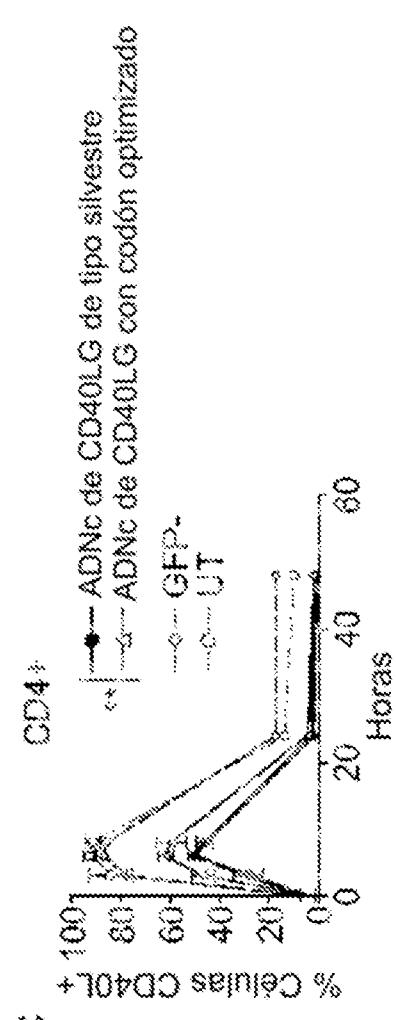
**FIGURA 4B Células T α** **FIGURA 4C**

FIGURA 4D

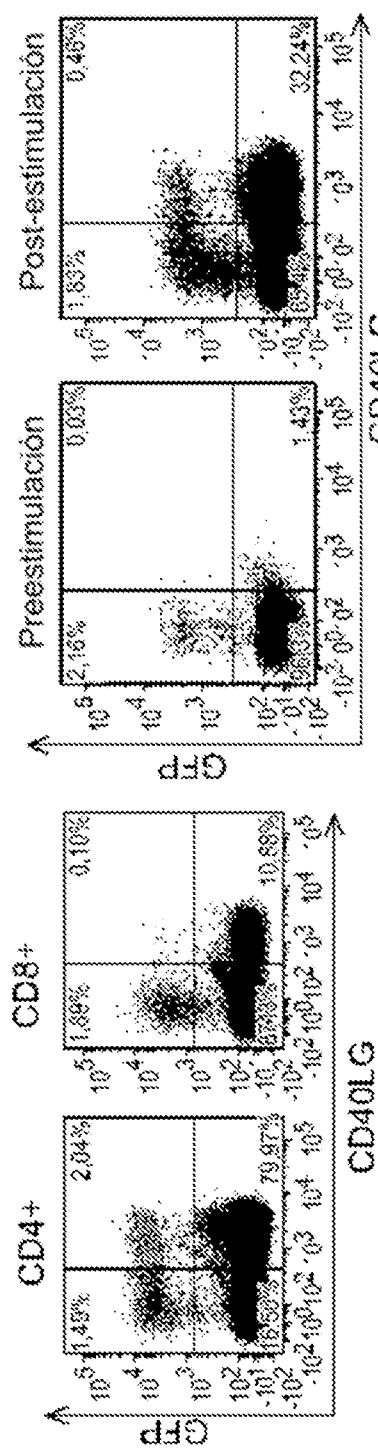


FIG. 4F

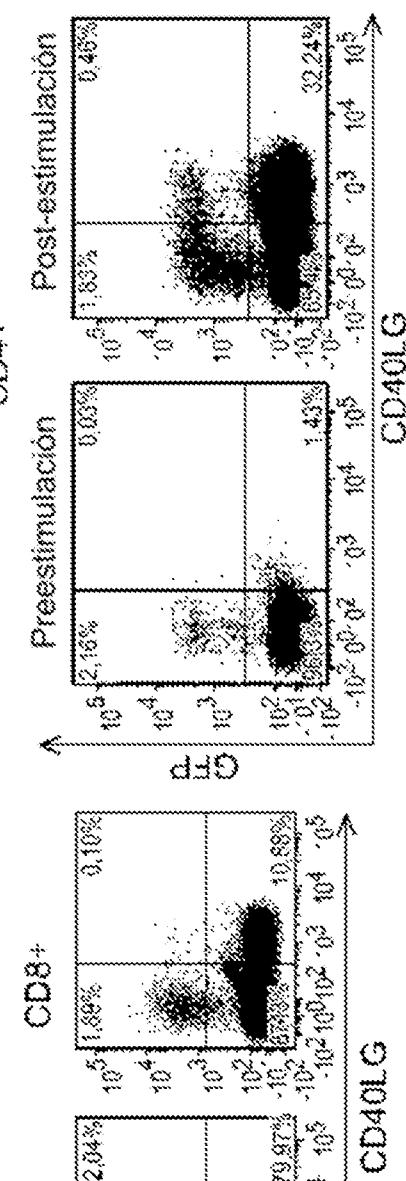


FIGURA 4E

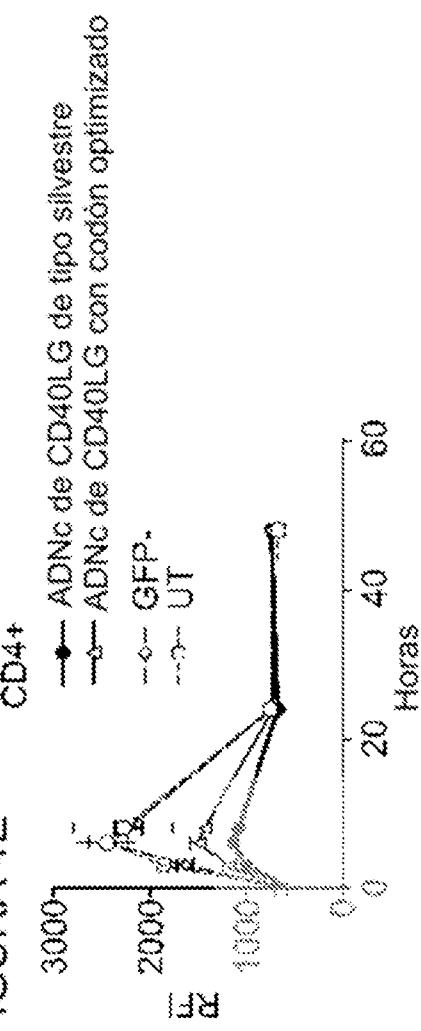


FIGURA 5A

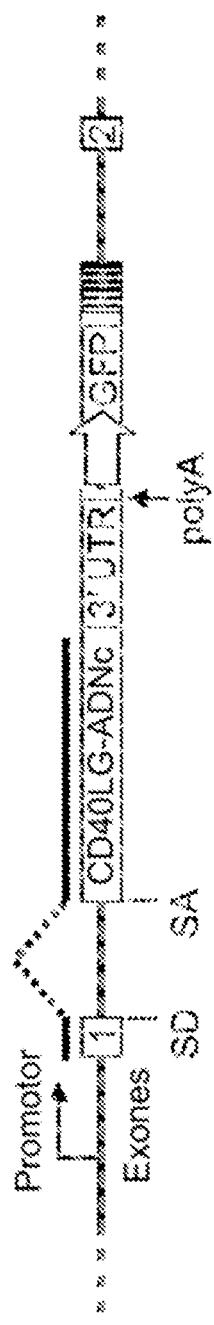


FIGURA 5B

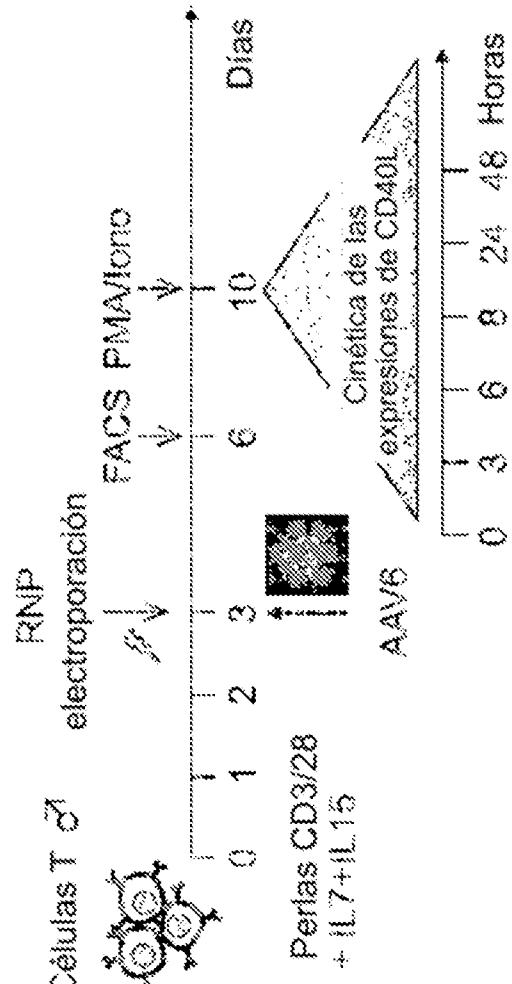


FIGURA 5C

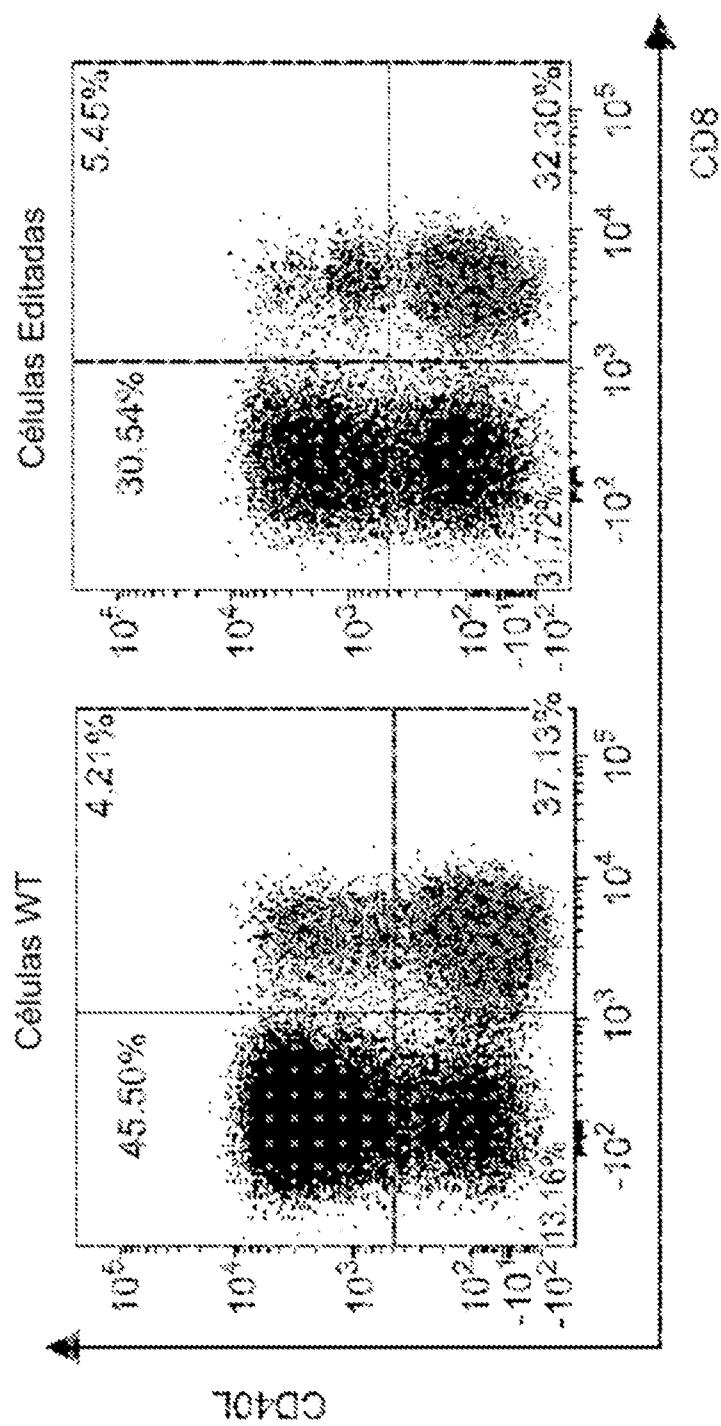


FIGURA 5D

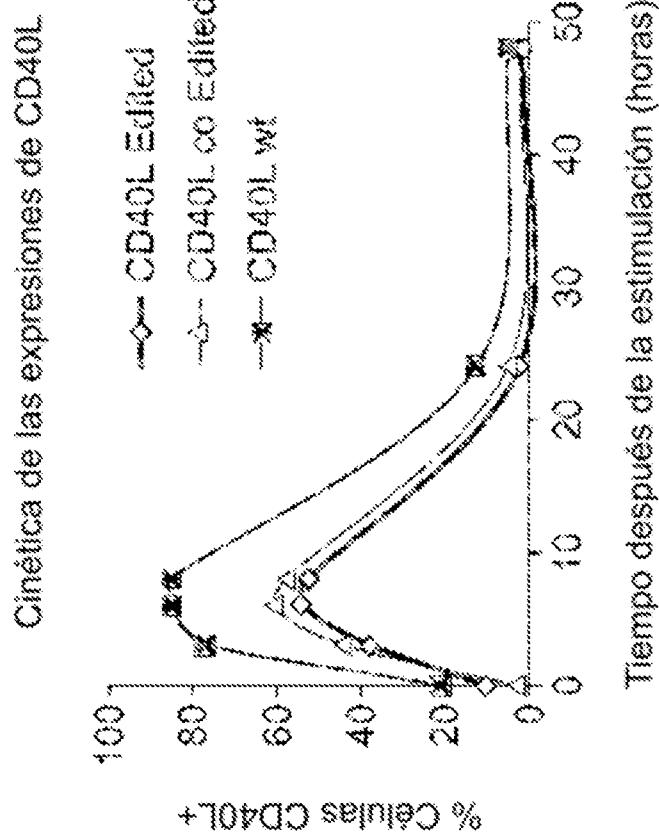


FIG. 5E

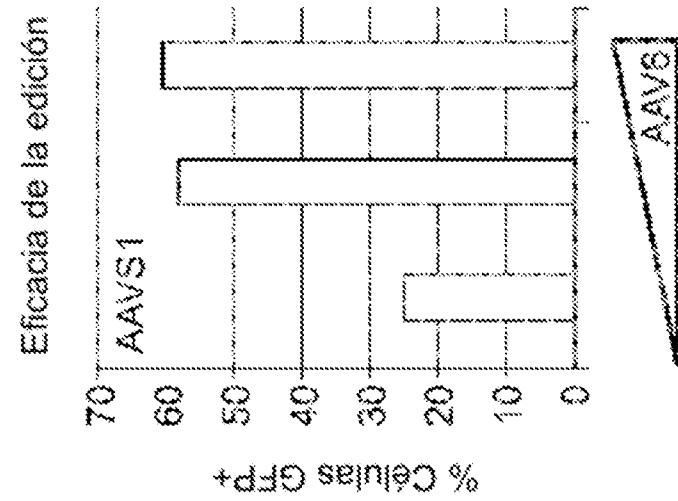


FIGURA 6A

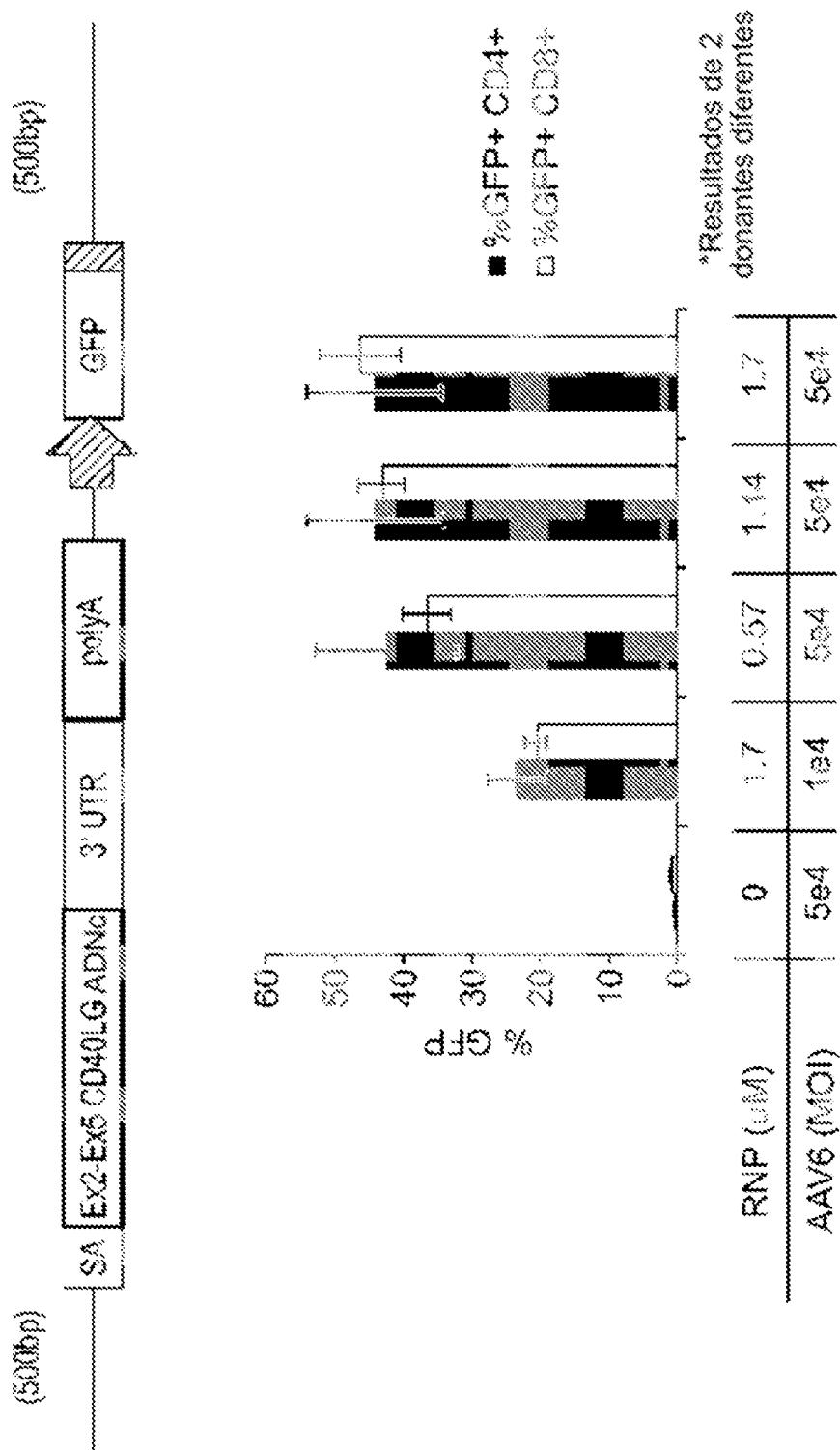
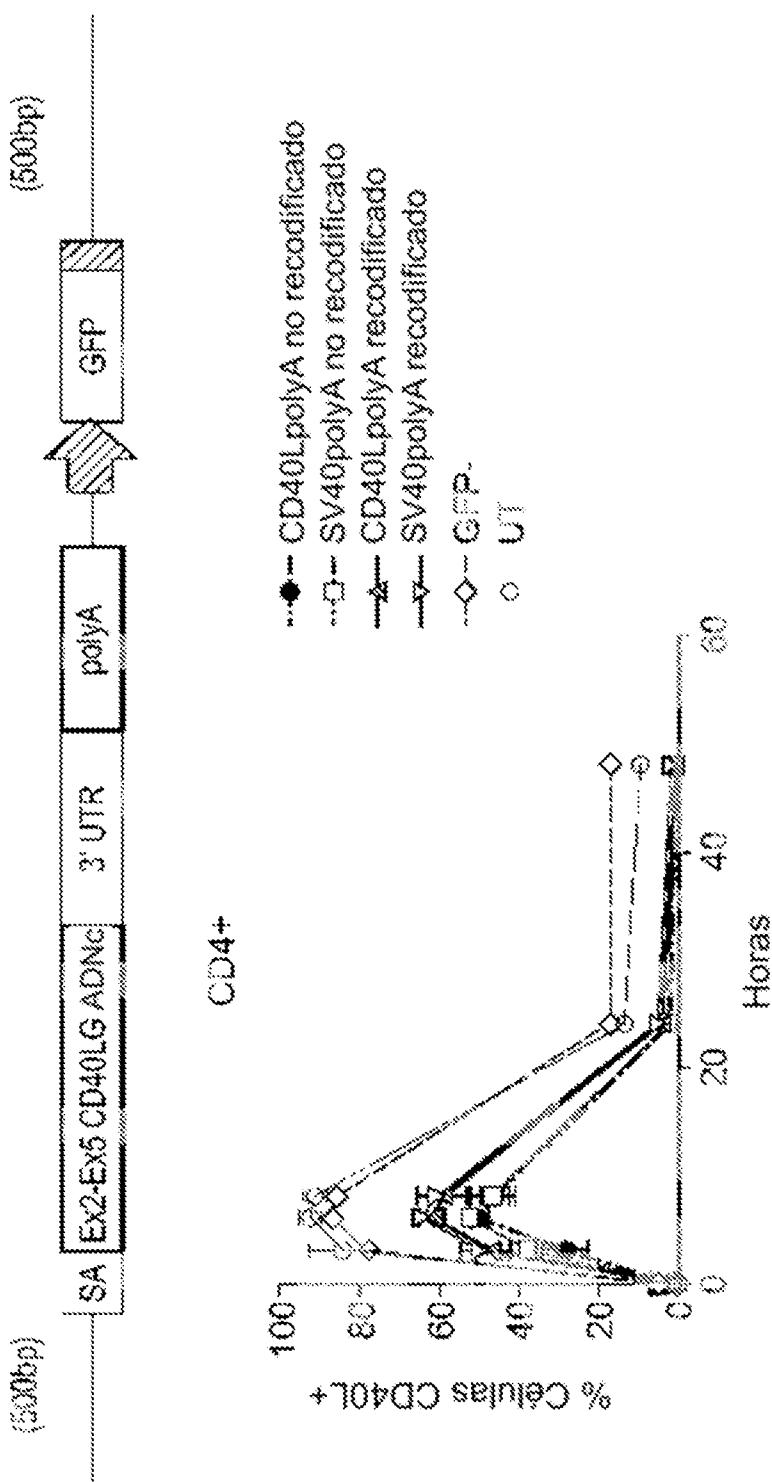


FIGURA 6B



Alcanzado el 60% de la expresión exógena de CD40L

FIGURA 7A

| Nombre Guía | Cass | #Lecturas On-Target (Bio-Rep1) | #Lecturas On-Target (Bio-Rep2) | #Lecturas On-Target (Bio-Rep3) | #Off-Targets Detectados | #Lecturas Off-Target (Bio-Rep1) | #Lecturas Off-Target (Bio-Rep2) | #Lecturas Off-Target (Bio-Rep3) |
