

UNIÓN DE EXTREMOS MEDIADA POR HOMOLOGÍA NO VIRAL

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

5 [0001]La presente solicitud reivindica el beneficio de la solicitud provisional estadounidense número 63/174,468, presentada el 13 de abril de 2021, que se incorpora a la presente en su totalidad mediante referencia a todos los efectos.

ANTECEDENTES DE LA DESCRIPCIÓN

10 [0002]La aplicación de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) y proteínas asociadas a CRISPR (Cas) ha revolucionado la biología molecular al hacer posible la edición del genoma. La edición de genes mediada por CRISPR es una herramienta poderosa y práctica con potencial para crear nuevas herramientas científicas, corregir mutaciones clínicamente relevantes y diseñar nuevas inmunoterapias basadas en células. Los vectores de administración viral y la electroporación han surgido como dos estrategias para la edición basada en CRISPR en células inmunitarias. Sin embargo, el campo de la terapia génica ha estado repleto de complicaciones de los vectores de administración virales.

15 [0003]La capacidad de manipular linfocitos T, células madre hematopoyéticas y células madre pluripotentes inducidas proporciona tanto una herramienta científica como una vía potencialmente terapéutica para la para la modificación genética inmunitaria. Hasta el descubrimiento de la edición de genes CRISPR-Cas en células humanas, la terapia génica dirigida al sistema inmunitario se basaba casi exclusivamente en vectores virales como retrovirus y lentivirus para insertar un
20 transgén mediante una integración genómica (semi)aleatoria. Sin embargo, los ensayos clínicos han tenido consecuencias no deseadas, como la leucemia, debido a la inserción en protooncogenes y respuestas inmunitarias sistémicas mortales al vector viral. A pesar de las mejoras en los vectores virales para minimizar el daño genético y las respuestas inmunitarias, las inserciones virales (como los lentivirus utilizados clínicamente para generar linfocitos T-CAR) pueden resultar en una
25 sobreexpresión de genes, por ejemplo, sobreexpresión de genes que no están bajo regulación endógena, lo que restringe la utilidad clínica de este método.

[0004]Se ha descrito la edición de genes CRISPR-Cas no viral. Sin embargo, la eficiencia y/o la

viabilidad de la edición utilizando dichos sistemas son generalmente bajas. Por ejemplo, una mayor eficiencia de edición utilizando los métodos actuales tiene el costo de una menor viabilidad celular, y viceversa. Se necesitan métodos mejorados de edición de genes CRISPR-Cas no viral, como métodos que tengan una mayor eficiencia de la edición y viabilidad.

5 COMPENDIO

[0005] En el presente documento se proporciona una composición para modificar un ácido nucleico diana, que comprende: (a) una proteína nucleasa dirigible; y (b) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR, por sus siglas en inglés); (ii) una primera secuencia diana de linealización de coadministración (CDL, por sus siglas en inglés), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante la proteína nucleasa dirigible o un complejo que comprende la proteína nucleasa dirigible, y en donde la composición se formula para administración no viral a una célula.

[0006] También se proporciona en la presente una composición para modificar un ácido nucleico diana, que comprende: (a) una nucleasa guiada por ARN CRISPR-CAS; (b) un ARN guía (ARNg) donante; y (c) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM, por sus siglas en inglés) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL, y en donde la composición se formula para administración no viral a una célula.

[0007] También se proporciona en la presente un método para modificar un ácido nucleico diana de una célula, el método comprende: proporcionar la célula, e introducir o haber introducido en la

célula una composición formulada para administración no viral, la composición comprende: (a) una proteína nucleasa dirigible; y (b) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante la proteína nucleasa dirigible o un complejo que comprende la proteína nucleasa dirigible.

10 **[0008]** También se proporciona en la presente un método para modificar un ácido nucleico diana de una célula, el método comprende: proporcionar la célula, e introducir o haber introducido en la célula una composición formulada para administración no viral, la composición comprende: (a) una nucleasa guiada por ARN CRISPR-CAS; (b) un ARN guía (ARNg) donante; y (c) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR);
15 (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias
20 diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL.

[0009] También se proporciona en la presente un método para modificar una secuencia diana genómica de una célula, el método comprende: proporcionar la célula, e introducir o haber
25 introducido en la célula una composición formulada para administración no viral, la composición comprende: (a) una nucleasa guiada por ARN CRISPR-CAS; (b) un ARN guía (ARNg) donante; y (c) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR) que comprende un ácido nucleico para inserción flanqueada por brazos de homología; (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde
30 la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y

(iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL, en donde el ARNg donante comprende al menos 17 nucleótidos que son complementarios a la secuencia diana genómica de la célula, y en donde los brazos de homología son complementarios a secuencias de ácido nucleico que flanquean la secuencia diana genómica de la célula, en donde el ácido nucleico para inserción se configura para la inserción en la secuencia diana genómica de la célula.

[0010]En algunos aspectos, la proteína nucleasa dirigible es una nucleasa guiada por ARN. En algunos aspectos, la composición comprende además un ARN que comprende al menos 17 nucleótidos que son complementarios a la secuencia diana de CDL. En algunos aspectos, la nucleasa guiada por ARN es una proteína Cas.

[0011]En algunos aspectos, la composición comprende además un ARN guía (ARNg) donante configurado para formar el complejo que comprende la proteína nucleasa dirigible, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL están unidas operativamente a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL. En algunos aspectos, la composición comprende además un primer ARN guía (ARNg) donante que comprende al menos 17 nucleótidos que son complementarios a la primera secuencia diana de CDL, un segundo ARNg donante que comprende al menos 17 nucleótidos que son complementarios a la segunda secuencia diana de CDL, en donde cada ARNg donante se configura para formar un complejo distinto que comprende la proteína nucleasa dirigible, y en donde cada una de la primera y la segunda secuencias diana de CDL están unidas operativamente a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL. En algunos aspectos, uno o más de los ARNg donantes comprenden al menos 17 nucleótidos que son complementarios a una secuencia diana genómica de la célula. En algunos aspectos, la plantilla de HDR comprende brazos de homología que son complementarios a secuencias de ácido nucleico que flanquean la

secuencia diana genómica de la célula. En algunos aspectos, los brazos de homología se seleccionan cada uno independientemente de una longitud de al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, 1000 pb, al menos 1100 pb, al menos 1200 pb, al menos 1300 pb, 1400 pb, al menos 1500 pb, al menos 1600 pb, al menos 1700 pb, al menos 1800 pb, al menos 1900 pb o al menos 2000 pb de longitud.

