



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



⑪ Número de publicación: **2 983 043**

⑮ Int. Cl.:

A61K 35/15	(2015.01)	C12N 5/073	(2010.01)
A61K 35/17	(2015.01)	C12N 5/0735	(2010.01)
A61K 35/545	(2015.01)	C12N 5/074	(2010.01)
C07K 14/705	(2006.01)	C12N 15/52	(2006.01)
C12N 9/22	(2006.01)		
C12N 15/62	(2006.01)		
C12N 5/0783	(2010.01)		
C12N 5/0789	(2010.01)		
C12N 5/10	(2006.01)		
C12N 9/16	(2006.01)		

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑥ Fecha de presentación y número de la solicitud internacional: **15.12.2016 PCT/US2016/066975**

⑦ Fecha y número de publicación internacional: **22.06.2017 WO17106528**

⑨ Fecha de presentación y número de la solicitud europea: **15.12.2016 E 16876699 (6)**

⑨ Fecha y número de publicación de la concesión europea: **26.06.2024 EP 3389677**

⑮ Título: **Alteración dirigida del receptor de células T**

⑩ Prioridad:

**18.12.2015 US 201562269365 P
10.03.2016 US 201662306500 P**

⑮ Fecha de publicación y mención en BOPI de la traducción de la patente:
21.10.2024

⑮ Titular/es:

**SANGAMO THERAPEUTICS, INC. (100.0%)
7000 Marina Blvd
Brisbane, CA 94005, US**

⑮ Inventor/es:

**LEE, GARY K.;
PASCHON, DAVID y
ZHANG, LEI**

⑮ Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 983 043 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

Alteración dirigida del receptor de células T

5 CAMPO TÉCNICO

[0001] La presente divulgación se encuentra en el campo de la modificación del genoma de células humanas, incluidos linfocitos y células madre.

10 ANTECEDENTES

[0002] La terapia génica tiene un enorme potencial para una nueva era de la terapéutica humana. Estas metodologías permitirán el tratamiento de afecciones que no han sido abordadas por la práctica médica estándar. La terapia génica puede incluir muchas variaciones de las técnicas de edición del genoma, como la alteración o corrección de un locus del gen y la inserción de un transgén expresable que puede controlarse mediante un promotor exógeno específico fusionado al transgén o mediante el promotor endógeno que se encuentra en el sitio de inserción en el genoma.

[0003] La entrega y la inserción del transgén son ejemplos de obstáculos que deben resolverse para cualquier implementación real de esta tecnología. Por ejemplo, aunque hay una variedad de métodos de administración de genes potencialmente disponibles para uso terapéutico, todos implican importantes compensaciones entre seguridad, durabilidad y nivel de expresión. Los métodos que proporcionan el transgén como un episoma (por ejemplo, adenovirus básico (Ad), virus adenoasociado (VAA) y sistemas basados en plásmidos) son generalmente seguros y pueden producir altos niveles de expresión inicial; sin embargo, estos métodos carecen de una replicación episomal robusta, lo que puede limitar la duración de la expresión en tejidos mitóticamente activos. Por el contrario, los métodos de administración que dan como resultado la integración aleatoria del transgén deseado (por ejemplo, la integración de lentivirus (VL)) proporcionan una expresión más duradera, pero, debido a la naturaleza no dirigida de la inserción aleatoria, pueden provocar un crecimiento no regulado en las células receptoras, lo que podría provocar un crecimiento no regulado en las células receptoras, potencialmente llevando a la malignidad mediante la activación de oncogenes en las proximidades del casete transgénico integrado aleatoriamente. Además, aunque la integración del transgén evita la pérdida impulsada por la replicación, no previene el silenciamiento eventual del promotor exógeno fusionado al transgén. Con el tiempo, dicho silenciamiento da como resultado una expresión transgénica reducida para la mayoría de los eventos de inserción no específicos. Además, la integración de un transgén rara vez ocurre en cada célula diana, lo que puede dificultar alcanzar un nivel de expresión suficientemente alto del transgén de interés para lograr el efecto terapéutico deseado.

[0004] En los últimos años, se ha desarrollado una nueva estrategia para la integración transgénica que utiliza escisión con nucleasas específicas de sitio (por ejemplo, nucleasas con dedos de zinc (ZFN), nucleasas de dominio efector tipo activador de la transcripción (TALEN), sistema CRISPR/Cas con un ARNcr/ARNtracr diseñado ("ARN guía única") para guiar la escisión específica, etc.) para sesgar la inserción en un locus genómico elegido. Véanse, por ejemplo, las patentes estadounidenses números 9.255.250; 9.045.763; 9.005.973; 8.956.828; 8.945.868; 8.703.489; 8.586.526; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.067.317; 7.262.054; 7.888.121; 7.972.854; 7.914.796; 7.951.925; 8.110.379; 8.409.861; publicaciones de patentes estadounidenses 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y 20150056705. Además, se están desarrollando nucleasas dirigidas basadas en el sistema Argonaute (por ejemplo, de *T thermophilus*, conocido como 'TtAgo', ver Swarts et al. (2014) *Nature* 507(7491): 258-261), que También puede tener potencial para usos en la edición del genoma y la terapia génica. Este enfoque mediado por nucleasas para la integración de transgenes ofrece la perspectiva de una expresión transgénica mejorada, mayor seguridad y durabilidad de la expresión, en comparación con los enfoques de integración clásicos, ya que permite el posicionamiento exacto del transgén para un riesgo mínimo de silenciamiento génico o activación de oncogenes cercanos.

[0005] El receptor de células T (TCR) es una parte esencial de la activación selectiva de las células T. Con cierta semejanza con un anticuerpo, la parte de reconocimiento de antígeno del TCR suele estar formada por dos cadenas, α y β , que se coensamblan para formar un heterodímero. La semejanza de los anticuerpos radica en la forma en que se combina un único gen que codifica un complejo TCR alfa y beta. Las cadenas TCR alfa (TCR α) y beta (TCR β) están compuestas cada una de dos regiones, una región constante C-terminal y una región variable N-terminal. Los loci genómicos que codifican las cadenas alfa y beta de TCR se parecen a los loci que codifican anticuerpos en que el gen TCR α comprende segmentos V y J, mientras que el locus de cadena comprende segmentos D además de segmentos V y J. Para el locus TCR, hay además dos regiones constantes diferentes que se seleccionan durante el proceso de selección. Durante el desarrollo de las células T, los diversos segmentos se recombinan de manera que cada célula T comprende una porción variable de TCR única en las cadenas alfa y beta, llamada región determinante de la complementariedad (CDR), y el cuerpo tiene un gran repertorio de células T que, debido a sus CDR únicas, son capaces de interactuar con antígenos únicos mostrados por células presentadoras de antígenos. Una vez que se ha producido un reordenamiento genético TCR α o TCR β , la expresión del segundo TCR α o TCR β correspondiente se reprime de modo que cada célula T solo expresa una estructura de TCR única en un proceso llamado "exclusión alélica del receptor de antígeno" (ver Brady et al, (2010) *J Immunol* 185:3801-3808).

5 [0006] Durante la activación de las células T, el TCR interactúa con antígenos presentados como péptidos en el complejo principal de histocompatibilidad (MHC) de una célula presentadora de antígeno. El reconocimiento del complejo antígeno-MHC por parte del TCR conduce a la estimulación de las células T, que a su vez conduce a la diferenciación tanto de las células T colaboradoras (CD4⁺) como de los linfocitos T citotóxicos (CD8⁺) en los linfocitos de memoria y efectores. A continuación, estas células pueden expandirse de manera clonal para dar una subpoblación activada dentro de toda la población de células T capaz de reaccionar ante un antígeno particular.

10 [0007] La terapia celular adoptiva (ACT) es una forma en desarrollo de terapia contra el cáncer basada en la administración de células inmunitarias específicas del tumor a un paciente para que las células administradas ataquen y 15 eliminan el cáncer del paciente. ACT puede implicar el uso de linfocitos infiltrantes de tumores (TIL), que son células T que se aíslan de las masas tumorales del propio paciente y se expanden ex vivo para reinfundirlas en el paciente. Este enfoque ha sido prometedor en el tratamiento del melanoma metastásico, donde en un estudio se observó una tasa de respuesta a largo plazo de >50 % (ver, por ejemplo, Rosenberg et al. (2011) Clin Cane Res 17(13): 4550). Los TIL son una fuente prometedora de células porque son un conjunto mixto de células propias del paciente que tienen receptores 20 de células T (TCR) específicos para los antígenos asociados al tumor (TAA) presentes en el tumor (Wu et al. (2012) Cancer J 18 (2):160). Otros enfoques implican editar células T aisladas de la sangre de un paciente de modo que estén diseñadas para responder a un tumor de alguna manera (Kalos et al. (2011) Sci Transl Med 3(95):95ra73).

25 [0008] Los receptores de antígenos químicos (CAR) son moléculas diseñadas para dirigir las células inmunes a dianas moleculares específicas expresadas en las superficies celulares. En su forma más básica, son receptores introducidos en una célula que acoplan un dominio de especificidad expresado en el exterior de la célula con vías de señalización en el interior de la célula, de modo que cuando el dominio de especificidad interactúa con su objetivo, la célula se activa. A menudo, los CAR se crean a partir de la emulación de los dominios funcionales de los receptores de células T (TCR), donde un dominio específico de antígeno, como un scFv o algún tipo de receptor, se fusiona con el dominio de señalización, como los ITAM y otros dominios coestimuladores. A continuación, estas construcciones se introducen en una célula T ex vivo permitiendo que la célula T se active en presencia de una célula que expresa el antígeno objetivo, lo que resulta en el ataque a la célula objetivo por parte de la célula T activada en un sistema no MHC. 30 manera dependiente (ver Chicaybam et al. (2011) Int Rev Immunol 30:294-311) cuando la célula T se reintroduce en el paciente. Por lo tanto, la terapia celular adoptiva que utiliza células T alteradas ex vivo con un TCR o CAR diseñado es un enfoque clínico muy prometedor para varios tipos de enfermedades. Por ejemplo, los cánceres y sus antígenos a los que se dirige incluyen el linfoma folicular (CD20 o GD2), el neuroblastoma (CD171), el linfoma no Hodgkin (CD19 y CD20), el linfoma (CD19), el glioblastoma (IL13Ra2), la leucemia linfocítica crónica o LLC y leucemia linfocítica aguda o LLA (ambas CD19). También 35 se han desarrollado CAR específicos de virus para atacar las células que albergan virus como el VIH. Por ejemplo, se inició un ensayo clínico utilizando un CAR específico para Gp100 para el tratamiento del VIH (Chicaybam, ibid).

35 [0009] Los ACTR (receptores de células T acoplados a anticuerpos) son componentes de células T diseñados que son capaces de unirse a un anticuerpo suministrado exógenamente. La unión del anticuerpo al componente ACTR prepara a la célula T para interactuar con el antígeno reconocido por el anticuerpo, y cuando se encuentra ese antígeno, la célula T que comprende ACTR se activa para interactuar con el antígeno (consulte la publicación de patente de EE. UU. 40 20150139943).

45 [0010] Sin embargo, uno de los inconvenientes de la terapia celular adoptiva es que la fuente del producto celular debe ser específica del paciente (autólogo) para evitar el posible rechazo de las células trasplantadas. Esto ha llevado a los investigadores a desarrollar métodos de edición de las propias células T del paciente para evitar este rechazo. Por ejemplo, las células T o las células madre hematopoyéticas de un paciente pueden manipularse ex vivo con la adición de un CAR, ACTR y/o receptor de células T (TCR) diseñado y luego tratarse adicionalmente con nucleasas diseñadas para eliminar los inhibidores de los puntos de control de las células T, tales como PD1 y/o CTLA4 (véase la publicación PCT WO2014/059173). Para la aplicación de esta tecnología a una población de pacientes más grande, sería ventajoso desarrollar una población universal de células (alogénicas). Además, la desactivación del TCR dará como resultado que 50 las células sean incapaces de generar una respuesta de enfermedad de injerto contra huésped (EICH) una vez introducidas en un paciente.

55 [0011] Por lo tanto, sigue existiendo una necesidad de métodos y composiciones que puedan usarse para modificar (por ejemplo, eliminar) la expresión de TCR en células T.

55 [0012] El documento US 2015/164954 se refiere a métodos y composiciones para inactivar genes de TCR usando ZFN. El documento WO 2015/136001 se refiere a células T modificadas genéticamente, en las que se inhibe la expresión de beta 2-microglobulina (B2M) y/o transactivador del complejo mayor de histocompatibilidad de clase II (CIITA), y al menos puede inactivarse un gen que codifica un componente del receptor de la célula T. Polic et al. ((2001) Proc Natl Acad Sci EE. UU. 98(15): 8744-9) analiza cómo las células T alfa beta abordan la ablación alfa de TCR inducida.

RESUMEN

65 [0013] La presente invención es tal como se describe en las reivindicaciones adjuntas. Específicamente, la presente invención proporciona una célula aislada en la que la expresión de un gen del receptor de células T (TCR) se modula

mediante modificación dentro del exón c2 del gen del receptor de células T alfa (TCRA), en donde la modificación es una inserción y/o delección dentro de TTGAAA del exón c2 del gen TCRA, y en donde la célula aislada comprende un par de nucleasas con dedos de zinc (ZFN) que escinden el gen TCRA, comprendiendo el par de ZFN primera y ZFN segunda, en donde la ZFN primera comprende un primer dominio de escisión o semidominio de escisión, y una proteína de dedos de zinc (ZFP) que comprende seis dedos de zinc designados y ordenados de F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como sigue: F1: QSSDLR (SEQ ID NO:57); F2: QSGNRTT (SEQ ID NO:58); F3: RSANLAR (SEQ ID NO:59); F4: DRSLALAR (SEQ ID NO:33); F5: RSDVLSE (SEQ ID NO:60); y F6: KHSTRRV (SEQ ID NO:61); y la segunda ZFN comprende un segundo dominio de escisión o semidominio de escisión, y una ZFP que comprende seis dedos de zinc designados y ordenados de F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como sigue: F1: TMHQRVE (SEQ ID NO:62); F2: TSGHLSR (SEQ ID NO:63); F3: RSDHLTQ (SEQ ID NO:64); F4: DSANLSR (SEQ ID NO:65); F5: QSGSLTR (SEQ ID NO:66); y F6: AKWNLDA (SEQ ID NO:67), y además en donde la inserción y/o eliminación en el gen TRAC se realiza mediante el par de ZFN y la expresión del gen TRAC se modula mediante la inserción y/o eliminación. También se proporciona una composición farmacéutica que comprende una célula según la invención.

15 **[0014]** La invención también proporciona un método de diferenciación 2 *in vitro* que comprende: tomar la célula según la invención, en el que la célula es una célula madre; y diferenciar la célula para generar una célula diferenciada que desciende de dicha célula madre.

20 **[0015]** La invención también proporciona una molécula de fusión que comprende un dominio de unión al ADN que se une al exón c1, c2 o c3 de un gen TCR y un dominio regulador transcripcional o un dominio de nucleasa, en donde el dominio de unión al ADN comprende una proteína con dedos de zinc (ZFP) que tiene cinco o seis dedos de zinc designados y ordenados F1 a F5 o F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como se muestra en una sola fila de la siguiente tabla:

25

5

10

15

20

25

30

35

40

45

50

55 [0016] También se proporciona un polinucleótido que codifica la molécula de fusión de la invención.

[0017] La invención también proporciona la célula según la invención o la composición farmacéutica según la invención para su uso en un método para tratar o prevenir un cáncer, comprendiendo el método introducir la célula o la composición farmacéutica a un sujeto con cáncer.

60

[0018] La invención también proporciona la célula según la invención, la composición farmacéutica según la invención, la molécula de fusión según la invención o el polinucleótido según la invención, para uso en el tratamiento de un sujeto con cáncer.

65

F1	F2	F3	F4	F5	F6
DRSNLSE	QKYLRA	DRSALSR	TSGNLTR	YRSSELR	TSGNLTR
(SEQ ID NO:23)	(SEQ ID NO:23)	(SEQ ID NO:24)	(SEQ ID NO:25)	(SEQ ID NO:26)	(SEQ ID NO:25)
QKQVLLN	QNATPTK	QSGHLAR	DRYDINT	RSRSLLR	QSSDLTR
(SEQ ID NO:37)	(SEQ ID NO:28)	(SEQ ID NO:30)	(SEQ ID NO:31)	(SEQ ID NO:31)	(SEQ ID NO:32)
DRSALSR	QSGHLAR	HRSTLQG	QSGDLTR	TSGGITE	NA
(SEQ ID NO:23)	(SEQ ID NO:24)	(SEQ ID NO:25)	(SEQ ID NO:26)	(SEQ ID NO:27)	
QKQVLLN	QNATPTK	QSGHLSP	DRGDLSP	RSRULAP	NA
(SEQ ID NO:38)	(SEQ ID NO:28)	(SEQ ID NO:30)	(SEQ ID NO:40)	(SEQ ID NO:41)	
DRSMLRA	TSGNEST	RSSTPLKT	QSGNLAR	RSDDLSK	TMSNRKR
(SEQ ID NO:42)	(SEQ ID NO:43)	(SEQ ID NO:44)	(SEQ ID NO:34)	(SEQ ID NO:45)	(SEQ ID NO:46)
RSDHLST	DRSHLAR	LKQHLNE	TSGNLTR	RSFGITE	NA
(SEQ ID NO:47)	(SEQ ID NO:48)	(SEQ ID NO:49)	(SEQ ID NO:25)	(SEQ ID NO:50)	
RSQNLRA	TSGNEST	LOQFLAD	QSGNLAP	RREDLIT	TSSMLSP
(SEQ ID NO:42)	(SEQ ID NO:43)	(SEQ ID NO:51)	(SEQ ID NO:34)	(SEQ ID NO:52)	(SEQ ID NO:53)
RSQNLSE	DRSHLAR	LKQHLNE	QSGNLAR	RSNSLSD	NA
(SEQ ID NO:47)	(SEQ ID NO:48)	(SEQ ID NO:49)	(SEQ ID NO:34)	(SEQ ID NO:54)	
RSDHLST	DRSHLAR	LNRSHLQG	QSGNLAR	RSKISLSD	NA
(SEQ ID NO:47)	(SEQ ID NO:48)	(SEQ ID NO:55)	(SEQ ID NO:34)	(SEQ ID NO:56)	
QSSDLSE	QSGNLKT	RSANLAR	RSALAR	RSFVISE	KRSTRV
(SEQ ID NO:57)	(SEQ ID NO:58)	(SEQ ID NO:33)	(SEQ ID NO:33)	(SEQ ID NO:60)	(SEQ ID NO:61)
TMNQFVE	TSGNLSP	RSQDLTQ	DRANLSP	QSGSLTR	AFNNLSE
(SEQ ID NO:62)	(SEQ ID NO:63)	(SEQ ID NO:64)	(SEQ ID NO:65)	(SEQ ID NO:66)	(SEQ ID NO:67)
TMNQFVE	TEGHLSE	RSQSLKT	DSQNLSP	QKATPTT	ENASRTR
(SEQ ID NO:62)	(SEQ ID NO:63)	(SEQ ID NO:68)	(SEQ ID NO:69)	(SEQ ID NO:70)	(SEQ ID NO:72)
RSDSLLR	QSSDLSP	RSQDLSE	RSANRNS	RSENLAR	QKVNLMN
(SEQ ID NO:31)	(SEQ ID NO:73)	(SEQ ID NO:74)	(SEQ ID NO:75)	(SEQ ID NO:76)	(SEQ ID NO:77)
RSDSLSP	QSSDLSP	RSQNLSE	RSANRNS	RSQNLAP	QKVNLR
(SEQ ID NO:31)	(SEQ ID NO:73)	(SEQ ID NO:74)	(SEQ ID NO:75)	(SEQ ID NO:76)	(SEQ ID NO:78)
RSQNLSE	TEGHLSE	RSQSLKT	TSGNEST	RSQNLSE	WHSLRNV
(SEQ ID NO:79)	(SEQ ID NO:37)	(SEQ ID NO:47)	(SEQ ID NO:71)	(SEQ ID NO:74)	(SEQ ID NO:83)
RSDHLST	RSSESLPR	RSQDLST	TSANLSE	RSENLSE	WHSSLPV
(SEQ ID NO:80)	(SEQ ID NO:81)	(SEQ ID NO:47)	(SEQ ID NO:82)	(SEQ ID NO:74)	(SEQ ID NO:83)
RSQNLSE	RSQSLSP	RSQVLSY	QNNRKT	RSQVLSE	SPSSRPT
(SEQ ID NO:484)	(SEQ ID NO:80)	(SEQ ID NO:85)	(SEQ ID NO:86)	(SEQ ID NO:69)	(SEQ ID NO:87)
RSQNLSP	RSQSLSP	RSQSLSV	QANPKT	RSQVLSE	SPSSRPT
(SEQ ID NO:84)	(SEQ ID NO:48)	(SEQ ID NO:88)	(SEQ ID NO:89)	(SEQ ID NO:69)	(SEQ ID NO:87)
TMNQFVE	TEGHLSE	RSQSLST	DRANPLK	QKATPTT	ENASRTR
(SEQ ID NO:62)	(SEQ ID NO:63)	(SEQ ID NO:80)	(SEQ ID NO:91)	(SEQ ID NO:79)	(SEQ ID NO:72)

- [0019]** En el presente documento se describen composiciones y métodos para la inactivación o alteración parcial o completa de un gen de TCR y composiciones y métodos para introducir y expresar a niveles deseados de transgenes de TCR exógenos en linfocitos T, después o simultáneamente con la alteración del TCR endógeno. Además, en el presente documento se proporcionan métodos y composiciones para eliminar (inactivar) o reprimir un gen de TCR para producir 5 células T nulas para TCR, células madre, tejidos u organismos completos, por ejemplo, una célula que no expresa uno o más receptores de células T en su superficie. En determinadas formas de realización, las células o tejidos nulos de TCR son células o tejidos humanos que son ventajosos para su uso en trasplantes. En formas de realización preferidas, las células T nulas de TCR se preparan para su uso en terapia con células T adoptivas.
- [0020]** En la célula aislada (por ejemplo, una célula eucariota tal como una célula de mamífero que incluye una célula linfoide, una célula madre (por ejemplo, iPSC, MSC o HSC), o una célula progenitora) la expresión de un gen de TCR se modula mediante modificación del exón c2 del gen TCR. La modificación está dentro de TTGAAA. La célula puede incluir 10 modificaciones adicionales, por ejemplo, un gen del receptor de células T inactivado, un gen PD1 y/o CTLA4 y/o un transgén, un transgén que codifica un receptor de antígeno químérico (CAR), un transgén que codifica un receptor de células T acoplado a anticuerpo (ACTR) y/o un transgén que codifica un anticuerpo. También se proporcionan 15 composiciones farmacéuticas que comprenden cualquier célula como se describe en el presente documento, así como métodos para usar las células y composiciones farmacéuticas en terapias ex vivo para el tratamiento de un trastorno (por ejemplo, un cáncer) en un sujeto.
- [0021]** Por tanto, en un aspecto, en el presente documento se describen células en las que la expresión de un gen de TCR está modulada (por ejemplo, activada, reprimida o inactivada). Se modula el exón c2 de un gen TCR. La modulación se realiza mediante modificación de la secuencia del gen TCR utilizando una nucleasa que escinde el gen TCR y modifica 20 la secuencia del gen mediante inserciones y/o eliminaciones. En algunas formas de realización, se describen células que comprenden una nucleasa diseñada para provocar la desactivación de un gen TCR. En algunas formas de realización, las células son células T. Se describen además células en las que se modula la expresión de un gen TCR y en las que 25 las células se modifican adicionalmente para que comprendan al menos un transgén exógeno y/o una eliminación adicional de al menos un gen endógeno (por ejemplo, microglobulina beta 2 (B2M) y/o gen de punto de control inmunológico tal como PD1 y/o CTLA4) o combinaciones de los mismos. El transgén exógeno puede integrarse en un gen TCR (por ejemplo, cuando el gen TCR se desactiva) y/o puede integrarse en un gen que no es TCR tal como un gen de 30 puerto seguro. En algunos casos, el transgén exógeno codifica un ACTR y/o un CAR. La construcción transgénica puede insertarse mediante procesos impulsados por HDR o NHEJ. En algunos aspectos, las células con expresión de TCR modulada comprenden al menos un ACTR exógeno y un CAR exógeno. Algunas células que comprenden un modulador de TCR comprenden además una desactivación de uno o más genes inhibidores de puntos de control. En algunas formas 35 de realización, el inhibidor del punto de control es PD1. En otras formas de realización, el inhibidor del punto de control es CTLA4. En aspectos adicionales, la célula modulada por TCR comprende una desactivación de PD1 y una desactivación de CTLA4. En algunas formas de realización, el gen de TCR modulado es un gen que codifica TCR β (TCRB). En algunas formas de realización, esto se logra mediante la escisión dirigida de la región constante de este gen (región constante de TCR β o TRBC). En determinadas formas de realización, el gen de TCR modulado es un gen que codifica TCR α (TCRA). En formas de realización adicionales, la inserción se logra mediante la escisión dirigida de la 40 región constante de un gen TCR, incluida la escisión dirigida de la región constante de un gen TCR α (denominada en el presente documento secuencias "TRAC"). En algunas formas de realización, las células modificadas con el gen TCR se modifican adicionalmente en el gen B2M, los genes HLA-A, -B, -C o el gen TAP, o cualquier combinación de los mismos. En otras formas de realización, también se modifica el regulador para HLA clase II, CTLA.
- [0022]** Las células descritas en el presente documento comprenden una modificación (delección y/o inserción) de un gen TCRA mediante modificación del exón c2. Las células comprenden una modificación (unión, escisión, inserciones y/o eliminaciones) dentro de TTGAAA dentro de un gen TCRA (exón c2, véase la Figura 1B). La modificación de un gen TCRA está dentro de la secuencia TTGAAA de un gen TCR (exón c2 y/o c3), por ejemplo, una modificación de 1 o más pares de bases a dicha secuencia. La modificación genética mediada por nucleasa se produce entre sitios diana 45 emparejados (es decir, cuando se usa un dímero para escindir la diana). Las modificaciones genéticas mediadas por nucleasas pueden incluir inserciones y/o eliminaciones de cualquier número de pares de bases, incluidas inserciones de secuencias no codificantes de cualquier longitud y/o transgenes de cualquier longitud y/o eliminaciones de 1 par de bases hasta más de 1000 kb (o cualquier valor entre ellos, incluyendo, entre otros, 1-100 pares de bases, 1-50 pares de bases, 1-30 pares de bases, 1-20 pares de bases, 1-10 pares de bases o 1-5 pares de bases).
- [0023]** Las células modificadas de la invención pueden ser una célula eucariota, incluyendo un mamífero no humano y 50 una célula humana tal como una célula linfoide (por ejemplo, una célula T), una célula madre/progenitora (por ejemplo, una célula madre pluripotente inducida), célula (iPSC), una célula madre mesenquimal (MSC), o una célula madre hematopoyética (HSC). Las células madre pueden ser pluripotentes (por ejemplo, parcialmente diferenciadas tales como una HSC que es una célula madre mieloide o linfoide pluripotente en otras). En algunas formas de realización, se proporcionan métodos para producir células que tienen un genotipo nulo para la expresión de TCR. Cualquiera de las 55 células madre modificadas descritas en el presente documento (modificadas en el locus TCRA) puede luego diferenciarse para generar una célula diferenciada (*in vivo* o *in vitro*) descendiente de una célula madre como se describe en el presente documento con expresión del gen TCRA modificado.