|-------------|------|--------------------------------|--------------------------------|--------------------------------|-------------------------|---------------------------------|---------------------------------|---------------------------------|
| A | Spy | 13275 | 5466 | 8778 | 1 | 380* | 12* | 124 |
| B | Sau | 13155 | 8893 | 5813 | 0 | N/A | N/A | N/A |

*Lecturas unidireccionales

FIGURA 7B

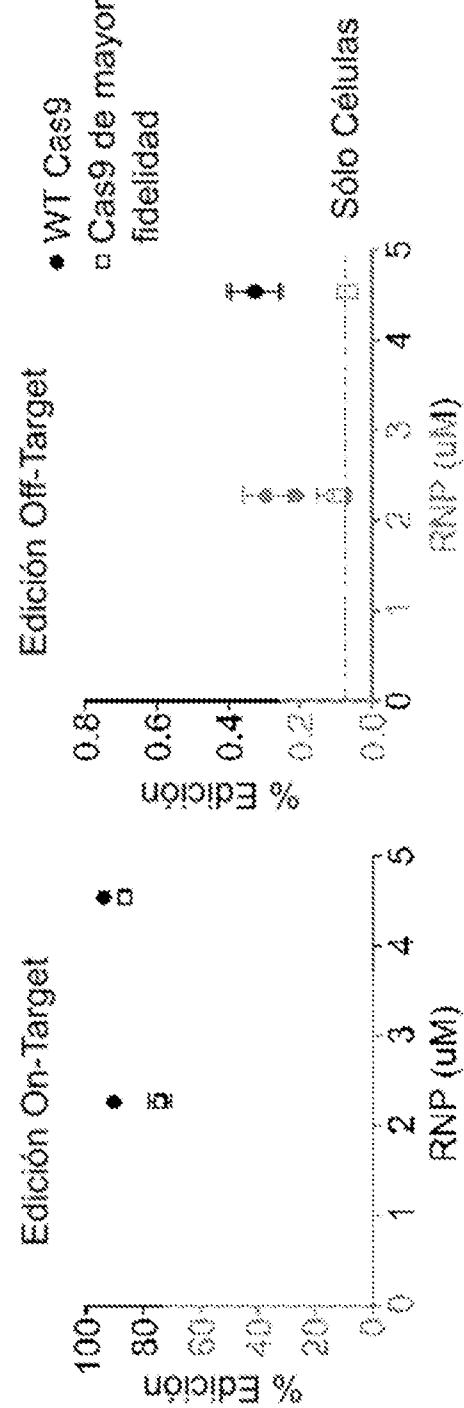


FIGURA 7C

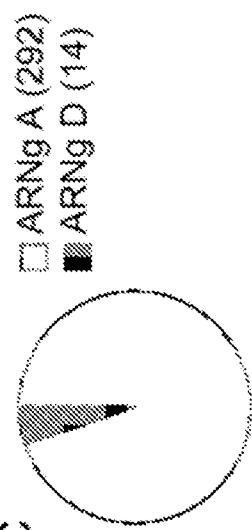


FIGURA 7D

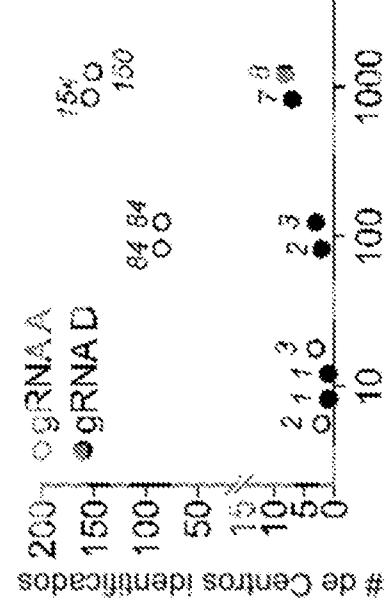
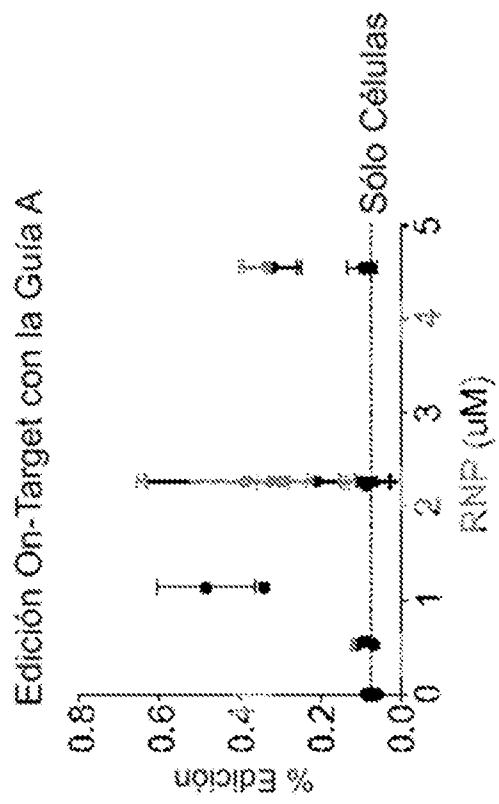


FIGURA 7E



Edición On-Target con la Guía A

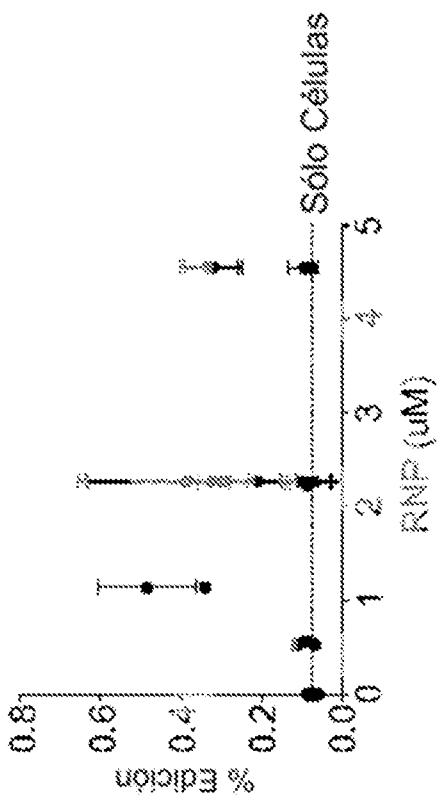
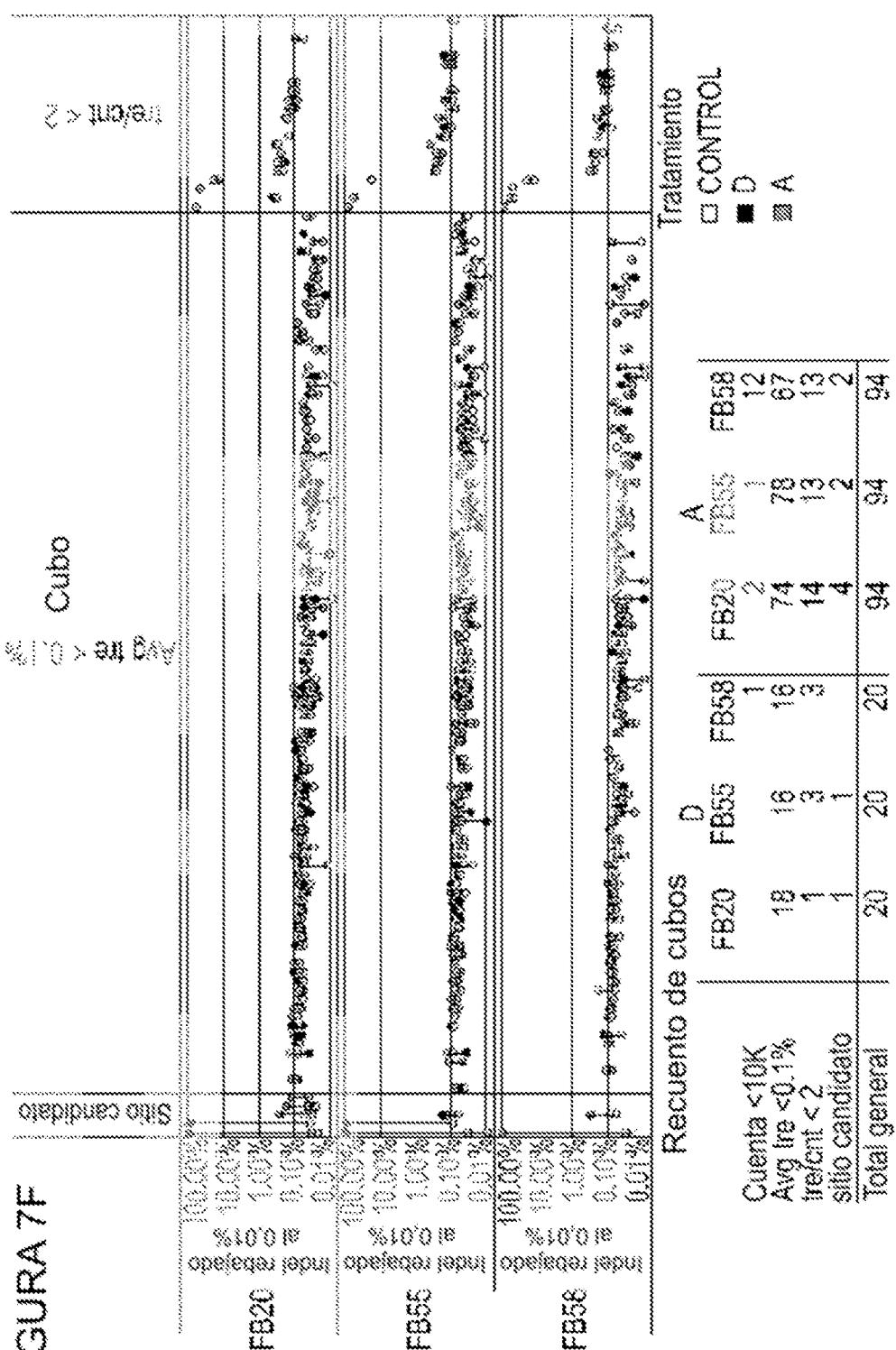


FIGURE 7E



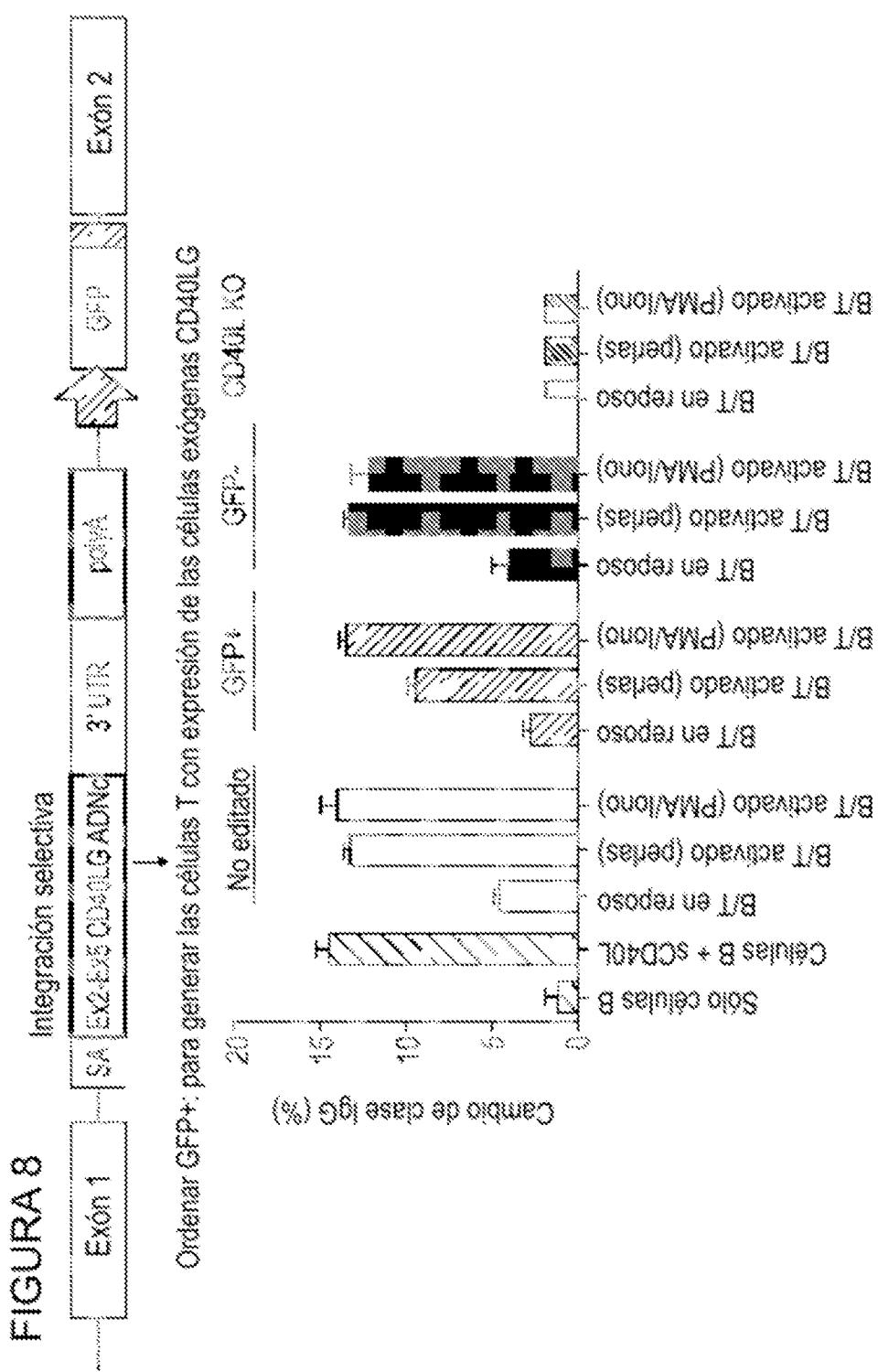
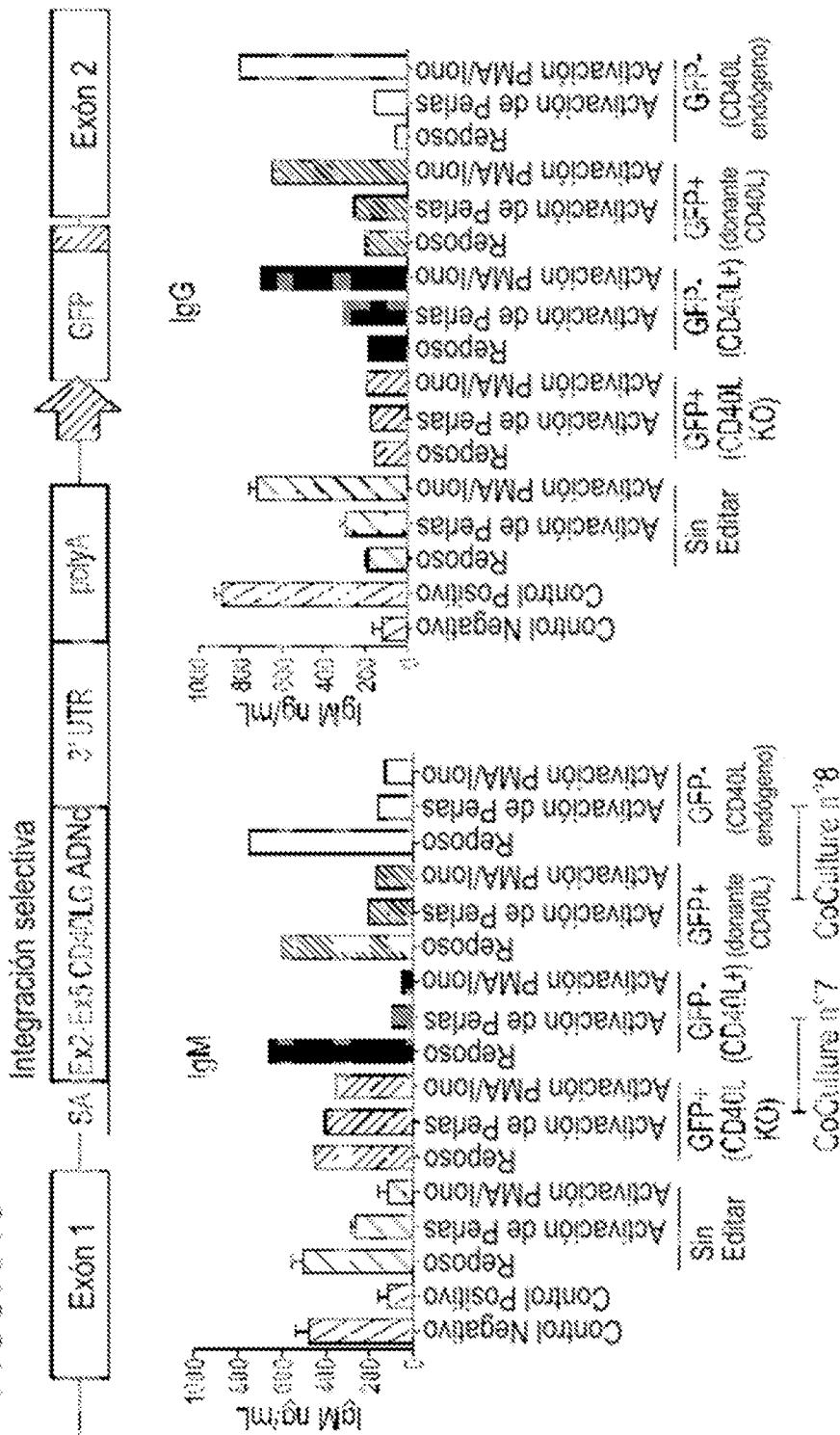