[0012]En algunos aspectos: (a) la nucleasa dirigible comprende una nucleasa guiada por ARN, en donde la nucleasa guiada por ARN comprende CRISPR-CAS; (b) la composición comprende además un ARN guía (ARNg) donante configurado para formar el complejo que comprende la proteína nucleasa dirigible, y (c) en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL están unidas operativamente a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL.

[0013]En algunos aspectos, la composición comprende además una segunda proteína nucleasa dirigible, en donde la segunda proteína nucleasa dirigible o un complejo que comprende la segunda proteína nucleasa dirigible es capaz de escindir una secuencia diana genómica de la célula. En algunos aspectos, la segunda proteína nucleasa dirigible es una nucleasa guiada por ARN. En algunos aspectos, la composición comprende además un segundo ARN que comprende al menos 17 nucleótidos que son complementarios a la secuencia diana genómica. En algunos aspectos, la nucleasa guiada por ARN es una proteína Cas. En algunos aspectos, la composición comprende además un ARN guía (ARNg) diana configurado para formar el complejo que comprende la segunda proteína nucleasa dirigible, y en donde (1) el ARNg diana comprende al menos 17 nucleótidos que son complementarios a la secuencia diana genómica, y (2) la secuencia diana genómica está unida operativamente a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de la secuencia diana genómica.

[0014]En algunos aspectos, la primera secuencia diana de CDL, la segunda secuencia diana de CDL y la secuencia diana genómica comprenden la misma secuencia de ácido nucleico.

[0015]En algunos aspectos, la secuencia diana genómica comprende una secuencia de ácido nucleico de puerto seguro. En algunos aspectos, la secuencia de ácido nucleico de puerto seguro

comprende la secuencia de ácido nucleico GAGCCATGCTTGGCTTACGA. En algunos aspectos, una, ambas o ninguna de las secuencias PAM están codificadas entre las secuencias diana de CDL y la plantilla de HDR.

5 **[0016]**En algunos aspectos, la nucleasa dirigible y cada uno de los ARNg están en una relación molar de entre 1:10 y 2:1, respectivamente. En algunos aspectos, la nucleasa dirigible y la plantilla donante están en una relación molar de entre: 10:1 y 1000:1, respectivamente.

[0017]En algunos aspectos, la proteína nucleasa dirigible y/o la segunda proteína nucleasa dirigible comprende una proteína efectora de unión al ADN de tipo activador de la transcripción (TAL) y una nucleasa.

10 **[0018]**En algunos aspectos, la proteína nucleasa dirigible y/o la segunda proteína nucleasa dirigible comprende una proteína de unión al ADN con dedos de cinc y una nucleasa.

[0019]En algunos aspectos, la proteína nucleasa dirigible y/o la segunda proteína nucleasa dirigible se fusiona a una secuencia de señal de localización nuclear (NLS, por sus siglas en inglés).

15 **[0020]**En algunos aspectos, la proteína nucleasa dirigible y/o la segunda proteína nucleasa dirigible es una proteína Cas9.

20 **[0021]**También se proporciona en el presente documento un método para modificar el ácido nucleico diana en la célula, que comprende introducir o haber introducido de forma no viral en la célula cualquiera de las composiciones descritas en el presente documento, en donde la plantilla de HDR se integra en el ácido nucleico diana. En algunos aspectos, la introducción comprende electroporación. En algunos aspectos, la célula es una célula primaria. En algunos aspectos, la célula primaria es un linfocito T primario.

[0022]También se proporciona en el presente documento un complejo de ribonucleoproteína para modificar un ácido nucleico diana, que comprende cualquiera de las composiciones descritas en el presente documento.

25 **[0023]**También se proporciona en el presente documento un complejo de ribonucleoproteína para modificar un ácido nucleico diana, (a) una nucleasa guiada por ARN CRISPR-CAS; y (b) un ARN guía donante (ARNg), en donde el ARNg donante comprende al menos 17 nucleótidos que son

complementarios a una secuencia diana de linealización de coadministración (CDL), y en donde la composición se formula para la administración no viral a una célula. En algunos aspectos, la composición comprende, además, una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL.

[0024] También se proporciona en el presente documento un método para modificar el ácido nucleico diana en la célula, que comprende introducir en la célula cualquiera de los complejos de ribonucleoproteína descritos en el presente documento. En algunos aspectos, la introducción comprende electroporación. En algunos aspectos, la célula es una célula primaria. En algunos aspectos, la célula primaria es un linfocito T primario. En algunos aspectos, el método se realiza in vivo, in vitro o ex vivo.

[0025] También se proporciona en el presente documento un método para formar un complejo de ribonucleoproteína (RNP), que comprende incubar o haber incubado (a) una nucleasa guiada por ARN CRISPR-CAS; y (b) un ARN guía (ARNg) donante, en donde el ARNg donante comprende al menos 17 nucleótidos que son complementarios a una secuencia diana de linealización de coadministración (CDL). En algunos aspectos, la proteína Cas y el ARNg se incuban juntos a 37 °C durante al menos 17 minutos. En algunos aspectos, la relación molar de ARNg:proteína Cas está entre 0,25:1 y 4:1. En algunos aspectos, el complejo RNP tiene un tamaño inferior a 100 nm. En algunos aspectos, el complejo RNP tiene un tamaño de entre 20 nm y 90 nm.

[0026] También se proporciona en el presente documento una composición que comprende una plantilla donante de plásmido, que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL),

en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante una proteína nucleasa
5 dirigible o un complejo que comprende la proteína nucleasa dirigible. En algunos aspectos, la composición comprende además un ARN guía (ARNg) donante configurado para formar el complejo que comprende la proteína nucleasa dirigible. En algunos aspectos, la composición comprende la proteína nucleasa dirigible.