- [0024]** En otro aspecto, las composiciones (células modificadas) descritas en el presente documento se pueden usar, por ejemplo, en el tratamiento o prevención o mejora de un trastorno. Normalmente, tal uso comprende (a) escindir o regular negativamente un gen TCR endógeno en una célula aislada (por ejemplo, células T o linfocitos) usando una nucleasa (es decir, ZFN) o usando un factor de transcripción diseñado (es decir, ZFN-TF) de manera que el gen TCR está inactivado o modulado negativamente; y (b) introducir la célula en el sujeto, tratando o previniendo así el trastorno. En formas de realización preferidas, el gen que codifica TCR α (TCRA) está inactivado o modulado negativamente. En formas de realización preferidas adicionales, el trastorno es un cáncer o una enfermedad infecciosa. En otras formas de realización preferidas, la inactivación se logra mediante la escisión dirigida de la región constante de este gen (TCR, una región constante, o abreviada como TRAC). En algunas formas de realización, los genes adicionales están modulados (eliminados), por ejemplo, B2M, PD1 y/o CTLA4 y/o uno o más transgenes terapéuticos están presentes en la célula (episomal, integrada aleatoriamente o integrada mediante integración dirigida tal como integración mediada por nucleasas).
- [0025]** Los factores de transcripción y/o las nucleasas se pueden introducir en una célula o en los medios de cultivo circundantes como ARNm, en forma de proteína y/o como una secuencia de ADN que codifica la(s) nucleasa(s). En ciertas formas de realización, la célula aislada introducida en el sujeto comprende, además modificación genómica adicional, por ejemplo, una secuencia exógena integrada (en el gen TCR escindido o un gen diferente, por ejemplo, un gen o locus de puerto seguro) y/o inactivación (ej., mediada por nucleasa) de genes adicionales, por ejemplo, uno o más genes HLA. La secuencia o proteína exógena se puede introducir mediante un vector (por ejemplo, Ad, VAA, VL) o utilizando una técnica como la electroporación. En algunas formas de realización, las proteínas se introducen en la célula mediante compresión celular (véase Kollmannsperger et al. (2016) *Nat Comm* 7, 10372 doi:10.1038/ncomms10372). En algunos aspectos, la composición puede comprender fragmentos celulares aislados y/o células diferenciadas (parcial o totalmente).
- [0026]** En algunos aspectos, las células maduras se pueden usar para terapia celular, por ejemplo, para transferencia de células adoptivas. En otras formas de realización, las células para uso en trasplante de células T contienen otra modificación genética de interés. En un aspecto, las células T contienen un receptor de antígeno químérico (CAR) insertado específico para un marcador de cáncer. En un aspecto adicional, el CAR insertado es específico para el marcador CD19 característico de las células B, incluidas las neoplasias malignas de células B. Dichas células serían útiles en una composición terapéutica para tratar pacientes sin tener que coincidir con HLA y, por lo tanto, podrían usarse como un producto terapéutico "disponible en el mercado" para cualquier paciente que lo necesite.
- [0027]** En otro aspecto, las células T moduladas (modificadas) por TCR contienen una secuencia donante del receptor de células T acoplado a anticuerpo (ACTR) insertada. En algunas formas de realización, la secuencia donante de ACTR se inserta en un gen de TCR para alterar la expresión de ese gen de TCR después de la escisión inducida por nucleasa. En otras formas de realización, la secuencia donante se inserta en un locus de "puerto seguro", tal como los genes VAAS1, HPRT, albúmina y CCR5. En algunas formas de realización, la secuencia ACTR se inserta mediante integración dirigida donde la secuencia donante de ACTR comprende brazos de homología flanqueantes que tienen homología con la secuencia que flanquea el sitio de escisión de la nucleasa diseñada. En algunas formas de realización, la secuencia donante de ACTR comprende, además un promotor y/u otras secuencias reguladoras de la transcripción. En otras formas de realización, la secuencia donante de ACTR carece de un promotor. En algunas formas de realización, el donante de ACTR se inserta en un gen codificador de TCR (TCRB). En algunas formas de realización, la inserción se logra mediante la escisión dirigida de la región constante de este gen (región constante de TCR o TRBC). En formas de realización preferidas, el donante de ACTR se inserta en un gen codificador de TCR (TCRA). En otras formas de realización preferidas, la inserción se logra mediante la escisión dirigida de la región constante de este gen (TCR, una región constante, abreviada TRAC). En algunas formas de realización, el donante se inserta en una secuencia de exón en TCRA, mientras que, en otras, el donante se inserta en una secuencia intrónica en TCRA. En algunas formas de realización, las células moduladas por TCR comprenden además un CAR. En aún otras formas de realización, las células moduladas por TCR se modulan adicionalmente en un gen HLA o un gen inhibidor de punto de control.
- [0028]** También se proporcionan composiciones farmacéuticas que comprenden las células modificadas como se describen en el presente documento (por ejemplo, células T o células madre con gen TCR inactivado), o composiciones farmacéuticas que comprenden una o más de las moléculas de unión al gen TCR (por ejemplo, factores de transcripción diseñados y/o nucleasas) como se describe en el presente documento. En determinadas formas de realización, las composiciones farmacéuticas comprenden además uno o más excipientes farmacéuticamente aceptables. Las células modificadas, las moléculas de unión al gen TCR (o los polinucleótidos que codifican estas moléculas) y/o las composiciones farmacéuticas que comprenden estas células o moléculas se introducen en el sujeto mediante métodos conocidos en la técnica, por ejemplo, mediante infusión intravenosa, infusión en un recipiente específico tal como el arteria hepática, o mediante inyección directa en tejido (por ejemplo, músculo). En algunas formas de realización, el sujeto es un ser humano adulto con una enfermedad o afección que puede tratarse o mejorarse con la composición. En otras formas de realización, el sujeto es un sujeto pediátrico en el que la composición se administra para prevenir, tratar o mejorar la enfermedad o afección (por ejemplo, cáncer, enfermedad de injerto contra huésped, etc).
- [0029]** En algunos aspectos, la composición (células moduladas por TCR que comprenden un ACTR) comprende, además un anticuerpo exógeno. Véase también la solicitud de EE. UU. nº 15/357.772. En algunos aspectos, el anticuerpo es útil para armar una célula T que comprende ACTR para prevenir o tratar una afección. En algunas formas de realización,

el anticuerpo reconoce un antígeno asociado con una célula tumoral o con procesos asociados al cáncer tales como EpCAM, CEA, gpA33, mucinas, TAG-72, CAIX, PSMA, anticuerpos de unión a folato, CD19, EGFR, ERBB2, ERBB3, MET, IGFIR, EPHA3, TRAILR1, TRAILR2, RANKL, FAP, VEGF, VEGFR, integrinas aV β 3 y a5 β 1, CD20, CD30, CD33, CD52, CTLA4 y enascina (Scott et al. (2012) *Nat Rev Cancer* 12:278). En otras formas de realización, el anticuerpo reconoce un antígeno asociado con una enfermedad infecciosa tal como VIH, VHC y similares.

[0030] En otro aspecto, en el presente documento se describen dominios de unión al ADN del gen TCR (ZFP) que se unen a un sitio diana en un gen TCR. El dominio de unión al ADN comprende una ZFP con las regiones de hélice de reconocimiento en el orden que se muestra en una sola fila de la Tabla 1. Estas proteínas de unión al ADN pueden asociarse con dominios reguladores de la transcripción para formar factores de transcripción diseñados que modulen la expresión de TCR. Alternativamente, estas proteínas de unión al ADN se pueden asociar con uno o más dominios de nucleasa para formar nucleasas con dedos de zinc (ZFN) diseñadas que se unen y escinden un gen TCR. Las ZFN pueden unirse a sitios diana en un gen TCR humano. El dominio de unión al ADN del factor de transcripción o nucleasa (ZFP) puede unirse a un sitio diana en un gen TCRA que comprende 9, 10, 11, 12 o más (p. ej., 13, 14, 15, 16, 17, 18, 19, 20 o más) nucleótidos de cualquiera de las SEQ ID NO:8-21 y/o 92-103. Las proteínas de dedos de zinc pueden incluir 1, 2, 3, 4, 5, 6 o más dedos de zinc, teniendo cada dedo de zinc una hélice de reconocimiento que contacta específicamente con un subsitio diana en el gen diana. Las proteínas con dedos de zinc comprenden 5 o 6 dedos (designados F1, F2, F3, F4, F5 y F6 y ordenados F1 a F5 o F6 desde el extremo N al extremo C), como se muestra en la Tabla 1. Las nucleasas descritas en el presente documento (que comprenden un dominio de unión a ZFPDNA) son capaces de realizar modificaciones genéticas dentro de un gen TCRA que comprende cualquiera de las SEQ ID NO:8-21 y/o 92-103, incluidas modificaciones (inserciones y/o eliminaciones) dentro de cualquiera de estas secuencias (SEQ ID NO:8-21 y/o 92-103) y/o modificaciones a las secuencias del gen TCRA que flanquean las secuencias del sitio diana mostradas en SEQ ID NO:8-21 y/o 92-103, por ejemplo, modificaciones dentro del exón c1, c2 y/o c3 de un gen TCR dentro de una o más de las siguientes secuencias: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT y AACCTC.

[0031] Cualquiera de las proteínas descritas en el presente documento puede comprender además un dominio de escisión y/o un medio dominio de escisión (por ejemplo, un medio dominio de escisión *FokI* de tipo salvaje o modificado genéticamente). Por tanto, en cualquiera de las nucleasas (ZFN) descritas en el presente documento, el dominio de nucleasa puede comprender un dominio de nucleasa de tipo salvaje o un medio dominio de nucleasa (por ejemplo, un medio dominio de escisión de *FokI*). En otras formas de realización, las nucleasas (ZFN) comprenden dominios o semidominios de nucleasa diseñados, por ejemplo, semidominios de escisión *FokI* diseñados que forman heterodímeros obligados. Véanse, por ejemplo, las patentes de EE. UU. nº 7.914.796 y 8.034.598.

[0032] En otro aspecto, la divulgación proporciona un polinucleótido que codifica cualquiera de las moléculas de fusión descritas en el presente documento. El polinucleótido puede ser parte de un vector viral, un vector no viral (por ejemplo, plásmido) o estar en forma de ARNm. Cualquiera de los polinucleótidos descritos en el presente documento también puede comprender secuencias (donante, brazos de homología o secuencias de parche) para la inserción dirigida en el gen TCR α y/o el TCR β . Un vector de administración de genes puede comprender cualquiera de los polinucleótidos descritos en el presente documento. En determinadas formas de realización, el vector es un vector adenoviral (p. ej., un vector Ad5/F35) o un vector lentiviral (VL) que incluye vectores lentivirales competentes en integración o de integración defectuosa o un vector adenoasociado (VAA). Por lo tanto, también se proporcionan en el presente documento vectores virales que comprenden una secuencia que codifica una nucleasa (ZFN) y/o una secuencia donante para la integración dirigida en un gen diana. En algunas formas de realización, la secuencia donadora y las secuencias que codifican la nucleasa están en vectores diferentes. En otras formas de realización, las nucleasas se suministran como polipéptidos. En formas de realización preferidas, los polinucleótidos son ARNm. En algunos aspectos, el ARNm puede modificarse químicamente (véase, por ejemplo, Kormann et al. (2011) *Nature Biotechnology* 29(2):154-157). En otros aspectos, el ARNm puede comprender una tapa ARCA (véanse las patentes estadounidenses 7.074.596 y 8.153.773). En algunos aspectos, el ARNm puede comprender una cubierta introducida mediante modificación enzimática. La tapa introducida enzimáticamente puede comprender Cap0, Cap1 o Cap2 (véase, por ejemplo, Smietanski et al. (2014) *Nature Communications* 5:3004). En aspectos adicionales, el ARNm puede estar protegido mediante modificación química. En formas de realización adicionales, el ARNm puede comprender una mezcla de nucleótidos modificados y no modificados (consulte la publicación de patente de EE. UU. 2012-0195936). En aún otras formas de realización, el ARNm puede comprender un elemento WPRE (ver la solicitud de patente de EE. UU. 15/141.333). En algunas formas de realización, el ARNm es bicatenario (véase, por ejemplo, Kariko et al. (2011) *Nucl Acid Res* 39:e142).

[0033] Una célula aislada puede comprender cualquiera de las proteínas, polinucleótidos y/o vectores descritos en el presente documento. En determinadas formas de realización, la célula se selecciona del grupo que consiste en una célula madre/progenitora o una célula T (por ejemplo, célula T CD4 $^{+}$). Se divulga una célula o línea celular que desciende de una célula o línea que comprende cualquiera de las proteínas, polinucleótidos y/o vectores descritos en el presente documento, concretamente una célula o línea celular que desciende (por ejemplo, en cultivo) de una célula en la que se ha cultivado TCR. Inactivado por uno o más ZFN y/o en el que un polinucleótido donante (por ejemplo, ACTR y/o CAR) se ha integrado de forma estable en el genoma de la célula. Por lo tanto, los descendientes de células como se describen en el presente documento pueden no comprender en sí mismos las proteínas, polinucleótidos y/o vectores descritos en el presente documento, pero, en estas células, un gen TCR está inactivado y/o un polinucleótido donante se integra en el genoma y/o se expresa.

5 [0034] En el presente documento se describen métodos para inactivar un gen TCR en una célula mediante la introducción de una o más proteínas, polinucleótidos y/o vectores en la célula como se describe en el presente documento. En cualquiera de los métodos descritos en el presente documento, las nucleasas pueden inducir mutagénesis dirigida, eliminaciones de secuencias de ADN celular y/o facilitar la recombinación dirigida en un locus cromosómico predeterminado. Por tanto, las nucleasas pueden eliminar y/o insertar uno o más nucleótidos desde o dentro del gen diana. El gen TCR puede inactivarse mediante escisión con nucleasa seguida de unión de extremos no homólogos. Se puede reemplazar una secuencia genómica en el gen diana, por ejemplo, usando una nucleasa (o vector que codifica dicha nucleasa) como se describe en el presente documento y una secuencia "donante" que se inserta en el gen después 10 de la escisión dirigida con la nucleasa. La secuencia donante puede estar presente en el vector de nucleasa, presente en un vector separado (por ejemplo, vector de VAA, Ad o VL) o, alternativamente, puede introducirse en la célula usando un mecanismo de administración de ácido nucleico diferente.

15 [0035] Además, cualquiera de los métodos descritos en el presente documento se puede practicar *in vitro*, *in vivo* y/o ex vivo. Los métodos se pueden practicar ex vivo, por ejemplo, para modificar células T, para hacerlas útiles como terapéutica en un entorno alógénico para tratar a un sujeto (por ejemplo, un sujeto con cáncer). Los ejemplos no limitantes de cánceres que pueden tratarse y/o prevenirse incluyen carcinomas de pulmón, cánceres de páncreas, cánceres de hígado, cánceres de huesos, cánceres de mama, cánceres colorrectales, leucemias, cánceres de ovario, linfomas, 20 cánceres de cerebro y similares.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

25 [0036] Las **Figuras 1A y 1B** son una representación del gen TCRA que muestra las ubicaciones de los sitios objetivo de las nucleasas. La **Figura 1A** es una ilustración del procesamiento del gen TCRA desde la forma de la línea germinal hasta la de una célula T madura e indica el objetivo general de las nucleasas. La **Figura 1B** (SEQ ID NO:116 (exón c1), 117 (exón c2) y 118 (exón c3)) muestra las regiones entre los sitios diana en la secuencia de la región constante. La secuencia que se muestra en letras negras mayúsculas es la secuencia de la secuencia de exones indicada, mientras que la secuencia en letras grises minúsculas es la secuencia de intrones contigua.

30 Las **Figuras 2A y 2B** son gráficos que representan el porcentaje de cada sitio modificado en células T tratadas con ZFN específicas para los sitios A, B y D de TCRA (**Figura 2A**) y los sitios E, F y G (**Figura 2B**). Muchos de los pares dieron tasas de modificación del 80% o más.

35 La **Figura 3** representa el porcentaje de células T CD3 negativas después del tratamiento con los pares de ZFN específicos de TCRA analizados mediante análisis FACS.

40 La **Figura 4** es un gráfico que muestra el alto grado de correlación en células T entre los niveles de modificación de la secuencia de TCRA medidos mediante secuenciación de alto rendimiento y la pérdida de expresión de CD3 medida mediante clasificación de células activadas por fluorescencia.

Las **Figuras 5A a 5D** son gráficos que representan el crecimiento de células T después del tratamiento con ZFN específica de TCRA agrupadas según el sitio diana en el gen TCRA.

45 La **Figura 6** muestra los resultados de la doble eliminación de TRAC (TCRA) y B2M y la integración dirigida de un donante en el locus TRAC (TCRA) o B2M.

DESCRIPCIÓN DETALLADA

50 [0037] En el presente documento se describen composiciones y métodos para generar células en las que la expresión de un gen de TCR se modula de manera que las células ya no comprenden un TCR en sus superficies celulares. Las células modificadas de esta manera pueden usarse como terapéutica, por ejemplo, en trasplantes, ya que la falta de un complejo TCR previene o reduce una respuesta inmune basada en HLA. Además, se pueden insertar otros genes de interés en células en las que se ha manipulado el gen TCR y/o se pueden eliminar otros genes de interés.

55 General

60 [0038] La práctica general de los métodos, así como la preparación y el uso de las composiciones descritas en el presente documento emplean, a menos que se indique lo contrario, técnicas convencionales en biología molecular, bioquímica, estructura y análisis de la cromatina, química computacional, cultivo celular, ADN recombinante y campos relacionados como están dentro de la habilidad de la técnica. Estas técnicas se explican completamente en la literatura. Véase, por ejemplo, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edición, Cold Spring Harbor Laboratory Press, 1989 y Tercera edición, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, 1987 y actualizaciones periódicas; la serie METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Tercera edición, Academic Press,

San Diego, 1998; METHODS IN ENZYMOLOGY, vol. 304, "Chromatin" (P.M. Wassarman y A.P. Wolffe, eds.), Academic Press, San Diego, 1999; y METHODS IN MOLECULAR BIOLOGY, vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

5 **Definiciones**

[0039] Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan indistintamente y se refieren a un polímero de desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma monocatenaria o bicitánea. Para los fines de la presente divulgación, estos términos no deben interpretarse como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en los restos de base, azúcar y/o fosfato (por ejemplo, cadenas principales de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo de A formará un par de bases con T.

10 [0040] Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente para referirse a un polímero de residuos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de los correspondientes aminoácidos naturales.

15 [0041] "Unión" se refiere a una interacción no covalente, específica de secuencia, entre macromoléculas (por ejemplo, entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión necesitan ser específicos de una secuencia (p. ej., contactos con residuos de fosfato en el esqueleto del ADN), siempre y cuando la interacción en su conjunto sea específica de una secuencia. Tales interacciones se caracterizan generalmente por una constante de disociación (K_d) de 10^{-6} M⁻¹ o inferior. "Afinidad" se refiere a la fuerza de unión: una mayor afinidad de unión se correlaciona con una K_d más baja. "Unión no específica" se refiere a interacciones no covalentes que se producen entre cualquier molécula de interés (por ejemplo, una nucleasa diseñada) y una macromolécula (por ejemplo, ADN) que no dependen de la secuencia diana.

20 [0042] Una "molécula de unión a ADN" es una molécula que puede unirse al ADN. Dicha molécula de unión a ADN puede ser un polipéptido, un dominio de una proteína, un dominio dentro de una proteína más grande o un polinucleótido. 30 En algunas formas de realización, el polinucleótido es ADN, mientras que, en otras formas de realización, el polinucleótido es ARN. En algunas formas de realización, la molécula de unión a ADN es un dominio proteico de una nucleasa (por ejemplo, el dominio FokI), mientras que, en otras formas de realización, la molécula de unión a ADN es un componente de ARN guía de una nucleasa guiada por ARN (por ejemplo, Cas9 o Cfp1).

25 [0043] Una "proteína de unión" es una proteína que es capaz de unirse de forma no covalente a otra molécula. Una proteína de unión puede unirse, por ejemplo, a una molécula de ADN (una proteína de unión a ADN), una molécula de ARN (una proteína de unión a ARN) y/o una molécula de proteína (una proteína de unión a proteínas). En el caso de una proteína de unión a proteínas, puede unirse a sí misma (para formar homodímeros, homotímeros, etc.) y/o puede unirse a una o más moléculas de una proteína o proteínas diferentes. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas con dedos de zinc tienen actividad de unión al ADN, unión al ARN y unión a proteínas.

35 [0044] Una "proteína de unión a ADN con dedos de zinc" (o dominio de unión) es una proteína, o un dominio dentro de una proteína más grande, que se une al ADN de una manera específica de secuencia a través de uno o más dedos de zinc, que son regiones de secuencia de aminoácidos dentro del dominio de unión cuya estructura se estabiliza mediante la coordinación de un ion zinc. El término proteína de unión al ADN con dedos de zinc a menudo se abrevia como proteína con dedos de zinc o ZFP.

40 [0045] Un "dominio de unión a ADN de TALE" o "TALE" es un polipéptido que comprende uno o más dominios/unidades de repetición de TALE. Los dominios repetidos, cada uno de los cuales comprende un residuo variable repetido (RVR), participan en la unión de TALE a su secuencia de ADN diana afín. Una única "unidad de repetición" (también denominada "repetición") tiene normalmente una longitud de 33 a 35 aminoácidos y muestra al menos cierta homología de secuencia con otras secuencias repetidas de TALE dentro de una proteína TALE de origen natural. Las proteínas TALE pueden diseñarse para unirse a un sitio objetivo utilizando RVR canónicos o no canónicos dentro de las unidades repetidas. 45 Véanse, por ejemplo, las patentes de EE. UU. n° 8.586.526 y 9.458.205. Los dominios de unión a ADN de dedos de zinc y TALE pueden "diseñarse" para unirse a una secuencia de nucleótidos predeterminada, por ejemplo, mediante ingeniería (alterando uno o más aminoácidos) de la región de hélice de reconocimiento de una proteína de dedos de zinc de origen natural o mediante ingeniería de los aminoácidos implicados en la unión del ADN (el residuo variable repetido o región RVR). Por lo tanto, las proteínas con dedos de zinc modificadas genéticamente o proteínas TALE son proteínas que no se producen de forma natural. Ejemplos no limitantes de métodos para diseñar proteínas con dedos de zinc y TALE son el diseño y la selección. Una proteína diseñada es una proteína que no se encuentra en la naturaleza y cuya diseño/composición resulta principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y algoritmos computarizados para procesar información en una base de datos que almacena información de diseños ZFP o TALE existentes (RVR canónicos y no canónicos) y datos vinculantes. Véanse,

por ejemplo, las patentes estadounidenses números 9.458.205; 8.586.526; 6.140.081; 6.453.242; y 6.534.261; véase también WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

5 **[0046]** Una proteína de dedos de zinc "seleccionada", proteína TALE o sistema CRISPR/Cas no se encuentra en la naturaleza y cuya producción resulta principalmente de un proceso empírico tal como presentación en fagos, trampa de interacción o selección de híbridos. Véanse, por ejemplo, los documentos EE. UU. 5.789.538; EE. UU. 5.925.523; EE. UU. 6.007.988; EE. UU. 6.013.453; EE. UU. 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197 y WO 02/099084.

10 **[0047]** "TtAgo" es una proteína Argonauta procariótica que se cree que está implicada en el silenciamiento de genes. TtAgo se deriva de la bacteria *Thermus thermophilus*. Véase, por ejemplo, Swarts et al, ibid, G. Sheng et al., (2013) Proc. Natl. Acad. Sci. EE. UU. 111, 652). Un "sistema TtAgo" son todos los componentes necesarios, incluidos, por ejemplo, ADN guía para la escisión mediante una enzima TtAgo.

15 **[0048]** "Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los fines de esta divulgación, "recombinación homóloga (RH)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble hebra en células mediante mecanismos de reparación dirigidos por homología. Este proceso requiere homología de secuencia de nucleótidos, utiliza una molécula "donante" para reparar una molécula "objetivo" (es decir, la que experimentó la rotura de la doble cadena) y se conoce como "conversión genética sin cruce" o "conversión de genes del tracto gastrointestinal", porque conduce a la transferencia de información genética del donante al objetivo. Sin desear quedar vinculado a ninguna teoría particular, dicha transferencia puede implicar la corrección de errores de coincidencia del ADN heterodúplex que se forma entre la diana rota y el donante, y/o el "recocido de hebras dependiente de la síntesis", en el que el donante se utiliza para resintetizar información genética que pasarán a formar parte del objetivo y/o procesos relacionados. Dicha RH especializada a menudo da como resultado una alteración de la secuencia de la molécula diana de modo que parte o toda la secuencia del polinucleótido donante se incorpora al polinucleótido diana.

20 **[0049]** En los métodos de la divulgación, una o más nucleasas objetivo como se describe en el presente documento crean una rotura bicanalicular (RBC) en la secuencia objetivo (por ejemplo, cromatina celular) en un sitio predeterminado (por ejemplo, un gen o locus de interés), y se puede introducir en la célula un polinucleótido "donante", que tenga homología con la secuencia de nucleótidos en la región de la rotura. Se ha demostrado que la presencia del RBC facilita la integración de la secuencia del donante. Opcionalmente, la construcción tiene homología con la secuencia de nucleótidos en la región de la ruptura. La secuencia donante puede integrarse físicamente o, alternativamente, el polinucleótido donante se usa como plantilla para reparar la rotura mediante recombinación homóloga, lo que da como resultado la introducción de toda o parte de la secuencia de nucleótidos como en el donante en la cromatina celular. Por tanto, se puede alterar una primera secuencia en la cromatina celular y, en determinadas formas de realización, se puede convertir en una secuencia presente en un polinucleótido donante. Por tanto, puede entenderse que el uso de los términos "reemplazar" o "reemplazo" representa la sustitución de una secuencia de nucleótidos por otra (es decir, sustitución de una secuencia en el sentido informativo), y no requiere necesariamente la sustitución física o química de un polinucleótido por otro.

30 **[0050]** En cualquiera de los métodos descritos en el presente documento, se pueden usar pares adicionales de proteínas con dedos de zinc para una escisión bicanalicular adicional de sitios diana adicionales dentro de la célula.

45 **[0051]** En ciertas formas de realización de métodos para la recombinación dirigida y/o el reemplazo y/o la alteración de una secuencia en una región de interés en la cromatina celular, una secuencia cromosómica se altera mediante recombinación homóloga con una secuencia de nucleótidos "donante" exógena. Esta recombinación homóloga es estimulada por la presencia de una rotura bicanalicular en la cromatina celular, si están presentes secuencias homólogas a la región de la rotura.

50 **[0052]** En cualquiera de los métodos descritos en el presente documento, la primera secuencia de nucleótidos (la "secuencia donante") puede contener secuencias que son homólogas, pero no idénticas, a secuencias genómicas en la región de interés, estimulando así la recombinación homóloga para insertar una secuencia no idéntica en la región de interés. Por lo tanto, en ciertas formas de realización, porciones de la secuencia donante que son homólogas a secuencias en la región de interés exhiben entre aproximadamente 80 a 99 % (o cualquier número entero entre ellas) de identidad de secuencia con la secuencia genómica que se reemplaza. En otras formas de realización, la homología entre el donante y la secuencia genómica es superior al 99%, por ejemplo, si solo 1 nucleótido difiere entre las secuencias donante y genómica de más de 100 pares de bases contiguas. En ciertos casos, una porción no homóloga de la secuencia donante puede contener secuencias no presentes en la región de interés, de modo que se introducen nuevas secuencias en la región de interés. En estos casos, la secuencia no homóloga generalmente está flanqueada por secuencias de 50 a 1000 pares de bases (o cualquier valor integral entre ellas) o cualquier número de pares de bases mayor que 1000, que son homólogas o idénticas a secuencias en la región de interés. En otras formas de realización, la secuencia donante no es homóloga a la primera secuencia y se inserta en el genoma mediante mecanismos de recombinación no homólogos.

5

[0053] Cualquiera de los métodos descritos en el presente documento se puede utilizar para la inactivación parcial o completa de una o más secuencias diana en una célula mediante la integración dirigida de la secuencia donante que altera la expresión del gen o genes de interés. También se proporcionan líneas celulares con genes parcial o completamente inactivados.