FIGURA 9



ES 2 998 052 T3

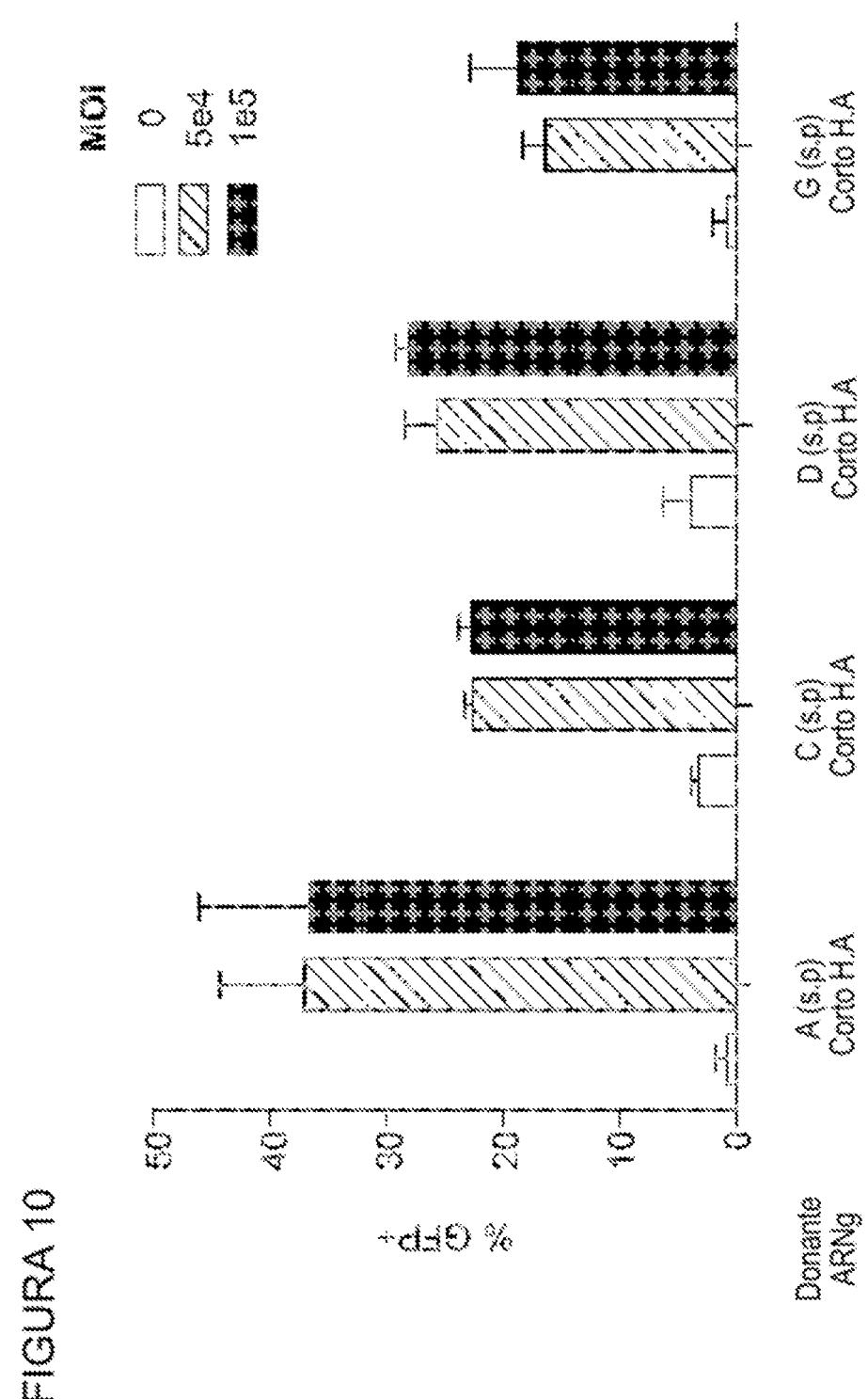


FIGURA 11

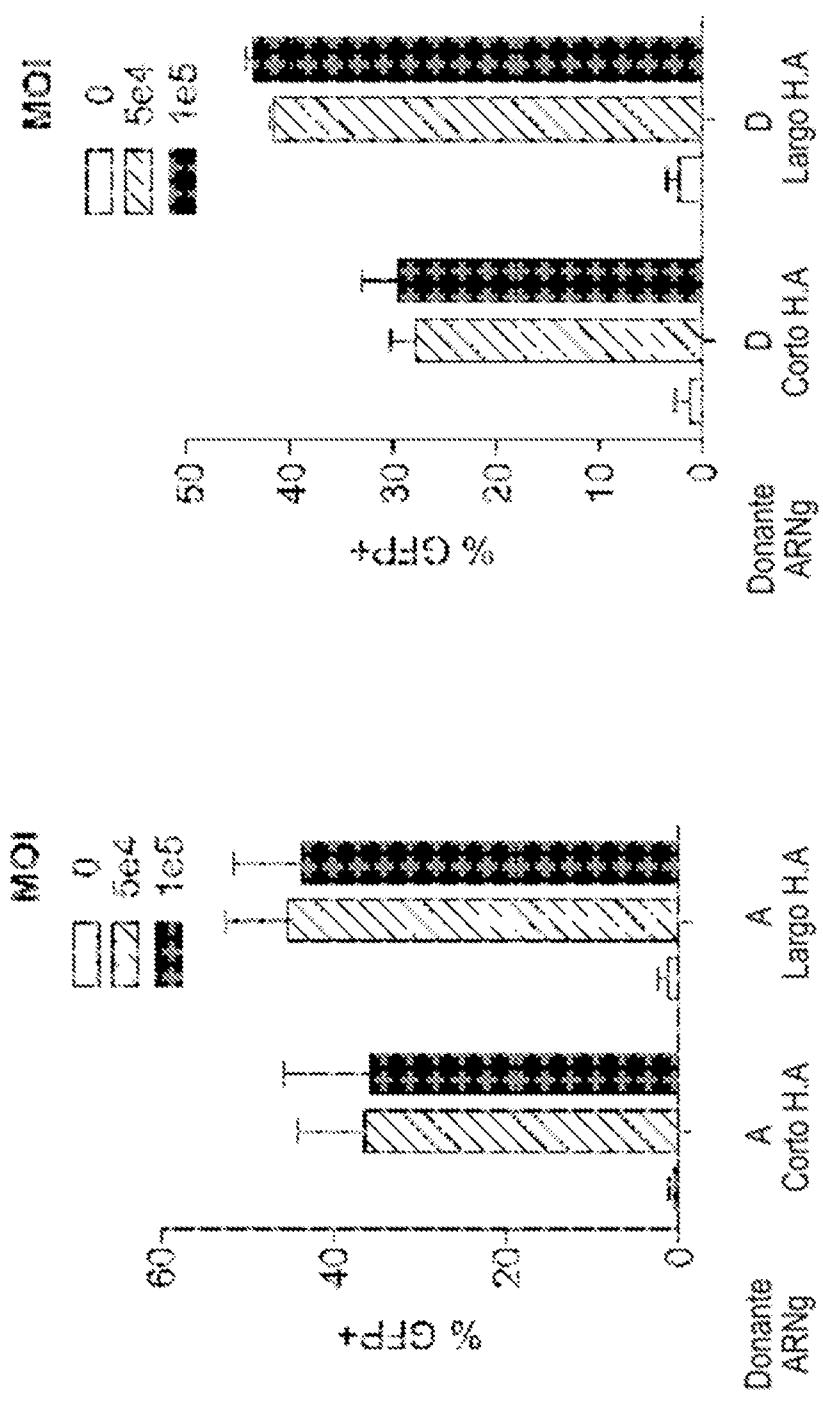
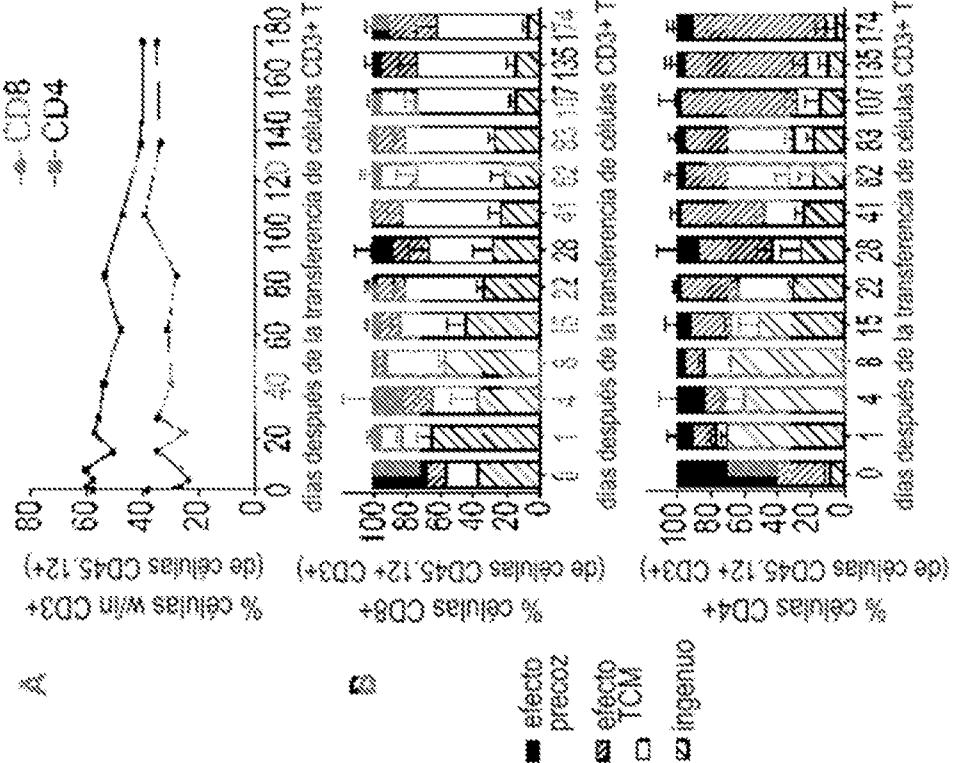