[0027]En algunos aspectos, la plantilla donante comprende de 5' a 3' la secuencia: P1a-N1-P2b-
10 H-P3c-N2-P4d, y en donde: (1) P1, P2, P3 y P4 son secuencias PAM; (2) N1 es la primera secuencia diana de CDL y N2 es la segunda secuencia diana de CDL; (3) H es la plantilla de HDR; (4) a es 0 y b es 1, o a es 1 y b es 0; y (5) c es 0 y d es 1; o c es 1 y d es 0.

[0028]En determinados aspectos, en la presente se describe un método para modificar un ácido nucleico diana de una célula, el método comprende: proporcionar la célula, e introducir o haber
15 introducido en la célula una composición formulada para administración no viral, la composición comprende: (a) una proteína nucleasa dirigible; y (b) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia
20 diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante la proteína nucleasa dirigible o un complejo que comprende la proteína nucleasa dirigible. En determinados aspectos, en la presente se describe un método para modificar un ácido nucleico diana de una célula, el método comprende: proporcionar la célula, e
25 introducir o haber introducido en la célula una composición formulada para administración no viral, la composición comprende: (a) una nucleasa guiada por ARN CRISPR-CAS; (b) un ARN guía (ARNg) donante; y (c) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida

hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL. En determinados aspectos, en la presente se describe un método para modificar una secuencia diana genómica de una célula, el método comprende: proporcionar la célula, e introducir o haber introducido en la célula una composición formulada para administración no viral, la composición comprende: (a) una nucleasa guiada por ARN CRISPR-CAS; (b) un ARN guía (ARNg) donante; y (c) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR) que comprende un ácido nucleico para inserción flanqueada por brazos de homología; (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde (1) el ARNg donante comprende al menos 17 nucleótidos que son complementarios a cada una de la primera y la segunda secuencias diana de CDL, y (2) cada una de la primera y la segunda secuencias diana de CDL está operativamente unida a un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL, en donde el ARNg donante comprende al menos 17 nucleótidos que son complementarios a la secuencia diana genómica de la célula, y en donde los brazos de homología son complementarios a secuencias de ácido nucleico que flanquean la secuencia diana genómica de la célula, en donde el ácido nucleico para inserción se configura para la inserción en la secuencia diana genómica de la célula. En determinadas modalidades, la célula es una célula humana. En determinadas modalidades, la célula es una célula inmunitaria. En determinadas modalidades, la célula inmunitaria es un linfocito T. En determinadas modalidades, el linfocito T es un linfocito T primario. En determinadas modalidades, la introducción de la composición formulada para administración no viral comprende electroporación. En determinadas modalidades, la cantidad de plantilla donante es al menos alrededor de 80, 10-120, 10, 20, 30, 40, 50, 60, 70, 90, 100, 110 o 120 µg. En determinadas modalidades, el número de células para una reacción de electroporación

individual es al menos alrededor de 5×10^{-10} , 1×10^{-2} , 3×10^{-4} , 6×10^{-7} , 8×10^{-8} , 9×10^{-9} o 10×10^{-7} . En determinadas modalidades, el número total de células proporcionadas es al menos superior a 10×10^7 y se realiza más de una reacción de electroporación. En determinadas modalidades, el volumen total de suspensión celular es de alrededor de 1 mL. En determinadas modalidades, el método da como resultado una mayor inserción de plantilla en la secuencia diana genómica de las células con respecto a una composición de control por lo demás idéntica pero que carece de las secuencias diana de CDL, opcionalmente en donde la inserción de plantilla aumenta en al menos alrededor de 1×10^{-5} , 1×10^{-2} , 3×10^{-4} o 5 veces con respecto al control.

Definiciones

10 **[0029]** Tal como se usa en la presente memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares «un/una» y «el/la», incluyen la referencia plural a menos que el contexto indique claramente lo contrario.

[0030] Como se usa en el presente documento, el sistema «CRISPR-Cas» hace referencia a una clase de sistemas bacterianos para la defensa contra ácidos nucleicos extraños. Los sistemas
15 CRISPR-Cas se encuentran en una amplia gama de organismos eubacterianos y arqueanos. Los sistemas CRISPR-Cas incluyen los subtipos tipo I, II y III. Los sistemas CRISPR-Cas de tipo II de tipo salvaje utilizan una nucleasa mediada por ARN, por ejemplo, la proteína Cas9, en complejo con ARN guía y activador (por ejemplo, ARN guía simple o ARNgs) para reconocer y escindir ácidos nucleicos extraños, por ejemplo, ácidos nucleicos extraños que incluyen nucleótidos
20 naturales o modificados.

[0031] Como se usa en el presente documento, la expresión «nucleasa dirigible» hace referencia a una proteína que puede reconocer una secuencia de una secuencia de ácido nucleico afín (por ejemplo, un gen diana dentro de un genoma y/o una secuencia diana de CDL) y unirse a la secuencia de ácido nucleico afín. En algunas modalidades, la nucleasa dirigible puede modificar la secuencia
25 de ácido nucleico afín. En algunas modalidades, una nucleasa dirigible puede ser una nucleasa guiada por ARN, por ejemplo, una proteína Cas. En otras modalidades, una nucleasa dirigible puede ser una proteína de fusión que incluye una proteína que puede unirse a una secuencia de ácido nucleico afín (por ejemplo, una proteína efectora de unión al ADN de tipo activador de la transcripción (TAL) o una proteína de unión al ADN con dedos de cinc) y una proteína que puede

modificar una secuencia de ácido nucleico afín (por ejemplo, una nucleasa, un activador o represor de la transcripción). En algunas modalidades, una nucleasa dirigible tiene actividad nucleasa. En otras modalidades, una nucleasa dirigible no tiene actividad nucleasa. En algunas modalidades, una nucleasa dirigible puede modificar una secuencia de ácido nucleico afín mediante escisión del ácido nucleico diana. El ácido nucleico diana escindido puede luego experimentar recombinación homóloga con una plantilla de reparación dirigida por homología (HDR) cercana, tal como mediante reparación dirigida por homología o unión de extremos mediada por homología (HMEJ). En otras modalidades, una nucleasa dirigible (por ejemplo, una nucleasa dirigible sin ninguna actividad nucleasa) puede regular la expresión de una secuencia de ácido nucleico afín. Por ejemplo, una nucleasa dirigible puede ser una proteína de fusión que contiene una proteína efectora de unión al ADN TAL y un activador de la transcripción.