10

[0054] Además, los métodos de integración dirigida como se describen en el presente documento también se pueden usar para integrar una o más secuencias exógenas. La secuencia de ácido nucleico exógena puede comprender, por ejemplo, uno o más genes o moléculas de ADNc, o cualquier tipo de secuencia codificante o no codificante, así como uno o más elementos de control (por ejemplo, promotores). Además, la secuencia de ácido nucleico exógena puede producir una o más moléculas de ARN (por ejemplo, ARN en horquilla pequeña (ARNhp), ARN inhibidores (ARNi), microARN (miARN), etc.).

15

[0055] "Escisión" se refiere a la rotura del esqueleto covalente de una molécula de ADN. La escisión puede iniciarse mediante una variedad de métodos que incluyen, entre otros, hidrólisis enzimática o química de un enlace fosfodiéster. Son posibles tanto la escisión 1 monocatenaria como la escisión bicatenaria, y la escisión bicatenaria puede ocurrir como resultado de dos eventos distintos de escisión monocatenaria. La escisión del ADN puede dar lugar a la producción de extremos romos o escalonados. En determinadas formas de realización, los polipéptidos de fusión se utilizan para la escisión dirigida del ADN bicatenario.

20

[0056] Un "semidominio de escisión" es una secuencia polipeptídica que, junto con un segundo polipéptido (ya sea idéntico o diferente) forma un complejo que tiene actividad de escisión (preferiblemente actividad de escisión de doble hebra). Los términos "primer y segundo semidominio de escisión"; "Los semidominios de escisión + y -" y los "semidominios de escisión derecho e izquierdo" se usan indistintamente para referirse a pares de semidominios de escisión que se dimerizan.

25

[0057] Un "medio dominio de escisión diseñado" es un medio dominio de escisión que se ha modificado para formar heterodímeros obligados con otro medio dominio de escisión (por ejemplo, otro medio dominio de escisión diseñado). Véanse también las patentes estadounidenses números 7.888.121; 7.914.796; 8.034.598; 8.623.618 y publicación de patente estadounidense n.º 2011/0201055.

30

[0058] El término "secuencia" se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser ADN o ARN; puede ser lineal, circular o ramificado y puede ser monocatenario o bicatenario. El término "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede tener cualquier longitud, por ejemplo, entre 2 y 10.000 nucleótidos de longitud (o cualquier valor entero entre ellos o superior), preferiblemente entre aproximadamente 100 y 1.000 nucleótidos de longitud (o cualquier valor entero entre ellos), más preferiblemente entre aproximadamente 200 y 500 nucleótidos de longitud.

35

[0059] "Cromatina" es la estructura de nucleoproteína que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteínas, incluidas histonas y proteínas cromosómicas no histonas. La mayor parte de la cromatina celular eucariota existe en forma de nucleosomas, en los que un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociados con un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y el ADN conector (de longitud variable según el organismo) se extiende entre los núcleos de los nucleosomas. Una molécula de histona H1 está generalmente asociada con el ADN conector. Para los fines de la presente divulgación, el término "cromatina" pretende abarcar todos los tipos de nucleoproteína celular, tanto procarióticas como eucariotas. La cromatina celular incluye tanto la cromatina cromosómica como la episomal.

45

[0060] Un "cromosoma" es un complejo de cromatina que comprende todo o una parte del genoma de una célula. El genoma de una célula suele caracterizarse por su cariotipo, que es el conjunto de todos los cromosomas que componen el genoma de la célula. El genoma de una célula puede comprender uno o más cromosomas.

50

[0061] Un "episoma" es un ácido nucleico replicante, complejo de nucleoproteína u otra estructura que comprende un ácido nucleico que no forma parte del cariotipo cromosómico de una célula. Ejemplos de episomas incluyen plásmidos y ciertos genomas virales.

55

[0062] Un "sitio objetivo" o "secuencia objetivo" es una secuencia de ácido nucleico que define una porción de un ácido nucleico a la que se unirá una molécula de unión, siempre que existan condiciones suficientes para la unión. Por ejemplo, la secuencia 5' GAATTC 3' es un sitio diana para la endonucleasa de restricción Eco RI.

60

[0063] Una molécula "exógena" es una molécula que normalmente no está presente en una célula, pero que puede introducirse en una célula mediante uno o más métodos genéticos, bioquímicos u otros. La "presencia normal en la célula" se determina con respecto a la etapa de desarrollo particular y las condiciones ambientales de la célula. Así, por ejemplo, una molécula que está presente sólo durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular adulta. De manera similar, una molécula inducida por un choque térmico es una molécula exógena con respecto a una célula que no sufre un choque térmico. Una molécula exógena puede comprender, por ejemplo, una

versión funcional de una molécula endógena que funciona mal o una versión que funciona mal de una molécula endógena que funciona normalmente.

5 [0064] Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como se genera mediante un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, carbohidrato, lípido, glicoproteína, lipoproteína, polisacárido, cualquier modificado derivado de las moléculas anteriores, o cualquier complejo que comprenda una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser monocatenarios o bicatenarios; puede ser lineal, ramificado o circular; y puede tener cualquier longitud. Véanse, por ejemplo, las patentes estadounidenses números 8.703.489 y 9.255.259. Los ácidos nucleicos incluyen aquellos capaces de formar dúplex, así como ácidos nucleicos que forman triplex. Véanse, por ejemplo, las patentes estadounidenses números 5.176.996 y 5.422.251. Las proteínas incluyen, entre otras, proteínas de unión a ADN, factores de transcripción, factores de remodelación de la cromatina, proteínas de unión a ADN metiladas, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, quinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.

10 [0065] Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína o ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma viral infectante, un plásmido o episoma introducido en una célula, o un cromosoma que normalmente no está presente en la célula. Los expertos en la técnica conocen métodos para la introducción de moléculas exógenas en células e incluyen, entre otros, transferencia mediada por lípidos (es decir, liposomas, incluidos lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación de fosfato cálcico, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores virales. Una molécula exógena también puede ser el mismo tipo de molécula que una molécula endógena pero derivada de una especie diferente a la de la célula. Por ejemplo, se puede introducir una secuencia de ácido nucleico humano en una línea celular derivada originalmente de un ratón o hámster.

15 [0066] Por el contrario, una molécula "endógena" es aquella que normalmente está presente en una célula particular en una etapa de desarrollo particular bajo condiciones ambientales particulares. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, cloroplasto u otro orgánulo, o un ácido nucleico episomal de origen natural. Las moléculas endógenas adicionales pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

20 [0067] Una molécula de " fusión" es una molécula en la que dos o más moléculas de subunidad están unidas, preferiblemente de forma covalente. Las moléculas de la subunidad pueden ser del mismo tipo químico de molécula o pueden ser de diferentes tipos químicos de molécula. Los ejemplos del primer tipo de molécula de fusión incluyen, entre otros, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión a ADN ZFP o TALE y uno o más dominios de activación) y ácidos nucleicos de fusión (por ejemplo, un ácido que codifica la proteína de fusión descrita anteriormente). Los ejemplos del segundo tipo de molécula de fusión incluyen, entre otros, una fusión entre un ácido nucleico formador de triplex y un polipéptido, y una fusión entre un aglutinante de surco menor y un ácido nucleico. El término también incluye sistemas en los que un componente polinucleotídico se asocia con un componente polipeptídico para formar una molécula funcional (por ejemplo, un sistema CRISPR/Cas en el que un único ARN guía se asocia con un dominio funcional para modular la expresión génica).

25 [0068] La expresión de una proteína de fusión en una célula puede resultar del suministro de la proteína de fusión a la célula o mediante el suministro de un polinucleótido que codifica la proteína de fusión a una célula, en donde el polinucleótido se transcribe y el transcripto se traduce, para generar la proteína de fusión. El trans-empalme, la escisión de polipéptidos y la ligación de polipéptidos también pueden participar en la expresión de una proteína en una célula. Los métodos para la administración de polinucleótidos y polipéptidos a las células se presentan en otra parte de esta divulgación.

30 [0069] Un "gen", para los fines de la presente divulgación, incluye una región de ADN que codifica un producto génico (ver más adelante), así como todas las regiones de ADN que regulan la producción del producto génico, ya sean o no tales secuencias reguladoras son adyacentes a secuencias codificantes y/o transcritas. Por consiguiente, un gen incluye, entre otros, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión a ribosomas y sitios de entrada internos a ribosomas, potenciadores, silenciadores, aislantes, elementos límite, orígenes de replicación, sitios de unión a matrices y regiones de locus de control.

35 [0070] Un locus de "puerto seguro" es un locus dentro del genoma en el que se puede insertar un gen sin ningún efecto perjudicial sobre la célula huésped. Lo más beneficioso es un locus de puerto seguro en el que la expresión de la secuencia del gen insertado no se ve perturbada por ninguna expresión de lectura directa de los genes vecinos. Los ejemplos no limitantes de loci de puerto seguro a los que se dirigen las nucleasas incluyen CCR5, CCR5, HPRT, VAAS1, Rosa y albúmina. Véanse, por ejemplo, las patentes estadounidenses números 8.771.985; 8.110.379; 7.951.925; Publicaciones de EE. UU. N°s 20100218264; 20110265198; 20130137104; 20130122591; 20130177983; 20130177960; 20150056705 y 20150159172).

40 [0071] "Expresión génica" se refiere a la conversión de la información contenida en un gen, en un producto génico. Un producto génico puede ser el producto transcripcional directo de un gen (por ejemplo, ARNm, ARNt, ARNr, ARN

antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida mediante traducción de un ARNm. Los productos genéticos también incluyen ARN que se modifican mediante procesos tales como protección, poliadenilación, metilación y edición, y proteínas modificadas mediante, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ribosilación de ADP, miristilación y glicosilación. "Modulación" o "modificación" de la expresión genética se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, entre otras, activación y represión genética, incluida la modificación del gen mediante la unión de una molécula exógena (por ejemplo, factor de transcripción diseñado). La modulación también se puede lograr mediante la modificación de la secuencia genética mediante edición del genoma (por ejemplo, escisión, alteración, inactivación, mutación aleatoria). La inactivación genética se refiere a cualquier reducción en la expresión genética en comparación con una célula que no ha sido modificada como se describe en el presente documento. Por tanto, la inactivación genética puede ser parcial o completa.

5 [0072] Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante dentro o adyacente a un gen, en la que es deseable unirse a una molécula exógena. La unión puede tener como objetivo la escisión dirigida del ADN y/o la recombinación dirigida. Una región de interés puede estar presente en 15 un cromosoma, un episoma, un genoma organelar (por ejemplo, mitocondrial, cloroplasto) o un genoma viral infectante, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias líder, secuencias finales o intrans, o dentro de regiones no transcritas, ya sea cadena arriba o cadena abajo de la región codificante. Una región de interés puede ser tan pequeña 20 como un único par de nucleótidos o hasta 2000 pares de nucleótidos de longitud, o cualquier valor integral de pares de nucleótidos.

[0073] Las células "eucariotas" incluyen, entre otras, células fúngicas (tales como levaduras), células vegetales, células animales, células de mamíferos y células humanas (por ejemplo, células T).

25 [0074] Los términos "enlace operativo" y "enlazado operativamente" (o "enlazado operativamente") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en la que los componentes están dispuestos de manera que ambos Los componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora transcripcional, tal como un promotor, está unida operativamente a una secuencia codificante si la secuencia reguladora transcripcional controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores transcripcionales. Una secuencia reguladora transcripcional generalmente está unida operativamente en cis con una secuencia codificante, pero no es necesario que esté directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora transcripcional que está operativamente unida a una secuencia codificante, aunque no sean contiguas.

35 [0075] Con respecto a los polipéptidos de fusión, el término "unido operativamente" puede referirse al hecho de que cada uno de los componentes realiza la misma función en enlace con el otro componente que lo haría si no estuviera así unido. Por ejemplo, con respecto a un polipéptido de fusión en el que un dominio de unión al ADN (por ejemplo, ZFP, TALE) está fusionado con un dominio de activación, el dominio de unión al ADN y el dominio de activación están en enlace operativo si, en el polipéptido de fusión, la porción del dominio de unión al ADN es capaz de unirse a su sitio diana y/o su sitio de unión, mientras que el dominio de activación es capaz de regular positivamente la expresión génica. Cuando un polipéptido de fusión en el que un dominio de unión al ADN está fusionado con un dominio de escisión, el dominio de unión al ADN y el dominio de escisión están en enlace operativo si, en el polipéptido de fusión, la porción del dominio de unión al ADN es capaz de unirse a su sitio objetivo y/o su sitio de unión, mientras que el dominio de escisión es capaz de escindir el ADN en las proximidades del sitio objetivo. De manera similar, con respecto a un polipéptido de fusión en el que un dominio de unión a ADN está fusionado con un dominio de activación o represión, el dominio de unión a ADN y el dominio de activación o represión están en enlace operativo si, en el polipéptido de fusión, el dominio de unión a ADN está fusionado con un dominio de unión a ADN. La porción del dominio es capaz de unirse a su sitio objetivo y/o su sitio de unión, mientras que el dominio de activación es capaz de regular positivamente la expresión génica o el dominio de represión es capaz de regular negativamente la expresión génica.

55 [0076] Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero conserva la misma función que la proteína, polipéptido o ácido nucleico completo, longitud de proteína, polipéptido o ácido nucleico. Un fragmento funcional puede poseer más, menos o el mismo número de residuos que la molécula nativa correspondiente y/o puede contener una o más sustituciones de aminoácidos o nucleótidos. Los métodos para determinar la función de un ácido nucleico (por ejemplo, función codificante, capacidad de hibridarse con otro ácido nucleico) son bien conocidos en la técnica. De manera similar, son bien conocidos métodos para determinar la función de las proteínas. Por ejemplo, la función de unión al ADN de un polipéptido se puede determinar, por ejemplo, mediante ensayos de unión a filtro, de cambio de movilidad electroforética o de inmunoprecipitación. La escisión del ADN se puede analizar mediante electroforesis en gel. Véase Ausubel *et al.*, *supra*. La capacidad de una proteína para interactuar con otra proteína se puede determinar, por ejemplo, mediante coinmunoprecipitación, ensayos de dos híbridos o complementación, tanto genética como bioquímica. Véase, por ejemplo, Fields *et al.* (1989) *Naturaleza* 340:245-246; Patente de EE. UU. nº 5.585.245 y PCT WO 98/44350.

5 [0077] Un "vector" es capaz de transferir secuencias genéticas a células diana. Normalmente, "construcción de vector", "vector de expresión" y "vector de transferencia génica" significan cualquier construcción de ácido nucleico capaz de dirigir la expresión de un gen de interés y que puede transferir secuencias genéticas a células diana. Por tanto, el término incluye vehículos de clonación y expresión, así como vectores de integración.

10 5 [0078] Un "gen informador" o "secuencia informadora" se refiere a cualquier secuencia que produzca un producto proteico que se pueda medir fácilmente, preferiblemente, aunque no necesariamente en un ensayo de rutina. Los genes indicadores adecuados incluyen, entre otros, secuencias que codifican proteínas que median la resistencia a los antibióticos (por ejemplo, resistencia a la ampicilina, resistencia a la neomicina, resistencia a G418, resistencia a la puromicina), secuencias que codifican proteínas coloreadas o fluorescentes o luminiscentes (por ejemplo, proteína verde fluorescente, proteína fluorescente verde mejorada, proteína fluorescente roja, luciferasa) y proteínas que median el crecimiento celular mejorado y/o la amplificación genética (por ejemplo, dihidrofolato reductasa). Las etiquetas de epítopos incluyen, por ejemplo, una o más copias de FLAG, His, myc, Tap, HA o cualquier secuencia de aminoácidos detectable. Las "etiquetas de expresión" incluyen secuencias que codifican indicadores que pueden unirse operativamente a una secuencia genética deseada para controlar la expresión del gen de interés.

15 20 [0079] Los términos "sujeto" y "paciente" se usan indistintamente y se refieren a mamíferos tales como pacientes humanos y primates no humanos, así como a animales de experimentación tales como conejos, perros, gatos, ratas, ratones y otros animales. Por consiguiente, el término "sujeto" o "paciente" como se usa en el presente documento significa cualquier paciente o sujeto mamífero al que se le pueden administrar los cassetes de expresión de la invención. Los sujetos de la presente invención incluyen aquellos con un trastorno o aquellos en riesgo de desarrollar un trastorno.

25 25 [0080] Los términos "tratar" y "tratamiento" como se usan en el presente documento se refieren a la reducción de la gravedad y/o la frecuencia de los síntomas, la eliminación de los síntomas y/o la causa subyacente, la prevención de la aparición de los síntomas y/o su causa subyacente, y mejora o reparación del daño. El cáncer y la enfermedad de injerto contra huésped son ejemplos no limitantes de afecciones que pueden tratarse usando las composiciones y métodos descritos en el presente documento. Por tanto, "tratar" y "tratamiento" incluyen:

- 30 35 (i) prevenir que la enfermedad o afección se produzca en un mamífero, en particular, cuando dicho mamífero está predispuesto a la afección, pero aún no se le ha diagnosticado que la padece;
- (ii) inhibir la enfermedad o afección, es decir, detener su desarrollo;
- (iii) aliviar la enfermedad o afección, es decir, provocar la regresión de la enfermedad o afección o
- (iv) aliviar los síntomas resultantes de la enfermedad o afección, es decir, aliviar el dolor sin abordar el problema; enfermedad o condición subyacente.

40 45 [0081] Como se usan en el presente documento, los términos "enfermedad" y "afección" pueden usarse indistintamente o pueden ser diferentes en el sentido de que la enfermedad o afección particular puede no tener un agente causante conocido (de modo que la etiología aún no se ha determinado) y, por lo tanto, todavía no se reconoce como una enfermedad, sino sólo como una condición o síndrome indeseable, en el que los médicos han identificado un conjunto más o menos específico de síntomas.

50 55 [0082] Una "composición farmacéutica" se refiere a una formulación de un compuesto de la invención y un medio generalmente aceptado en la técnica para la administración del compuesto biológicamente activo a mamíferos, por ejemplo, seres humanos. Dicho medio incluye todos los vehículos, diluyentes o excipientes farmacéuticamente aceptables para el mismo.

60 65 [0083] "Cantidad eficaz" o "cantidad terapéuticamente eficaz" se refiere a la cantidad de un compuesto de la invención que, cuando se administra a un mamífero, preferentemente un ser humano, es suficiente para efectuar el tratamiento en el mamífero, preferentemente un ser humano. La cantidad de una composición de la invención que constituye una "cantidad terapéuticamente eficaz" variará dependiendo del compuesto, la afección y su gravedad, la forma de administración y la edad del mamífero a tratar, pero puede determinarse de forma rutinaria mediante un experto en la técnica teniendo en cuenta su propio conocimiento y esta divulgación.

55 Dominios de unión al ADN

60 65 [0084] En el presente documento se describen composiciones que comprenden un dominio de unión a ADN que se une específicamente a un sitio diana en cualquier gen que comprenda un gen HLA o un regulador de HLA. Se puede usar cualquier dominio de unión a ADN en las composiciones y métodos descritos en el presente documento, incluidos, entre otros, un dominio de unión a ADN con dedos de zinc, un dominio de unión a ADN TALE, la porción de unión a ADN (ARNgu) de una nucleasa CRISPR/Cas, o un dominio de unión al ADN de una meganucleasa. El dominio de unión al ADN puede unirse a cualquier secuencia diana dentro del gen, incluida, entre otras, una secuencia diana de 12 o más nucleótidos como se muestra en cualquiera de los sitios diana descritos en el presente documento (SEQ ID NO:8-21 y/o 92-103).

- [0085]** En determinadas formas de realización, el dominio de unión al ADN comprende una proteína con dedos de zinc. Preferiblemente, la proteína con dedos de zinc no se produce de forma natural porque está diseñada para unirse a un sitio diana de elección. Véase, por ejemplo, Beerli et al. (2002) *Nature Biotechnol.* 20:135-141; Pabo et al. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al. (2001) *Nature Biotechnol.* 19:656-660; Segal et al. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al. (2000) *Curr. Opin. Struct. Biol.* 10:411-416; Patentes de EE. UU. N°s 6.453.242; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.030.215; 6.794.136; 7.067.317; 7.262.054; 7.070.934; 7.361.635; 7.253.273; y Publicaciones de Patentes de EE. UU. N°s 2005/0064474; 2007/0218528; 2005/0267061. En determinadas formas de realización, el dominio de unión al ADN comprende una proteína con dedos de zinc descrita en la publicación de patente estadounidense n.º 2012/0060230 (p. ej., Tabla 1).
- [0086]** Un dominio de unión con dedos de zinc diseñado puede tener una especificidad de unión novedosa, en comparación con una proteína con dedos de zinc natural. Los métodos de ingeniería incluyen, entre otros, diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos tripletes (o cuádruples) y secuencias de aminoácidos de dedos de zinc individuales, en las que cada secuencia de nucleótidos triplete o cuádruple está asociada con una o más secuencias de aminoácidos de dedos de zinc que se unen al triplete particular o secuencia cuádruple. Véanse, por ejemplo, las patentes estadounidenses 6.453.242 y 6.534.261.
- [0087]** Se describen métodos de selección ejemplares, que incluyen presentación en fagos y sistemas de dos híbridos, en las patentes de EE. UU. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568, así como el documento WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, se ha descrito una mejora de la especificidad de unión para los dominios de unión con dedos de zinc, por ejemplo, en la patente de EE. UU. n.º 6.794.136.
- [0088]** Además, como se describe en estas y otras referencias, los dominios de dedos de zinc y/o las proteínas de dedos de zinc múltiples pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véanse también las patentes estadounidenses números 6.479.626; 6.903.185; y 7.153.949 para secuencias enlazadoras ejemplares de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína. Además, se ha descrito una mejora de la especificidad de unión para los dominios de unión con dedos de zinc, por ejemplo, en la patente de EE. UU. n.º 6.794.136.
- [0089]** Selección de sitios objetivo; los expertos en la técnica conocen las ZFP y los métodos para el diseño y construcción de proteínas de fusión (y los polinucleótidos que las codifican) y se describen en detalle en las patentes estadounidenses números 6.140.081; 5.789.538; 6.453.242; 6.534.261; 5.925.523; 6.007.988; 6.013.453; 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.
- [0090]** Además, como se describe en estas y otras referencias, los dominios de dedos de zinc y/o las proteínas de dedos de zinc múltiples pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véanse también las patentes estadounidenses números 6.479.626; 6.903.185; y 7.153.949 para secuencias enlazadoras ejemplares de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína.
- [0091]** En determinadas formas de realización, el dominio de unión al ADN es una proteína con dedos de zinc diseñada que se une (de una manera específica de secuencia) a un sitio diana en un gen TCR o gen regulador de TCR y modula la expresión de un gen TCR. En algunas formas de realización, la proteína con dedos de zinc se une a un sitio objetivo en TCRA, mientras que, en otras formas de realización, el dedo de zinc se une a un sitio objetivo en TRBC.
- [0092]** Normalmente, las ZFP incluyen al menos tres dedos. Algunas de las ZFP incluyen cuatro, cinco o seis dedos. Las ZFP que incluyen tres dedos normalmente reconocen un sitio objetivo que incluye 9 o 10 nucleótidos; las ZFP que incluyen cuatro dedos normalmente reconocen un sitio objetivo que incluye de 12 a 14 nucleótidos; mientras que las ZFP que tienen seis dedos pueden reconocer sitios objetivo que incluyen de 18 a 21 nucleótidos. Las ZFP también pueden ser proteínas de fusión que incluyen uno o más dominios reguladores, cuyos dominios pueden ser dominios de activación o represión transcripcional.
- [0093]** En algunas formas de realización, el dominio de unión al ADN puede derivarse de una nucleasa. Por ejemplo, las secuencias de reconocimiento de endonucleasas y meganucleasas homing tales como I-SceI, 1-Ceul, PI-Pspl, PI-Sce, 1-ScelV, 1-Csml, 1-PanI, 1-Scell, 1-Ppol, I-Scell, 1-Crel, 1-TevI, 1-TevII y 1-TevIII se conocen. Véanse también la patente de EE. UU. n.º 5.420.032; Patente de EE. UU. n.º 6.833.252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jaslin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J Mol. Biol.* 263:163-180; Argast et al. (1998) *J Mol. Biol.* 280:345-353 y el catálogo de New England Biolabs. Además, la especificidad de unión al ADN de las endonucleasas y meganucleasas dirigidas se puede diseñar para unirse a sitios objetivo no naturales. Véase, por ejemplo, Chevalier et al. (2002) *Molec. Cell* 10:895-905;

Epinat et al. (2003) Nucleic Acids Res. 31:2952-2962; Ashworth et al. (2006) Nature 441:656-659; Paques et al. (2007) Current Gene Therapy 7:49-66; Publicación de patente estadounidense nº 20070117128.