FIGURA 12

NO CONDICIONADO



CONDICIONADO

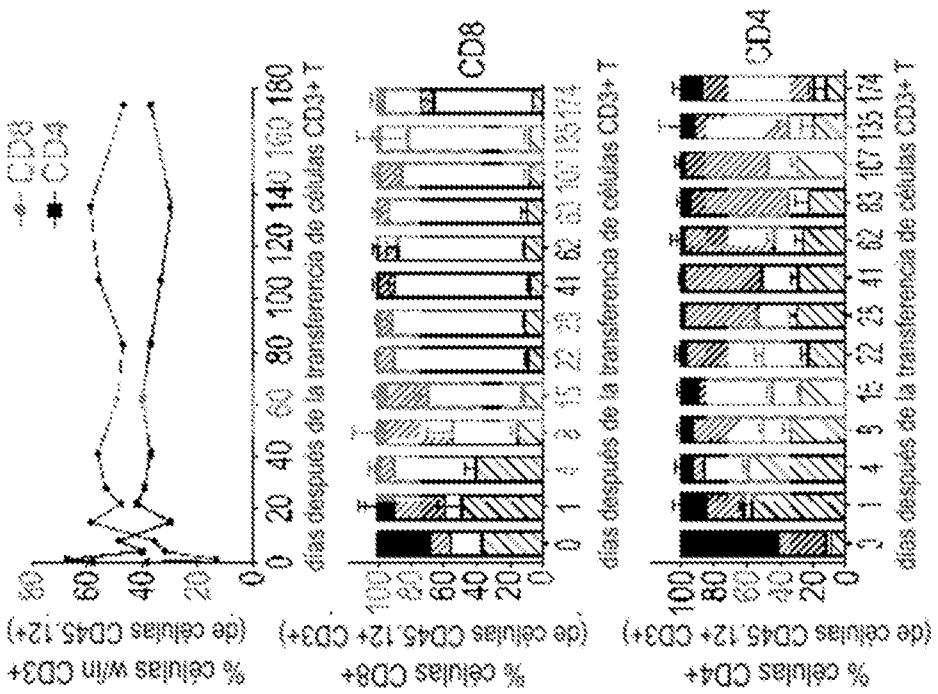


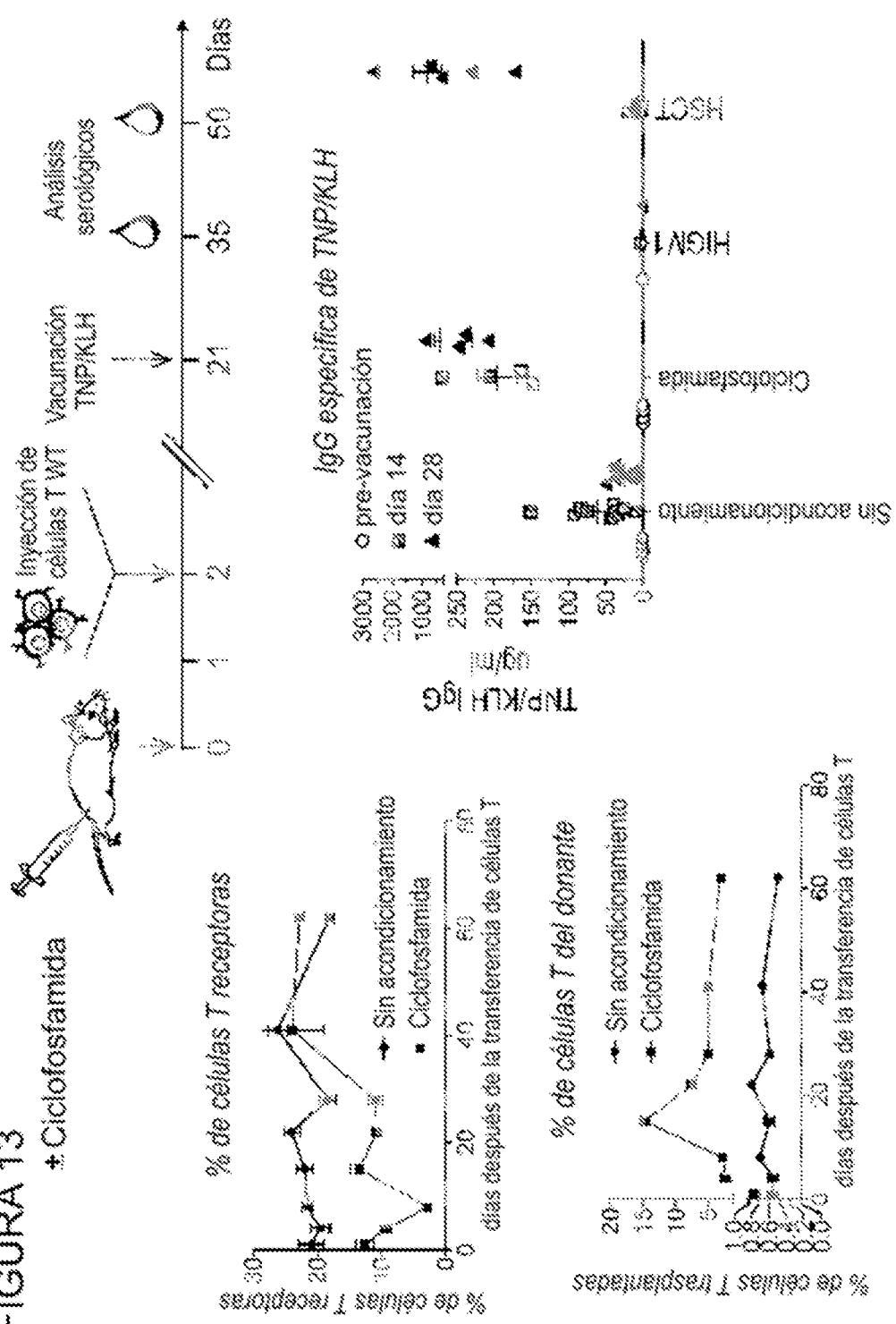
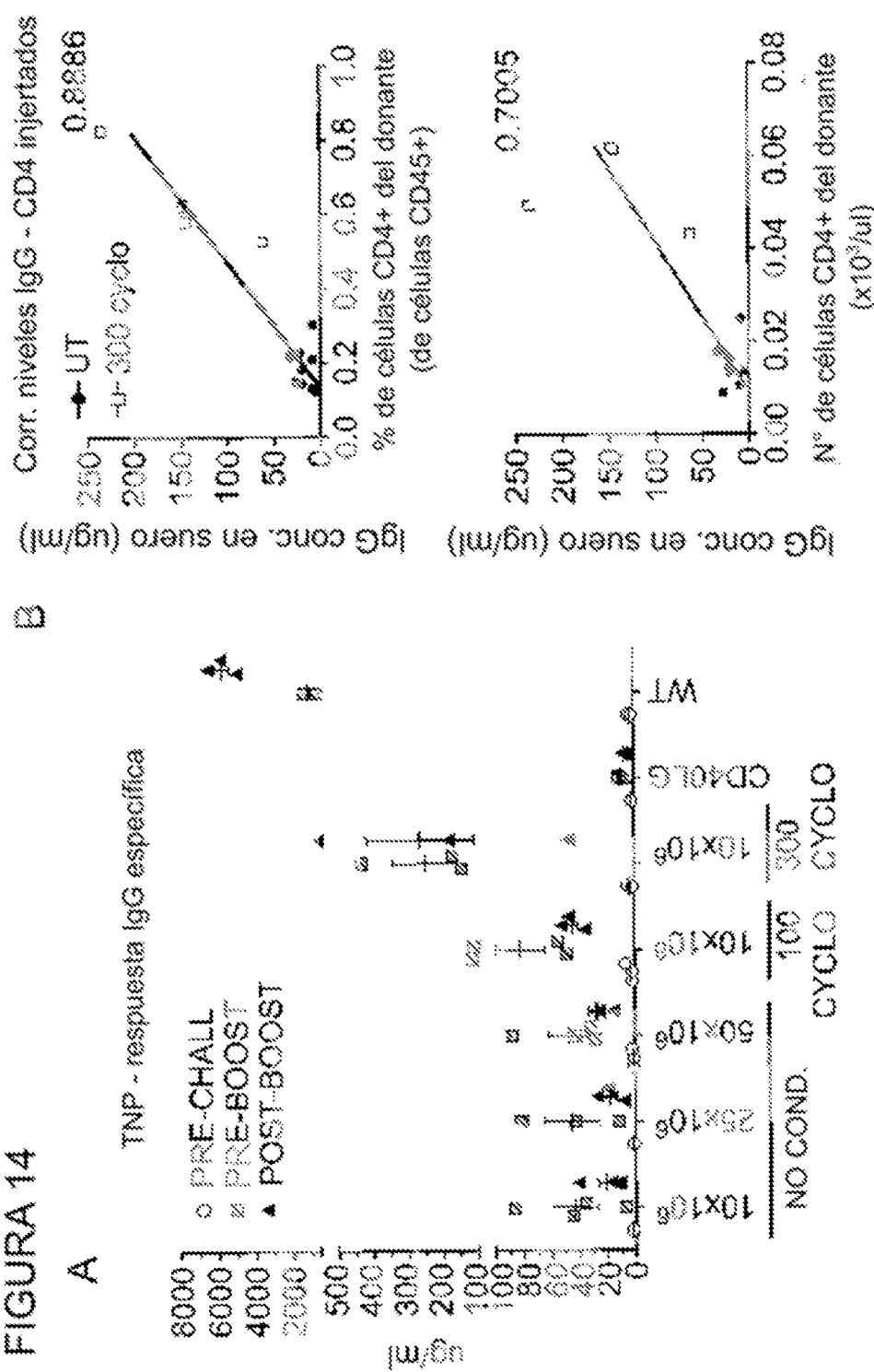
FIGURA 13

FIGURA 14



**EFFECTO CONDICIONANTE
EN EL RECEPTOR**

FIGURA 14

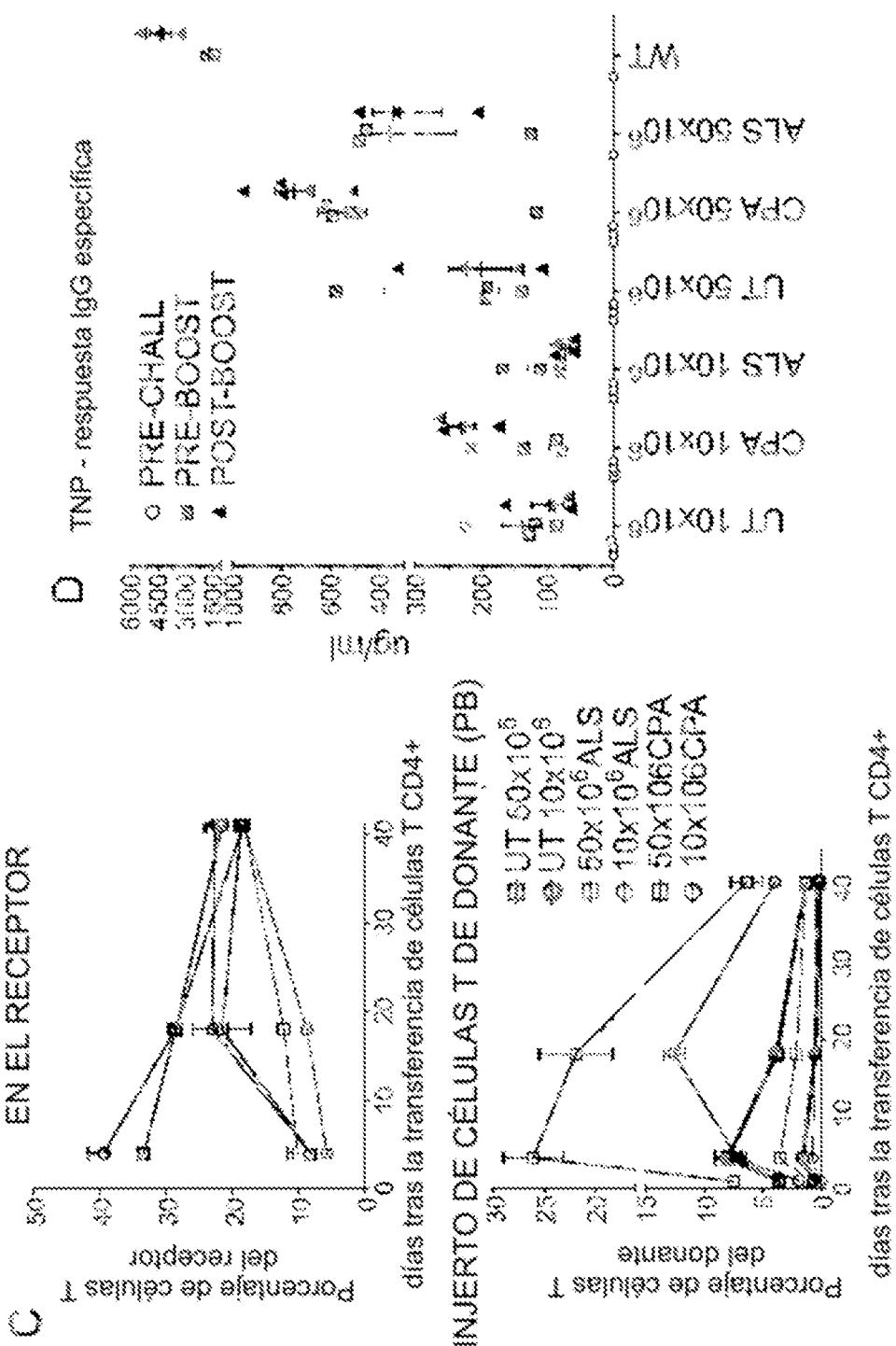


FIGURA 14

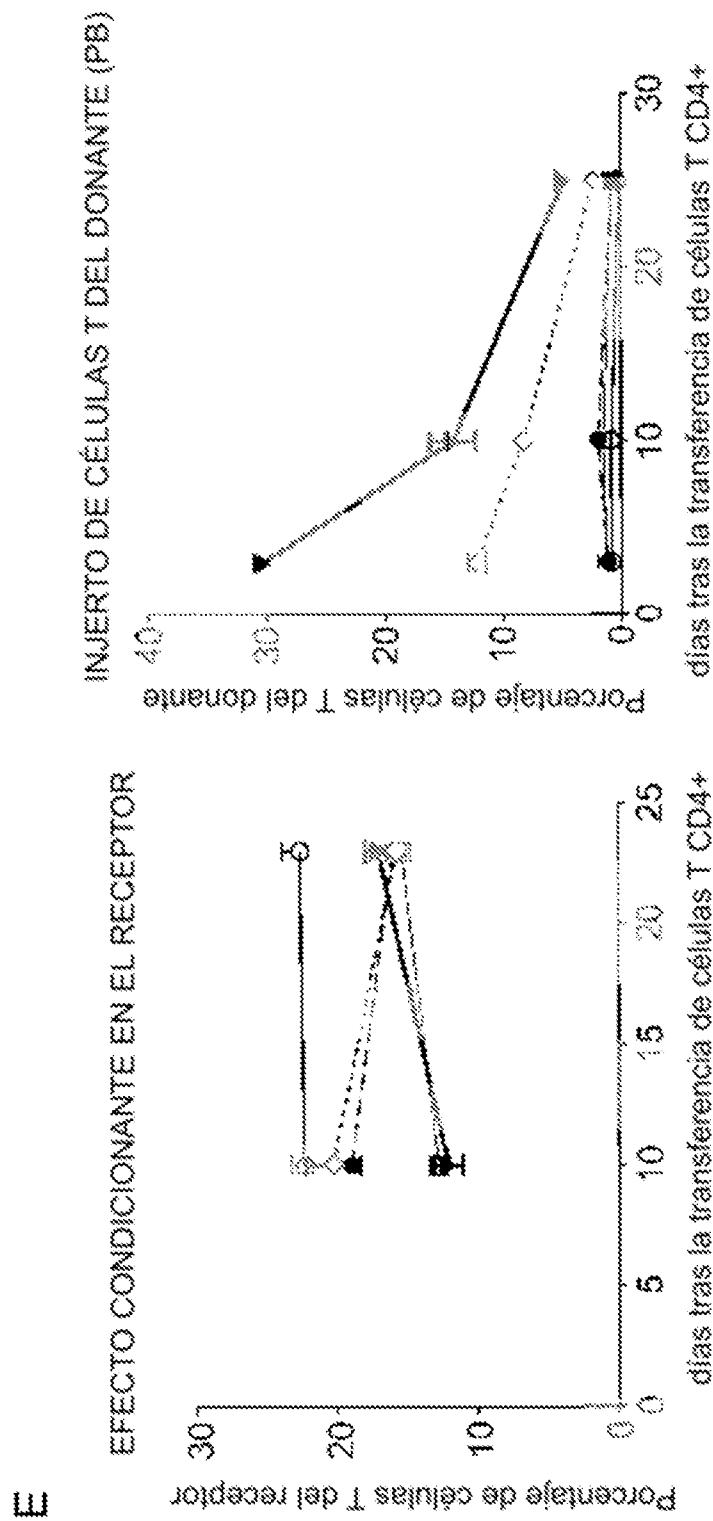
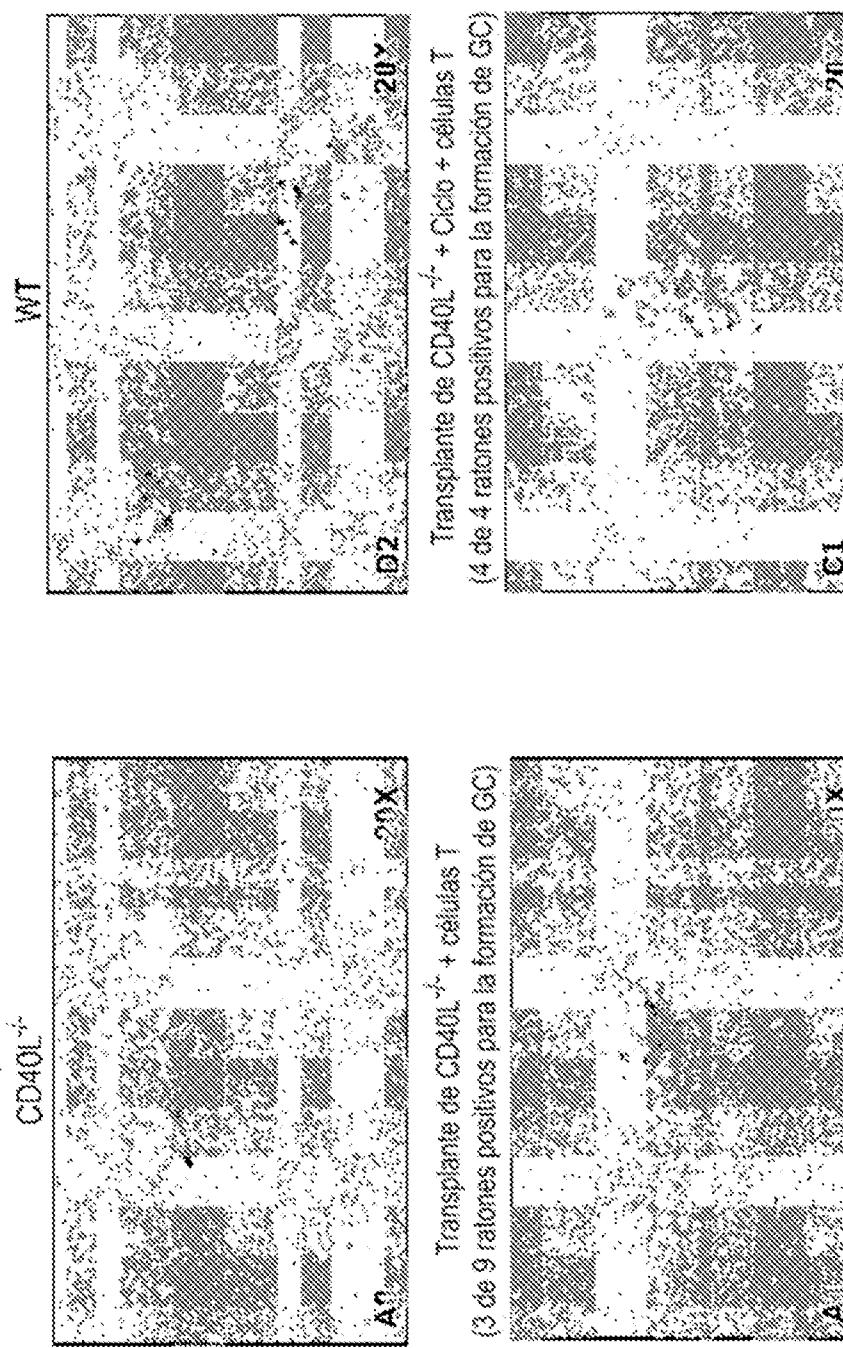


FIGURA 15A

Inmunohistoquímica de células SC B en secciones de bazo de ratones inmunizados con ThP-XLH



La inyección con PHA (agujaña de cacahuate) se muestra en marrón y la contrasteación con hematoxilina en azul.