[0032] Como se usa en el presente documento, la expresión «plantilla donante de plásmido» hace referencia a un polinucleótido que incluye una plantilla de reparación dirigida por homología (HDR) y secuencias diana de CDL. Una plantilla de HDR puede incluir un brazo de homología 5', una inserción de nucleótidos (por ejemplo, una secuencia exógena y/o una secuencia que codifica una proteína heteróloga o un fragmento de esta) y un brazo de homología 3' (por ejemplo, véase la Figura 1). Como se describe más detalladamente en el presente documento, la preincubación del complejo RNP que contiene una nucleasa dirigible (por ejemplo, una proteína Cas) y ARNg donante y una plantilla donante de plásmido antes de la electroporación mejora el rendimiento de activación, por ejemplo, a través de una eficiencia de activación mejorada y/o citotoxicidad reducida.

[0033] Como se usa en el presente documento, la expresión «secuencia diana de linealización de coadministración (CDL)» hace referencia a una secuencia de nucleótidos que es reconocida y a la que se une una nucleasa dirigible. En algunas modalidades, una nucleasa dirigible, por ejemplo, una proteína efectora de unión al ADN de tipo activador de la transcripción (TAL) o una proteína de unión al ADN con dedos de cinc, puede reconocer y unirse directamente a una secuencia diana de CDL. En otras modalidades, una nucleasa dirigible, por ejemplo, una nucleasa guiada por ARN, puede reconocer y unirse indirectamente a una secuencia diana de CDL a través de un ARNg donante. Una nucleasa guiada por ARN se une al ARNg donante, mientras que el ARNg donante se hibrida con una secuencia diana de CDL. En algunas modalidades, una secuencia diana de CDL

es una parte de ácido nucleico diana genómico.

[0034] Como se usa en el presente documento, la «nucleasa guiada por ARN» hace referencia a una nucleasa que se une a un ARN guía (ARNg) y utiliza el ARNg para unir selectivamente regiones dentro de un polinucleótido de ADN. En general, una nucleasa guiada por ARN puede unirse selectivamente a casi cualquier secuencia dentro de un polinucleótido de ADN que sea complementaria al ARNg. En algunas modalidades, una nucleasa guiada por ARN tiene actividad nucleasa y puede escindir el enlace (por ejemplo, enlaces fosfodiéster) entre nucleótidos en el polinucleótido de ADN. En otras modalidades, una nucleasa guiada por ARN no tiene actividad nucleasa y puede usarse para unir selectivamente y/o localizar otras proteínas (por ejemplo, activadores o represores transcripcionales) que están fusionadas a la nucleasa guiada por ARN a la región de interés dentro de el polinucleótido de ADN.

[0035] Como se usa en el presente documento, el término «ARN guía» o «ARNg» hace referencia a un ARN dirigido a ADN que puede guiar una nucleasa guiada por ARN (por ejemplo, una proteína Cas) a una secuencia de ácido nucleico afín mediante hibridación con la secuencia de ácido nucleico afín. En algunas modalidades, un ARN guía puede ser un ARN guía simple (ARNgs), que contiene (1) una secuencia guía (por ejemplo, parte equivalente a ARNcr del ARN guía simple) que guía una nucleasa guiada por ARN hacia una secuencia de ácido nucleico afín y (2) una secuencia de andamiaje (por ejemplo, parte equivalente a ARNtracr del ARN guía simple) que interactúa con la nucleasa guiada por ARN. En otras modalidades, un ARN guía puede contener dos componentes, (1) una secuencia guía (por ejemplo, parte equivalente a ARNcr del ARN guía simple) que guía una nucleasa guiada por ARN hacia una secuencia de ácido nucleico afín y (2) una secuencia de andamiaje (por ejemplo, parte equivalente a ARNtracr del ARN guía simple) que interactúa con la nucleasa guiada por ARN. Una parte de la secuencia guía puede hibridarse con una parte de la secuencia de andamiaje para formar el ARN guía de dos componentes.

[0036] Como se usa en el presente documento, la expresión «ARN guía diana» o «ARNg diana» hace referencia a un ARNg que puede hibridarse con una secuencia de ácido nucleico afín para modificarse, por ejemplo, en una ubicación en un polinucleótido de ADN donde se desea la integración de una plantilla de HDR, como el genoma de un linfocito T y/o ubicaciones genómicas

de puerto seguro.

[0037] Como se usa en el presente documento, la expresión «ARN guía donante» o «ARNg donante» hace referencia a un ARNg que puede hibridar una secuencia diana de CDL dentro de una plantilla de plásmido donante. En algunas modalidades, una secuencia diana de CDL puede ser complementaria (por ejemplo, parcialmente complementaria o completamente complementaria) a una parte de igual longitud de la secuencia de un ARNg donante.

[0038] Como se usa en el presente documento, el término «ARN guía simple» o «ARNgs» hace referencia a un ARN dirigido a ADN que incluye (1) una secuencia guía (por ejemplo, parte equivalente a ARNcr del ARN guía simple) que dirige una proteína Cas a una secuencia de ácido nucleico afín y (2) una secuencia de andamiaje (por ejemplo, parte equivalente a ARNtracr del ARN guía simple) que interactúa con una proteína Cas.