- 5 **[0094]** En otras formas de realización, el dominio de unión al ADN comprende un dominio diseñado a partir de un efecto TAL similar a los derivados de los patógenos vegetales *Xanthomonas* (véase Boch et al, (2009) Science 326: 1509-1512 y Moscou y Bogdanove, (2009) Science 326: 1501) y *Ralstonia* (véase Heuer et al. (2007) Applied and Environmental Microbiology 73(13): 4379-4384); Solicitudes de patente de EE. UU. N°s 20110301073 y 20110145940. Se sabe que las bacterias fitopatógenas del género *Xanthomonas* causan muchas enfermedades en importantes plantas de cultivo. La patogenicidad de *Xanthomonas* depende de un sistema de secreción conservado de tipo III (T3S) que inyecta más de 25 proteínas efectoras diferentes en la célula vegetal. Entre estas proteínas inyectadas se encuentran efectores similares a activadores de la transcripción (TALE) que imitan los activadores transcripcionales de las plantas y manipulan el transcriptoma de las plantas (véase Kay et al. (2007) Science 318:648-651). Estas proteínas contienen un dominio de unión al ADN y un dominio de activación transcripcional. Uno de los TALE mejor caracterizados es AvrBs3 de *Xanthomonas campestris* pv. *Vesicatoria* (ver Bonas et al. (1989) Mol Gen Genet 218: 127-136 y WO2010079430). Los TALE contienen un dominio centralizado de repeticiones en tandem, cada repetición contiene aproximadamente 34 aminoácidos, que son clave para la especificidad de unión al ADN de estas proteínas. Además, contienen una secuencia de localización nuclear y un dominio de activación transcripcional ácido (para una revisión, véase Schornack S, et al. (2006) J Plant Physiol 163(3): 256-272). Además, en la bacteria fitopatógena *Ralstonia solanacearum* se han encontrado dos genes, denominados *brgl* 1 y *hpxl* 7, que son homólogos a la familia AvrBs3 de *Xanthomonas* en la cepa GMI1000 de biovar 1 de *R. solanacearum* y en la cepa RS1000 de biovar 4 (véase Heuer et al. al (2007) Appl y Envir Micro 73(13): 4379-4384). Estos genes son 98,9% idénticos en la secuencia de nucleótidos entre sí, pero se diferencian por una eliminación de 1.575 pb en el dominio repetido de *hpxl* 7. Sin embargo, ambos productos genéticos tienen menos del 40% de identidad de secuencia con las proteínas de la familia AvrBs3 de *Xanthomonas*.
- 10 25 **[0095]** La especificidad de estos efectores TAL depende de las secuencias encontradas en las repeticiones en tandem. La secuencia repetida comprende aproximadamente 102 pares de bases y las repeticiones son típicamente homólogas entre sí en un 91-100% (Bonas et al, ibid). El polimorfismo de las repeticiones generalmente se ubica en las posiciones 12 y 13 y parece haber una correspondencia uno a uno entre la identidad de los diresiduos hipervariables (el diresiduo variable repetido o región RVR) en las posiciones 12 y 13 con la identidad de los nucleótidos contiguos en la secuencia diana del efecto TAL (véase Moscou y Bogdanove, (2009) Science 326:1501 y Boch et al. (2009) Science 326:1509-1512). Experimentalmente, se ha determinado el código natural para el reconocimiento del ADN de estos efectores TAL de modo que una secuencia HD en las posiciones 12 y 13 (residuo variable repetido o RVR) conduce a una unión a la citosina (C), NG se une a T, NI a A, C, G o T, NN se une a A o G, y ING se une a T. Estas repeticiones de unión al ADN se han ensamblado en proteínas con nuevas combinaciones y números de repeticiones, para crear factores de transcripción artificiales que puedan interactuar con nuevas secuencias y activan la expresión de un gen informador no endógeno en células vegetales (Boch et al, ibid). Las proteínas TAL diseñadas se han vinculado a un medio dominio de escisión de *FokI* para producir una fusión de nucleasa del dominio efecto TAL (TALEN), incluidas las TALEN con RVR atípicos. Véase, por ejemplo, la patente de EE. UU. nº 8.586.526.
- 30 35 40 **[0096]** En algunas formas de realización, TALEN comprende un dominio de escisión o medio dominio de escisión de endonucleasa (p. ej., *FokI*). En otras formas de realización, la nucleasa TALE es una mega TAL. Estas nucleasas mega TAL son proteínas de fusión que comprenden un dominio de unión al ADN TALE y un dominio de escisión de meganucleasa. El dominio de escisión de la meganucleasa es activo como monómero y no requiere dimerización para su actividad. (Ver Boissel et al., (2013) Nucl Acid Res: 1-13, doi: 10.1093/nar/gkt1224).
- 45 **[0097]** En aún otras formas de realización, la nucleasa comprende una TALEN compacta. Se trata de proteínas de fusión de cadena sencilla que unen un dominio de unión al ADN de TALE a un dominio de nucleasa *TevI*. La proteína de fusión puede actuar como una nickasa localizada por la región TALE o puede crear una rotura de doble hebra, dependiendo de dónde se encuentre el dominio de unión al ADN de TALE con respecto al dominio de nucleasa *TevI* (ver Beurdeley et al. (2013) Nat Comm: 1-8 DOI: 10.1038/ncomms2782). Además, el dominio de nucleasa también puede exhibir funcionalidad de unión al ADN. Cualquier TALEN se puede utilizar en combinación con TALEN adicionales (por ejemplo, uno o más TALEN (cTALEN o *FokI*-TALEN) con uno o más mega-TALE).
- 50 55 **[0098]** Además, como se describe en estas y otras referencias, los dominios de dedos de zinc y/o proteínas de dedos de zinc con múltiples dedos o TALE se pueden unir entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos en longitud. Véanse también las patentes estadounidenses números 6.479.626; 6.903.185; y 7.153.949 para secuencias enlazadoras ejemplares de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína. Además, se ha descrito una mejora de la especificidad de unión para los dominios de unión con dedos de zinc, por ejemplo, en la patente de EE. UU. nº 6.794.136.
- 60 65 **[0099]** En ciertas formas de realización, el dominio de unión al ADN es parte de un sistema de nucleasa CRISPR/Cas, que incluye un ARN guía único (ARNsg) que se une al ADN. Véase, por ejemplo, la patente de EE. UU. n.º 8.697.359 y las publicaciones de patente de EE. UU. n.º 20150056705 y 20150159172. El locus CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente espaciadas), que codifica los componentes de ARN del sistema, y el locus cas (asociado

a CRISPR), que codifica proteínas (Jansen et al., 2002. Mol. Microbiol. 43: 1565-1575; Makarova et al., 2002. Nucleic Acids Res. 30: 482-496;) forman las secuencias genéticas del sistema de nucleasa CRISPR/Cas. Los loci CRISPR en huéspedes microbianos contienen una combinación de genes asociados a CRISPR (Cas), así como elementos de ARN no codificantes capaces de programar la especificidad de la escisión de ácidos nucleicos mediada por CRISPR.

5

[0100] El CRISPR de tipo II es uno de los sistemas mejor caracterizados y lleva a cabo la rotura dirigida de la doble hebra del ADN en cuatro pasos secuenciales. Primero, se transcriben dos ARN no codificantes, la matriz de pre-ARNcr y el ARNtracr, del locus CRISPR. En segundo lugar, el ARNtracr se hibrida con las regiones repetidas del pre-ARNcr y media el procesamiento del pre-ARNcr en ARNcr maduros que contienen secuencias espaciadoras individuales. En tercer lugar, el complejo ARNcr:ARNtracr maduro dirige el dominio funcional (p. ej., nucleasa como Cas) al ADN objetivo mediante el emparejamiento de bases Watson-Crick entre el espaciador del ARNcr y el protoespaciador del ADN objetivo junto al motivo adyacente al protoespaciador (PAM), un requisito adicional para el reconocimiento de objetivos. Finalmente, Cas9 media la escisión del ADN objetivo para crear una rotura bicanalicular dentro del protoespaciador. La actividad del sistema CRISPR/Cas consta de tres pasos: (i) inserción de secuencias de ADN extrañas en la matriz CRISPR para evitar futuros ataques, en un proceso llamado "adaptación", (ii) expresión de las proteínas relevantes, así como expresión y procesamiento de la matriz, seguido de (iii) interferencia mediada por ARN con el ácido nucleico extraño. Por lo tanto, en la célula bacteriana, varias de las llamadas proteínas 'Cas' están involucradas en la función natural del sistema CRISPR/Cas y desempeñan funciones como la inserción del ADN extraño, etc.

10

15

20

25

30

35

40

45

50

55

[0101] En determinadas formas de realización, la proteína Cas puede ser un "derivado funcional" de una proteína Cas de origen natural. Un "derivado funcional" de un polipéptido de secuencia nativa es un compuesto que tiene una propiedad biológica cualitativa en común con un polipéptido de secuencia nativa. Los "derivados funcionales" incluyen, entre otros, fragmentos de una secuencia nativa y derivados de un polipéptido de secuencia nativa y sus fragmentos, siempre que tengan una actividad biológica en común con un polipéptido de secuencia nativa correspondiente. Una actividad biológica contemplada en el presente documento es la capacidad del derivado funcional de hidrolizar un sustrato de ADN en fragmentos. El término "derivado" abarca variantes de secuencia de aminoácidos del polipéptido, modificaciones covalentes y fusiones de los mismos, tales como proteínas Cas derivadas. Los derivados adecuados de un polipéptido Cas o un fragmento del mismo incluyen, entre otros, mutantes, fusiones, modificaciones covalentes de la proteína Cas o un fragmento de la misma. La proteína Cas, que incluye la proteína Cas o un fragmento de la misma, así como derivados de la proteína Cas o un fragmento de la misma, puede obtenerse a partir de una célula o sintetizarse químicamente o mediante una combinación de estos dos procedimientos. La célula puede ser una célula que produce de forma natural proteína Cas, o una célula que produce de forma natural proteína Cas y está diseñada genéticamente para producir la proteína Cas endógena a un nivel de expresión más alto o para producir una proteína Cas a partir de un ácido nucleico introducido exógenamente, cuyo ácido nucleico El ácido codifica un Cas que es igual o diferente al Cas endógeno. En algunos casos, la célula no produce proteína Cas de forma natural y está diseñada genéticamente para producir una proteína Cas. En algunas formas de realización, la proteína Cas es un pequeño ortólogo de Cas9 para la administración a través de un vector VAA (Ran et al. (2015) Nature 510, p. 186).

[0102] En algunas formas de realización, el dominio de unión al ADN es parte de un sistema TtAgo (ver Swarts et al., ibid; Sheng et al, ibid). En eucariotas, el silenciamiento de genes está mediado por la familia de proteínas Argonautas (Ago). En este paradigma, Ago está unido a ARN pequeños (19-31 nt). Este complejo silenciador de proteína-ARN reconoce los ARN objetivo mediante el emparejamiento de bases de Watson-Crick entre el ARN pequeño y el objetivo y escinde endonucleolíticamente el ARN objetivo (Vogel (2014) Science 344:972-973). Por el contrario, las proteínas Ago procarióticas se unen a pequeños fragmentos de ADN monocatenario y probablemente funcionan para detectar y eliminar ADN extraño (a menudo viral) (Yuan et al., (2005) Mol. Cell 19,405; Olovnikov, et al. (2013) Mol. Cell 51, 594; las proteínas Ago procarióticas ejemplares incluyen aquellas de *Aquifex aeolicus*, *Rhodobacter sphaeroides* y *Thermus thermophilus*.

[0103] Una de las proteínas Ago procariotas mejor caracterizadas es la de *T. thermophilus* (TtAgo; Swarts et al. ibid). TtAgo se asocia con fragmentos de ADN monocatenario de 15 nt o de 13-25 nt con grupos fosfato 5'. Este "ADN guía" unido por TtAgo sirve para dirigir el complejo proteína-ADN para que se una a una secuencia de ADN complementaria de Watson Crick en una molécula de ADN de terceros. Una vez que la información de secuencia en estos ADN guía ha permitido la identificación del ADN objetivo, el complejo de ADN guía TtAgo escinde el ADN objetivo. Este mecanismo también está respaldado por la estructura del complejo de ADN guía TtAgo mientras está unido a su ADN objetivo (G. Sheng et al., ibid). Ago de *Rhodobacter sphaeroides* (RsAgo) tiene propiedades similares (Olivnikov et al. ibid).

[0104] Se pueden cargar ADN guía exógenos de secuencia de ADN arbitraria en la proteína TtAgo (Swarts et al. ibid). Dado que la especificidad de la escisión de TtAgo está dirigida por el ADN guía, un complejo TtAgo-ADN formado con un ADN guía exógeno especificado por el investigador dirigirá la escisión del ADN objetivo de TtAgo a un ADN objetivo complementario especificado por el investigador. De esta manera, se puede crear una rotura de doble cadena dirigida en el ADN. El uso del sistema de ADN guía TtAgo (o sistemas de ADN guía Ago ortólogos de otros organismos) permite la escisión dirigida del ADN genómico dentro de las células. Dicha escisión puede ser monocatenaria o bicanalicular. Para la escisión del ADN genómico de mamíferos, sería preferible utilizar una versión del codón TtAgo optimizada para la expresión en células de mamíferos. Además, podría ser preferible tratar las células con un complejo TtAgo-ADN formado *in vitro* donde la proteína TtAgo se fusiona con un péptido que penetra en las células. Además, podría ser preferible utilizar una versión de la proteína TtAgo que haya sido alterada mediante mutagénesis para tener una actividad mejorada a 37°C.

La escisión de ADN mediada por ago-ARN podría usarse para afectar una variedad de resultados que incluyen eliminación de genes, adición de genes dirigida, corrección de genes, eliminación de genes dirigida utilizando técnicas estándar en la técnica para la explotación de roturas de ADN.

- 5 [0105] Por tanto, se puede utilizar cualquier dominio de unión a ADN.

Moléculas de fusión

10 [0106] También se proporcionan moléculas de fusión que comprenden dominios de unión a ADN (por ejemplo, ZFP o TALE, componentes CRISPR/Cas tales como ARN guía únicos) como se describe en el presente documento asociados con un dominio regulador (funcional) heterólogo (o fragmento funcional del mismo). Los dominios comunes incluyen, por ejemplo, dominios de factores de transcripción (activadores, represores, coactivadores, correpresores), silenciadores, oncogenes (por ejemplo, miembros de la familia myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos, etc.); enzimas reparadoras del ADN y sus factores y modificadores asociados; enzimas de reordenamiento del ADN y sus factores y modificadores asociados; proteínas asociadas a la cromatina y sus modificadores (por ejemplo, quinasas, acetilasas y desacetilasas); y enzimas modificadoras del ADN (por ejemplo, metiltransferasas, topoisomerasas, helicasas, ligasas, quinasas, fosfatases, polimerasas, endonucleasas) y sus factores y modificadores asociados. Dichas moléculas de fusión incluyen factores de transcripción que comprenden los dominios de unión al ADN descritos en el presente documento y un dominio regulador transcripcional, así como nucleasas que comprenden los dominios de unión al ADN y uno o más dominios de nucleasa.

15 [0107] Los dominios adecuados para lograr la activación (dominios de activación transcripcional) incluyen el dominio de activación de HSV VP16 (ver, por ejemplo, Hagmann et al., J. Virol. 71, 5952-5962 (1997)) receptores de hormonas nucleares (ver, por ejemplo, Torchia et al., Curr. Opin. Cell. Biol. 10:373-383 (1998)); la subunidad p65 del factor nuclear kappa B (Bitko & Barik, J Virol. 72:5610-5618 (1998) y Doyle & Hunt, Neuroreport 8:2937-2942 (1997)); Liu et al., Cancer Gene Ther. 5:3-28 (1998)), o dominios funcionales químicos artificiales tales como VP64 (Beerli et al., (1998) Proc. Natl. Acad. Sci. EE. UU. 95:14623-33), y degron (Molinari et al., (1999) EMBO J. 18, 6439-6447). Dominios de activación ejemplares adicionales incluyen Oct 1, Oct-2A, Sp1, AP-2 y CTF1 (Seipel et al., EMBO J. 11, 4961-4968 (1992) así como p300, CBP, P3CAF, SRC1 PvALF, AtHD2A y ERF-2 Véase, por ejemplo, Robyr et al. (2000) Mol. 14:329-347; Leo et al. 20 (2000) Gene 245:1-11; Manteuffel-Cymborowska (1999) Acta Biochim. Pol. 46:77-89; (2000) Trends Biochem. 25:277-283; y Lemon et al. (1999) Curr. Opin. Genet. Dev 9:499-504. Los dominios de activación ejemplares adicionales incluyen, pero no se limitan a OsGAI, HALF-1, C1, AP1, ARF-5,-6,-7 y -8, CPRF1, CPRF4, MYC-RP/GP y TRAB1. Véase, por ejemplo, Ogawa et al. (2000) Gene 245:21-29; Okanami et al. (1996) Genes Cells 1:87-99; 429; Ulmason et al. (1999) Proc. Natl. Acad. Sci. EE. UU. 96:5844-5849; Gong et al. (1999) Plant Mol. Biol. 41:33-44; y Hobo et al. (1999) Proc. Natl. Acad. Sci. EE. UU. 96:15,348-15,353.

25 [0108] Será claro para los expertos en la técnica que, en la formación de una proteína de fusión (o un ácido nucleico que codifica la misma) entre un dominio de unión al ADN y un dominio funcional, ya sea un dominio de activación o una molécula que interactúa con un dominio de activación y es adecuado como dominio funcional. Básicamente, cualquier 30 molécula capaz de reclutar un complejo activador y/o una actividad activadora (tal como, por ejemplo, acetilación de histonas) en el gen diana es útil como dominio activador de una proteína de fusión. Los dominios aislantes, los dominios de localización y las proteínas remodeladoras de la cromatina tales como los dominios que contienen ISWI y/o las proteínas del dominio de unión a metilo adecuadas para su uso como dominios funcionales en moléculas de fusión se describen, por ejemplo, en la patente de EE. UU. nº 7.053.264.

35 [0109] Los dominios de represión ejemplares incluyen, entre otros, KRAB A/B, KOX, gen temprano inducible por TGF-beta (TIEG), v-erbA, SID, MBD2, MBD3, miembros de la familia DNMT (p. ej., DNMT1, DNMT3A, DNMT3B), Rb y MeCP2. Véase, por ejemplo, Bird et al. (1999) Cell 99:451-454; Tyler et al. (1999) Cell 99:443-446; Knoepfli et al. (1999) Cell 99:447-450; y Robertson et al. (2000) Naturaleza Genet. 25:338-342. Los dominios de represión ejemplares adicionales incluyen, entre otros, ROM2 y AtHD2A. Véase, por ejemplo, Chem et al. (1996) Plant Cell 8:305-321; y Wu et al. (2000) Planta J. 22:19-27.

40 [0110] Las moléculas de fusión se construyen mediante métodos de clonación y conjugación bioquímica que son bien conocidos por los expertos en la técnica. Las moléculas de fusión comprenden un dominio de unión al ADN (p. ej., ZFP, TALE, ARN_{gu}) asociado con un dominio funcional (p. ej., un dominio de activación o represión transcripcional). Las moléculas de fusión también comprenden opcionalmente señales de localización nuclear (tales como, por ejemplo, las del antígeno T medio SV40) y etiquetas de epitopos (tales como, por ejemplo, FLAG y hemaglutinina). Las proteínas de fusión (y los ácidos nucleicos que las codifican) están diseñadas de manera que el marco de lectura traduccional se conserve entre los componentes de la fusión.

45 [0111] Fusiones entre un componente polipeptídico de un dominio funcional (o un fragmento funcional del mismo) por un lado, y un dominio de unión al ADN no proteico (por ejemplo, antibiótico, intercalador, aglutinante de surcos menores, ácido nucleico) por el otro, se construyen mediante métodos de conjugación bioquímica conocidos por los expertos en la técnica. Véase, por ejemplo, el catálogo de Pierce Chemical Company (Rockford, IL). Se han descrito métodos y composiciones para realizar fusiones entre un aglutinante de surcos menores y un polipéptido. Mapp et al. (2000) Proc.

Natl. Acad. Sci. EE. UU. 97:3930-3935. Además, los ARN guía únicos del sistema CRISPR/Cas se asocian con dominios funcionales para formar reguladores transcripcionales activos y nucleasas.

5 [0112] En determinadas formas de realización, el sitio diana está presente en una región accesible de la cromatina celular. Las regiones accesibles se pueden determinar como se describe, por ejemplo, en las patentes estadounidenses números 7.217.509 y 7.923.542. Si el sitio objetivo no está presente en una región accesible de la cromatina celular, se pueden generar una o más regiones accesibles como se describe en las patentes estadounidenses números 7.785.792 y 8.071.370. En formas de realización adicionales, el dominio de unión al ADN de una molécula de fusión es capaz de unirse a la cromatina celular independientemente de si su sitio objetivo está en una región accesible o no. Por ejemplo, dichos dominios de unión al ADN son capaces de unirse al ADN conector y/o al ADN nucleosomal. Ejemplos de este tipo de dominio de unión al ADN "pionero" se encuentran en ciertos receptores de esteroides y en el factor nuclear 3 de los hepatocitos (HNF3) (Cordingley et al. (1987) Cell 48:261-270; Pina et al. (1990) Cell 60:719-731; y Cirillo et al. (1998) EMBO J. 17:244-254).

10 15 [0113] La molécula de fusión se puede formular con un vehículo farmacéuticamente aceptable, como saben los expertos en la técnica. Véase, por ejemplo, Remington's Pharmaceutical Sciences, 17^a ed., 1985; y las patentes estadounidenses números 6.453.242 y 6.534.261.

20 25 [0114] El componente/dominio funcional de una molécula de fusión se puede seleccionar entre cualquiera de una variedad de componentes diferentes capaces de influir en la transcripción de un gen una vez que la molécula de fusión se une a una secuencia diana a través de su dominio de unión al ADN. Por tanto, el componente funcional puede incluir, entre otros, diversos dominios de factores de transcripción, tales como activadores, represores, coactivadores, correpresores y silenciadores.

30 35 [0115] Se describen dominios funcionales ejemplares adicionales, por ejemplo, en las patentes estadounidenses números 6.534.261 y 6.933.113.

[0116] También se pueden seleccionar dominios funcionales que están regulados por pequeñas moléculas o ligandos exógenos. Por ejemplo, se puede emplear la tecnología RheoSwitch® en la que un dominio funcional solo asume su conformación activa en presencia del ligando externo RheoChem™ (véase, por ejemplo, el documento US 20090136465). Por tanto, la ZFP puede estar operativamente unida al dominio funcional regulable en el que la actividad resultante de la ZFP-TF está controlada por el ligando externo.

Nucleasas

40 45 [0117] En determinadas formas de realización, la molécula de fusión comprende un dominio de unión de unión a ADN asociado con un dominio de escisión (nucleasa). Como tal, la modificación genética se puede lograr usando una nucleasa, por ejemplo, una nucleasa modificada genéticamente. La tecnología de nucleasas diseñadas se basa en la ingeniería de proteínas de unión al ADN naturales. Por ejemplo, se ha descrito la ingeniería de endonucleasas dirigidas con especificidades de unión al ADN adaptadas. Chames et al. (2005) Nucleic Acids Res 33(20):e178; Arnould et al. (2006) J Mol. Biol. 355:443-458. Además, también se ha descrito la ingeniería de ZFP. Véanse, por ejemplo, las patentes estadounidenses números 6.534.261; 6.607.882; 6.824.978; 6.979.539; 6.933.113; 7.163.824; y 7.013.219.

50 55 [0118] Además, las ZFP y/o TALE se pueden fusionar con dominios de nucleasa para crear ZFN y TALEN - una entidad funcional que es capaz de reconocer su objetivo de ácido nucleico previsto a través de su dominio de unión al ADN diseñado (ZFP o TALE) y provocar el ADN se cortará cerca del sitio de unión del ADN mediante la actividad de la nucleasa.

[0119] Por tanto, los métodos y composiciones descritos en el presente documento son ampliamente aplicables y pueden implicar cualquier nucleasa de interés. Ejemplos no limitantes de nucleasas incluyen meganucleasas, TALEN y nucleasas con dedos de zinc. La nucleasa puede comprender dominios de unión y escisión de ADN heterólogos (por ejemplo, nucleasas con dedos de zinc; dominios de unión a ADN de meganucleasa con dominios de escisión heterólogos) o, alternativamente, el dominio de unión a ADN de una nucleasa natural puede alterarse para unirse a un sitio diana seleccionado (por ejemplo, una meganucleasa que ha sido diseñada para unirse a un sitio diferente al sitio de unión afín).

60 65 [0120] En cualquiera de las nucleasas descritas en el presente documento, la nucleasa puede comprender un dominio de unión a ADN TALE diseñado y un dominio de nucleasa (por ejemplo, dominio de endonucleasa y/o meganucleasa), también denominados TALEN. Se han publicado métodos y composiciones para diseñar estas proteínas TALEN para una interacción robusta y específica del sitio con la secuencia objetivo elegida por el usuario (consulte la patente de EE. UU. N° 8.586.526). En algunas formas de realización, TALEN comprende un dominio de escisión o medio dominio de escisión de endonucleasa (p. ej., *FokI*). En otras formas de realización, la nucleasa TALE es una mega TAL. Estas meganucleasas TAL son proteínas de fusión que comprenden un dominio de unión al ADN TALE y un dominio de escisión de meganucleasa. El dominio de escisión de la meganucleasa es activo como monómero y no requiere dimerización para su actividad. (Ver Boissel et al., (2013) Nucl Acid Res: 1-13, doi: 10.1093/nar/gkt1224). Además, el dominio de nucleasa también puede exhibir funcionalidad de unión al ADN.

- 5 [0121] En aún otras formas de realización, la nucleasa comprende una TALEN compacta (cTALEN). Se trata de proteínas de fusión de cadena sencilla que unen un dominio de unión al ADN de TALE a un dominio de nucleasa Tevl. La proteína de fusión puede actuar como una nickasa localizada por la región TALE o puede crear una rotura de doble hebra, dependiendo de dónde se encuentre el dominio de unión al ADN de TALE con respecto al dominio de nucleasa Tevl (ver Beurdeley et al. (2013) *Nat Comm*: 1-8 DOI: 10.1038/ncomms2782). Se puede usar cualquier TALEN en combinación con TALEN adicionales (por ejemplo, uno o más TALEN (cTALEN o *FokI*-TALEN) con uno o más mega-TAL) u otras enzimas de escisión de ADN.
- 10 [0122] En ciertas formas de realización, la nucleasa comprende una meganucleasa (endonucleasa localizada) o una porción de la misma que exhibe actividad de escisión. Las meganucleasas naturales reconocen sitios de escisión de 15 a 40 pares de bases y comúnmente se agrupan en cuatro familias: la familia LAGLIDADG, la familia GIY-YIG, la familia His-Cyst box y la familia HNH. Las endonucleasas de localización ejemplares incluyen I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PspI, I-CreI, I-TevI y I-TevIII. Se conocen sus secuencias de reconocimiento. Véanse también la patente de EE. UU. nº 5.420.032; Patente de EE. UU. nº 6.833.252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J Mol. Biol.* 263:163-180; Argast et al. (1998) *J Mol. Biol.* 280:345-353 y el catálogo de New England Biolabs.
- 15 [0123] Se han utilizado dominios de unión a ADN de meganucleasas naturales, principalmente de la familia LAGLIDADG, para promover la modificación genómica específica de sitio en plantas, levaduras, *Drosophila*, células de mamíferos y ratones, pero este enfoque se ha limitado a la modificación de genes homólogos que conservan la secuencia de reconocimiento de meganucleasa (Monet et al. (1999), *Biochem. Biophysics. Res. Common.* 255: 88-93) o de genomas prediseñados en los que se ha introducido una secuencia de reconocimiento (Route et (1994), *Mol Cell.* 14: 8096-106, *Proc Natl. EE. UU.* 93: 5055-60; Rong et al. (2002), *Genes Dev.* 16: 1568-81; En consecuencia, se han realizado intentos de diseñar meganucleasas para que exhiban una nueva especificidad de unión en sitios médica o biotecnológicamente relevantes (Porteus et al. (2005), *Nat. Biotechnol.* 23: 967-73; Sussman et al. (2004), *J Mol. Biol.* 342: 31-41; Epinat et al. (2003), *Nucleic Acids Res.* 31: 2952-62; *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66; Publicación de patente de EE. UU. N°s 20070117128; 20060206949; 20060153826; 20060078552; y 20040002092). Además, los dominios de unión a ADN de meganucleasas naturales o diseñados pueden unirse operativamente con un dominio de escisión de una nucleasa heteróloga (por ejemplo, *FokI*) y/o los dominios de escisión de meganucleasas pueden unirse operativamente con un dominio de unión a ADN heterólogo (ej., ZFP o TALE).
- 20 [0124] En otras formas de realización, la nucleasa es una nucleasa con dedos de zinc (ZFN) o una fusión de nucleasa y dominio de unión a ADN TALE (TALEN). Los ZFN y TALEN comprenden un dominio de unión al ADN (proteína con dedos de zinc o dominio de unión al ADN TALE) que ha sido diseñado para unirse a un sitio objetivo en un gen de elección y un dominio de escisión o un semidominio de escisión (p. ej., de una restricción y/o o meganucleasa como se describe en el presente documento).
- 25 [0125] Como se describió en detalle anteriormente, los dominios de unión a dedos de zinc y los dominios de unión a ADN TALE pueden diseñarse para unirse a una secuencia de elección. Véase, por ejemplo, Beerli et al. (2002) *Nature Biotechnol.* 20:135-141; Pabo et al. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al. (2001) *Nature Biotechnol.* 19:656-660; Segal et al. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al. (2000) *Curr. Opin. Struct. Biol.* 10:411-416. Un dominio de unión con dedos de zinc diseñado o proteína TALE puede tener una especificidad de unión novedosa, en comparación con una proteína natural. Los métodos de ingeniería incluyen, entre otros, diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos tripletes (o cuádruples) y secuencias de aminoácidos de dedos de zinc o TALE individuales, en las que cada secuencia de nucleótidos triplete o cuádruple está asociada con una o más secuencias de aminoácidos de dedos de zinc o unidades de repeticiones TALE que se unen a la secuencia particular triplete o cuádruple. Véanse, por ejemplo, las patentes estadounidenses 6.453.242 y 6.534.261.
- 30 [0126] Selección de sitios diana; y los expertos en la técnica conocen métodos para el diseño y construcción de proteínas de fusión (y polinucleótidos que las codifican) y se describen en detalle en las patentes estadounidenses números 7.888.121 y 8.409.861.
- 35 [0127] Además, como se describe en estas y otras referencias, los dominios de dedos de zinc, TALE y/o proteínas de dedos de zinc múltiples pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos en longitud. Véanse, por ejemplo, las patentes estadounidenses números 6.479.626; 6.903.185; y 7.153.949 para secuencias enlazadoras ejemplares de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína. Véase también la patente de EE. UU. nº 8.772.453.
- 40 [0128] Por tanto, las nucleasas tales como ZFN, TALEN y/o meganucleasas pueden comprender cualquier dominio de unión a ADN y cualquier dominio (de escisión) de nucleasa (dominio de escisión, semidominio de escisión). Como se señaló anteriormente, el dominio de escisión puede ser heterólogo al dominio de unión a ADN, por ejemplo, un dominio
- 45
- 50
- 55
- 60

de unión a ADN con dedos de zinc o efecto TAL y un dominio de escisión de una nucleasa o un dominio de unión a ADN de meganucleasa y un dominio de escisión de un dominio de una nucleasa diferente. Los dominios de escisión heterólogos pueden obtenerse a partir de cualquier endonucleasa o exonucleasa. Las endonucleasas ejemplares de las que se puede derivar un dominio de escisión incluyen, entre otras, endonucleasas de restricción y endonucleasas homing. Véase, por ejemplo, Catalogue 2002-2003, New England Biolabs, Beverly, MA; y Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388. Se conocen enzimas adicionales que escinden el ADN (p. ej., nucleasa S1; nucleasa de frijol mungo; ADNasa I pancreática; nucleasa microcócica; endonucleasa HO de levadura; véase también Linn et al. (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o fragmentos funcionales de las mismas) pueden usarse como fuente de dominios de escisión y semidominios de escisión.