FIGURA 15B

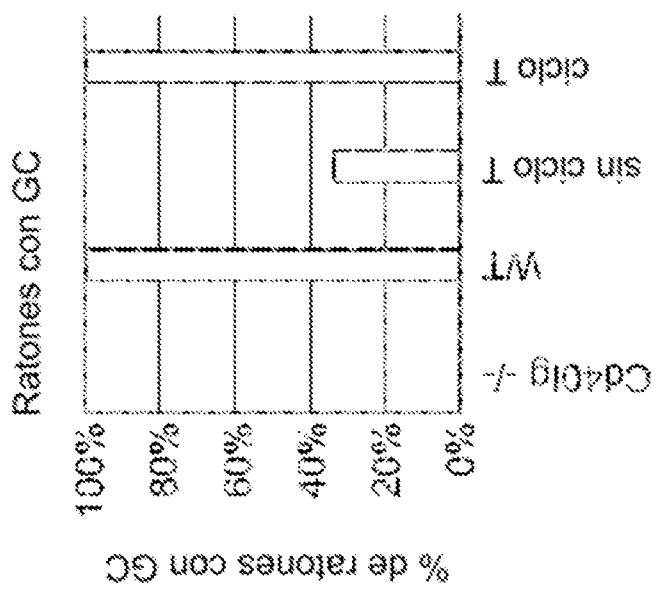
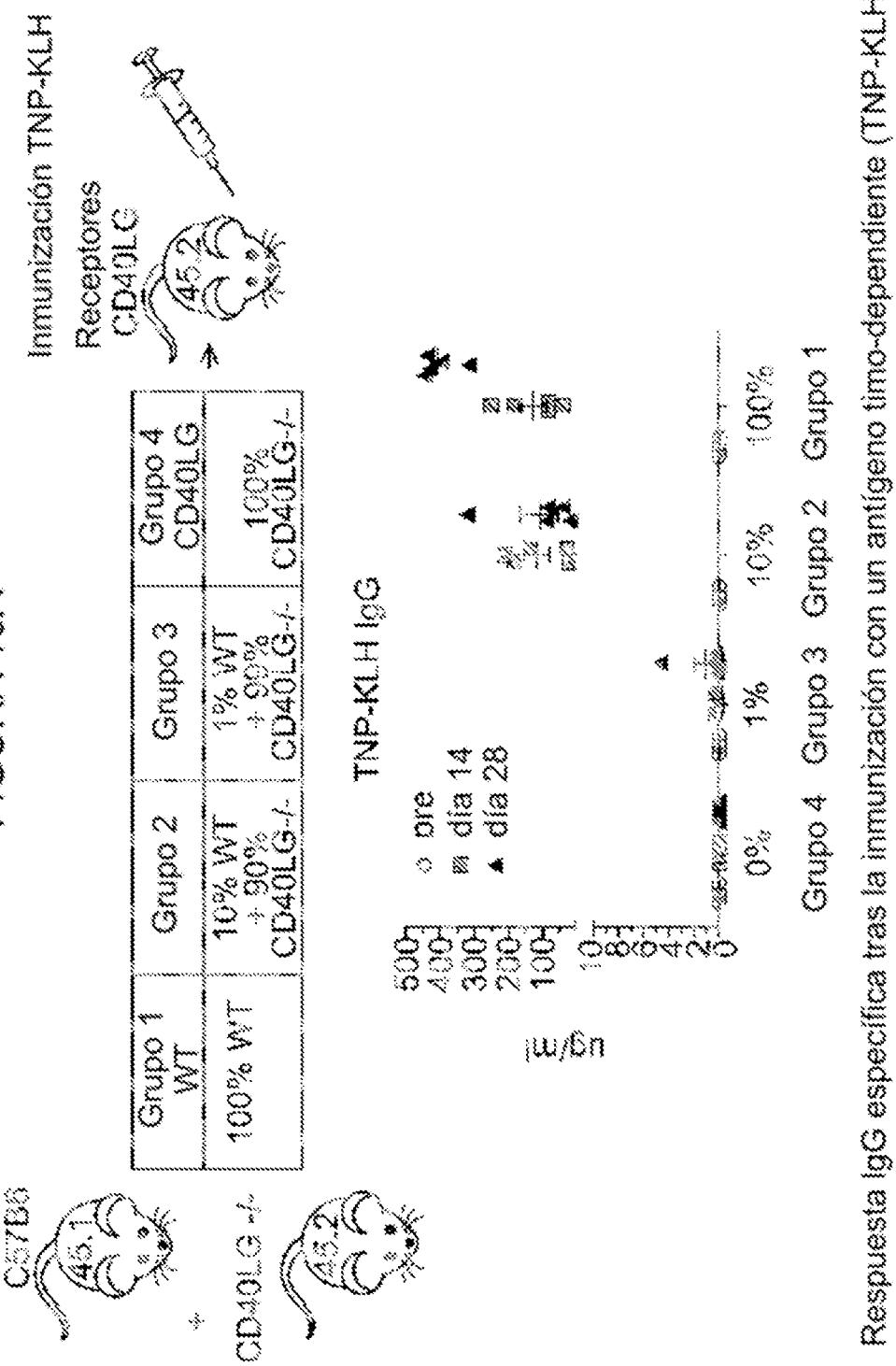


FIGURA 16A



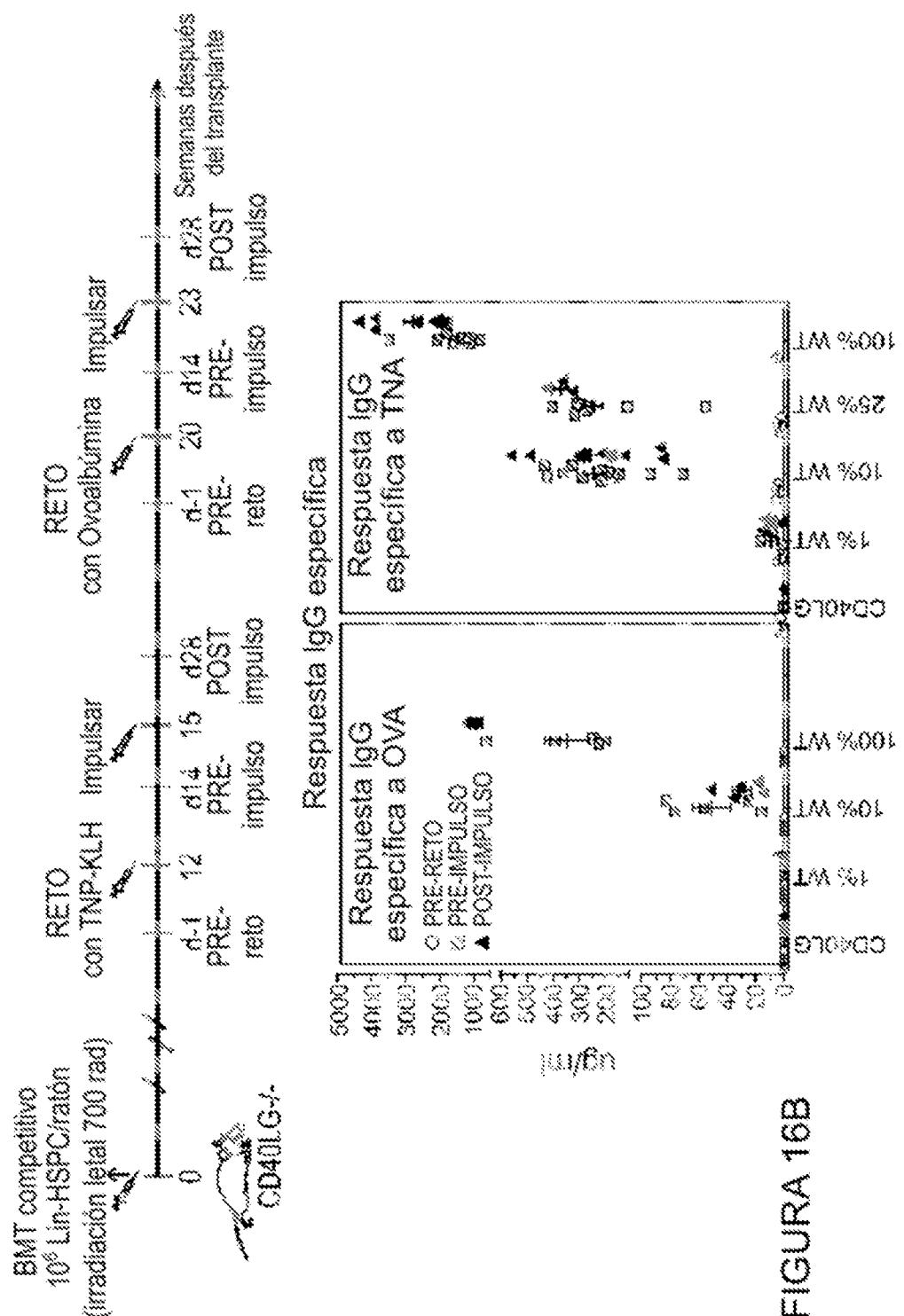


FIGURA 16B

FIGURA 16C

Injerto de células T del donante en la PB

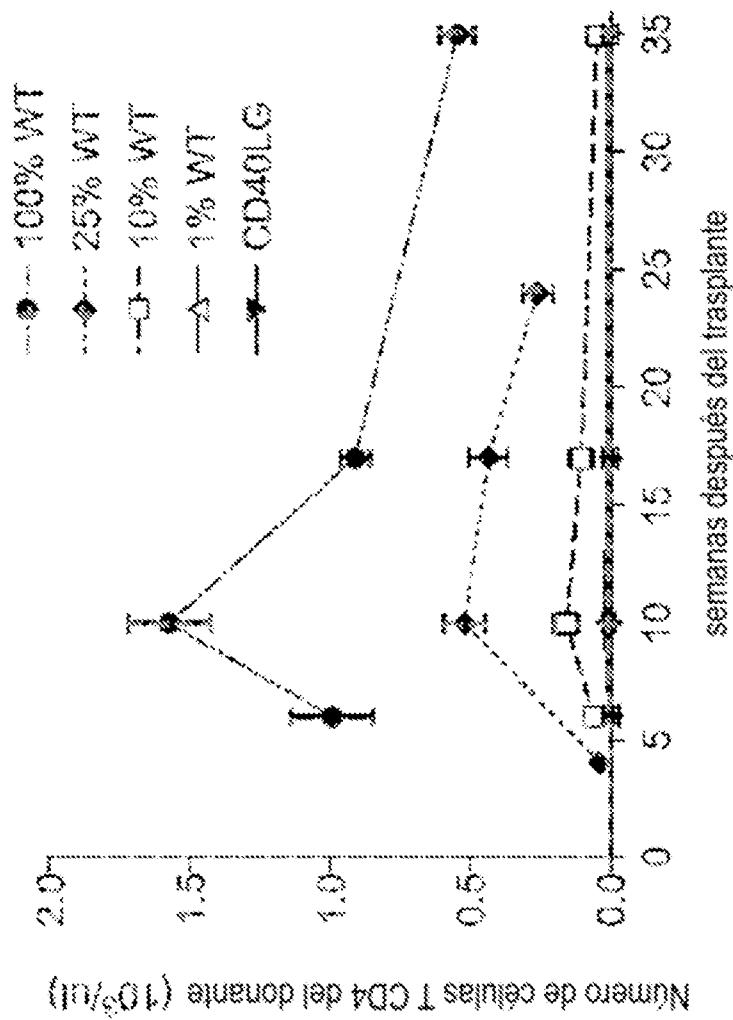


FIGURA 17

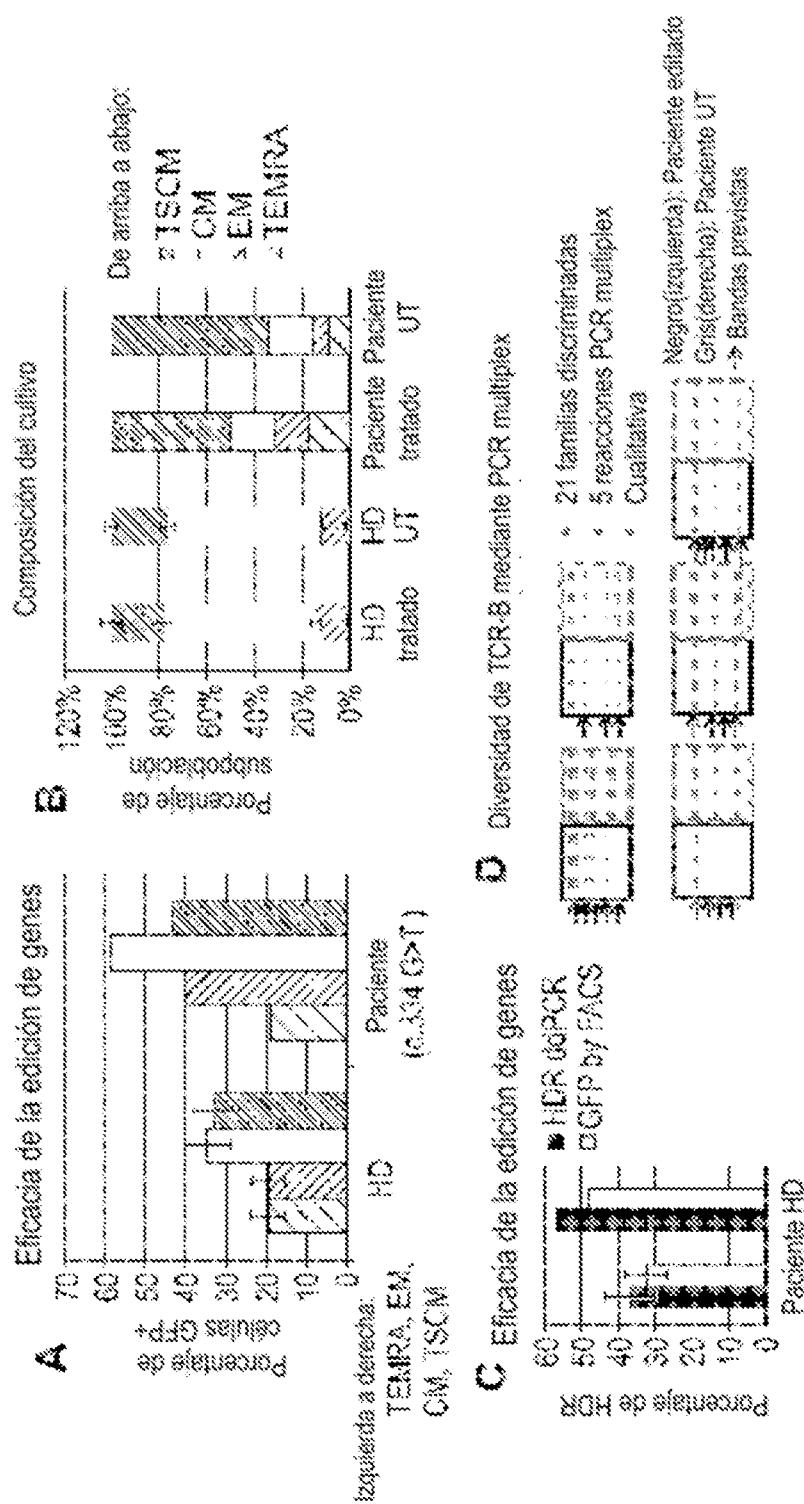


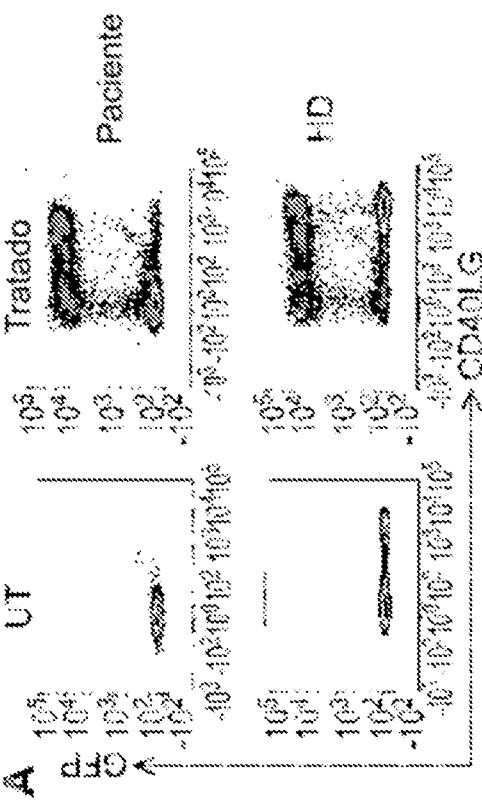
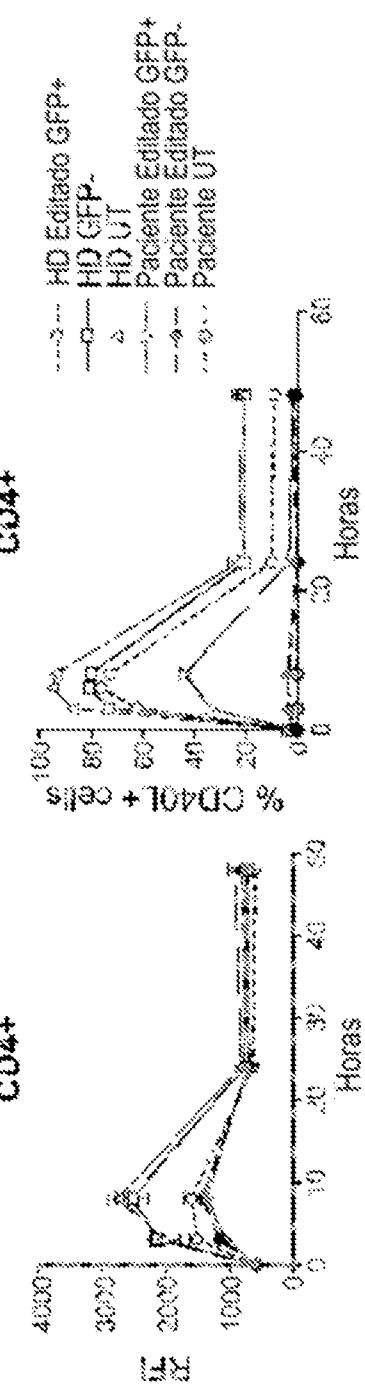
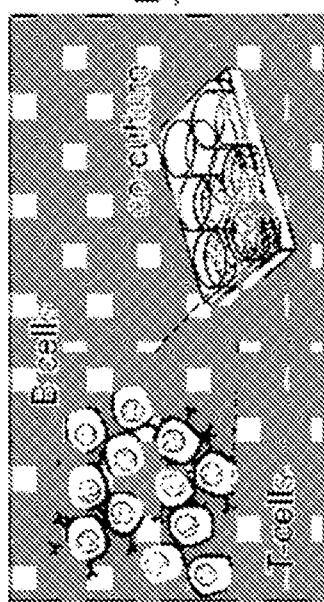
FIGURA 18 Expresión de CD40LG tras la estimulación**3** Niveles de expresión de CD40LG