[0039] Como se usa en el presente documento, el término «complementario» o «complementariedad» hace referencia a la capacidad de emparejamiento de bases entre nucleobases, nucleósidos o nucleótidos, así como a la capacidad de emparejamiento de bases entre un polinucleótido y otro polinucleótido. En algunas modalidades, un polinucleótido puede tener «complementariedad completa» o ser «completamente complementario» con otro polinucleótido, lo que significa que cuando los dos polinucleótidos están opcionalmente alineados, cada nucleótido en un polinucleótido puede interactuar en el emparejamiento de bases de Watson-Crick con su nucleótido correspondiente en el otro polinucleótido. En otras modalidades, un polinucleótido puede tener «complementariedad parcial» o ser «parcialmente complementario» con otro polinucleótido, lo que significa que cuando los dos polinucleótidos están opcionalmente alineados, al menos 60 % (por ejemplo, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 97 %) pero menos de 100 % de los nucleótidos en un polinucleótido pueden interactuar en el emparejamiento de bases de Watson-Crick con sus nucleótidos correspondientes en el otro polinucleótido. En otras palabras, hay al menos un (por ejemplo, uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve o diez) par de bases de nucleótidos no coincidentes cuando se hibridan los dos polinucleótidos. Los pares de nucleótidos que interactúan en el emparejamiento de bases de Watson-Crick incluyen, por ejemplo, adenina y timina, citosina y guanina, y adenina y uracilo, todos los cuales se emparejan mediante la formación de enlaces de hidrógeno. Los ejemplos de bases no coincidentes incluyen un

emparejamiento de guanina y uracilo, guanina y timina, y adenina y citosina.

[0040] Como se utiliza en el presente documento, la expresión «proteína Cas» hace referencia a una proteína o nucleasa asociada a repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas. Una proteína Cas puede ser una proteína Cas de tipo salvaje o una variante de la proteína Cas. La proteína Cas9 es un ejemplo de una proteína Cas que pertenece al sistema CRISPR-Cas de tipo II (por ejemplo, Rath et al., *Biochimie* 117:119, 2015). Otros ejemplos de proteínas Cas se describen con más detalle en el presente documento. Una proteína Cas de tipo II de origen natural generalmente requiere tanto un ARNcr como un ARNtracr para el reconocimiento y la escisión de ADN con especificidad de sitio. El ARNcr se asocia, a través de una región de complementariedad parcial, con el ARNtracr para guiar la proteína Cas a una región homóloga al ARNcr en el ADN diana que se denomina «protoespaciador». Una proteína Cas de tipo II de origen natural escinde el ADN para generar extremos romos en la rotura de la cadena doble en sitios especificados por una secuencia guía contenida dentro de una transcripción de ARNcr. En algunas modalidades de las composiciones y métodos descritos en el presente documento, una proteína Cas se asocia con un ARNg diana o un ARNg donante para formar un complejo de ribonucleoproteína (RNP). En algunas modalidades de las composiciones y métodos descritos en el presente documento, la proteína Cas tiene actividad nucleasa. En otras modalidades, la proteína Cas no tiene actividad nucleasa.

[0041] Como se usa en el presente documento, el término «variante de proteína Cas» hace referencia a una proteína Cas que tiene al menos una sustitución de aminoácido (por ejemplo, una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más sustituciones de aminoácidos) con respecto a la secuencia de una proteína Cas de tipo salvaje y/o es una versión truncada o fragmento de una proteína Cas de tipo salvaje. En algunas modalidades, una variante de la proteína Cas tiene al menos un 75 % de identidad de secuencia (por ejemplo, al menos un 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia) con respecto a la secuencia de una proteína Cas de tipo salvaje. En algunas modalidades, una variante de la proteína Cas es un fragmento de una proteína Cas de tipo salvaje y tiene al menos una sustitución de aminoácido con respecto a la secuencia de la proteína Cas de tipo salvaje. Una variante de la proteína Cas puede ser una variante de la proteína Cas9. En algunas modalidades, una variante de la proteína Cas tiene actividad nucleasa. En otras modalidades, la variante de la

proteína Cas no tiene actividad nucleasa.

[0042] Como se usa en el presente documento, la expresión «complejo de ribonucleoproteína» o «complejo RNP» hace referencia a un complejo que comprende una proteína Cas o variante (por ejemplo, una proteína Cas9 o variante) y un ARNg.

5 **[0043]** Como se usa en el presente documento, el término «modificar» en el contexto de modificar un ácido nucleico diana en el genoma de una célula hace referencia a inducir un cambio (por ejemplo, escisión) en el ácido nucleico diana. En algunas modalidades, el cambio puede ser un cambio estructural en la secuencia del ácido nucleico diana. Por ejemplo, la modificación puede tomar la forma de insertar una secuencia de nucleótidos en el ácido nucleico diana. Por ejemplo,
10 se puede insertar una secuencia de nucleótidos exógena en el ácido nucleico diana. El ácido nucleico diana también puede extirparse y reemplazarse con una secuencia de nucleótidos exógena. En otro ejemplo, la modificación puede tomar la forma de escindir el ácido nucleico diana sin insertar una secuencia de nucleótidos en el ácido nucleico diana. Por ejemplo, el ácido nucleico diana puede escindirse y extirparse. Dicha modificación se puede realizar, por ejemplo, mediante
15 la inducción de una rotura de doble cadena dentro del ácido nucleico diana, o un par de cortes monocatenarios en cadenas opuestas y flanqueando el ácido nucleico diana. Los métodos para inducir roturas de cadena simple o doble en o dentro de un ácido nucleico diana incluyen el uso de una nucleasa dirigible (por ejemplo, una proteína Cas) como se describe en el presente documento dirigida al ácido nucleico diana. En otras modalidades, modificar un ácido nucleico diana incluye
20 dirigir otra proteína al ácido nucleico diana y no incluye escindir el ácido nucleico diana.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0044] La presente solicitud incluye las siguientes figuras. Se pretende que las figuras ilustren determinadas modalidades y/o características de las composiciones y métodos, y complementen
25 cualquier descripción de las composiciones y métodos. Las figuras no limitan el alcance de las composiciones y métodos, a menos que la descripción escrita indique expresamente que tal es el caso.