[0129] De manera similar, un semidominio de escisión puede derivarse de cualquier nucleasa o porción de la misma, como se establece anteriormente, que requiere dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Alternativamente, se puede utilizar una única proteína que comprende dos semidominios de escisión. Los dos semidominios de escisión pueden derivarse de la misma endonucleasa (o fragmentos funcionales de la misma), o cada semidominio de escisión puede derivarse de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios diana para las dos proteínas de fusión están dispuestos preferentemente, uno con respecto al otro, de manera que la unión de las dos proteínas de fusión a sus respectivos sitios diana coloca los semidominios de escisión en una orientación espacial entre sí que permite la semidominios de escisión para formar un dominio de escisión funcional, por ejemplo, mediante dimerización. Por tanto, en determinadas formas de realización, los bordes cercanos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número integral de nucleótidos o pares de nucleótidos puede intervenir entre dos sitios diana (por ejemplo, de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios objetivo, pero puede estar a 1 o más kilobases de distancia del sitio de escisión, incluyendo entre 1 y 50 pares de bases (o cualquier valor entre ellos, incluyendo 1-5, 1-10 y 1-20 pares de bases), 1-100 pares de bases (o cualquier valor intermedio), 100-500 pares de bases (o cualquier valor intermedio), 500 a 1000 pares de bases (o cualquier valor intermedio) o incluso más de 1 kb del sitio de escisión.

[0130] Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unirse de forma específica a una secuencia al ADN (en un sitio de reconocimiento) y escindir el ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (p. ej., tipo IIS) escinden el ADN en sitios eliminados del sitio de reconocimiento y tienen dominios de unión y escisión separables. Por ejemplo, la enzima Fok I de tipo IIS cataliza la escisión bicitaria del ADN, a 9 nucleótidos de su sitio de reconocimiento en una cadena y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véanse, por ejemplo, las patentes estadounidenses 5.356.802; 5.436.150 y 5.487.994; así como Li et al. (1992) Proc. Natl. Acad. Sci. EE. UU. 89:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. EE. UU. 90:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. EE. UU. 91:883-887; Kim et al. (1994b) J Biol. Chem. 269:31,978-31,982. Por tanto, en una forma de realización, las proteínas de fusión comprenden el dominio de escisión (o medio dominio de escisión) de al menos una enzima de restricción de tipo IIS y uno o más dominios de unión a dedos de zinc, que pueden diseñarse o no.

[0131] Una enzima de restricción de tipo IIS ejemplar, cuyo dominio de escisión es separable del dominio de unión, es Fok I. Esta enzima particular es activa como un dímero. Bítnaíte et al. (1998) Proc. Natl. Acad. Sci. EE. UU. 95: 10.570-10.575. Por consiguiente, para los fines de la presente divulgación, la porción de la enzima Fok I utilizada en las proteínas de fusión divulgadas se considera un semidominio de escisión. Por lo tanto, para la escisión bicitaria dirigida y/o el reemplazo dirigido de secuencias celulares usando fusiones de dedos de zinc-Fok I, se pueden usar dos proteínas de fusión, cada una de las cuales comprende un medio dominio de escisión *FokI*, para reconstituir un dominio de escisión catalíticamente activo. Alternativamente, también se puede usar una única molécula polipeptídica que contiene un dominio de unión con dedos de zinc y dos semidominios de escisión de Fok I. Los parámetros para la escisión dirigida y la alteración de la secuencia dirigida usando fusiones de dedos de zinc-Fok I se proporcionan en otra parte de esta divulgación.

[0132] Un dominio de escisión o semidominio de escisión puede ser cualquier porción de una proteína que conserve la actividad de escisión, o que conserve la capacidad de multimerizarse (por ejemplo, dimerizar) para formar un dominio de escisión funcional.

[0133] En la publicación internacional WO 07/014275 se describen enzimas de restricción de tipo IIS ejemplares. Las enzimas de restricción adicionales también contienen dominios de unión y escisión separables, y estos están contemplados en la presente divulgación. Véase, por ejemplo, Roberts et al. (2003) Nucleic Acids Res. 31:418-420.

[0134] En ciertas formas de realización, el dominio de escisión comprende uno o más semidominios de escisión diseñados (también denominados mutantes del dominio de dimerización) que minimizan o previenen la homodimerización, como se describe, por ejemplo, en las patentes de EE. UU. N°s 7.914.796; 8.034.598 y 8.623.618; y publicación de patente estadounidense n.º 20110201055. Residuos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de Fok I son todos objetivos para influir en la dimerización de los semidominios de escisión de Fok I.

[0135] Los semidominios de escisión diseñados a modo de ejemplo de Fok I que forman heterodímeros obligados incluyen un par en el que un primer semidominio de escisión incluye mutaciones en residuos de aminoácidos en las posiciones 490 y 538 de Fok I y un segundo semidominio de escisión incluye mutaciones en los residuos de aminoácidos 486 y 499.

5

[0136] Así, en una forma de realización, una mutación en 490 reemplaza Glu (E) con Lys (K); la mutación en 538 reemplaza Iso (I) con Lys (K); la mutación en 486 reemplazó Gln (Q) con Glu (E); y la mutación en la posición 499 reemplaza Iso (I) con Lys (K). Específicamente, los semidominios de escisión diseñados descritos en este documento se prepararon mutando las posiciones 490 (E→K) y 538 (I→K) en un semidominio de escisión para producir un medio dominio de escisión diseñado designado "E490K:I538K" y mutando las posiciones 486 (Q→E) y 499 (I→L) en otro medio dominio de escisión para producir un medio dominio de escisión diseñado designado "Q486E:1499L". Los semidominios de escisión diseñados descritos en este documento son mutantes heterodímeros obligados en los que se minimiza o elimina la escisión aberrante. Véanse, por ejemplo, las patentes de EE. UU. nº 7.914.796 y 8.034.598. En determinadas formas de realización, el medio dominio de escisión diseñado comprende mutaciones en las posiciones 486, 499 y 496 (numeradas en relación con *FokI* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo Gln (Q) de tipo salvaje en la posición 486 con un Glu (E) residuo, el residuo Iso (I) de tipo salvaje en la posición 499 con un residuo Leu (L) y el residuo Asn (N) de tipo salvaje en la posición 496 con un residuo Asp (D) o Glu (E) (también denominado dominios "ELD" y "ELE", respectivamente). En otras formas de realización, el semidominio de escisión diseñado comprende mutaciones en las posiciones 490, 538 y 537 (numeradas en relación con *FokI* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo Glu (E) de tipo salvaje en la posición 490 con un Lys (K), el residuo Iso (I) de tipo salvaje en la posición 538 con un residuo Lys (K) y el residuo His (H) de tipo salvaje en la posición 537 con un residuo Lys (K) o un residuo Arg (R) (también conocidos como dominios "KKK" y "KKR", respectivamente). En otras formas de realización, el semidominio de escisión diseñado comprende mutaciones en las posiciones 490 y 537 (numeradas en relación con *FokI* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo Glu (E) de tipo salvaje en la posición 490 con un residuo Lys (K) y el residuo His (H) de tipo salvaje en la posición 537 con un residuo Lys (K) o un residuo Arg (R) (también denominados dominios "KIK" y "KIR", respectivamente). Véanse, por ejemplo, las patentes estadounidenses números 7.914.796; 8.034.598 y 8.623.618. En otras formas de realización, el medio dominio de escisión 2 diseñado comprende las mutaciones "Sharkey" y/o "Sharkey" (véase Guo et al, (2010) *J Mol. Biol.* 400(1):96-107).

10

[0137] Alternativamente, las nucleasas pueden ensamblarse *in vivo* en el sitio diana del ácido nucleico utilizando la denominada tecnología de "enzima dividida" (ver, por ejemplo, la publicación de patente de EE. UU. nº 20090068164). Los componentes de tales enzimas divididas pueden expresarse en construcciones de expresión separadas o pueden unirse en un marco de lectura abierto donde los componentes individuales están separados, por ejemplo, mediante un péptido 2A de autoescisión o una secuencia IRES. Los componentes pueden ser dominios de unión a dedos de zinc individuales o dominios de un dominio de unión a ácido nucleico de meganucleasa.

15

[0138] Las nucleasas (por ejemplo, ZFN y/o TALEN) se pueden seleccionar para determinar su actividad antes de su uso, por ejemplo, en un sistema cromosómico basado en levadura como se describe en la patente de EE. UU. nº 8.563.314.

20

[0139] En ciertas formas de realización, la nucleasa comprende un sistema CRISPR/Cas. El locus CRISPR (repeticiones palindrómicas cortas agrupadas regularmente interespaciadas), que codifica los componentes de ARN del sistema, y el locus Cas (asociado a CRISPR), que codifica proteínas (Jansen et al., 2002. *Mol. Microbiol.* 43: 1565-1575 Makarova et al., 2002. *Nucleic Acids Res.* 30: 482-496; Makarova et al., 2006. *Biol Direct* 1: 7; Haft et al., 2005. *PLoS Comput. Biol.* 1: e60) forman las secuencias genéticas del sistema de nucleasa CRISPR/Cas. Los loci CRISPR en huéspedes microbianos contienen una combinación de genes asociados a CRISPR (Cas), así como elementos de ARN no codificantes capaces de programar la especificidad de la escisión de ácidos nucleicos mediada por CRISPR.

25

[0140] El CRISPR de tipo II es uno de los sistemas mejor caracterizados y lleva a cabo la rotura dirigida de la doble cadena del ADN en cuatro pasos secuenciales. Primero, se transcriben dos ARN no codificantes, la matriz de pre-ARNcr y el ARNtracr, del locus CRISPR. En segundo lugar, el ARNtracr se hibrida con las regiones repetidas del pre-ARNcr y media el procesamiento del pre-ARNcr en ARNcr maduros que contienen secuencias espaciadoras individuales. En tercer lugar, el complejo ARNcr:ARNtracr maduro dirige Cas9 al ADN objetivo mediante el emparejamiento de bases Watson-Crick entre el espaciador del ARNcr y el protoespaciador en el ADN objetivo junto al motivo adyacente al protoespaciador (PAM), un requisito adicional para el reconocimiento del objetivo. Finalmente, Cas9 media la escisión del ADN objetivo para crear una rotura bicitánera dentro del protoespaciador. La actividad del sistema CRISPR/Cas consta de tres pasos: (i) inserción de secuencias de ADN extrañas en la matriz CRISPR para evitar futuros ataques, en un proceso llamado "adaptación", (ii) expresión de las proteínas relevantes, así como la expresión y procesamiento de la matriz, seguido de (iii) interferencia mediada por ARN con el ácido nucleico extraño. Así, en la célula bacteriana, varias de las llamadas proteínas 'Cas' están implicadas en la función natural del sistema CRISPR/Cas y desempeñan funciones como la inserción del ADN extraño, etc.

30

[0141] En determinadas formas de realización, la proteína Cas puede ser un "derivado funcional" de una proteína Cas de origen natural. Un "derivado funcional" de un polipéptido de secuencia nativa es un compuesto que tiene una propiedad biológica cualitativa en común con un polipéptido de secuencia nativa. Los "derivados funcionales" incluyen, entre otros,

fragmentos de una secuencia nativa y derivados de un polipéptido de secuencia nativa y sus fragmentos, siempre que tengan una actividad biológica en común con un polipéptido de secuencia nativa correspondiente. Una actividad biológica contemplada en el presente documento es la capacidad del derivado funcional de hidrolizar un sustrato de ADN en fragmentos. El término "derivado" abarca tanto variantes de secuencia de aminoácidos del polipéptido, como modificaciones covalentes y fusiones de los mismos. Los derivados adecuados de un polipéptido Cas o un fragmento del mismo incluyen, entre otros, mutantes, fusiones, modificaciones covalentes de la proteína Cas o un fragmento de la misma. La proteína Cas, que incluye la proteína Cas o un fragmento de la misma, así como derivados de la proteína Cas o un fragmento de la misma, puede obtenerse a partir de una célula o sintetizarse químicamente o mediante una combinación de estos dos procedimientos. La célula puede ser una célula que produce de forma natural proteína Cas, o una célula que produce de forma natural proteína Cas y está diseñada genéticamente para producir la proteína Cas endógena a un nivel de expresión más alto o para producir una proteína Cas a partir de un ácido nucleico introducido exógenamente, cuyo ácido nucleico codifica un Cas que es igual o diferente al Cas endógeno. En algunos casos, la célula no produce proteína Cas de forma natural y está diseñada genéticamente para producir una proteína Cas.

5 [0142] Sistemas de nucleasa CRISPR/Cas ejemplares dirigidos a genes TCR y otros genes se describen, por ejemplo, en la publicación de EE. UU. n.º 20150056705. La(s) nucleasa(s) pueden realizar uno o más cortes bicatenarios y/o monocatenarios en el sitio de destino. En determinadas formas de realización, la nucleasa comprende un dominio de escisión catalíticamente inactivo (por ejemplo, proteína *FokI* y/o Cas). Véanse, por ejemplo, la patente de EE. UU. n.º 9.200.266; 8.703.489 y Guillinger et al. (2014) *Nature Biotechnol.* 32(6):577-582. El dominio de escisión catalíticamente inactivo puede, en combinación con un dominio catalíticamente activo, actuar como una nickasa para realizar un corte monocatenario. Por lo tanto, se pueden usar dos nickasas en combinación para realizar un corte bicatenario en una región específica. También se conocen en la técnica nickasas adicionales, por ejemplo, McCaffery et al. (2016) *Nucleic Acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Publicación electrónica del 19 de octubre de 2015.

25 Entrega

30 [0143] Las proteínas (por ejemplo, factores de transcripción, nucleasas, moléculas TCR y CAR), polinucleótidos y/o composiciones que comprenden las proteínas y/o polinucleótidos descritos en el presente documento pueden administrarse a una célula diana mediante cualquier medio adecuado, incluyendo, por ejemplo, mediante inyección de la proteína y/o componentes de ARNm. En algunas formas de realización, las proteínas se introducen en la célula mediante compresión celular (véase Kollmannsperger et al. (2016) *Nat Comm* 7, 10372 doi:10.1038/ncomms10372).

35 [0144] Las células adecuadas incluyen, entre otras, células y/o líneas celulares eucarióticas y procarióticas. Los ejemplos no limitantes de dichas células o líneas celulares generadas a partir de dichas células incluyen células T, COS, CHO (p. ej., CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (p. ej., HEK293-F, HEK293-H, HEK293-T) y células perC6, así como células de insecto como *Spodoptera frugiperda* (Sf), o células fúngicas como *Saccharomyces*, *Pichia* y *Schizosaccharomyces*. En determinadas formas de realización, la línea celular es una línea celular CHO-K1, MDCK o HEK293. Las células adecuadas también incluyen células madre tales como, a modo de ejemplo, células madre pluripotentes inducidas (células iPS), células madre hematopoyéticas, células madre neuronales y células madre mesenquimales.

40 [0145] Los métodos para administrar proteínas que comprenden dominios de unión al ADN como se describen en el presente documento se describen, por ejemplo, en las patentes de EE. UU. N.ºs 6.453.242; 6.503.717; 6.534.261; 6.599.692; 6.607.882; 6.689.558; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824.

45 [0146] Los dominios de unión a ADN y las proteínas de fusión que comprenden estos dominios de unión a ADN como se describe en el presente documento también pueden administrarse usando vectores que contienen secuencias que codifican una o más de las proteínas de unión a ADN. Además, también se pueden administrar ácidos nucleicos adicionales (por ejemplo, donantes) a través de estos vectores. Puede usarse cualquier sistema de vectores, incluidos, entre otros, vectores plásmidos, vectores retrovirales, vectores lentivirales, vectores de adenovirus, vectores de poxvirus; vectores de herpesvirus y vectores de virus adenoasociados, etc. Véanse también las patentes de EE. UU. números 6.534.261; 26.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824. Además, será evidente que cualquiera de estos vectores puede comprender una o más secuencias codificantes de proteínas de unión al ADN y/o ácidos nucleicos adicionales, según corresponda. Por lo tanto, cuando se introducen en la célula una o más proteínas de unión a ADN como se describen en el presente documento, y ADN adicionales según corresponda, pueden transportarse en el mismo vector o en vectores diferentes. Cuando se usan múltiples vectores, cada vector puede comprender una secuencia que codifica una o múltiples proteínas de unión a ADN y ácidos nucleicos adicionales según se deseé.

50 [0147] Se pueden usar métodos convencionales de transferencia de genes basados en virus y no virus para introducir ácidos nucleicos que codifican proteínas de unión a ADN diseñadas en células (por ejemplo, células de mamífero) y tejidos diana y para co-introducir secuencias de nucleótidos adicionales según se deseé. Dichos métodos también se pueden usar para administrar ácidos nucleicos (por ejemplo, proteínas que codifican proteínas de unión a ADN y/o donantes) a células *in vitro*. En determinadas formas de realización, los ácidos nucleicos se administran para usos de terapia génica *in vivo* o *ex vivo*. Los sistemas de administración de vectores no virales incluyen plásmidos de ADN, ácido nucleico

- desnudo y ácido nucleico complejado con un vehículo de administración tal como un liposoma, una nanopartícula lipídica o un poloxámero. Los sistemas de administración de vectores virales incluyen virus de ADN y ARN, que tienen genomas episomales o integrados después de su administración a la célula. Para una revisión de los procedimientos de terapia génica, véase Anderson, *Science* 256:808-813 (1992); Nabel y Felgner, *TIBTECH* 11:211-217 (1993); Mitani y Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer y Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al., en *Current Topics in Microbiology and Immunology* Doerfler y Bohm (eds.) (1995); y Yu et al., *Gene Therapy* 1:13-26 (1994).
- 10 [0148] Los métodos de administración no viral de ácidos nucleicos incluyen electroporación, lipofección, microinyección, biolística, virosomas, liposomas, nanopartículas de lípidos, inmunoliposomas, policación o conjugados de lípido:ácido nucleico, ADN desnudo, ARNm, viriones artificiales y absorción mejorada por agentes de ADN. La sonoporación utilizando, por ejemplo, el sistema Sonitron 2000 (Rich-Mar) también puede usarse para la administración de ácidos nucleicos. En una forma de realización preferida, uno o más ácidos nucleicos se administran como ARNm. También se prefiere el uso de ARNm protegidos para aumentar la eficiencia traduccional y/o la estabilidad del ARNm. Se prefieren especialmente las tapas ARCA (análogo de tapa anti-inversión) o dos variantes de las mismas. Véanse las patentes estadounidenses números 7.074.596 y 8.153.773.
- 20 [0149] Los sistemas de administración de ácido nucleico ejemplares adicionales incluyen los proporcionados por Amixa Biosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) y Copernicus Therapeutics Inc, (ver, por ejemplo, el documento US6008336). La lipofección se describe, por ejemplo, en los documentos US 5.049.386, US 4.946.787; y US 4.897.355) y los reactivos de lipofección se venden comercialmente (por ejemplo, Transfectam™, Lipofectin™ y Lipofectamine™ RNAiMAX). Los lípidos catiónicos y neutros que son adecuados para una lipofección eficaz de polinucleótidos con reconocimiento de receptores incluyen los de Felgner, WO 91/17424, WO 91/16024. La administración puede ser a células (administración *ex vivo*) o tejidos diana (administración *in vivo*).
- 25 [0150] La preparación de complejos de lípido:ácido nucleico, incluidos liposomas dirigidos tales como complejos de inmunolípidos, es bien conocida por un experto en la técnica (véase, por ejemplo, Crystal, *Science* 270:404-410 (1995); Blaese et al., *Cancer Gene Ther.* 2:291-297 (1995); *Bioconjugate Chem.* 5:382-389 (1994); et al., *Gene Therapy* 2:710-722 (1995); Ahmad et al., *Cancer Res* 52:4817-4820 (1992, 4.485.054, 4.501,728, 4.774.085, 4.837.028 y 4.946.787).
- 30 [0151] Los métodos de administración adicionales incluyen el uso de empaquetar los ácidos nucleicos que se administrarán en vehículos de administración (EDV) EnGeneIC. Estos EDV se administran específicamente a los tejidos diana utilizando anticuerpos biespecíficos en los que un brazo del anticuerpo tiene especificidad por el tejido diana y el otro tiene especificidad por el EDV. El anticuerpo lleva los EDV a la superficie de la célula diana y luego el EDV ingresa a la célula mediante endocitosis. Una vez en la célula, el contenido se libera (ver, MacDiarmid et al. (2009) *Nature Biotechnology* 27(7) p. 643).
- 35 [0152] El uso de sistemas basados en ARN o ADN viral para la administración de ácidos nucleicos que codifican proteínas de unión a ADN diseñadas y/o donantes (por ejemplo, CAR o ACTR), según se desee, aprovecha procesos altamente evolucionados para dirigir un virus a células específicas en el cuerpo y transportando la carga viral al núcleo. Los vectores virales se pueden administrar directamente a los pacientes (*in vivo*) o se pueden usar para tratar células *in vitro* y las células modificadas se administran a los pacientes (*ex vivo*). Los sistemas basados en virus convencionales para la administración de ácidos nucleicos incluyen, entre otros, vectores de virus retrovirales, lentivirus, adenovirales, adenoasociados, vaccinia y herpes simple para transferencia de genes. La integración en el genoma del huésped es posible con los métodos de transferencia de genes de retrovirus, lentivirus y virus adenoasociados, lo que a menudo da como resultado una expresión a largo plazo del transgén insertado. Además, se han observado altas eficiencias de transducción en muchos tipos de células y tejidos diana diferentes.
- 40 [0153] El tropismo de un retrovirus se puede alterar incorporando proteínas de envoltura extrañas, ampliando la población diana potencial de células diana. Los vectores lentivirales son vectores retrovirales que pueden transducir o infectar células que no se dividen y normalmente producen títulos virales elevados. La selección de un sistema de transferencia de genes retrovirales depende del tejido diana. Los vectores retrovirales se componen de repeticiones terminales largas que actúan en *cis* y tienen capacidad de empaquetamiento para hasta 6-10 kb de secuencia extraña. Las LTR mínimas que actúan en *cis* son suficientes para la replicación y empaquetamiento de los vectores, que luego se usan para integrar el gen terapéutico en la célula diana para proporcionar una expresión transgénica permanente. Los vectores retrovirales ampliamente utilizados incluyen aquellos basados en el virus de la leucemia murina (MuLV), el virus de la leucemia del simio gibón (GaLV), el virus de la inmunodeficiencia simia (SIV), el virus de la inmunodeficiencia humana (VIH) y combinaciones de los mismos (véase, por ejemplo, Buchscher et al., *J Virol* 66:2731-2739 (1992); Johann et al., *J Virol* 66:1635-1640 (1992); *J Virol* 63:2374-2378 (1989); Miller et al., *J Virol* 65:2220-2224 (1991);
- 45 [0154] En aplicaciones en las que se prefiere la expresión transitoria, se pueden usar sistemas basados en adenovirus. Los vectores basados en adenovirus son capaces de lograr una eficiencia de transducción muy alta en muchos tipos de células y no requieren división celular. Con dichos vectores se han obtenido títulos elevados y niveles elevados de
- 50

- expresión. Este vector se puede producir en grandes cantidades en un sistema relativamente simple. Los vectores del virus asociado a Adena ("VAA") también se usan para transducir células con ácidos nucleicos diana, por ejemplo, en la producción *in vitro* de ácidos nucleicos y péptidos, y para procedimientos de terapia génica *in vivo* y *ex vivo* (véase, por ejemplo, West et al., *Virology* 160:38-47 (1987); Patente de EE. UU. nº 4.797.368; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzychka, *J. Clin. Invest.* 94:1351 (1994). La construcción de vectores VAA recombinantes se describe en varias publicaciones, incluyendo la patente de EE. UU. nº 5.173.414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzychka, *PNAS USA* 81:6466-6470 (1984); y Samulski et al., *J. Virol.* 63:03822-3828 (1989).
- 5 [0155] Actualmente están disponibles al menos seis enfoques de vectores virales para la transferencia de genes en ensayos clínicos, que utilizan enfoques que implican la complementación de vectores defectuosos mediante genes insertados en líneas celulares auxiliares para generar el agente transductor.
- 10 [0156] pLASN y MFG-S son ejemplos de vectores retrovirales que se han utilizado en ensayos clínicos (Dunbar et al., *Blood* 85:3048-305 (1995); Kohn et al., *Nat. Med.* 1: 1017-102 (1995); Malech et al., *PNAS USA* 94:22 12133-12138 (1997)). PA317/pLASN fue el primer vector terapéutico utilizado en un ensayo de terapia génica. (Blaese et al., *Science* 270:475-480 (1995)). Se han observado eficiencias de transducción del 50 % o más para los vectores empaquetados MFG-S. (Ellem et al., *Immunol Immunother.* 44(1):10-20 (1997); Dranoff et al., *Hum. Gene Ther.* 1:111-2 (1997)).
- 15 [0157] Los vectores de virus adenoasociados recombinantes (VAAr) son sistemas de administración de genes alternativos prometedores basados en el virus adenoasociado de tipo 2 parvovirus defectuoso y no patógeno. Todos los vectores derivan de un plásmido que retiene sólo las repeticiones terminales invertidas de 145 pb de VAA que flanquean el casete de expresión transgénica. La transferencia eficiente de genes y la entrega estable de transgenes debido a la integración en los genomas de la célula transducida son características clave de este sistema de vector. (Wagner et al., *Lancet* 351:9117 1702-3 (1998), Kearns et al., *Gene Ther.* 9:748-55 (1996)). También se pueden usar otros serotipos de VAA, incluidos VAA1, VAA3, VAA4, VAA5, VAA6, VAA8, VAA8.2, VAA9 y VAArh10 y VAA pseudotipificados tales como VAA2/8, VAA2/5 y VAA2/6 de acuerdo con la presente invención.
- 20 [0158] Los vectores adenovirales recombinantes (Ad) con replicación deficiente pueden producirse con títulos elevados e infectar fácilmente varios tipos de células diferentes. La mayoría de los vectores de adenovirus están diseñados de manera que un transgén reemplace los genes Ad E1a, E1b y/o E3; posteriormente, el vector de replicación defectuosa se propaga en células 293 humanas que suministran la función del gen eliminado en trans. Los vectores publicitarios pueden transducir múltiples tipos de tejidos *in vivo*, incluidas células diferenciadas que no se dividen, como las que se encuentran en el hígado, los riñones y los músculos. Los vectores publicitarios convencionales tienen una gran capacidad de carga. 25 Un ejemplo del uso de un vector Ad en un ensayo clínico implicó la terapia con polinucleótidos para la inmunización antitumoral con inyección intramuscular (Sterman et al., *Hum. Gene Ther.* 7:1083-9 (1998)). Ejemplos adicionales del uso de vectores de adenovirus para la transferencia de genes en ensayos clínicos incluyen Rosenecker et al., *Infection* 24:1 5-10 (1996); Sterman et al., *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh et al., *Hum. Gene Ther.* 2:205-18 (1995); Alvarez et al., *Hum. Gene Ther.* 5:597-613 (1997); Topf et al., *Gene Ther.* 5:507-513 (1998); Sterman et al., *Hum. Gene Ther.* 7:1083-1089 (1998).
- 30 [0159] Las células empaquetadoras se usan para formar partículas de virus que son capaces de infectar una célula huésped. Dichas células incluyen 293 células, que empaquetan adenovirus, y células ψ2 o células PA317, que empaquetan retrovirus. Los vectores virales utilizados en terapia génica generalmente son generados por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula viral. Los vectores normalmente contienen las secuencias virales mínimas requeridas para el empaquetamiento y la posterior integración en un huésped (si corresponde), siendo reemplazadas otras secuencias virales por un casete de expresión que codifica la proteína que se va a expresar. Las funciones virales que faltan son aportadas en trans por la línea celular empaquetadora. Por ejemplo, los vectores de VAA utilizados en terapia génica normalmente solo poseen secuencias de repetición terminal invertida (ITR) del genoma de VAA que son necesarias para el empaquetado y la integración en el genoma del huésped. El ADN viral está empaquetado en una línea celular que contiene un plásmido auxiliar que codifica los otros genes de VAA, a saber, rep y cap, pero que carece de secuencias ITR. La línea celular también está infectada con adenovirus como ayudante. El virus auxiliar promueve la replicación del vector VAA y la expresión de genes VAA a partir del plásmido auxiliar. El plásmido auxiliar no está empaquetado en cantidades significativas debido a la falta de secuencias ITR. La 35 contaminación con adenovirus se puede reducir, por ejemplo, mediante un tratamiento térmico al que el adenovirus es más sensible que el VAA. Además, el VAA se puede fabricar utilizando un sistema de baculovirus (véanse, por ejemplo, las patentes estadounidenses 6.723.551 y 7.271.002).
- 40 [0160] La purificación de partículas de VAA de un sistema 293 o baculovirus normalmente implica el crecimiento de las células que producen el virus, seguido de la recolección de las partículas virales del sobrenadante celular o la lisis de las células y la recolección del virus del lisado crudo. A continuación, el VAA se purifica mediante métodos conocidos en la técnica que incluyen cromatografía de intercambio iónico (p. ej., véanse las patentes de EE. UU. 7.419.817 y 6.989.264), cromatografía de intercambio iónico y centrifugación de densidad en CsCl (p. ej., publicación PCT WO2011094198A10), cromatografía de inmunoafinidad (p. ej., WO2016128408) o purificación usando AVB Sepharose (por ejemplo, GE Healthcare Life Sciences).
- 45
- 50
- 55
- 60
- 65

5 [0161] En muchas aplicaciones de terapia génica, es deseable que el vector de terapia génica se suministre con un alto grado de especificidad a un tipo de tejido particular. En consecuencia, se puede modificar un vector viral para que tenga especificidad por un tipo de célula determinado expresando un ligando como una proteína de fusión con una proteína de cubierta viral en la superficie exterior del virus. El ligando se elige para que tenga afinidad por un receptor que se sabe que está presente en el tipo de célula de interés. Por ejemplo, Han et al., (Proc. Natl. Acad. Sci. EE. UU. 92:9747-9751 (1995)), informaron que el virus de la leucemia murina de Moloney puede modificarse para expresar heregulina humana fusionada a gp70, y el virus recombinante infecta ciertas células de cáncer de mama humano que expresan el receptor del factor de crecimiento epidérmico humano. Este principio se puede extender a otros pares de células diana-virus, en 10 los que la célula diana expresa un receptor y el virus expresa una proteína de fusión que comprende un ligando para el receptor de la superficie celular. Por ejemplo, se pueden diseñar fagos filamentosos para que presenten fragmentos de anticuerpos (por ejemplo, FAB o Fv) que tengan afinidad de unión específica por prácticamente cualquier receptor celular elegido. Aunque la descripción anterior se aplica principalmente a vectores virales, se pueden aplicar los mismos principios 15 a vectores no virales. Dichos vectores pueden diseñarse para que contengan secuencias de captación específicas que favorezcan la captación por células diana específicas.