FIGURA 19 A

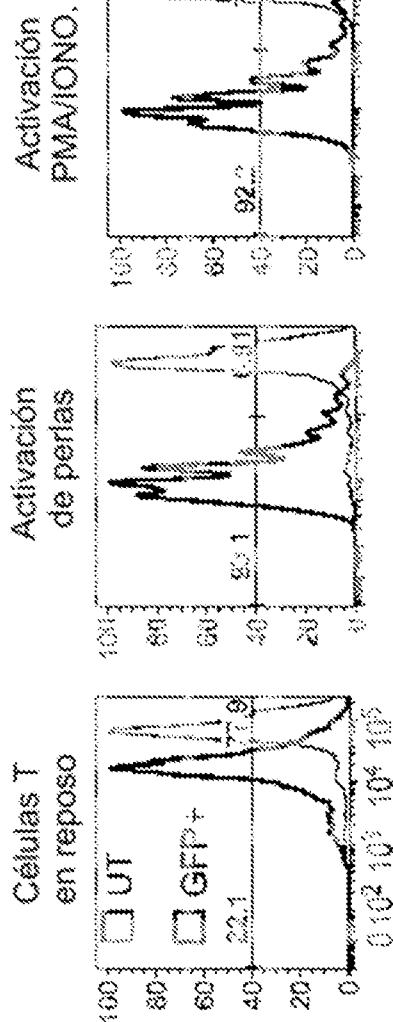
CITOQUINA

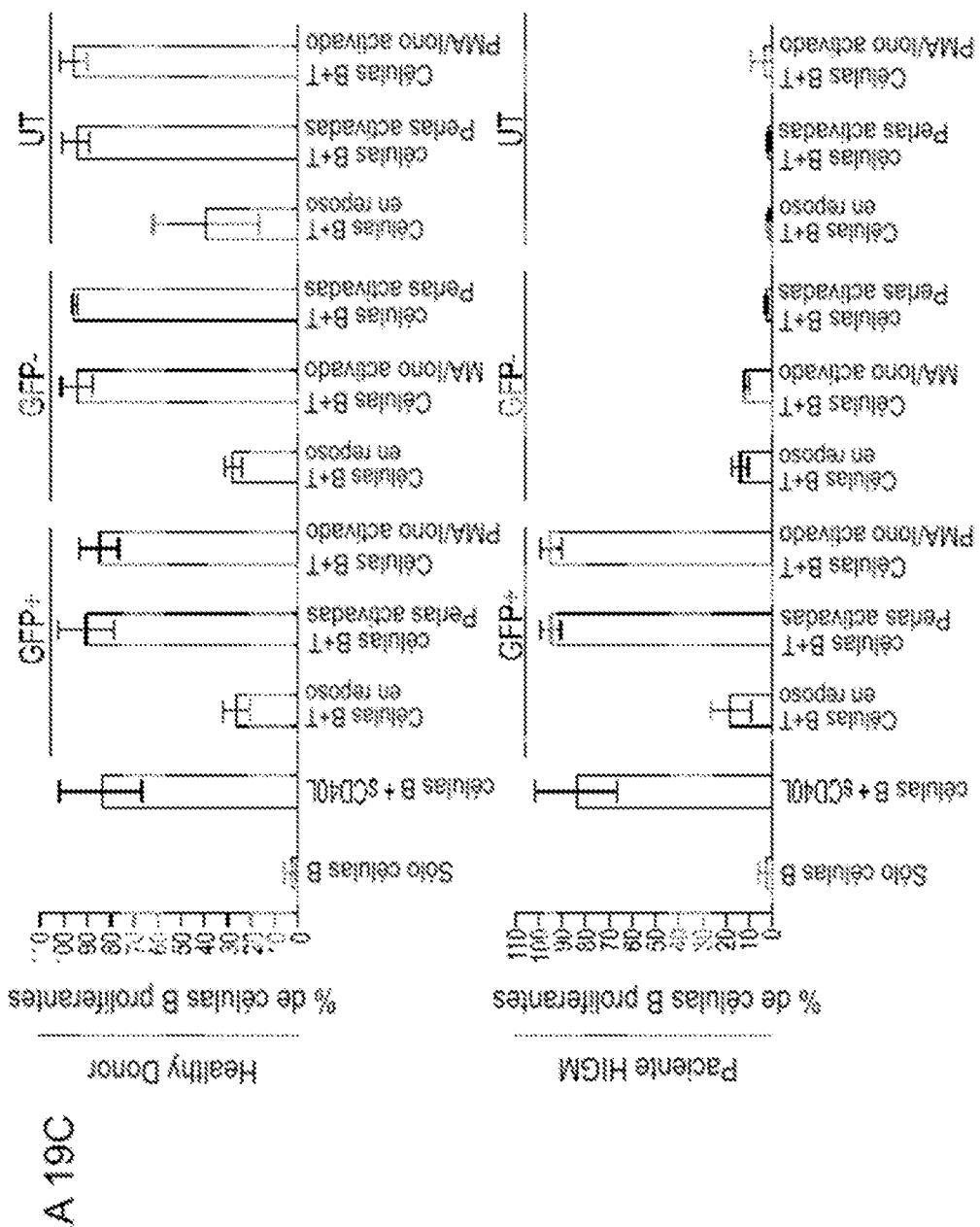


- | FACTORES ESTIMULANTES | |
|-----------------------|--|
| • Interleucina-2 | • Interleucina-21 |
| • Interleucina-7 | • cGMP + algG + algA |
| • Interleucina-16 | • CpG ODNs (Ligandos TLR9) |
| | • CD40L soluble como control positivo. |

B

Ensayo de proliferación de células B





EICURA 199C

FIGURA 20

A Ensayo de cambio de clase de células B

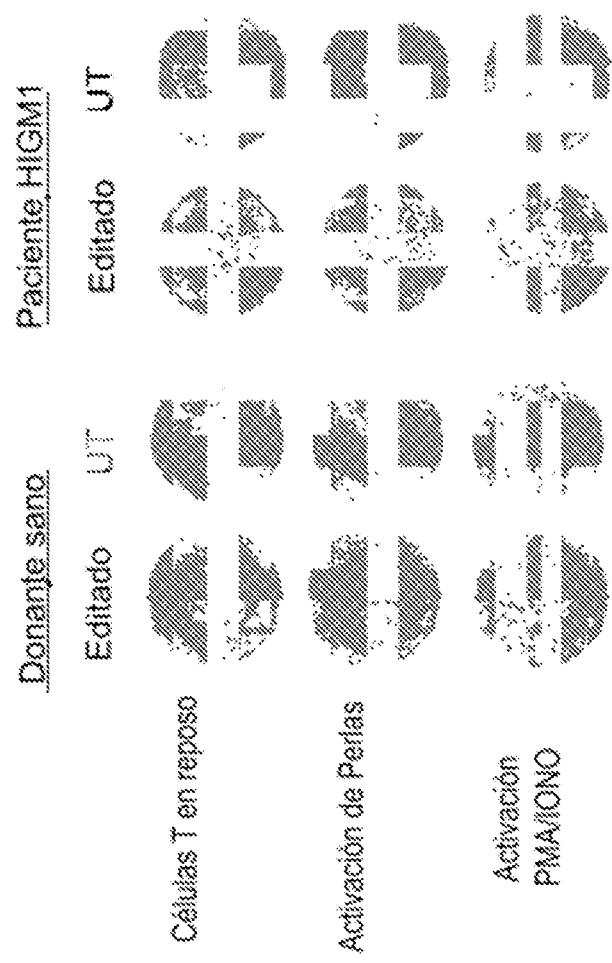


FIGURA 20

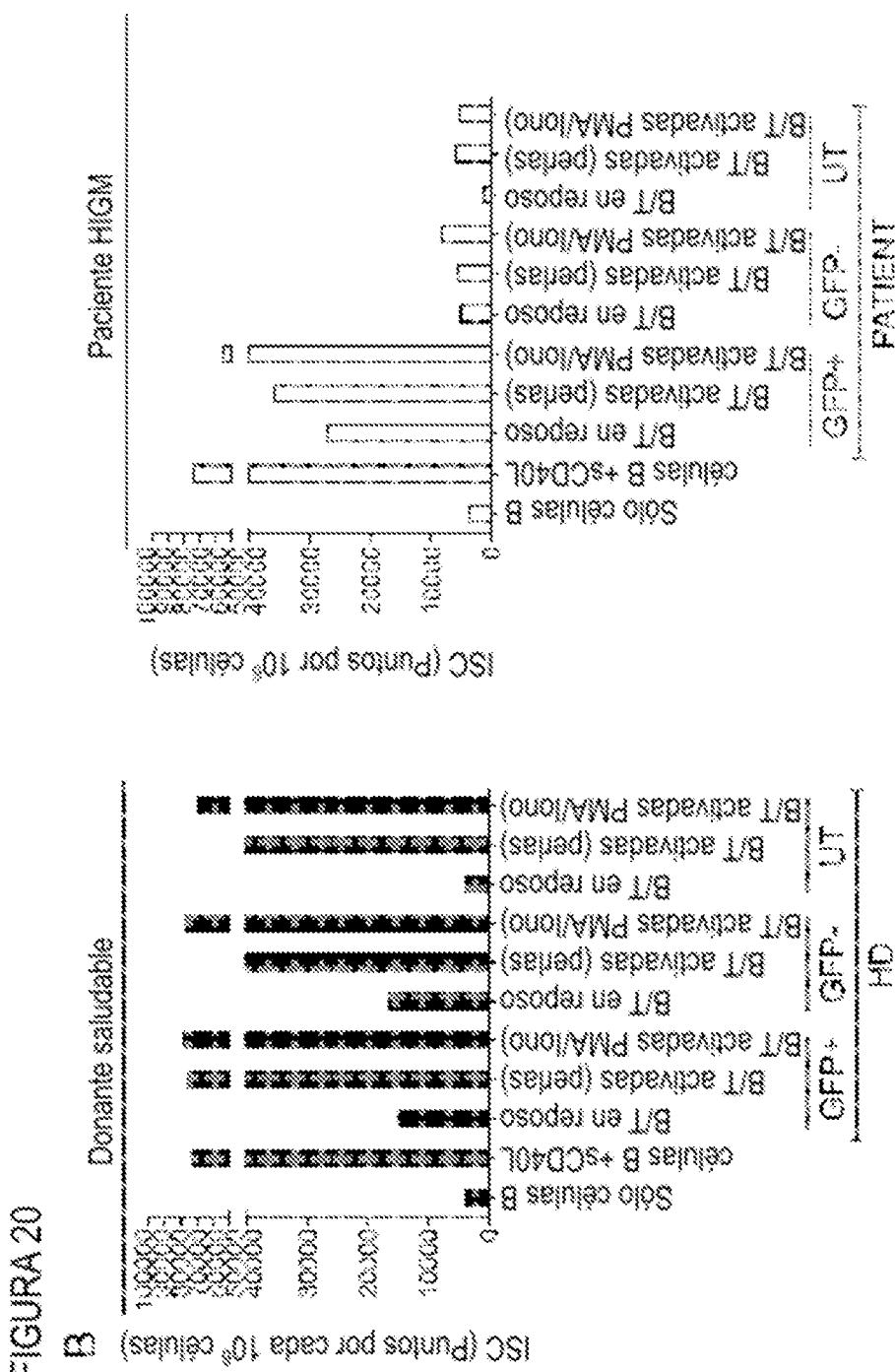


FIGURA 21

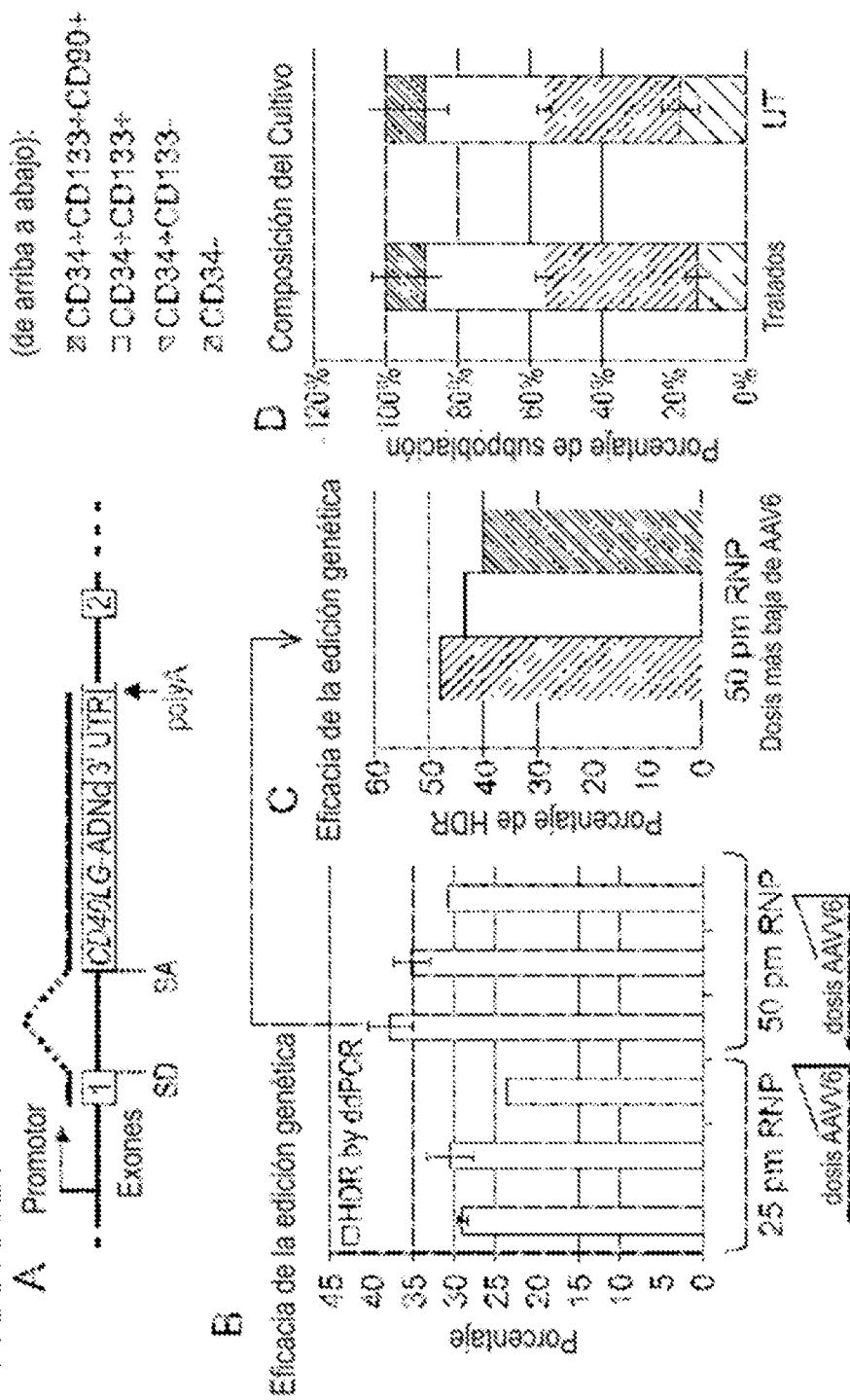
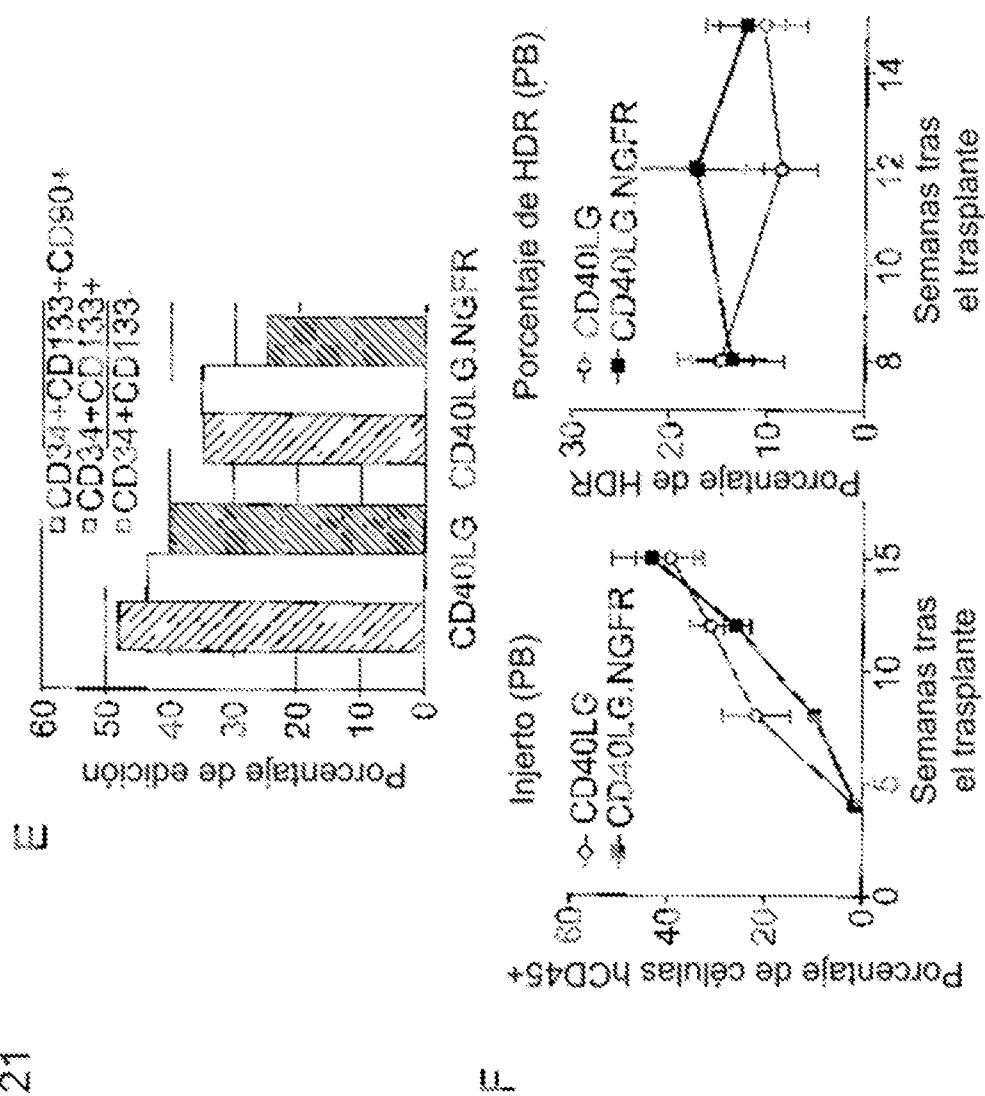