[0045] La Figura 1 es un dibujo conceptual que ilustra un mecanismo de ejemplo del diseño del

plásmido de CDL (linealización de coadministración) para mediar la inserción genómica de un casete. Como se muestra a la derecha, agregar la secuencia de CDL (una secuencia objetivo de Cas9 idéntica a la diana genómica de Cas9 para la integración de HDR) y un motivo adyacente de protoespaciador (PAM) de 3 pares de bases fuera de los brazos de homología que flanquean el casete permite que actividad endonucleasa mediada por ribonucleoproteína (RNP) de Cas9-ARNgs libere el casete más los brazos de homología del plásmido, además del corte mediado por RNP de Cas9-ARNgs en el genoma. Por el contrario, con el plásmido estándar, que se muestra a la izquierda, la RNP de Cas9-ARNgs solo corta el locus genómico, pero no el plásmido.

[0046]La Figura 2 son gráficas y diagramas de puntos de citometría de flujo que muestran la eficiencia de activación (KI) y el rendimiento de células KI de linfocitos T electroporadas con una RNP CRISPR y plantillas de reparación dirigida por homología (HDR) de plásmidos estándar o plásmidos CDL. La Figura 2A muestra una serie de diagramas de puntos de citometría de flujo de linfocitos T electroporados con una RNP CRISPR y diferentes dosis (que aumentan de izquierda a derecha y oscilan entre 0 y 100 mg/L) de plantillas de reparación dirigida por homología (HDR) de plásmidos estándar (paneles superiores) o plásmidos CDL (paneles inferiores) que codifican una proteína de superficie marcada con MYC. La Figura 2B es una gráfica que muestra el % de KI con dosis variables de plásmidos estándar o plásmidos CDL. La Figura 2C es un gráfico que muestra el rendimiento de células KI+ viables (tinción viva/muerta; ThermoFisher) por 1 millón de células iniciales con dosis variables de plásmidos estándar o plásmidos CDL.

[0047]La Figura 3 es una gráfica que muestra la relación CD8/CD4 con dosis variables de plásmidos estándar o plásmidos CDL.

[0048]La Figura 4 es una gráfica que muestra la expresión transgénica de linfocitos T electroporados con 3 plásmidos que contienen la secuencia CDL GAGCCATGCTTGGCTTACGA y 3 plásmidos que contienen secuencias no CDL el día 6 después de la electroporación.

DESCRIPCIÓN DETALLADA DE LA DESCRIPCIÓN

[0049]La siguiente descripción enumera diversos aspectos y modalidades de las composiciones y métodos del presente documento. No se pretende que ninguna modalidad particular defina el

alcance de las composiciones y métodos. En cambio, las modalidades simplemente proporcionan ejemplos no taxativos de diversas composiciones y métodos que al menos están incluidos dentro del alcance de las composiciones y métodos descritos. La descripción debe leerse desde la perspectiva de un experto en la técnica; por lo tanto, no se incluye necesariamente información bien conocida por el experto en la técnica.

I. Introducción

[0050] Los linfocitos T modificados con virus están aprobados para la inmunoterapia contra el cáncer, pero se necesitan modificaciones del genoma más versátiles y precisas para una gama más amplia de terapias celulares adoptivas (Yin et al., *Nat Rev Clin Oncol*, 16(5):281-295, 2019; Dunbar et al., *Science* 359:6372, 2018; Cornu et al., *Nat Med* 23:415-423, 2017; y David y Doherty, *Toxicol Sci* 155:315-325, 2017). El sistema CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas)-nucleasa Cas (proteína asociada a CRISPR) es un sistema de nucleasa diseñado que se basa en un sistema bacteriano que se puede utilizar para la modificación genómica. Se basa en parte en la respuesta inmunitaria adaptativa de muchas bacterias y archeas. Cuando un virus o plásmido invade una bacteria, segmentos del ADN del invasor se convierten en ARN de CRISPR (ARNcr) mediante la respuesta «inmunitaria». El ARNcr luego se asocia, a través de una región de complementariedad parcial, con otro tipo de ARN denominado ARNtracr para guiar la nucleasa Cas (por ejemplo, Cas9) a una región homóloga al ARNcr en el ADN diana que se denomina «protoespaciador». La nucleasa Cas (por ejemplo, Cas9) escinde el ADN para generar extremos romos en la rotura de cadena doble en sitios especificados por una secuencia guía de 20 nucleótidos contenida dentro de la transcripción de ARNcr. La nucleasa Cas (por ejemplo, Cas9) puede requerir tanto el ARNcr como el ARNtracr para el reconocimiento y la escisión de ADN con especificidad de sitio. Este sistema ahora se ha modificado genéticamente de modo que el ARNcr y el ARNtracr se pueden combinar en una molécula (el «ARN guía simple» o «ARNgs»), y la parte equivalente al ARNcr del ARNgs se puede modificar genéticamente para guiar la nucleasa Cas (por ejemplo, Cas9) para que se dirija a cualquier secuencia deseada (véase, por ejemplo, Jinek et al. (2012) *Science* 337:816-821; Jinek et al. (2013) *eLife* 2:e00471; Segal (2013) *eLife* 2:e00563). Por lo tanto, el sistema CRISPR-Cas puede modificarse para crear una rotura de cadena doble en una diana deseada en un genoma de una célula, y aprovechar los mecanismos endógenos de la

célula para reparar la rotura inducida mediante reparación dirigida por homología (HDR) o unión de extremos no homólogos (NHEJ).

[0051] Como se describe en el presente documento, los inventores descubrieron secuencias a las que se dirigen las nucleasas que, cuando se añaden en los extremos de la plantilla de reparación dirigida por homología (HDR), pueden mejorar la eficiencia de la modificación del ácido nucleico diana. Las estrategias no virales, como la electroporación, para introducir el sistema CRISPR-Cas en la célula para su modificación genética evitan muchas complicaciones asociadas con la administración viral, como respuestas inmunitarias sistémicas mortales a vectores virales, ineficiencia en la administración viral y sobreexpresión de genes relacionada con la inserción viral. Sin embargo, en algunos casos, puede surgir la necesidad de grandes cantidades de plantilla de HDR y citotoxicidad dependiente de la dosis cuando se emplean estrategias no virales para la administración del sistema CRISPR-Cas.