20 [0162] Los vectores de terapia génica pueden administrarse *in vivo* mediante administración a un paciente individual, típicamente mediante administración sistémica (por ejemplo, infusión intravenosa, intraperitoneal, intramuscular, subdérmica o intracranial) o aplicación tópica, como se describe a continuación. Alternativamente, los vectores pueden administrarse a células *ex vivo*, tales como células explantadas de un paciente individual (por ejemplo, linfocitos, aspirados de médula ósea, biopsia de tejido) o células madre hematopoyéticas de donante universal, seguido de la reimplantación de las células en un paciente normalmente después de la selección de células que han incorporado el vector.

25 [0163] La transfección de células *ex vivo* para diagnóstico, investigación, trasplante o terapia génica (por ejemplo, mediante reinfusión de las células transfectadas en el organismo huésped) es bien conocida por los expertos en la técnica. En una forma de realización preferida, las células se aislan del organismo en cuestión, se transfectan con un ácido nucleico de proteínas de unión a ADN (gen o ADNc) y se vuelven a infundir en el organismo en cuestión (por ejemplo, paciente). Los expertos en la técnica conocen bien diversos tipos de células adecuados para la transfección *ex vivo* (véase, por ejemplo, Freshney et al., Culture of Animal Cells, A Manual of Basic Technique (3^a ed. 1994)) y las referencias 30 citadas en el mismo para una discusión sobre cómo aislar y cultivar células de pacientes).

35 [0164] En una forma de realización, las células madre se usan en procedimientos *ex vivo* para transfección celular y terapia génica. La ventaja de utilizar células madre es que pueden diferenciarse en otros tipos de células *in vitro* o pueden introducirse en un mamífero (como el donante de las células) donde se injertarán en la médula ósea. Se conocen métodos para diferenciar células CD34+ *in vitro* en tipos de células inmunitarias clínicamente importantes utilizando citocinas tales como GM-CSF, IFN-γ y TNF-α (véase Inaba et al., J. Exp. Med. 176:1693-1702 (1992)).

40 [0165] Se aíslan células madre para transducción y diferenciación usando métodos conocidos. Por ejemplo, las células madre se aíslan de células de la médula ósea analizando las células de la médula ósea con anticuerpos que se unen a células no deseadas, como CD4+ y CD8+ (células T), CD45+ (células panB), GR-1 (granulocitos) y Iad (células presentadoras de antígeno diferenciadas) (véase Inaba et al., J. Exp. Med. 176:1693-1702 (1992)).

45 [0166] Las células madre que han sido modificadas también se pueden usar en algunas formas de realización. Por ejemplo, las células madre neuronales que se han hecho resistentes a la apoptosis pueden usarse como composiciones terapéuticas donde las células madre también contienen los TF ZFP de la invención. La resistencia a la apoptosis puede producirse, por ejemplo, eliminando BAX y/o BAK utilizando ZFN específicas de BAX o BAK (véase la solicitud de patente estadounidense nº 12/456.043) en las células madre, o aquellas que se alteran en un caspasa, nuevamente usando ZFN específicas de caspasa-6, por ejemplo. Estas células pueden transfectarse con los TF ZFP que se sabe que regulan el TCR.

50 [0167] Los vectores (por ejemplo, retrovirus, adenovirus, liposomas, etc.) que contienen proteínas terapéuticas de unión a ADN (o ácidos nucleicos que codifican estas proteínas) también se pueden administrar directamente a un organismo para la transducción de células *in vivo*. Alternativamente, se puede administrar ADN desnudo. La administración se realiza mediante cualquiera de las vías normalmente utilizadas para introducir una molécula en contacto final con células sanguíneas o tisulares, incluidas, entre otras, inyección, infusión, aplicación tópica y electroporación. Los métodos adecuados para administrar tales ácidos nucleicos están disponibles y son bien conocidos por los expertos en la técnica y, aunque se puede usar más de una ruta para administrar una composición particular, una ruta particular a menudo puede proporcionar una reacción más inmediata y más efectiva que otra ruta.

60 [0168] Los métodos para la introducción de ADN en células madre hematopoyéticas se describen, por ejemplo, en la patente de EE. UU. nº 5.928.638. Los vectores útiles para la introducción de transgenes en células madre hematopoyéticas, por ejemplo, células CD34+, incluyen adenovirus tipo 35.

65 [0169] Los vectores adecuados para la introducción de transgenes en células inmunitarias (por ejemplo, células T) incluyen vectores de lentivirus no integrantes. Véase, por ejemplo, Ory et al. (1996) Proc. Natl. Acad. Sci. EE. UU.

93:11382-11388; Dull et al. (1998) J. Virol. 72:8463-8471; Zuffery et al. (1998) J Virol. 72:9873-9880; Follenzi et al. (2000) Nature Genetics 25:217-222.

5 [0170] Los vehículos farmacéuticamente aceptables se determinan en parte por la composición particular que se administra, así como por el método particular usado para administrar la composición. Por consiguiente, existe una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas disponibles, como se describe a continuación (véase, por ejemplo, Remington's Pharmaceutical Sciences, 17^a ed., 1989).

10 [0171] Como se señaló anteriormente, los métodos y composiciones divulgados se pueden usar en cualquier tipo de célula, incluidas, entre otras, células procarióticas, células fúngicas, células de arqueas, células vegetales, células de insectos, células animales, células de vertebrados, células de mamíferos y células humanas, incluidas células T y células madre de cualquier tipo. Los expertos en la técnica conocen líneas celulares adecuadas para la expresión de proteínas e incluyen, entre otras, COS, CHO (por ejemplo, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11), VERO, MDCK, WI38, V79, 15 BI4AF28-G3, BHK, HaK, NS0, SP2/0-Agl4, HeLa, HEK293 (p. ej., HEK293-F, HEK293-H, HEK293-T), perC6, células de insectos como *Spodoptera frugiperda* (Sf), y células de hongos como *Saccharomyces*, *Pichia* y *Schizosaccharomyces*. También se pueden utilizar progenie, variantes y derivados de estas líneas celulares.

Aplicaciones

20 [0172] Las composiciones y métodos divulgados se pueden usar para cualquier aplicación en la que se desee modular la expresión y/o funcionalidad de TCR, incluidas, entre otras, aplicaciones terapéuticas y de investigación en las que es deseable la modulación de TCR. Por ejemplo, las composiciones divulgadas se pueden usar *in vivo* y/o *ex vivo* (terapias celulares) para alterar la expresión de TCR endógenos en células T modificadas para terapia celular adoptiva para expresar uno o más CAR exógenos, TCR exógenos u otras moléculas receptoras específicas a cáncer, tratando y/o previniendo así el cáncer. Además, en tales entornos, la anulación de la expresión de TCR dentro de una célula puede eliminar o reducir sustancialmente el riesgo de una reacción cruzada no deseada con tejido sano no objetivo (es decir, una respuesta de injerto contra huésped).

30 [0173] Los métodos y composiciones también incluyen composiciones de células madre en las que los genes TCRA y/o TCRB dentro de las células madre han sido modulados (modificados) y las células comprenden además un ACTR y/o un CAR y/o un TCR aislado o diseñado. Por ejemplo, se pueden introducir células madre hematopoyéticas alogénicas moduladas con desactivación o desactivación de TCR en un paciente compatible con HLA después de la ablación de la médula ósea. Estas HSC alteradas permitirían la recolonización del paciente, pero no causarían una posible GvHD. Las 35 células introducidas también pueden tener otras alteraciones para ayudar durante la terapia posterior (por ejemplo, resistencia a la quimioterapia) para tratar la enfermedad subyacente. Las células nulas de TCR también se utilizan como terapia "lista para usar" en situaciones de sala de emergencias con pacientes traumatizados.

[0174] Los métodos y composiciones de la invención también son útiles para el diseño e implementación de modelos *in vitro* e *in vivo*, por ejemplo, modelos animales de TCR o trastornos asociados, lo que permite el estudio de estos trastornos.

EJEMPLOS

Ejemplo 1: Diseño de nucleasas específicas de TCR

45 [0175] Se construyeron ZFN específicas de TCR para permitir la introducción específica de sitio de roturas de doble hebra en el gen TCRA (TCRA). Las ZFN se diseñaron esencialmente como se describe en Urnov et al. (2005) Nature 435(7042):646-651, Lombardo et al. (2007) Nat Biotechnol. Nov;25(11): 1298-306, y publicaciones de patentes de EE. UU. 2008-0131962, 2015-016495, 2014-0120622, 2014-0301990 y patente de EE. UU. 8,956,828. Los pares de ZFN se dirigieron a diferentes sitios en la región constante del gen TCRA (ver Figura 1). Las hélices de reconocimiento para pares de ZFN ejemplares, así como la secuencia objetivo, se muestran a continuación en la Tabla 1. Los sitios objetivo de los diseños de dedos de zinc TCRA se muestran en la primera columna. Los nucleótidos en el sitio objetivo a los que se dirigen las hélices de reconocimiento de ZFP se indican en letras mayúsculas; nucleótidos no dirigidos indicados en minúsculas. También se muestran los conectores utilizados para unir el dominio de nucleasa *FokI* y el dominio de unión al ADN de ZFP (consulte la publicación de patente de EE. UU. 20150132269). Por ejemplo, la secuencia de aminoácidos 50 del conector de dominio LO es dominio de unión a ADN-dominio de nucleasa QLVKS-*FokI* (SEQ ID NO:5). De manera similar, las secuencias de aminoácidos para el conector de dominio N7a es el dominio de unión a ADN SGTPHEVGVTL 55 del dominio de nucleasa *FokI* (SEQ ID NO:6), y N7c es el dominio de unión a ADN SGAIRCHDEFWF-dominio de nucleasa *FokI* (SEQ ID NO:7).

60

65

Tabla 1: Diseños con dedos de zinc de TCR- α (TCRA)

Secuencia Diana Nombre ZFN	F1	F2	F3	F4	F5	F6	Entazador de dominio
SSS55204 5' ttGCTG TTGAGTC cATAGACC tcatgt (SEQ ID NO:8)	DRSNLSR (SEQ ID NO:22)	QSVTLAA (SEQ ID NO:23)	DP.SALSR (SEQ ID NO:24)	TSGNLTR (SEQ ID NO:25)	YRSSLKE (SEQ ID NO:26)	TGNNLTP (SEQ ID NO:25)	LO
SBS53759 5' gtGCTG TGGCCTGG	CCNVLIN (SEQ ID NO:27)	QNATRTX (SEQ ID NO:28)	QSGHLAR (SEQ ID NO:29)	NRYULMT (SEQ ID NO:30)	KGSLLR (SEQ ID NO:31)	QSSDLTP (SEQ ID NO:32)	LO

ES 2 983 043 T3

(Continuación)

ACCAACAA atctga (SEQ ID NO:9)							
SSS55229 5' ctGTTG CTGTTGAA GTCCatag acccca (SEQ ID NO:10)	DRSALAR (SEQ ID NO:33)	QSGNLAR (SEQ ID NO:34)	HRSTFLQG (SEQ ID NO:35)	QSGDLTR (SEQ ID NO:36)	TSGSLTR (SEQ ID NO:37)	NA	LO
SSS55765 5' ctGTGG CCTGGAGC AACAAatc tgactt (SEQ ID NO:11)	QNQVILVR (SEQ ID NO:38)	QNATRTK (SEQ ID NO:28)	QSGRLSR (SEQ ID NO:39)	DRSDLSR (SEQ ID NO:40)	RSDALAR (SEQ ID NO:41)	NA	LO
SSS53810 5' agGATT CGGAACCC AATCACTg (SEQ ID NO:12)	DQSNLRA (SEQ ID NO:42)	TSSNRKT (SEQ ID NO:43)	DSSTRKT (SEQ ID NO:44)	QSGNLAR (SEQ ID NO:34)	RSDDLSE (SEQ ID NO:45)	TNSNRKR (SEQ ID NO:46)	LO
SSS55255 5' ctCCTG AAAGTGCC CGGgtta atctgc (SEQ ID NO:13)	RSDDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LKQHLNE (SEQ ID NO:49)	TSGNLTR (SEQ ID NO:25)	HRDSLTD (SEQ ID NO:50)	NA	LO
SSS55248 5' agGATT CGGAACCC AATCACTg acaggt (SEQ ID NO:14)	DQSNLRA (SEQ ID NO:42)	TSSNRKT (SEQ ID NO:43)	LQQTLAD (SEQ ID NO:51)	QSGNLAR (SEQ ID NO:34)	RREDLIT (SEQ ID NO:52)	TSSNLSE (SEQ ID NO:53)	LO
SSS55254 5' ctCCTG AAAGTGCC CGGgtta atctgc (SEQ ID NO:13)	RSDDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LKQHLNE (SEQ ID NO:49)	QSGNLAR (SEQ ID NO:34)	HNSSLKD (SEQ ID NO:54)	NA	LO
SSS55260 5' ctCCTG AAAGTGCC CGGgtta atctgc (SEQ ID NO:13)	RSDDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LNHHLQQ (SEQ ID NO:55)	QSGNLAR (SEQ ID NO:34)	HKTSLLKD (SEQ ID NO:56)	NA	LO
SSS55266 5' tcAAGC TGGTCGAG aAAAGCTt tgaaac (SEQ ID NO:15)	QSDLSR (SEQ ID NO:57)	QSGRRPT (SEQ ID NO:58)	RSANLAK (SEQ ID NO:59)	DRSALAR (SEQ ID NO:33)	RSIVLSE (SEQ ID NO:60)	KHSTRRV (SEQ ID NO:61)	NTc
SSS53853 5' aaCAGG	TNGQRVE (SEQ ID	TSGNLSE (SEQ ID	RSINLHQ (SEQ ID	DSANLSE (SEQ ID	QSGSLTR (SEQ ID	AKWNLDA (SEQ ID	LO

ES 2 983 043 T3

(Continuación)

TAGGACAG GGGTCTA ^g cctggg (SEQ ID NO:16)	NO:62)	NO:63)	NO:64)	NO:65)	NO:66)	NO:67)	
SBS53860 5' ctGTGC TAGACATG aGGTCTAt ggactt (SEQ ID NO:17)	TMHQRVE (SEQ ID NO:62)	TSGHLSR (SEQ ID NO:63)	RNDLSK ^t (SEQ ID NO:68)	DSSNLSR (SEQ ID NO:69)	QKATRTT (SEQ ID NO:70)	RNASRTE (SEQ ID NO:72)	N7a
SBS53863 5' ttCAAG ACCAACAG tGCTGTGg cctgga (SEQ ID NO:18)	RSDSLLR (SEQ ID NO:31)	QSSDLRK (SEQ ID NO:73)	RSDNLSE (SEQ ID NO:74)	ERANRNS (SEQ ID NO:75)	RSDNLAR (SEQ ID NO:76)	QKVNLMS (SEQ ID NO:77)	LG
SBS55287 5' ttCAAG ACCAACAG tGCTGTGg cctgga (SEQ ID NO:18)	R3DSLLR (SEQ ID NO:31)	QSSDLPR (SEQ ID NO:73)	RSDNLSE (SEQ ID NO:74)	ERANPNS (SEQ ID NO:75)	RSDNLAR (SEQ ID NO:76)	QKVNLRE (SEQ ID NO:78)	LG
SBS53853 5' ctGTGC TAGACATG aGGTCTAt ggactt (SEQ ID NO:17)	TMHQRVE (SEQ ID NO:62)	TSGHLSR (SEQ ID NO:63)	RSDTLSQ (SEQ ID NO:79)	RSDLSR (SEQ ID NO:40)	QKATRTT (SEQ ID NO:70)	RNASRTR (SEQ ID NO:72)	N7a
SBS52774 5' ccTGTG AGtGATTG GCTTCCG ^a atccctc (SEQ ID NO:19)	RKOTPTT (SEQ ID NO:80)	HRSSLPR (SEQ ID NO:81)	RSDHLS ^t (SEQ ID NO:47)	TSANLSR (SEQ ID NO:82)	RSDNLSE (SEQ ID NO:74)	WHSSLRV (SEQ ID NO:83)	N7a
SBS53809 5' tcCTCC TGAAAATG CCCGGGtt taatct (SEQ ID NO:20)	RSAAHLSR (SEQ ID NO:84)	DRSDLSR (SEQ ID NO:40)	RSDVLSV (SEQ ID NO:85)	QNNHRIT (SEQ ID NO:86)	RSDVLSE (SEQ ID NO:60)	SPSSERT (SEQ ID NO:87)	LG
SBS52742 5' tcCTCC TGAAAATG CCCGGGtt taatct (SEQ ID NO:20)	RSAAHLSR (SEQ ID NO:84)	DRSDLSR (SEQ ID NO:40)	RSDSLSV (SEQ ID NO:88)	QNANPKT (SEQ ID NO:89)	RSDVLSE (SEQ ID NO:60)	SPSSERT (SEQ ID NO:87)	LG
SBS53856 5' ctGTGC TAGACATG aGGTCTAt g (SEQ ID NO:21)	TMHQRVE (SEQ ID NO:62)	TSGHLSR (SEQ ID NO:63)	RSDSLS ^t (SEQ ID NO:90)	DRANRIK (SEQ ID NO:91)	QKATRTT (SEQ ID NO:70)	RNASRTR (SEQ ID NO:72)	N7a

[0176] Se ensayaron todas las ZFN y se encontró que se unían a sus sitios diana y que eran activas como nucleasas.

[0177] También se construyeron ARN guía para el sistema CRISPR/Cas9 de *S. pyogenes* para dirigirse al gen TCRA. Véase también la publicación de EE. UU. n.º 201500566705 para obtener ARN guía dirigidos alfa de TCR adicionales.

5 Las secuencias diana en el gen TCRA se indican, así como las secuencias de ARN guía en la Tabla 2 a continuación. Todos los ARN guía se prueban en el sistema CRISPR/Cas9 y se encuentran activos.

Tabla 2: ARN guía para la región constante de TCRA humano (TRAC)

Nombre	Hebra	Diana (5' → 3')	ARNg (5' → 3')
TRAC-Gr14	R	GCTGGTACACGGCAGGGTCAGGG (SEQ ID NO:92)	GCTGGTACACGGCAGGGTCAGGG (SEQ ID NO:104)
TRAC-Gr25	R	AGAGCTCTCTCAAGCTGGTACAGGG (SEQ ID NO:93)	AGAGCTCTCTCAAGCTGGTACAGGG (SEQ ID NO:105)
TRAC-Gr71	R	GAGAATCAAATCGGTCAATAGG (SEQ ID NO:94)	GAGAATCAAATCGGTCAATAGG (SEQ ID NO:106)
TRAC-Gr155	R	ACAAAAACTGCTAGACATGAGG (SEQ ID NO:95)	ACAAAAACTGCTAGACATGAGG (SEQ ID NO:107)
TRAC-Gr191	R	AGAGCAACAGTGTCTGTGGCCTGG (SEQ ID NO:96)	AGAGCAACAGTGTCTGTGGCCTGG (SEQ ID NO:108)
TRAC-Gr271	R	GACACCTTCTTCCCCAGCCCCAGG (SEQ ID NO:97)	GACACCTTCTTCCCCAGCCCCAGG (SEQ ID NO:109)
TRAC-Gr2146	R	CTCGGACCAAGCTTGACATCACAGG (SEQ ID NO:98)	CTCGGACCAAGCTTGACATCACAGG (SEQ ID NO:110)
TRAC-Gr2187	R	AAGTTCCPTGTGATGTCAGGCTGG (SEQ ID NO:99)	AAGTTCCPTGTGATGTCAGGCTGG (SEQ ID NO:111)
TRAC-Gr2179	R	GTCCGAGAAAAGCTTTGAACAGGG (SEQ ID NO:100)	GTCCGAGAAAAGCTTTGAACAGGG (SEQ ID NO:112)
TRAC-Gr3081	R	TTCGGAAUCCAACTCACTCACAGG (SEQ ID NO:101)	TTCGGAAUCCAACTCACTCACAGG (SEQ ID NO:113)
TRAC-Gr3099	R	CCACTTTCAAGGAGGAGGATTGG (SEQ ID NO:102)	CCACTTTCAAGGAGGAGGATTGG (SEQ ID NO:114)
TRAC-Gr3105	R	ACCCGGGGCACTTTCAAGGAGGAGG (SEQ ID NO:103)	ACCCGGGGCACTTTCAAGGAGGAGG (SEQ ID NO:115)

[0178] Por tanto, las nucleasas descritas en el presente documento (por ejemplo, nucleasas que comprenden una ZFP o un dominio de unión a ADN de ARNg) se unen a sus sitios diana y escinden el gen TCRA, realizando de este modo modificaciones genéticas dentro de un gen TCRA que comprende cualquiera de las SEQ ID NO:6-48 o 137-205, incluyendo modificaciones (inserciones y/o eliminaciones) dentro de cualquiera de estas secuencias (SEQ ID NO:8-21 y/o 92-103) y/o modificaciones dentro de las siguientes secuencias: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT y/o AATCCTC (ver Figura 1B). Las nucleasas TALE dirigidas a estos sitios objetivo también están diseñadas y resultan funcionales en términos de unión y actividad.

[0179] Además, todos los dominios de unión a ADN (ZFP y ARNg) se unen a sus sitios objetivo y los dominios de unión a ADN ZFP, TALE y ARNs que reconocen estos sitios objetivo también se formulan en factores de transcripción activos diseñados cuando se asocian con uno o más dominios reguladores transcripcionales.

Ejemplo 2: Actividad nucleasa *in vitro*

[0180] Las ZFN descritas en la Tabla 1 se usaron para probar la actividad nucleasa en células K562. Para ensayar la actividad de escisión, los plásmidos que codifican los pares de ZFN específicos de TCRA humano descritos anteriormente se transfecaron en células K562 con plásmido o ARNm. Las células K562 se obtuvieron de la American Type Culture Collection y se cultivaron según lo recomendado en medio RPMI (Invitrogen) suplementado con suero bovino fetal calificado al 10% (FBS, Cyclone). Para la transfección, se clonaron ORF para las nucleasas activas enumeradas en la Tabla 1 en un vector de expresión optimizado para la producción de ARNm que porta UTR 5' y 3' y una señal poliA sintética. Los ARNm se generaron utilizando el kit mMessage mMachine T7 Ultra (Ambion) siguiendo las instrucciones del fabricante. La síntesis *in vitro* de ARNm de nucleasa utilizó un vector basado en pVAX que contiene un promotor T7,

la nucleasa propiamente dicha y un motivo poliA para la adición enzimática de una cola poliA después de la reacción de transcripción *in vitro*, o un vector basado en pGEM que contiene un promotor T7, un 5'UTR, la nucleasa propiamente dicha, una 3'UTR y una extensión poliA de 64 pb, o un amplicón de PCR que contiene un promotor T7, una 5'UTR, la nucleasa propiamente dicha, una 3'UTR y una extensión poliA de 60 pb. Se mezclaron un millón de células K562 con 250 ng o 500 ng del ARNm que codifica ZFN. Las células se transfectaron en un Amaxa Nucleofector IITM usando el programa T-16 y se recuperaron en 1,4 mL de medio RPMI tibio + FBS al 10 %. La actividad nucleasa se evaluó mediante secuenciación profunda (MiSeq, Illumina) según los protocolos estándar tres días después de la transfección. Los resultados se presentan a continuación en la Tabla 3.

10

Tabla 3: Actividad de la nucleasa de dedos de zinc

Par N°	Par ZFN	NHEJ% (250ng/ZFN)	DE	NHEJ% (500ng/ZFN)	DE	Sitio
1	55204;55759	76.7	1.3	87.7	1	AC
2	55219;55760	91.4	1.5	93.6	1.7	B
3	55210;55255	81.6	0.6	91.5	1.3	C1
4	55248;55154	98.4	1.8	98.2	1.2	C2
5	55249;55280	87.8	1.3	93.8	1	D3
6	55266;55393	85.8	1.4	88.9	0.4	E
7	55360;55363	77.1	1.7	87.8	1.1	F1
8	55386;55207	83.6	3.2	74.8	1.3	F2
9	55885;55303	98.1	1.6	90.2	1.8	G1
10	52774;52780	78.6	0.8	84.3	2.8	G0
11	G07	0		0		

15

[0181] También se han descrito previamente TALEN específicos de TCRA altamente activos (ver documento WO2014153470).