FIGURA 21



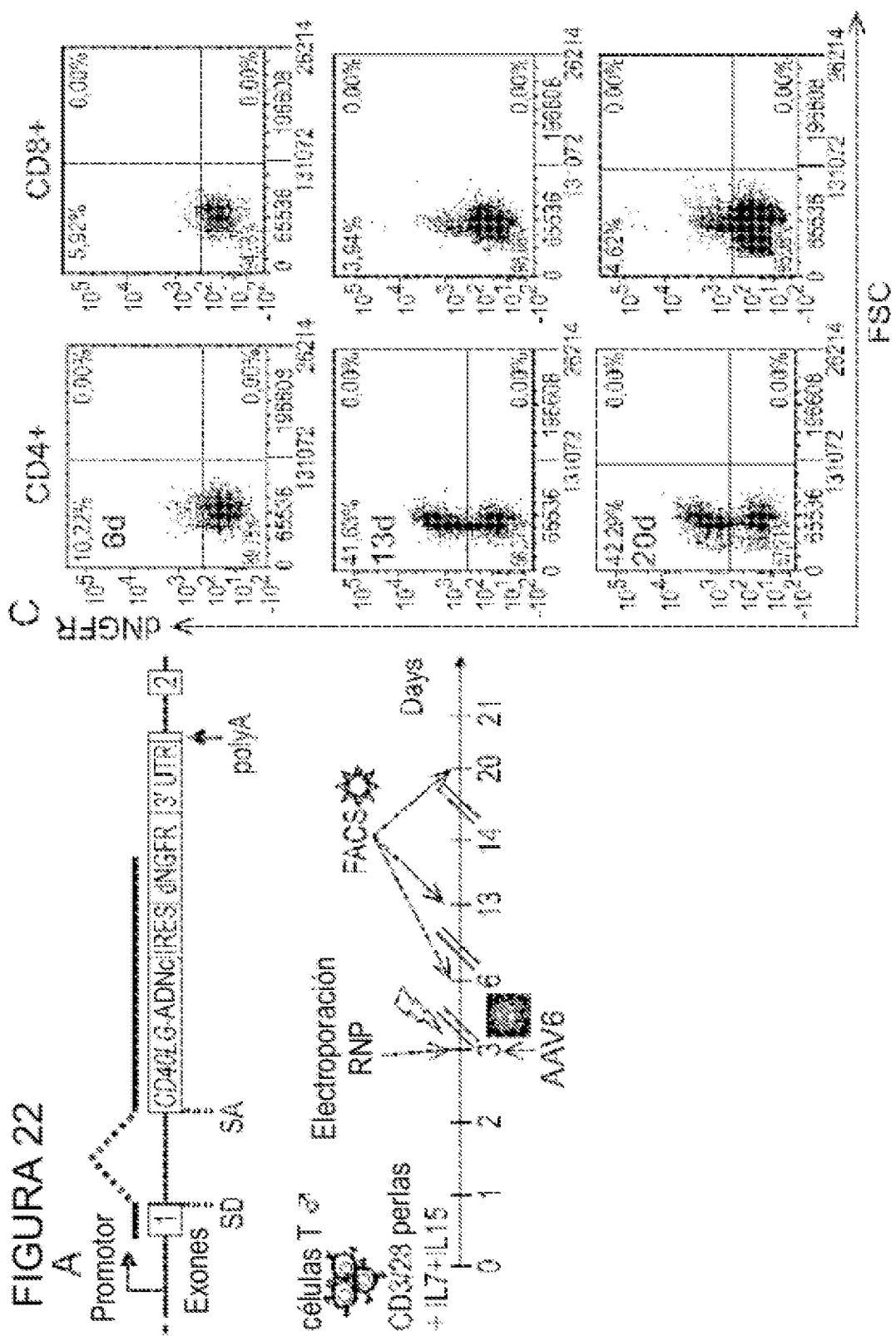


FIGURA 22

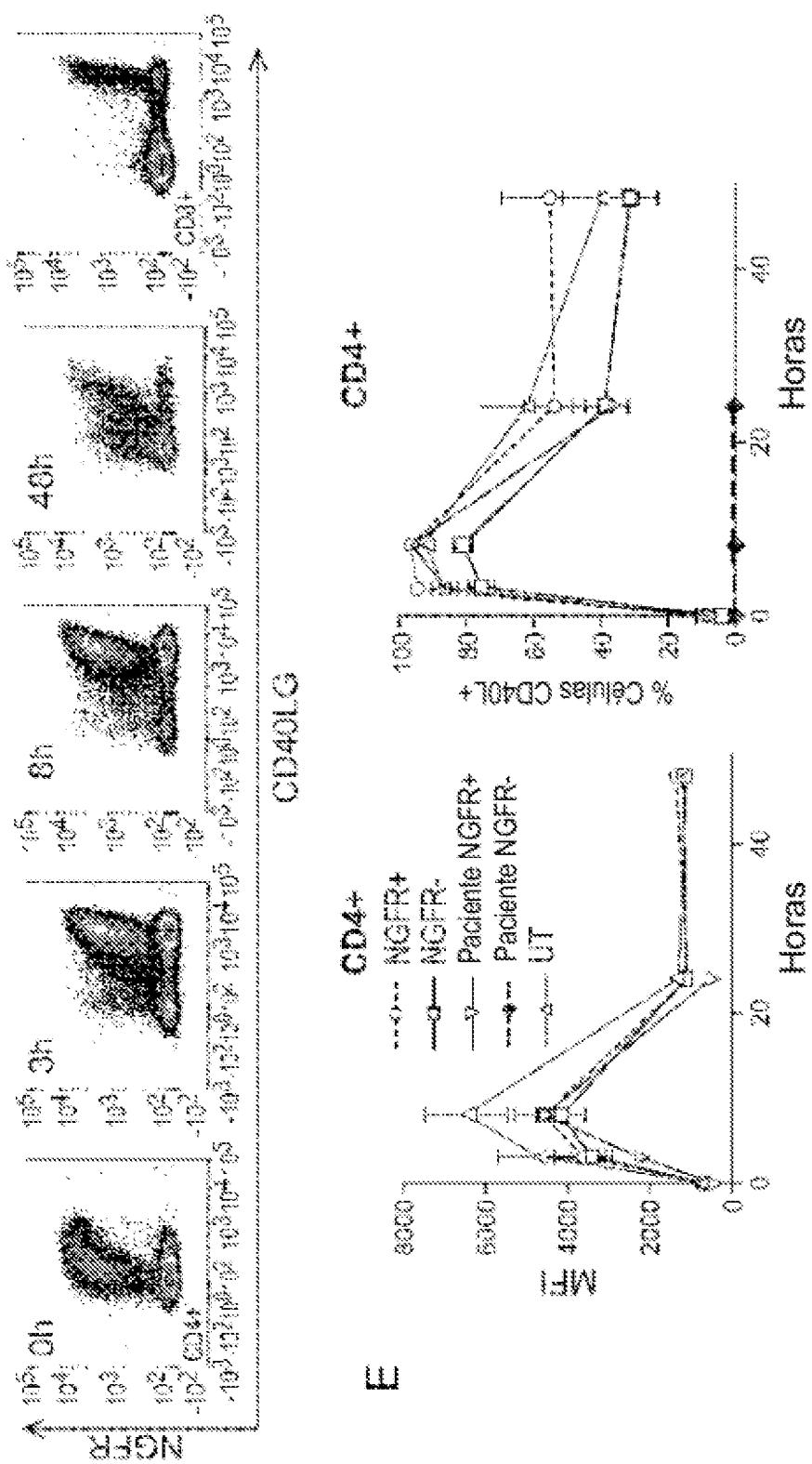
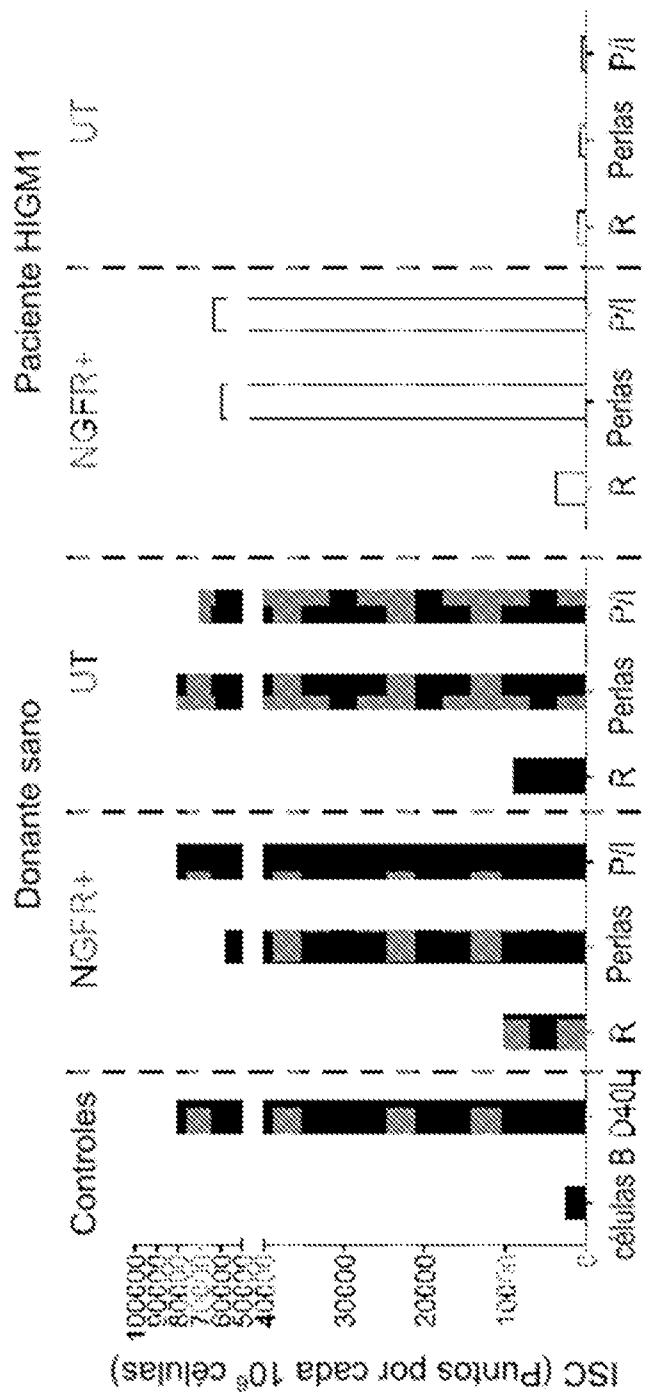