II. Composiciones

[0052] La descripción proporciona composiciones y métodos para modificar un ácido nucleico diana que incluyen: (a) una proteína nucleasa dirigible; y (b) una plantilla donante de plásmido que comprende: (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante la proteína nucleasa dirigible o un complejo que comprende la proteína nucleasa dirigible, y en donde la composición se formula para administración no viral a una célula. Como se describe en más detalle en el presente documento, en algunas modalidades, cuando la nucleasa dirigible es una efectora de tipo activador de la transcripción (TAL), la efectora TAL puede reconocer y unirse directamente a la secuencia diana de CDL. En algunas modalidades, cuando la nucleasa dirigible es un dedo de cinc, el dedo de cinc puede reconocer y unirse directamente a la secuencia diana de CDL. En otras modalidades, cuando la nucleasa dirigible es una nucleasa guiada por ARN (por ejemplo, una proteína Cas), la nucleasa guiada por ARN puede unirse indirectamente a una secuencia diana de CDL a través de un ARNg donante, que puede

hibridarse con la secuencia diana de CDL. Sin ceñirse a ninguna teoría, la nucleasa dirigible sirve para escindir la secuencia diana de CDL, extirpando así la plantilla de HDR de la plantilla donante de plásmido, lo que permite que la plantilla de HDR participe en la unión de extremos mediada por homología (HMEJ). La eficiencia de activación se puede mantener, o incluso aumentar, utilizando cantidades más bajas de plantilla donante de plásmido a través del proceso dirigido por HMEJ descrito en el presente documento, y también puede dar como resultado una citotoxicidad inducida por ADN reducida. Por lo tanto, las secuencias diana de CDL pueden mejorar la eficiencia de inserción de la plantilla de HDR en las células diana, al tiempo que reducen la toxicidad y aumentan el rendimiento general de las células editadas.

10 **III. CRISPR/Cas**

[0053]En algunas modalidades de las composiciones y métodos descritos en el presente documento, la nucleasa dirigible es una nucleasa guiada por ARN, y la plantilla donante comprende además motivos adyacentes de protoespaciador (PAM) inmediatamente adyacentes a las secuencias diana de CDL. La composición también puede comprender además un ARN guía diana (ARNg) que es complementario a un ácido nucleico diana a modificar, por ejemplo, una secuencia diana genómica, tal como un sitio deseado para la inserción del transgén en una célula (por ejemplo, un linfocito T). El ARNg diana puede formar un primer complejo RNP con la primera nucleasa guiada por ARN y guiar la primera nucleasa guiada por ARN (por ejemplo, proteína Cas) hacia el ácido nucleico diana. En algunas modalidades, una parte del ARNg diana (por ejemplo, una parte del ARNg diana que tiene al menos 17 nucleótidos (por ejemplo, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos) es complementaria al ácido nucleico diana.

[0054]La composición también puede comprender además un ARNg donante que sea complementario a la secuencia diana de CDL. El ARNg donante puede formar una segunda RNP con la segunda nucleasa guiada por ARN. La secuencia diana de CDL en la plantilla donante puede hibridarse con el ARNg donante o una parte de este. Por lo tanto, el complejo que contiene la segunda nucleasa guiada por ARN, el ARNg donante y la plantilla donante puede facilitar la extirpación de una plantilla de HDR y, sin desear ceñirse a ninguna teoría, promover que se produzca la recombinación homóloga en el sitio de integración en el ácido nucleico diana escindido por el ARNg diana. En algunas modalidades, las secuencias del ARNg diana y del ARNg donante

son iguales. En algunas modalidades, las secuencias del ARNg diana y del ARNg donante son diferentes. En algunas modalidades, la primera y la segunda nucleasas guiadas por ARN son iguales. En algunas modalidades, la primera y la segunda nucleasas guiadas por ARN son diferentes.

- 5 **[0055]**La composición puede contener una nucleasa dirigible y un ARNg diana y/o un ARNg donante en una relación molar de entre 1:10 y 2:1 (por ejemplo, entre 1:5 y 2:1, entre 2:5 y 2:1, entre 3:5 y 2:1, entre 4:5 y 2:1, entre 1:1 y 2:1, entre 1:10 y 1:1, entre 1:10 y 4:5, entre 1:10 y 3:5, entre 1:10 y 2:5 o entre 1:10 y 1:5), respectivamente. La composición puede contener una nucleasa dirigible y una plantilla donante de plásmido en una relación molar de entre 10:1 y 1000:1 (por ejemplo, entre 50:1 y 1000:1, entre 100:1 y 1000:1, entre 200:1 y 1000:1, entre 300:1 y 1000:1, entre 400:1 y 1000:1, entre 500:1 y 1000:1, entre 600:1 y 1000:1, entre 700:1 y 1000:1, entre 800:1 y 1000:1, entre 900:1 y 1000:1, entre 10:1 y 900:1, entre 10:1 y 800:1, entre 10:1 y 700:1, entre 10:1 y 600:1, entre 10:1 y 500:1, entre 10:1 y 400:1, entre 10:1 y 300:1, entre 10:1 y 200:1, entre 10:1 y 100:1 o entre 10:1 y 50:1), respectivamente.
- 10
- 15 **[0056]**También se puede fusionar una nucleasa guiada por ARN con un péptido o proteína de localización. Por ejemplo, la nucleasa guiada por ARN puede fusionarse con una o más secuencias de señal de localización nuclear (NLS), que pueden dirigir la nucleasa y los complejos RNP que forma hacia el núcleo para modificar el ácido nucleico diana. En la técnica se conocen ejemplos de secuencias de NLS, por ejemplo, como se describe en Lange et al., *J Biol Chem.* 282(8):5101-5, 2007, y también incluyen, de modo no taxativo, AVKRPAATKKAGQAKKKKLD, MSRRRKANPTKLSNAKKLAKEVEN, PAAKRVKLD, KLKIKRPVK y PKKKRKV. Ejemplos de otros péptidos o proteínas que pueden usarse para una nucleasa guiada por ARN, tales como péptidos que penetran en las células y péptidos que se dirigen a las células, están disponibles en la técnica y se describen, por ejemplo, en Vivès et al., *Biochim Biophys Acta.* 1786(2):126-38, 2008.
- 20
- 25