30

[0182] También se ensayaron los sistemas CRISPR/Cas9 específicos de TCRA humano. La actividad de los sistemas CRISPR/Cas9 en células K562 humanas se midió mediante análisis MiSeq. La escisión de la secuencia de ADN de TCRA endógena por Cas9 se analiza mediante secuenciación de alto rendimiento (MiSeq, Illumina).

35

[0183] En estos experimentos, Cas9 se suministró en un plásmido pVAX, y el ARN_g se suministra en un plásmido bajo el control de un promotor (por ejemplo, el promotor U6 o un promotor CMV). Los plásmidos se mezclaron a 100 ng de cada uno o 400 ng de cada uno y se mezclaron con 2e5 células por ejecución. Las células se transfectaron utilizando el sistema Amaxa. Brevemente, se utiliza un kit de transfección Amaxa y los ácidos nucleicos se transfieren utilizando un protocolo lanzadera Amaxa estándar. Después de la primera transfección, las células se dejan reposar durante 10 minutos a temperatura ambiente y luego se resuspenden en RPMI precalentado. A continuación, las células se cultivan en condiciones estándar a 37 °C. El ADN genómico se aisló 7 días después de la transfección y se sometió a análisis MiSeq.

40

[0184] Brevemente, se analizó la actividad de los ARN guía enumerados en la Tabla 2. Los ARN guía se probaron en tres configuraciones diferentes: G0 es la configuración descrita anteriormente. G1 usó un vector pVAX que comprende un promotor CMV que impulsa la expresión del gen Cas9 y un casete de expresión de trazador de ARN guía U6 donde la transcripción de ambos marcos de lectura está en la misma orientación. G2 es similar a G1 excepto que los cassetes de expresión Cas9 y U6-Guide están en orientaciones opuestas. Estas tres configuraciones se probaron utilizando 100 ng o 400 ng de ADN transfecido, y los resultados se presentan a continuación en la Tabla 4. Los resultados se expresan como 'porcentaje de indeles' o "NHEJ%", donde 'indeles' significa pequeñas inserciones y/o eliminaciones encontradas como resultado del proceso de reparación de NHEJ propenso a errores en el sitio de una escisión de doble cadena inducida por nucleasa.

Tabla 4: Actividad CRISPR/Cas

5	Guía utilizada	% total-indeles					
		GRO		GRI		GR2	
		NHEJ%	NHEJ%	NHEJ%	NHEJ%	NHEJ%	NHEJ%
10	TCRA-Gr14	6.4	25.8	0.6	12.4	0.5	10.2
15	TCRA-Gr25	14.6	26.9	2.4	21.7	1.1	21.6
20	TCRA-Gr71	3.7	13.8	0.3	4.2	0.3	7.8
25	TCRA-Gr155	6.0	19.9	1.2	12.7	0.8	15.9
30	TCRA-Gr191	1.0	6.9	0.3	2.3	0.4	4.5
35	TCRA-Gr271	4.7	21.5	0.8	10.3	0.7	15.2
	TCRA-Gr2146	1.1	6.8	0.3	1.7	0.3	2.6
	TCRA-Gr2157	3.8	22.2	0.6	9.6	0.6	12.0
	TCRA-Gr2179	0.8	4.9	0.2	1.8	0.2	1.4
	TCRA-Gr3081	5.9	23.6	0.7	11.5	0.8	12.6
	TCRA-Gr3099	2.1	21.1	0.4	7.1	0.3	6.2
	TCRA-Gr3105	12.1	45.9	2.2	22.0	1.0	7.6
	Controles ZFN						
	55248:55254	24.2	52.4				
	55229:53785	6.0	24.5				
	55266:53853	12.0	37.0				

40 [0185] Como se muestra, las nucleasas descritas en el presente documento inducen escisión y modificaciones genómicas en el sitio objetivo.

45 [0186] Por tanto, las nucleasas descritas en el presente documento (por ejemplo, nucleasas que comprenden una ZFP, un TALE o un dominio de unión a ADN de ARNg) se unen a sus sitios diana y escinden el gen TCRA, realizando de este modo modificaciones genéticas dentro de un gen TCRA que comprende cualquiera de las SEQ ID NO:8-21 o 92-103, incluidas modificaciones (inserciones y/o eliminaciones) dentro de cualquiera de estas secuencias (SEQ ID NO:8-21, 92-103); modificaciones dentro de 1-50 (p. ej., 1 a 10) pares de bases de estas secuencias de genes; modificaciones entre sitios objetivo de sitios objetivo emparejados (para dímeros); y/o modificaciones dentro de una o más de las siguientes secuencias: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT y/o AATCCTC (ver Figura 1B).

50 [0187] Además, todos los dominios de unión al ADN (ZFP, TALE y ARNg) se unen a sus sitios diana y también se formulan en factores de transcripción diseñados activos cuando se asocian con uno o más dominios reguladores de la transcripción.

55 Ejemplo 3: Actividad ZFN específica de TCRA en células T

60 [0188] Los pares de ZFN específicos de TCRA también se probaron en células T humanas para determinar la actividad nucleasa. Los ARNm que codifican las ZFN se transfecaron en células T purificadas. Brevemente, las células T se obtuvieron a partir del producto de leucoféresis y se purificaron utilizando el sistema Miltenyi CliniMACS (selección dual CD4 y CD8). A continuación, estas células se activaron utilizando Dynabeads (ThermoFisher) según el protocolo del fabricante. 3 días después de la activación, las células se transfecaron con tres dosis de ARNm (60, 120 y 250 µg/mL) usando un electroporador Maxcyte (Maxcyte), OC-100, 30e6 células/mL, volumen de 0,1 mL. Las células se analizaron para determinar la modificación de TCRA objetivo mediante secuenciación profunda (Miseq, Illumina) 10 días después de la transfección. La viabilidad celular y el crecimiento celular (duplicaciones celulares totales) se midieron durante los 13-14 días de cultivo. Además, se midió el TCR en la superficie celular de las células tratadas mediante análisis FACS estándar el día 10 de la tinción del cultivo para CD3.

[0189] Los pares de ZFN específicos de TCRA estaban todos activos en células T y algunos eran capaces de causar más del 80 % de modificación del alelo de TCRA en estas condiciones (véanse

5 [0190] las Figuras 2A y 2B). De manera similar, las células T tratadas con ZFN perdieron la expresión de CD3, donde el análisis FACS mostró que en algunos casos entre el 80 y el 90 % de las células T eran CD3 negativas (Figura 3). Una comparación entre el porcentaje de TCRA modificado por 2ZFN y la pérdida de CD3 en estas células demostró un alto grado de correlación (Figura 4). La viabilidad celular fue comparable a los controles de tratamiento simulados, y el crecimiento de células knockout para TCRA también fue comparable a los controles (ver Figura 5A-5D).

10

Ejemplo 4: Doble eliminación de B2M y TCRA con integración dirigida

[0191] Se usaron nucleasas como se describió anteriormente y la nucleasa dirigida B2M descrita en la Tabla 5 (ver también las solicitudes provisionales de EE. UU. N°s 62/269,410, presentada el 18 de diciembre de 2015; 62/305,097; y el número provisional de EE. UU. 62/329,439) para inactivar B2M y TCRA e introducir, mediante integración dirigida, un donante (transgén) en el locus TCRA o B2M. Las ZFN específicas de B2M se muestran a continuación en la Tabla 5:

20 Tabla 5: Diseños de ZFN específicos de B2M

Secuencia diana	F1	F2	F3	F4	F5	F6	Enlazador de dominio
Nombre ZFN							
SBS#7327 S' tcGCAATTTC AGGAAATTT GACTttccaa t (SEQ ID NO: 116)	DRSNLSR (SEQ ID NO: 22)	ARWYLDK (SEQ ID NO: 118)	QSGNLAR (SEQ ID NO: 34)	AKWNLDA (SEQ ID NO: 67)	QJNVLQN (SEQ ID NO: 119)	QMATNTK (SEQ ID NO: 28)	10
SBS#7332 S' tgTCGGAA TgGATGAAA CCCAGacac ata (SEQ ID NO: 117)	RSDNLSE (SEQ ID NO: 74)	ASKTNTN (SEQ ID NO: 120)	QEGNLAR (SEQ ID NO: 34)	TGANLSR (SEQ ID NO: 83)	TSGNLTR (SEQ ID NO: 25)	RTEGRLA (SEQ ID NO: 121)	NO.3

[0192] En este experimento, el par ZFN específico de TCRA fue SBS#55266/SBS#53853, que comprende la secuencia TTGAAA entre los sitios objetivo de ZFN específicos de TCRA (Tabla 1), y el par B2M fue SBS#57332/SBS#57327. (Tabla 5), que comprende la secuencia TCAAAT entre los sitios diana de ZFN específicos de B2M.

45

[0193] Brevemente, las células T (AC-TC-006) se descongelaron y activaron con dinabeads CD3/28 (proporción de células:perlas 1:3) en medio de cultivo de células T X-vivol5 (día 0). Despues de dos días en cultivo (día 2), se añadió al cultivo celular un donante de VAA (que comprende un transgén GFP y brazos de homología con el gen TCRA o B2M), excepto que también se mantuvieron los grupos de control sin donante. Al día siguiente (día 3), se agregaron TCRA y B2M ZFN mediante administración de ARNm en los siguientes 5 grupos:

- Grupo 1 (solo TCRA y B2M ZFN, sin donante): TCRA 120 ug/mL; B2M solo 60 ug/mL;
- Grupo 2 (TCRA y B2M ZFN y donante con brazos de homología de TCRA): TCRA 120 ug/mL; B2M 60 ug/mL y VAA (TCRA-Site E-hPGK-eGFP-Clone E2) 1E5vg/célula;
- Grupo 3 (TCRA y B2M ZFN y donante con brazos de homología de TCRA): TCRA 120 ug/mL; B2M 60 ug/mL; y VAA (TCRA-Sitio E-hPGK-eGFP-Clon E2) 3E4vg/célula;
- Grupo 4 (TCRA y B2M ZFN y donante con brazos de homología B2M): TCRA 120 ug/mL; B2M 60ug/mL y VAA (pVAA B2M-hPGK GFP) 1E5vg/célula;
- Grupo 5 (TCRA y B2M ZFN y donante con brazos de homología B2M): TCRA 120ug/mL; B2M 60ug/mL y VAA (pVAA B2M - hPGK GFP) 3E4vg/célula.

60

Todos los experimentos se realizaron a una densidad celular de 3,7 células/ml utilizando el protocolo descrito en la solicitud de EE. UU. n° 15/347.182 (choque por frío extremo) y se cultivaron hasta choque por frío a 30°C durante la noche después de la electroporación.

65

- [0194] Al día siguiente (día 4), las células se diluyeron a 0,5-6 células/ml y se transfirieron a cultivos a 37°C. Tres días después (día 7), las células se diluyeron nuevamente a 0,5-6 células/ml. Después de tres y siete días más en cultivo (días 10 y 14, respectivamente), se recogieron las células para análisis FACS y MiSeq (diluidas a 0,5-6 células/ml).
- 5 [0195] Como se muestra en la Figura 6, la expresión de GFP indicó que la integración objetivo fue exitosa y que las células modificadas genéticamente que comprenden modificaciones B2M y TCRA (inserciones y/o eliminaciones) dentro de los sitios objetivo de nucleasa (o dentro de 1 a 50, 1-20, Se obtuvieron 1-10 o 1-5 pares de bases de los sitios objetivo de nucleasa), incluidos dentro de TTGAAA y TCAAAT (entre los sitios objetivo emparejados) como se describe en el presente documento.
- 10 [0196] También se realizan experimentos en los que se integra un transgén CAR en el locus B2M y/o TCRA para crear inactivaciones dobles de B2M/TCRA que expresan un CAR.

REIVINDICACIONES

- 5 1. Una célula aislada en la que la expresión de un gen del receptor de células T (TCR) se modula mediante modificación dentro del exón c2 del gen del receptor de células T alfa (TCRA), en la que la modificación es una inserción y/o eliminación dentro de TTGAAA del exón c2 del gen TCRA, y en donde la célula aislada comprende un par de nucleasas con dedos de zinc (ZFN) que escinden el gen TCRA, comprendiendo el par de ZFN una primera y segunda ZFN, en donde la ZFN primera comprende un primer dominio de escisión o medio dominio de escisión, y una proteína de dedos de zinc (ZFP) que comprende seis dedos de zinc designados y ordenados de F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como sigue:

10 F1: QSSDLSR (SEQ ID NO:57);
F2: QSGNRTT (SEQ ID NO:58);
F3: RSANLAR (SEQ ID NO:59);
F4: DRSLALAR (SEQ ID NO:33);
F5: RSDVLSE (SEQ ID NO:60); y
F6: KHSTRRV (SEQ ID NO:61); y

15 la segunda ZFN comprende un segundo dominio de escisión o semidominio de escisión, y una ZFP que comprende seis dedos de zinc designados y ordenados de F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como sigue:

20 F1: TMHQRVE (SEQ ID NO:62);
F2: TSGHLSR (SEQ ID NO:63);
F3: RSDHLTQ (SEQ ID NO:64);
F4: DSANLSR (SEQ ID NO:65);
F5: QSGSLTR (SEQ ID NO:66); y
F6: AKWNLDA (SEQ ID NO:67)

25 y además en donde la inserción y/o eliminación en el gen TRAC se realiza mediante el par de ZFN y la expresión del gen TRAC se modula mediante la inserción y/o eliminación.

30 2. La célula de la reivindicación 1, en la que el semidominio de escisión es un dominio de escisión de célula *FokI* de tipo salvaje o modificada genéticamente.

35 3. La célula de la reivindicación 1 o la reivindicación 2, en la que la célula comprende una inserción y/o una eliminación dentro de una o más de SEQ ID NO:8-21, 92-103, AACAGT, AGTGCT, CTCCT, TGGACTT y/o AATCCTC.

40 4. La célula de cualquiera de las reivindicaciones 1 a 3, que comprende, además un gen inactivado de beta 2 microglobulina (B2M), PD1 y/o CTLA4.

45 5. La célula de cualquiera de las reivindicaciones 1 a 4, que comprende, además un transgén que codifica un receptor de antígeno químérico (CAR), un transgén que codifica un receptor de células T acoplado a anticuerpo (ACTR) y/o un transgén que codifica un TCR diseñado.

6. La célula de cualquiera de las reivindicaciones 1 a 5, en la que la célula es una célula linfoide, una célula madre o una célula progenitora.

45 7. La célula de la reivindicación 5, en la que la célula es una célula T, una célula madre pluripotente inducida (iPSC), una célula madre mesenquimal (MSC) o una célula madre hematopoyética (HSC).

50 8. Una composición farmacéutica que comprende una célula según cualquiera de las reivindicaciones 1 a 7.

9. Un método de diferenciación *in vitro*, comprendiendo el método: tomar una célula según cualquiera de las reivindicaciones 1-7, en el que la célula es una célula madre; y diferenciar la célula para generar una célula diferenciada que desciende de dicha célula madre.

55 10. Una molécula de fusión que comprende un dominio de unión al ADN que se une al exón c1, c2 o c3 de un gen TCR y un dominio regulador transcripcional o un dominio de nucleasa, en donde el dominio de unión al ADN comprende una proteína con dedos de zinc (ZFP) que tiene cinco o seis dedos de zinc designados y ordenados F1 a F5 o F1 a F6, comprendiendo cada dedo de zinc una región de hélice de reconocimiento como se muestra en una sola fila de la siguiente tabla:

ES 2 983 043 T3

5

10

15

20

25

30

35

40

45

50

55

60

65

F1	F2	F3	F4	F5	F6
DRSNLSP (SEQ ID NO:22)	QKVTLAA (SEQ ID NO:23)	DPSALSP (SEQ ID NO:24)	TSGNITR (SEQ ID NO:25)	YRSSEKE (SEQ ID NO:26)	TSGNLTR (SEQ ID NO:25)
QCONVLIN (SEQ ID NO:27)	QNATRTK (SEQ ID NO:28)	QSGHLAR (SEQ ID NO:29)	NRYDLMT (SEQ ID NO:30)	RSDSELLR (SEQ ID NO:31)	QSSDLTR (SEQ ID NO:32)
DESALAR (SEQ ID NO:33)	QSGNLAR (SEQ ID NO:34)	HRSTLQG (SEQ ID NO:35)	QSGDLTR (SEQ ID NO:36)	TSGSLTR (SEQ ID NO:37)	NA
QHOVVLVR (SEQ ID NO:38)	QNATRTK (SEQ ID NO:28)	QSGHLSR (SEQ ID NO:39)	RSDSLSR (SEQ ID NO:40)	RSDALAR (SEQ ID NO:41)	NA
DQSNLRA (SEQ ID NO:42)	TSSNRKT (SEQ ID NO:43)	DSSTRKT (SEQ ID NO:44)	QSGNLAR (SEQ ID NO:34)	RSDLILSE (SEQ ID NO:45)	TNSNRKR (SEQ ID NO:46)
RSDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LKQHLNE (SEQ ID NO:49)	TSGNLTR (SEQ ID NO:28)	HRSTLTD (SEQ ID NO:50)	NA
DQSNLRA (SEQ ID NO:42)	TSSNRKT (SEQ ID NO:43)	LOOTLAD (SEQ ID NO:51)	QSGNLAR (SEQ ID NO:34)	PREDLIT (SEQ ID NO:52)	TSSNLSP (SEQ ID NO:53)
RSDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LKQHLNE (SEQ ID NO:49)	QSGNLAR (SEQ ID NO:34)	HNSSLKD (SEQ ID NO:54)	NA
RSDHLST (SEQ ID NO:47)	DRSHLAR (SEQ ID NO:48)	LNRRRLQQ (SEQ ID NO:55)	QSGNLAR (SEQ ID NO:34)	HRSTLKD (SEQ ID NO:56)	NA
QSSDLAR (SEQ ID NO:57)	QSGNRRT (SEQ ID NO:58)	REANLAR (SEQ ID NO:59)	RDSALAR (SEQ ID NO:33)	RSDVLSR (SEQ ID NO:60)	RHSTRRV (SEQ ID NO:61)
TMHQRLVE (SEQ ID NO:62)	TSGHLSR (SEQ ID NO:63)	RSDHLTQ (SEQ ID NO:64)	DSANLSP (SEQ ID NO:65)	QSGSLTR (SEQ ID NO:66)	AKWNLDA (SEQ ID NO:67)
TMHQRLVE (SEQ ID NO:62)	TSGHLSR (SEQ ID NO:63)	RNDSLKT (SEQ ID NO:68)	DSSNLSPR (SEQ ID NO:69)	QKATRTT (SEQ ID NO:70)	RNASRTR (SEQ ID NO:72)
RSDSLLR (SEQ ID NO:31)	QSSDLRR (SEQ ID NO:73)	RSDNLSE (SEQ ID NO:74)	ERANRNS (SEQ ID NO:75)	RSDNLAR (SEQ ID NO:76)	QKVNLMS (SEQ ID NO:77)
RSDSLLR (SEQ ID NO:31)	QSSDLRR (SEQ ID NO:73)	RSDNLSE (SEQ ID NO:74)	ERANRNS (SEQ ID NO:75)	RSDNLAR (SEQ ID NO:76)	QKVNLRE (SEQ ID NO:78)
RSDTSLSE (SEQ ID NO:79)	TSGSLTR (SEQ ID NO:37)	RSDHLST (SEQ ID NO:47)	TSSNNTK (SEQ ID NO:71)	RSDNLSE (SEQ ID NO:74)	WHSSLRV (SEQ ID NO:83)
RKOTRTT (SEQ ID NO:80)	BRSSLRR (SEQ ID NO:81)	RSDHLST (SEQ ID NO:47)	TSANLSP (SEQ ID NO:82)	RSDNLSE (SEQ ID NO:74)	WHSSLRV (SEQ ID NO:83)
RSAHLSR (SEQ ID NO:84)	DRSULSR (SEQ ID NO:40)	RSDVLSV (SEQ ID NO:85)	QNNHPII (SEQ ID NO:86)	RSDVLSE (SEQ ID NO:60)	SPSSRPT (SEQ ID NO:87)
RSAHLSP (SEQ ID NO:84)	DRSULSP (SEQ ID NO:40)	RSDVLSV (SEQ ID NO:88)	QANANRKT (SEQ ID NO:89)	RSDVLSE (SEQ ID NO:60)	SPSSRRT (SEQ ID NO:87)
TMHQRLVE (SEQ ID NO:52)	TSGHLSP (SEQ ID NO:63)	RSDSLST (SEQ ID NO:90)	DRANRKK (SEQ ID NO:91)	QKATRTT (SEQ ID NO:70)	RNASRTR (SEQ ID NO:72)

11. Un polinucleótido que codifica la molécula de fusión de la reivindicación 10.

12. El polinucleótido de la reivindicación 11, en el que el polinucleótido es un vector viral, un plásmido o ARNm.
- 5 13. La célula según cualquiera de las reivindicaciones 1 a 7 o la composición farmacéutica según la reivindicación 8 para su uso en un método para tratar o prevenir un cáncer, comprendiendo el método introducir la célula o la composición farmacéutica a un sujeto con cáncer.
- 10 14. La célula según cualquiera de las reivindicaciones 1 a 7, la composición farmacéutica según la reivindicación 8, la molécula de fusión según la reivindicación 10 o el polinucleótido según la reivindicación 11, para uso en el tratamiento de un sujeto con cáncer.