FIGURA 22F



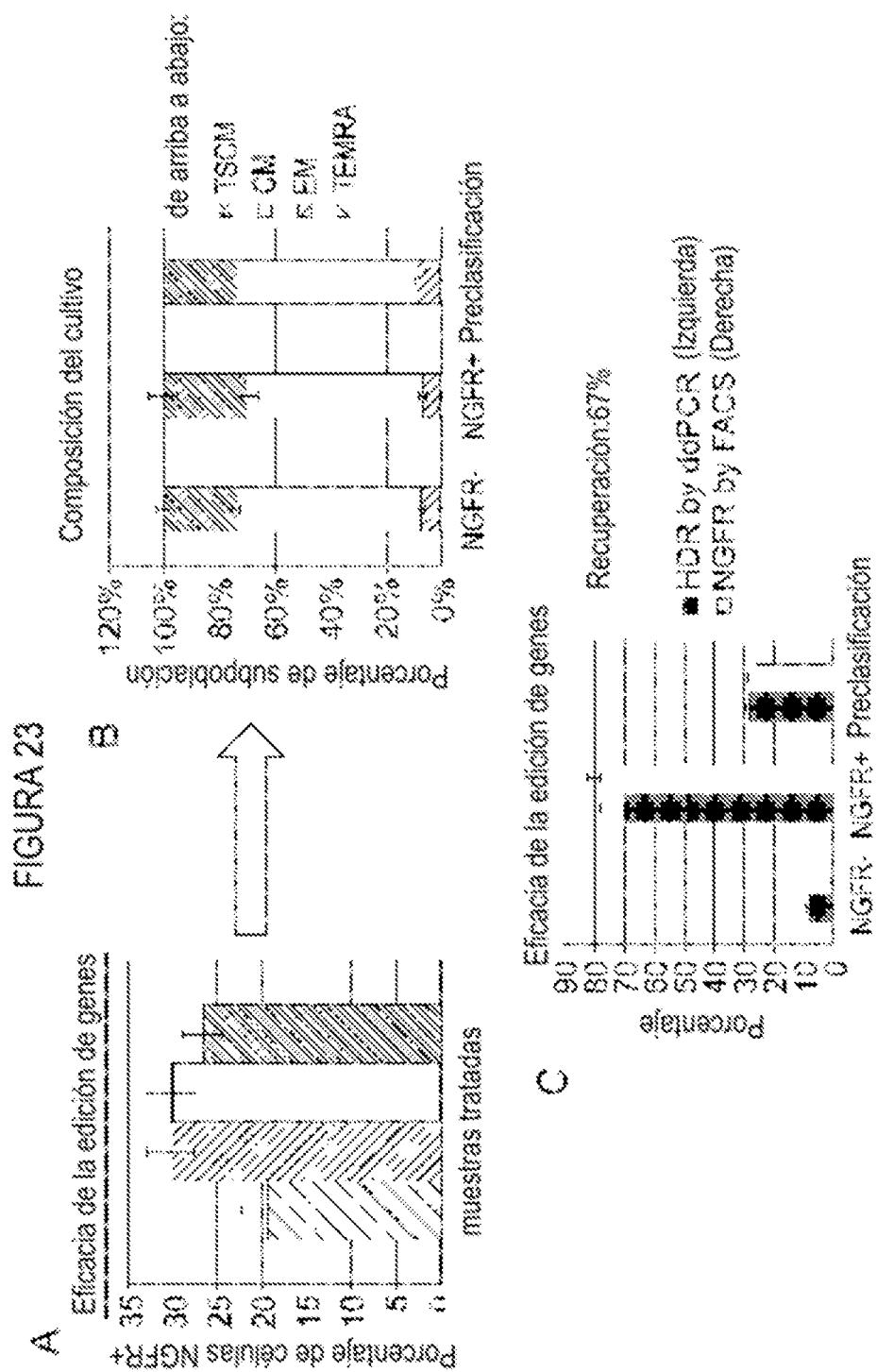


FIGURA 24

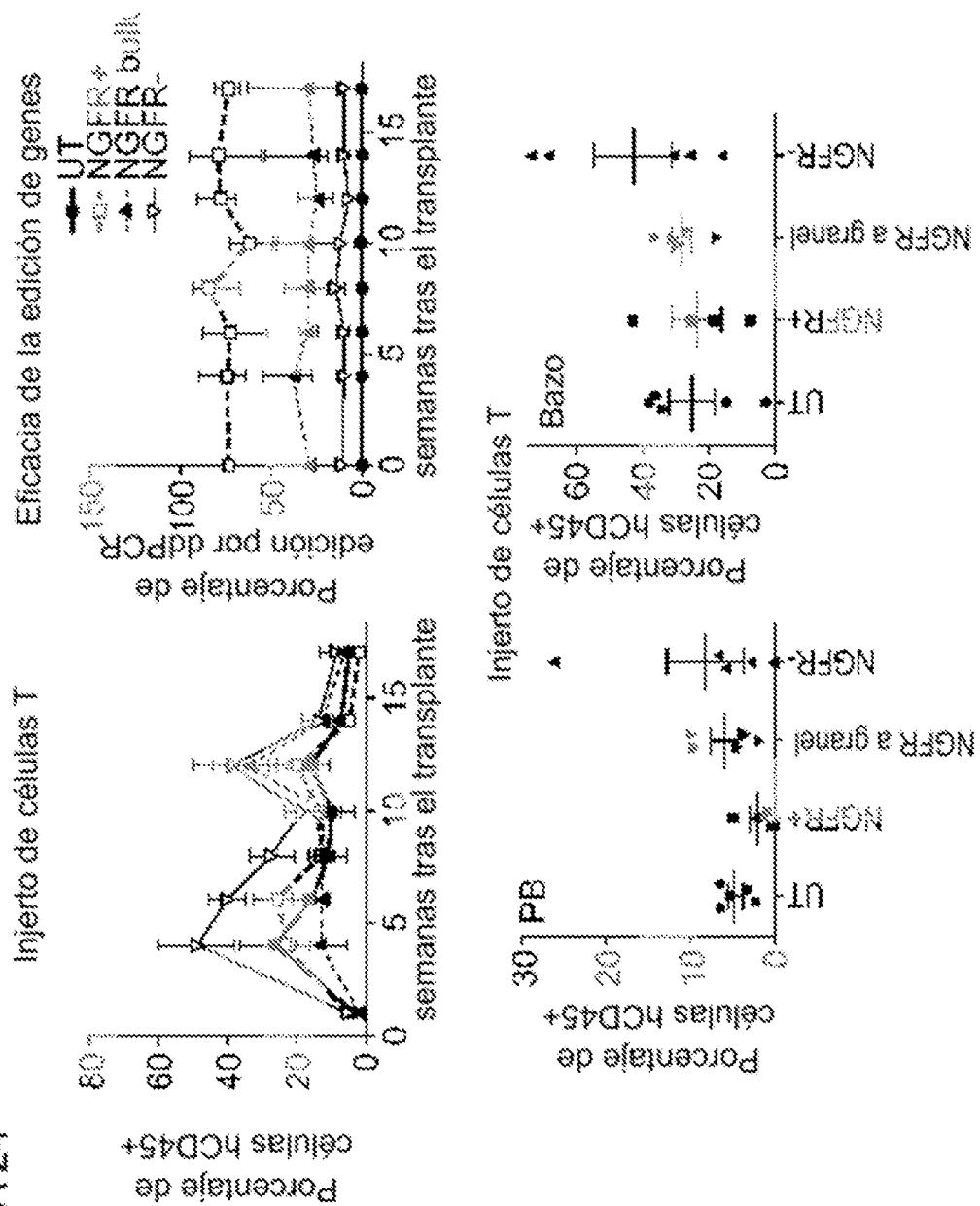


FIGURA 24

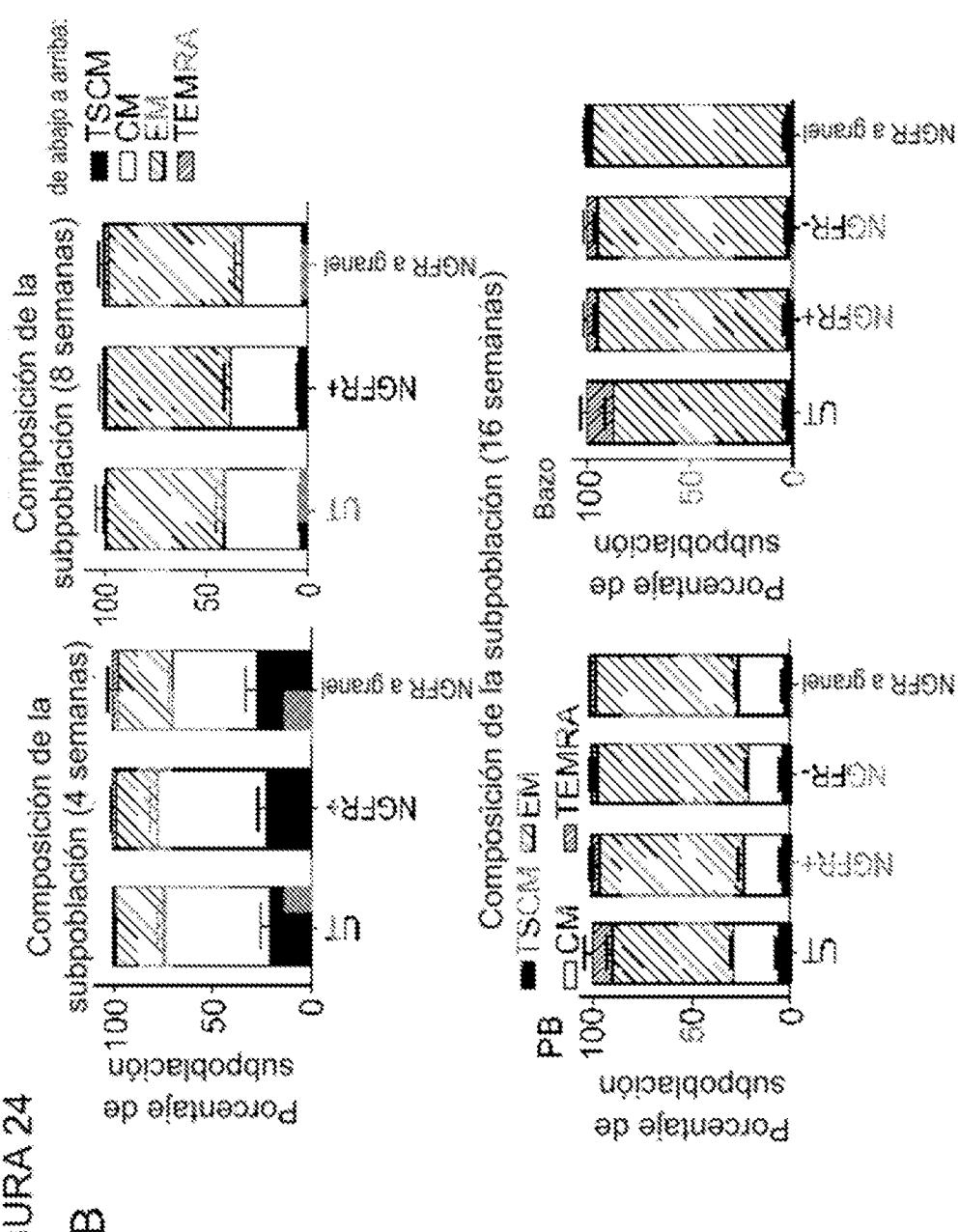


FIGURA 24

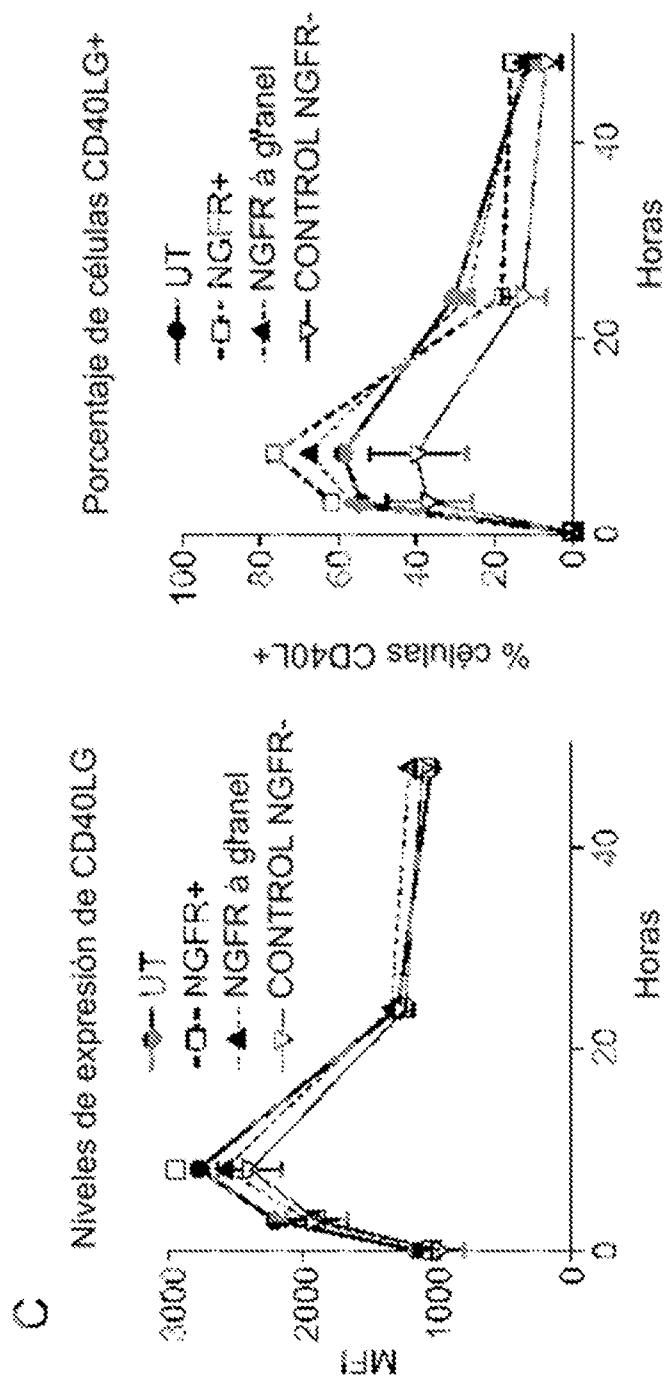
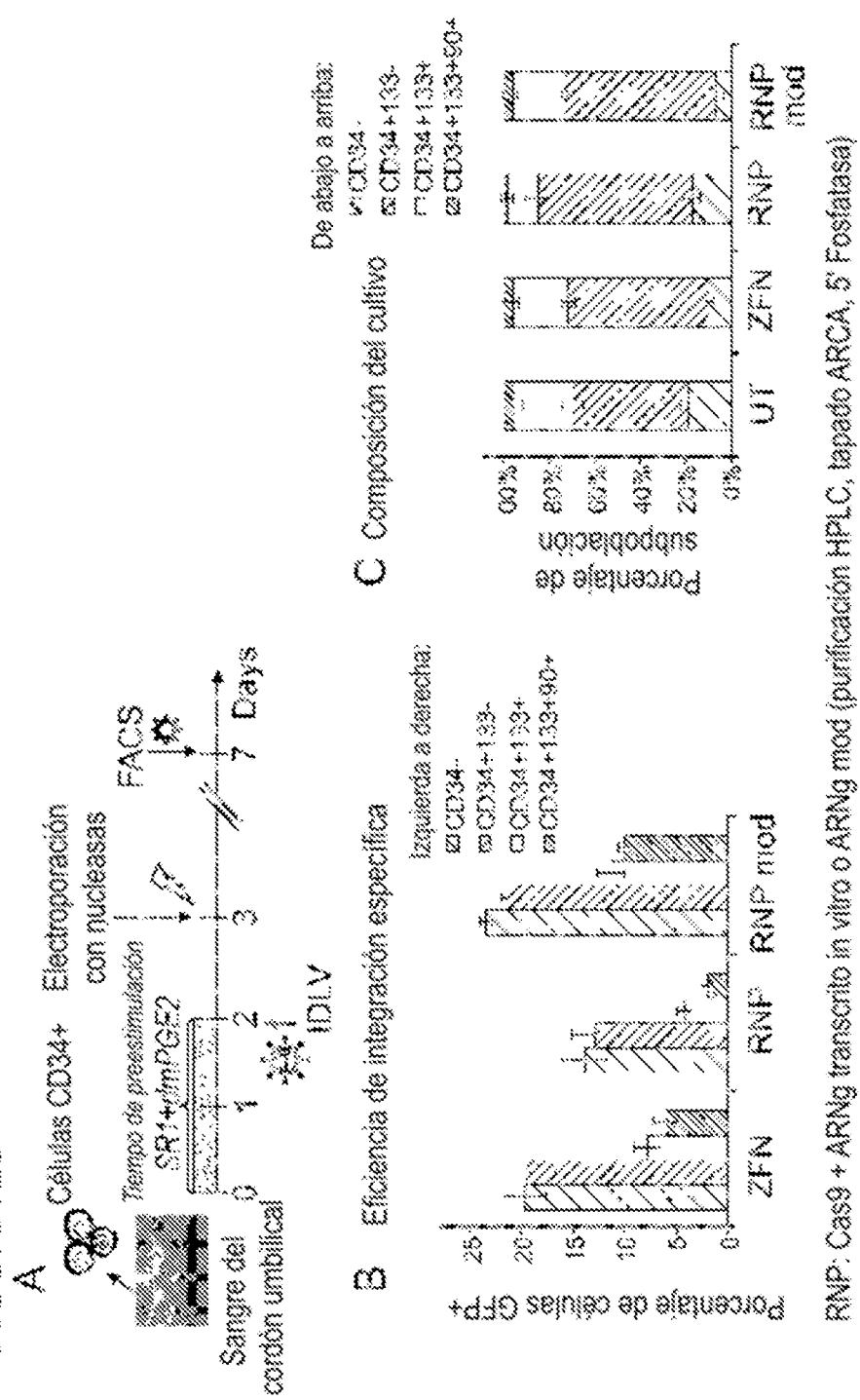


FIGURA 25



RNP: Cas9 + ARNg transcritto *in vitro* o ARNg mod (purificación HPLC, tapado ARCA, S Fosfatasa)

FIGURA 25

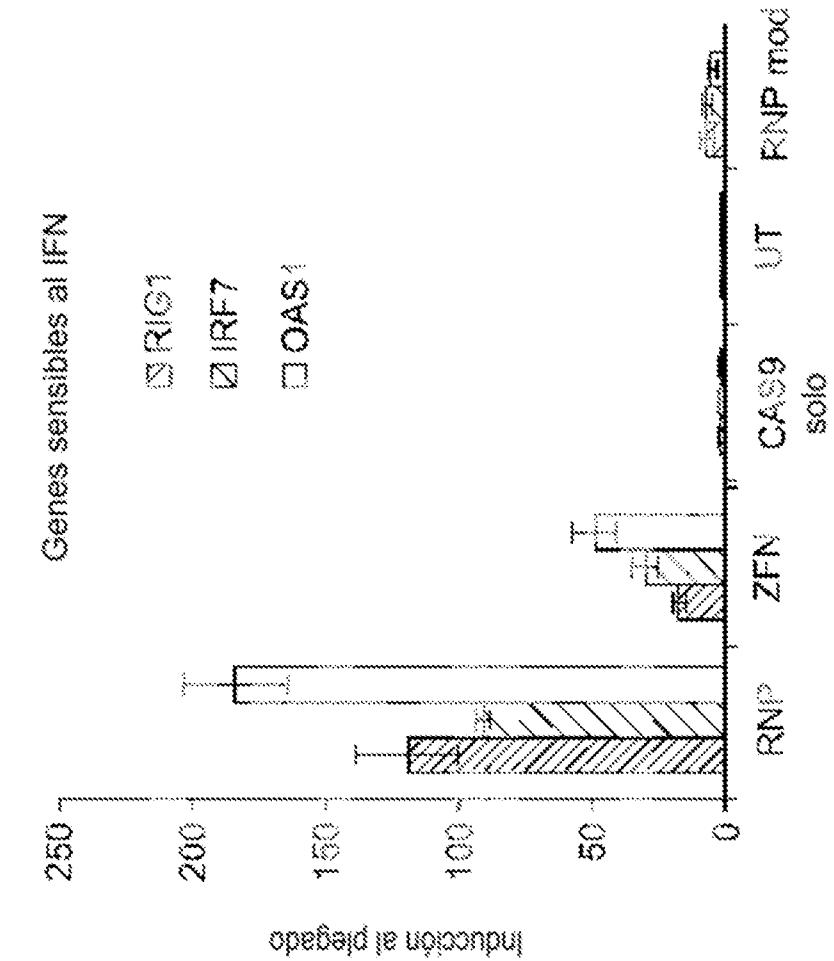


FIGURA 26

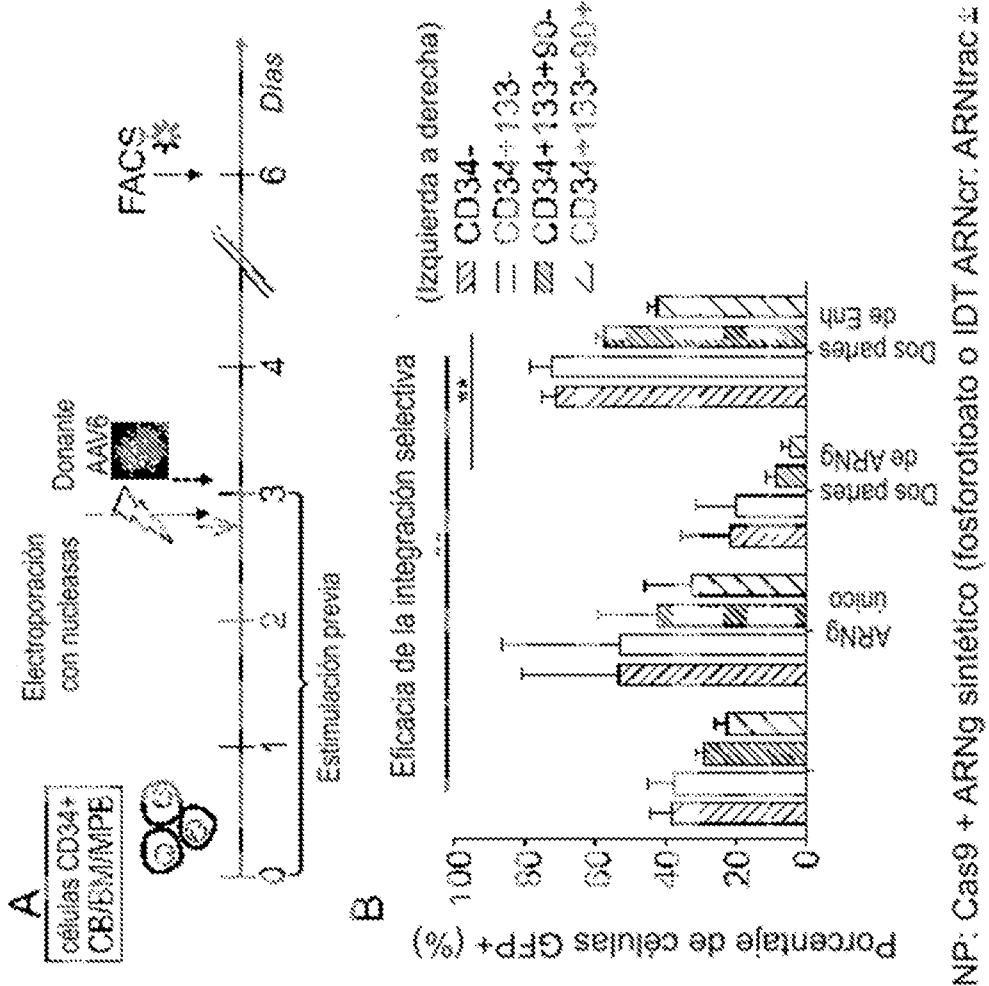


FIGURA 26 Composición del cultivo

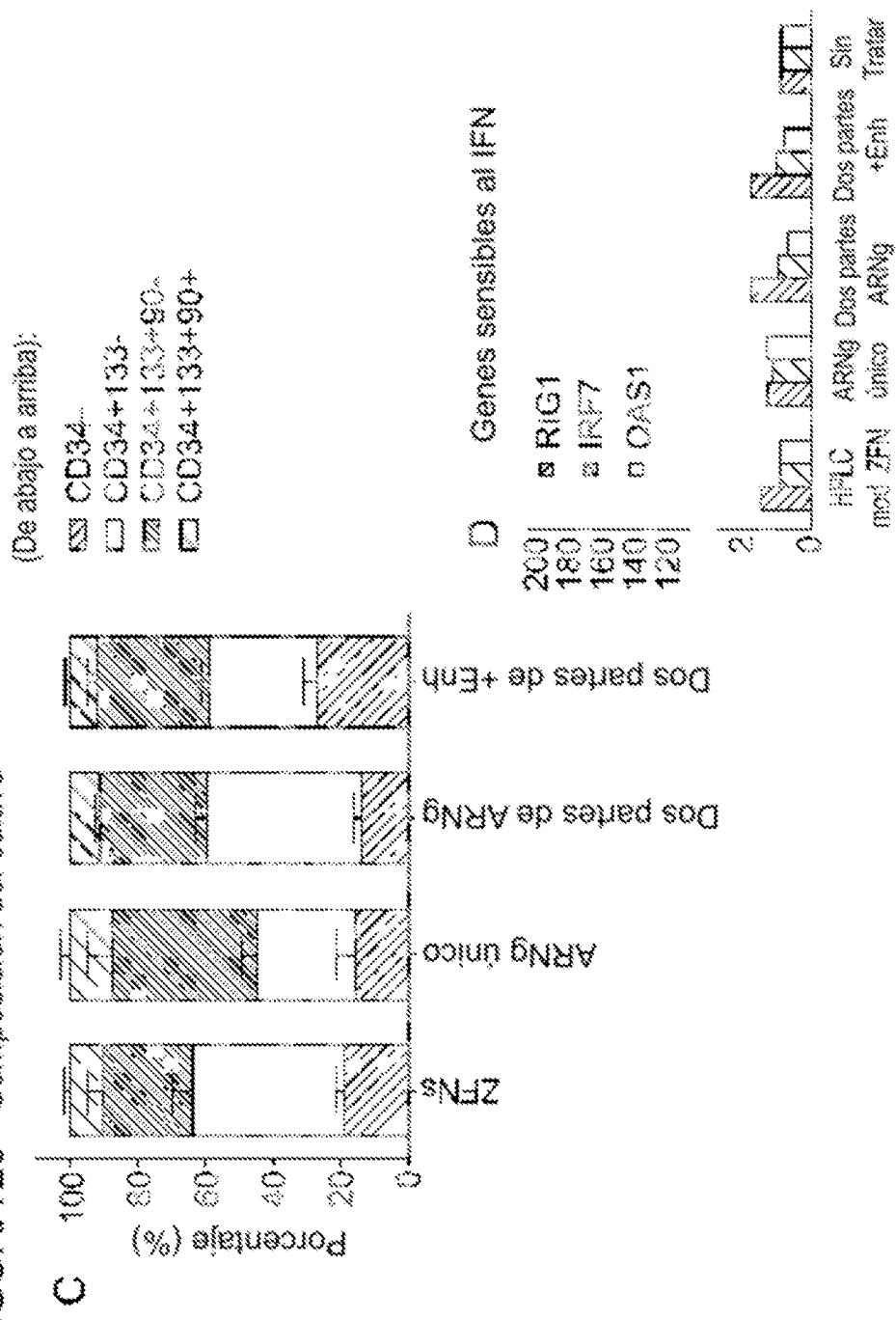


FIGURA 27