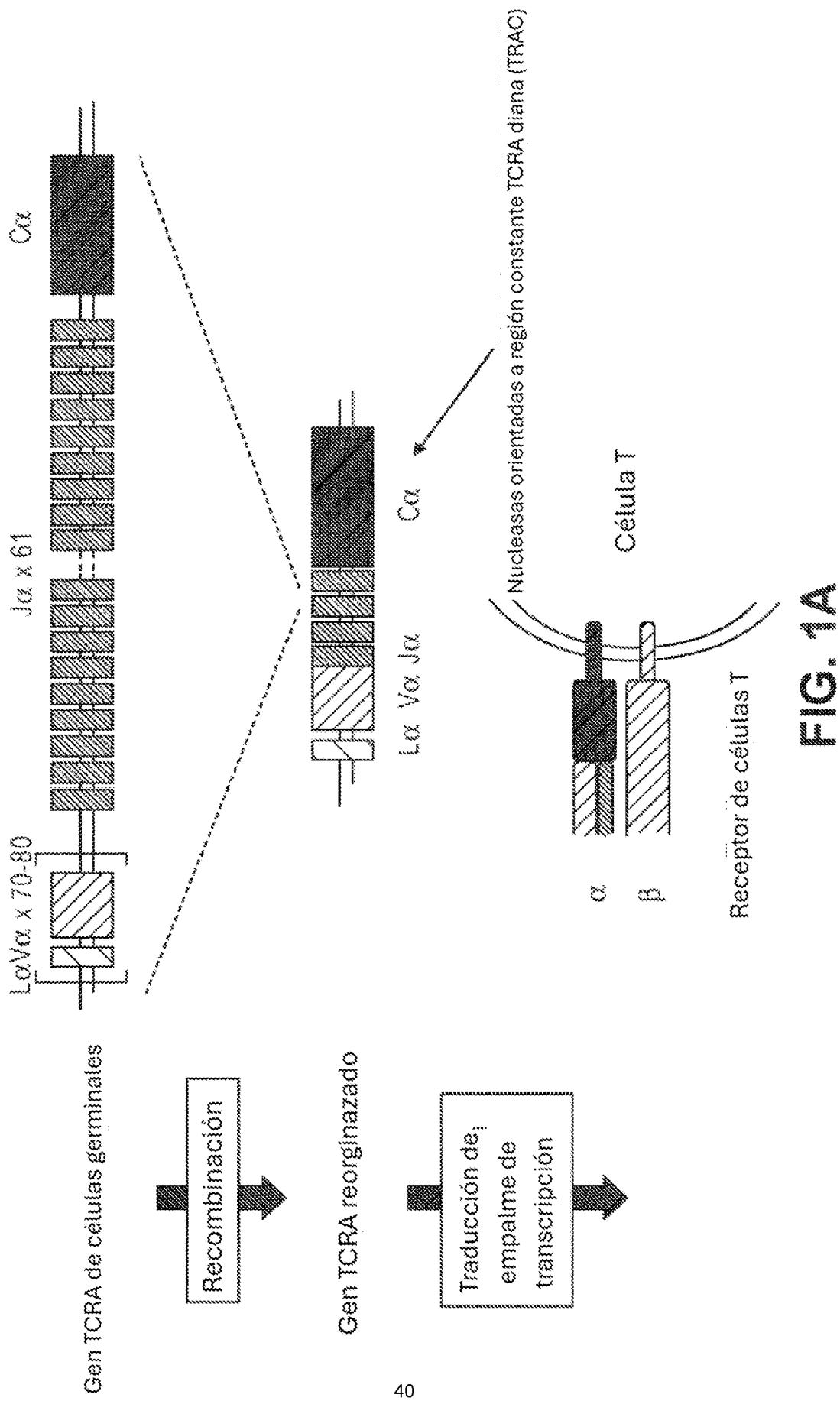


FIG. 1A

四

102 GTCACAAAGTAAGGATTCTGATGTTGATAACAGACAATACAGCTAGACATAGGTCAATGGACTTCAAGGCAAGGAACTCTGCTGAGCTGAGGAAAC

201 AAATCTGACTTGTGCAAAACGCCCTAACACGATTTACAGAGAACCTTCTTCCCAGGCCAGTTAAGCCGAGCTTGGCTTCCTAC

23

41

308

GESENIAVIACTECTATAGCTCTCCGCGCTTGCTGCCCTGCTGGCGCTGGCGCTGGCG

三
下

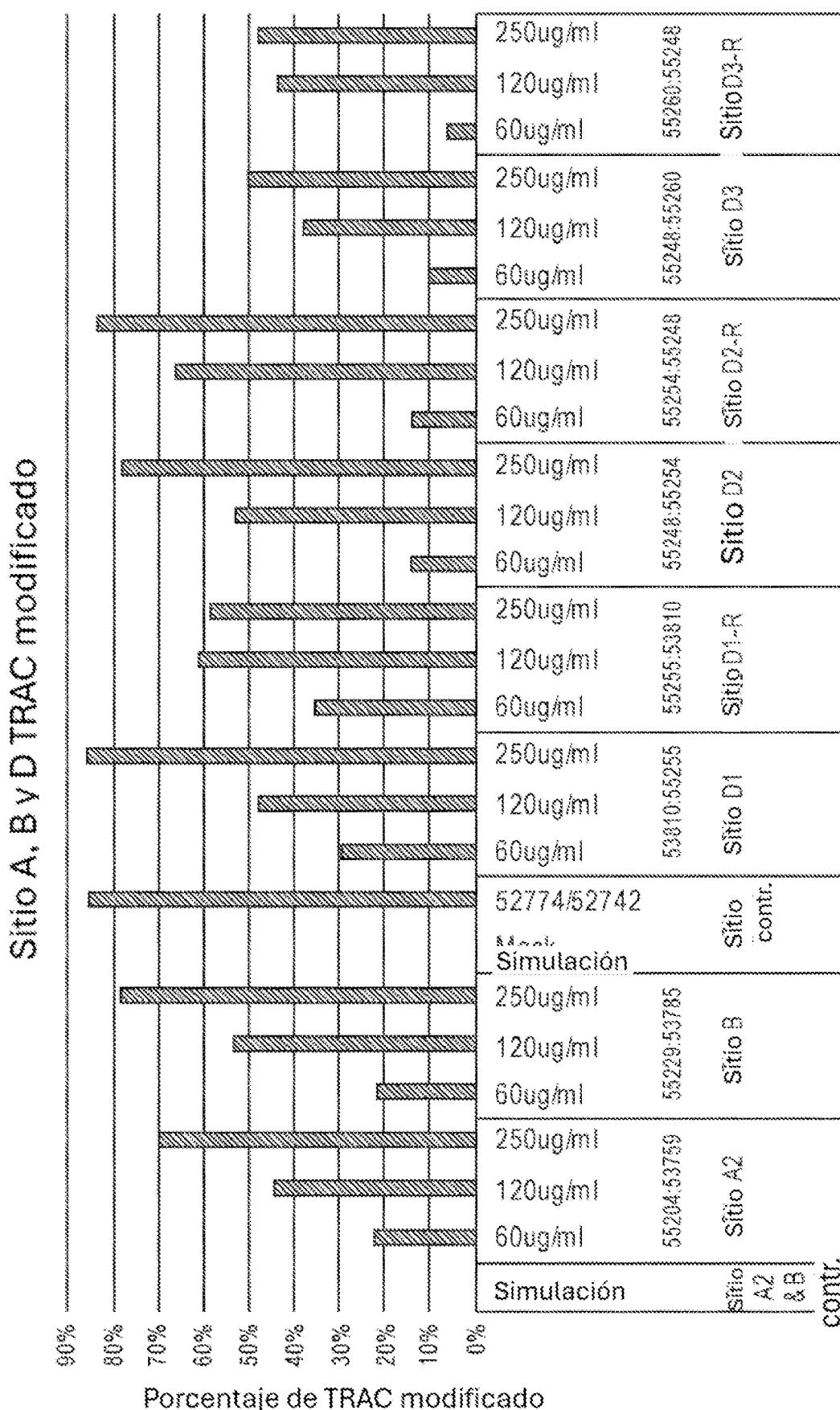


FIG. 2A

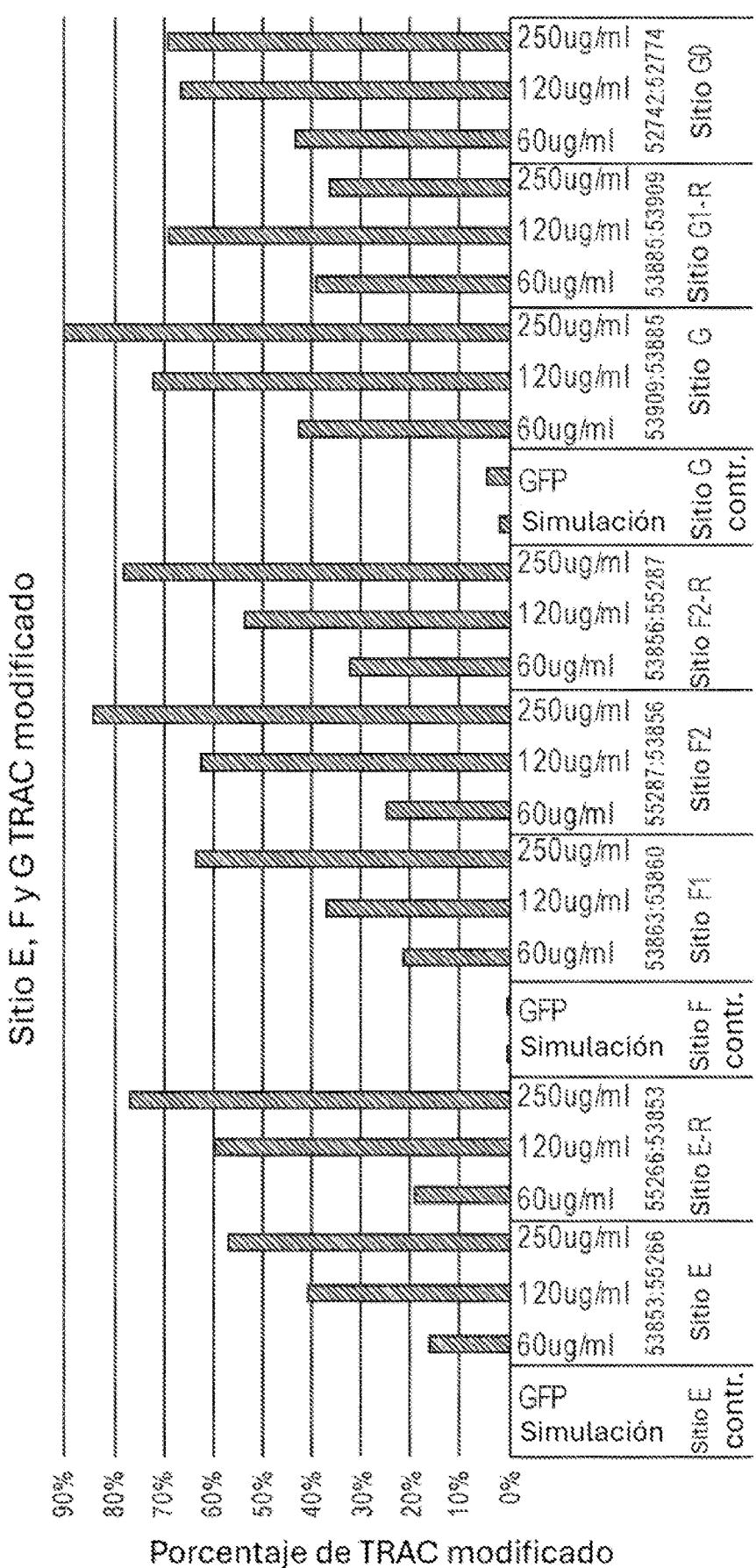


FIG. 2B

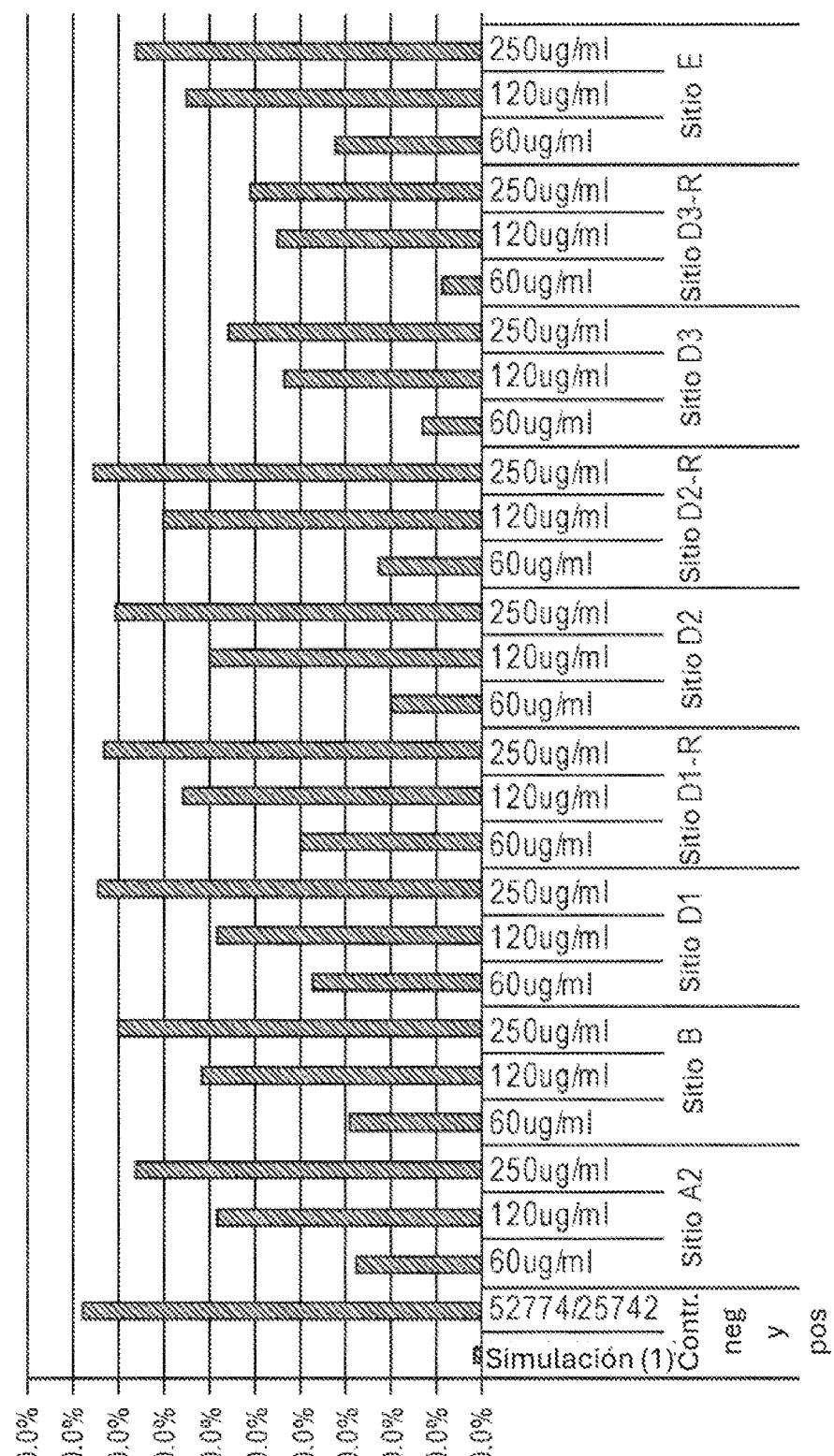
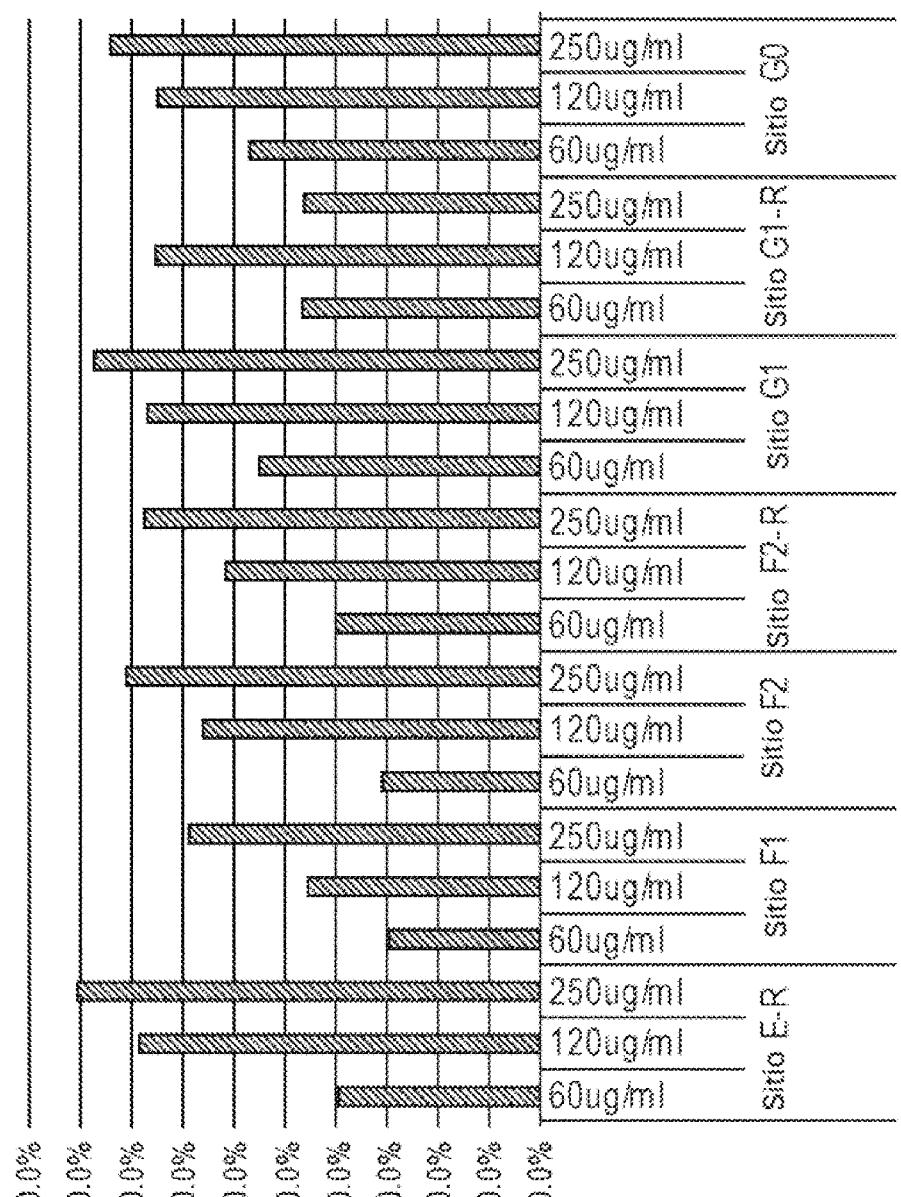


FIG. 3



Porcentaje de células T negativas por CD3

FIG. 3 (CONT.)

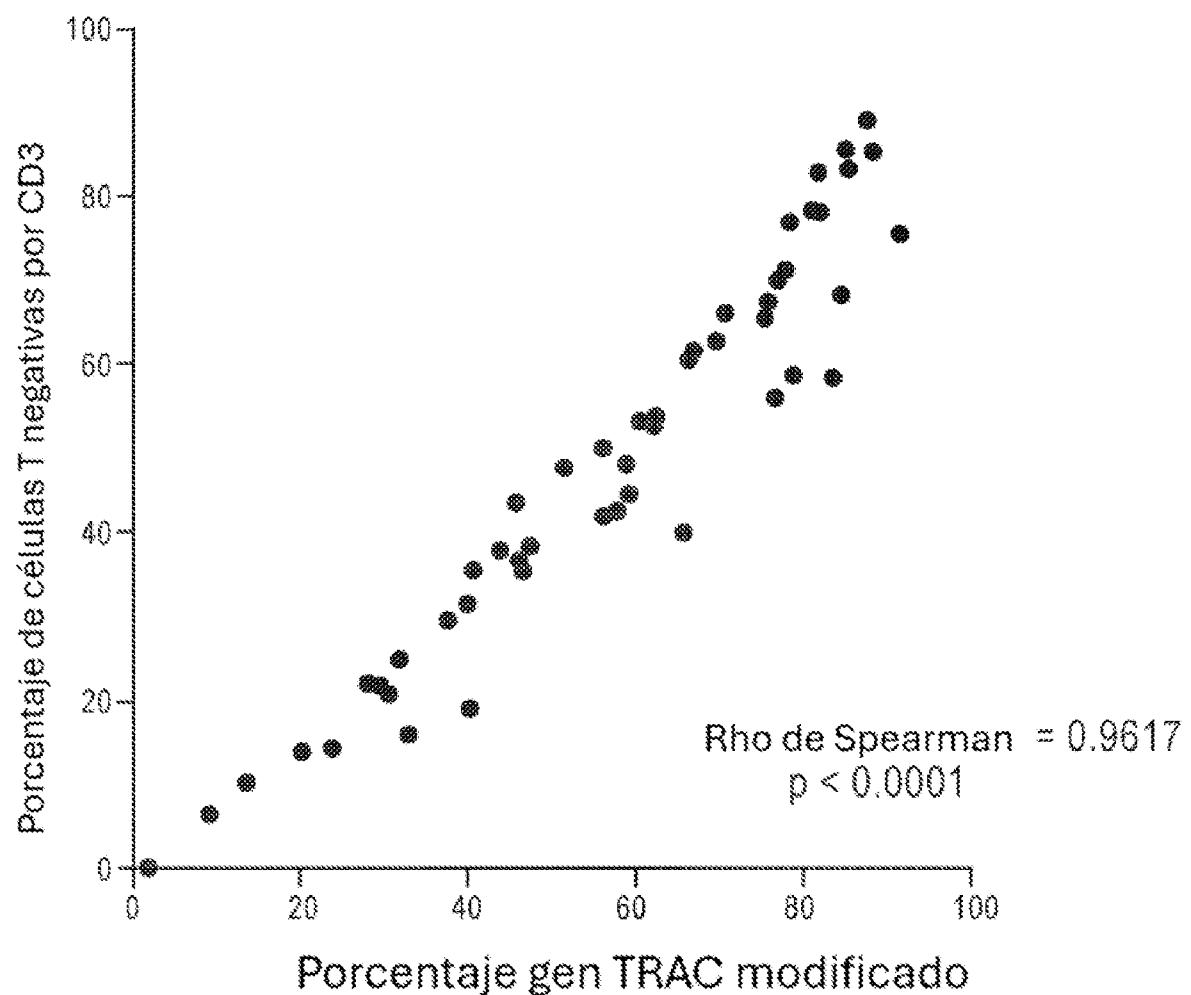


FIG. 4

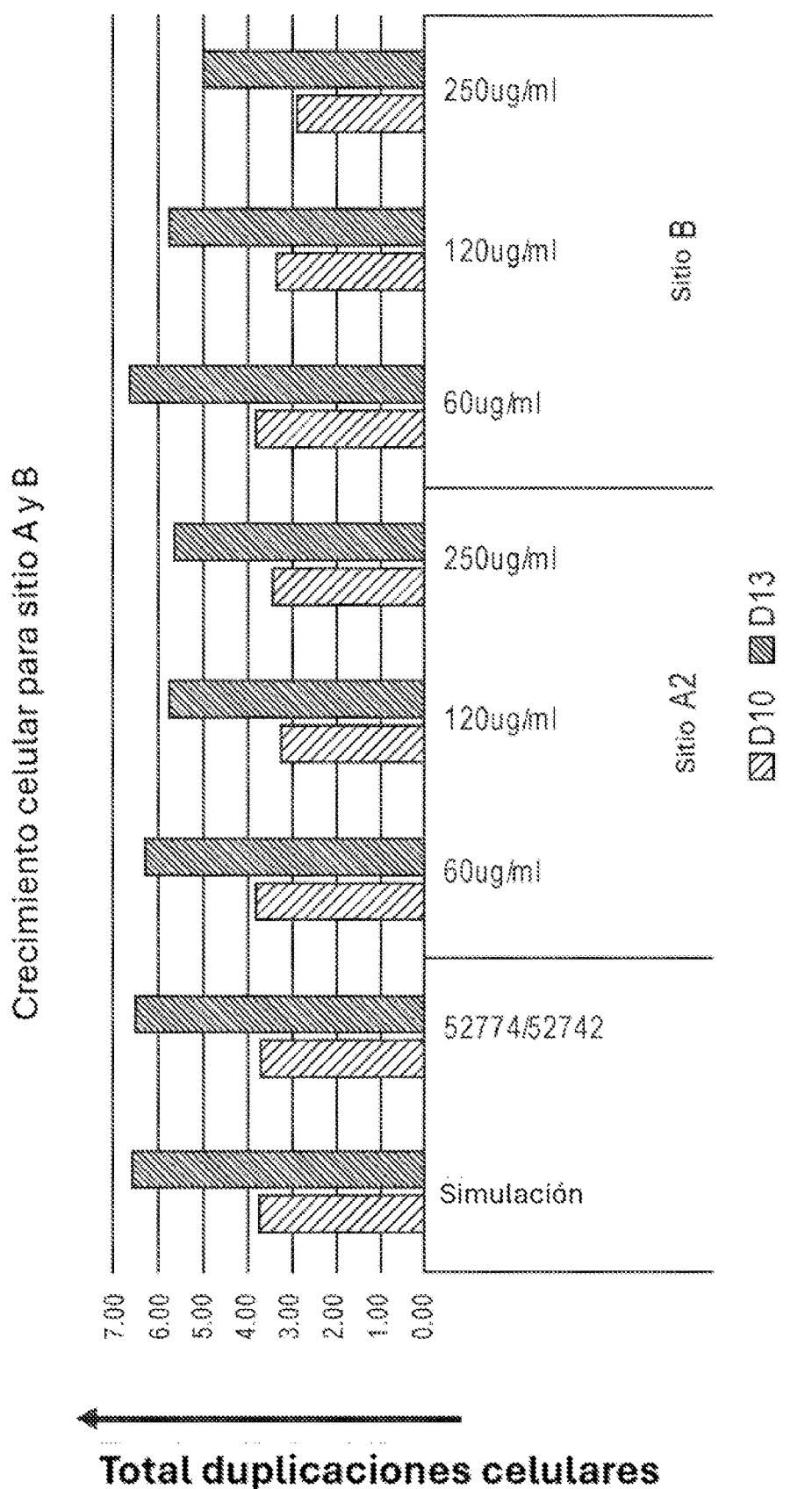


FIG. 5A

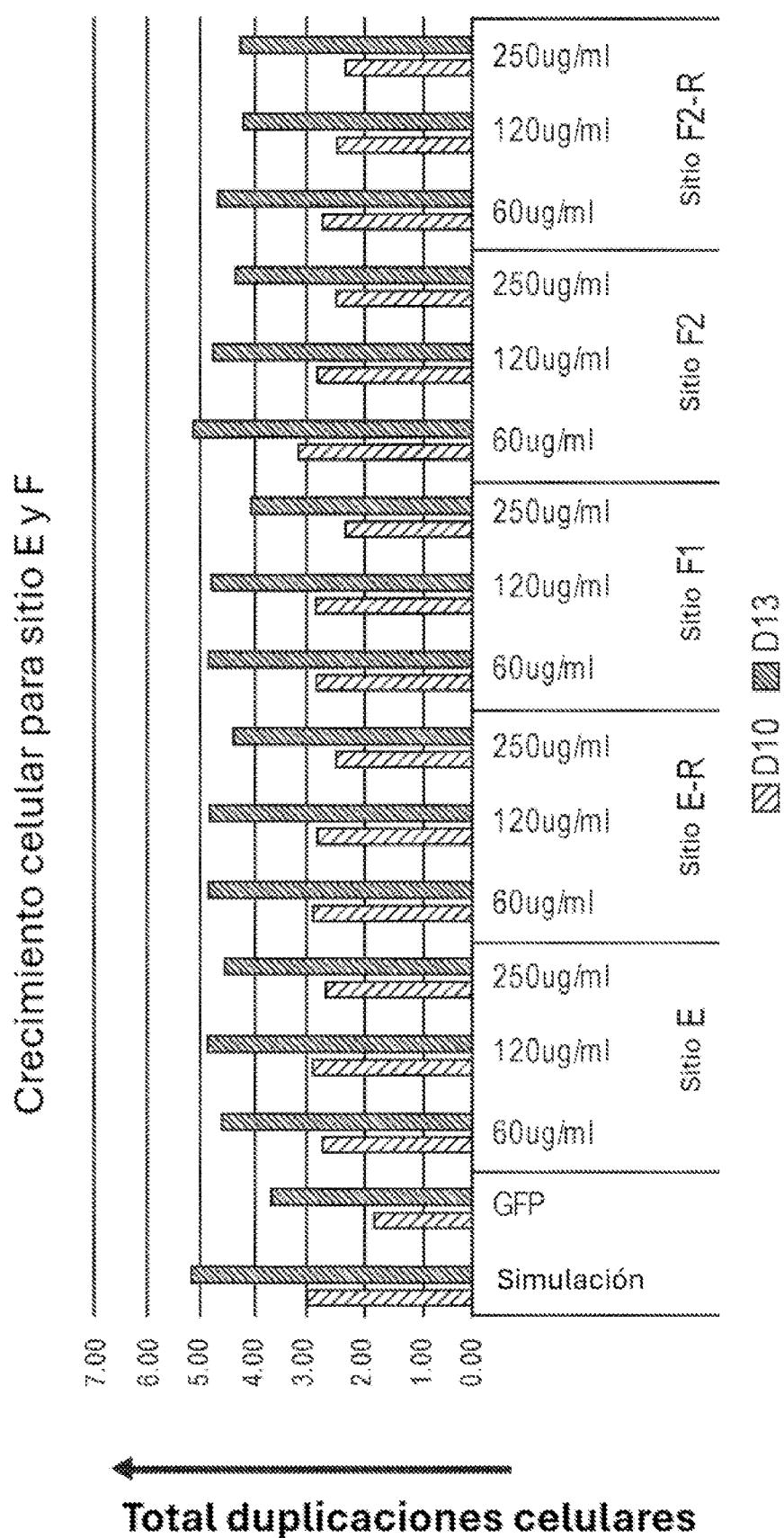
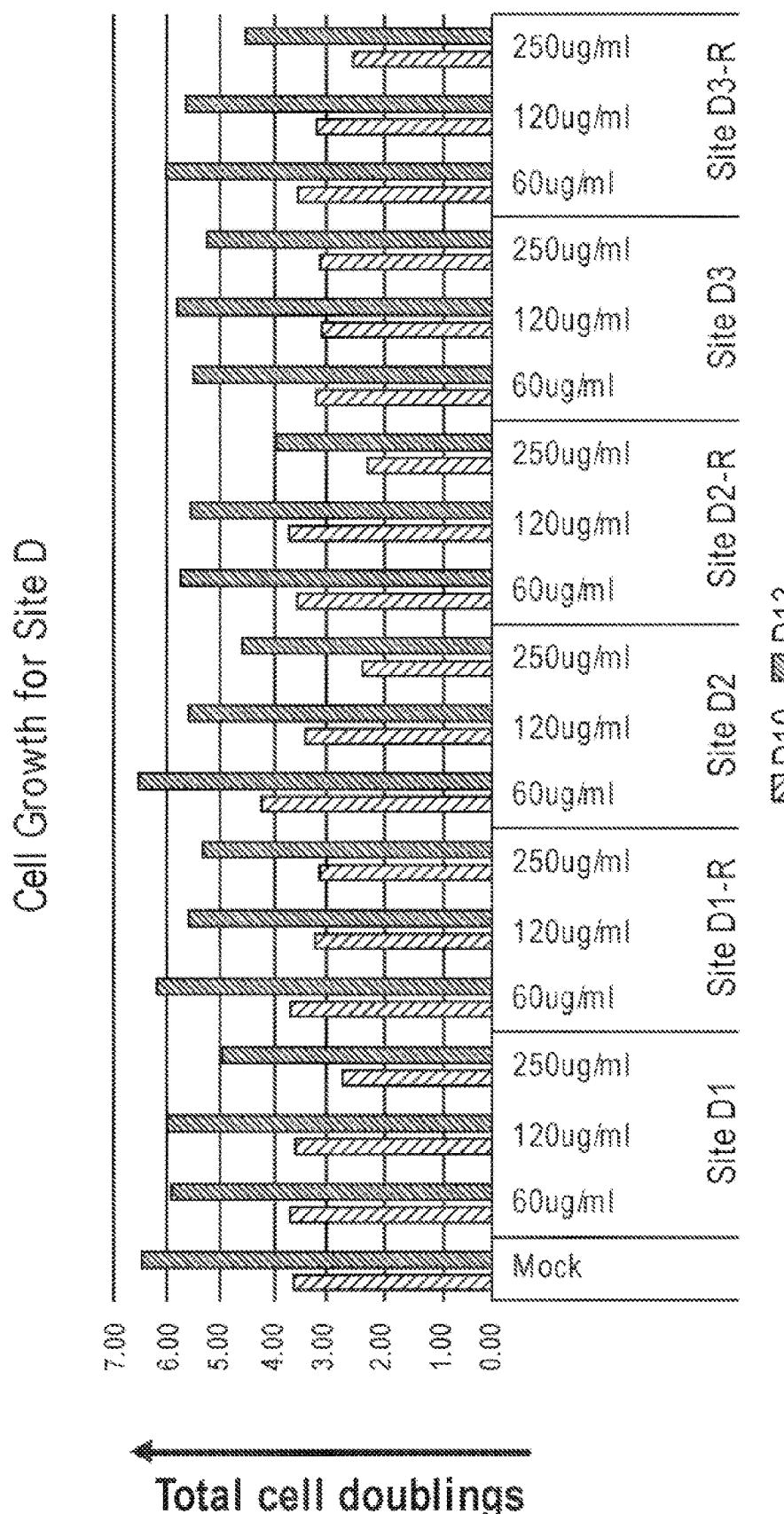


FIG. 5B



50