IV. ARN guía simple

[0057]Una proteína Cas puede guiarse a su ADN afín respectivo a escindir (por ejemplo, una secuencia diana de CDL y/o una secuencia diana genómica) mediante un ARN guía simple (ARNgs). Un ARNg es una versión del ARN guía de dos piezas de origen natural (ARNcr y

ARNtracr) modificado en una secuencia única y continua. Un ARNGs puede contener una secuencia guía (por ejemplo, la parte equivalente a ARNcr del ARNGs) que dirige la proteína Cas a la secuencia de ácido nucleico afín y una secuencia de andamiaje que interactúa con la proteína Cas (por *ejemplo*, la parte equivalente a ARNtracr del ARNGs). Se puede seleccionar un ARNGs usando un software. Como ejemplo no taxativo, las consideraciones para seleccionar un ARNGs pueden incluir, por ejemplo, la secuencia PAM para la proteína Cas9 que se va a usar y estrategias para minimizar modificaciones fuera de diana. Herramientas, como NUPACK® y CRISPR Design Tool, pueden proporcionar secuencias para preparar el ARNGs, evaluar la eficiencia de la modificación de la diana y/o evaluar la escisión en sitios fuera de la diana.

10 **Secuencia guía**

[0058]La secuencia guía en el ARNGs puede ser complementaria a una secuencia específica dentro de una secuencia de ácido nucleico afín (por ejemplo, una secuencia diana de CDL y/o una secuencia diana genómica). El extremo 3' de una secuencia de ácido nucleico afín puede ir seguido de una secuencia PAM. Una secuencia guía es generalmente complementaria a aproximadamente 20 nucleótidos secuencia arriba de la secuencia PAM. En general, una proteína Cas9 o una variante de esta escinde aproximadamente tres nucleótidos secuencia arriba de una secuencia PAM. La secuencia guía en el ARNGs puede ser complementaria a cualquiera de las cadenas de una secuencia de ácido nucleico afín.

[0059]En algunas modalidades, la secuencia guía de un ARNGs puede comprender de alrededor de 10 a alrededor de 2000 ácidos nucleicos, por ejemplo, de alrededor de 10 a alrededor de 100 ácidos nucleicos, de alrededor de 10 a alrededor de 500 ácidos nucleicos, de alrededor de 10 a alrededor de 1000 ácidos nucleicos, de alrededor de 10 a alrededor de 1500 ácidos nucleicos, de alrededor de 10 a alrededor de 2000 ácidos nucleicos, de alrededor de 50 a alrededor de 100 ácidos nucleicos, de alrededor de 50 a alrededor de 500 ácidos nucleicos, de alrededor de 50 a alrededor de 1000 ácidos nucleicos, de alrededor de 50 a alrededor de 1500 ácidos nucleicos, de alrededor de 50 a alrededor de 2000 ácidos nucleicos, de alrededor de 100 a alrededor de 500 ácidos nucleicos, de alrededor de 100 a alrededor de 1000 ácidos nucleicos, de alrededor de 100 a alrededor de 1500 ácidos nucleicos, de alrededor de 100 a alrededor de 2000 ácidos nucleicos, de alrededor de 500 a alrededor de 1000 ácidos nucleicos, de alrededor de 500 a alrededor de 1500 ácidos nucleicos, de

alrededor de 500 a alrededor de 2000 ácidos nucleicos, de alrededor de 1000 a alrededor de 1500 ácidos nucleicos, de alrededor de 1000 a alrededor de 2000 ácidos nucleicos o alrededor de 1500 a alrededor de 2000 ácidos nucleicos en el extremo 5' del ARNGs que puede dirigir la proteína Cas a una secuencia de ácido nucleico afín utilizando el emparejamiento de bases de complementariedad de ARN-ADN. En algunas modalidades, la secuencia guía de un ARNGs comprende alrededor de 100 ácidos nucleicos en el extremo 5' del ARNGs que pueden dirigir la proteína Cas a un sitio de secuencia de ácido nucleico afín usando el emparejamiento de bases de complementariedad de ARN-ADN. En algunas modalidades, la secuencia guía comprende 20 ácidos nucleicos en el extremo 5' del ARNGs que pueden dirigir la proteína Cas a un sitio de secuencia de ácido nucleico afín (por ejemplo, una secuencia diana de CDL y/o secuencia diana genómica) usando el emparejamiento de bases de complementariedad de ARN-ADN. En otras modalidades, la secuencia guía comprende menos de 20, por ejemplo, 19, 18, 17 o menos, ácidos nucleicos que son complementarios a una secuencia de ácido nucleico afín. En algunos casos, la secuencia guía en el ARNGs contiene al menos una falta de coincidencia de ácido nucleico en la región de complementariedad de una secuencia de ácido nucleico afín. En algunos casos, la secuencia guía contiene de alrededor de 1 a alrededor de 10 faltas de coincidencia de ácido nucleico en la región de complementariedad de una secuencia de ácido nucleico afín.

Secuencia de andamiaje

[0060] La secuencia de andamiaje en el ARNGs puede servir como una secuencia de unión a proteínas que interactúa con la proteína Cas o una variante de esta. En algunas modalidades, la secuencia de andamiaje en el ARNGs puede comprender dos tramos complementarios de nucleótidos que se hibridan entre sí para formar un dúplex de ARN bicatenario (dúplex de ARNbc). La secuencia de andamiaje puede tener estructuras tales como tallo inferior, bulto, tallo superior, nexo y/u horquilla. En algunas modalidades, la secuencia de andamiaje en el ARNGs puede ser de alrededor de 90 ácidos nucleicos a alrededor de 120 ácidos nucleicos, por ejemplo, de alrededor de 90 ácidos nucleicos a alrededor de 115 ácidos nucleicos, de alrededor de 90 ácidos nucleicos a alrededor de 110 ácidos nucleicos, de alrededor de 90 ácidos nucleicos a alrededor de 105 ácidos nucleicos, de alrededor de 90 ácidos nucleicos a alrededor de 100 ácidos nucleicos, de alrededor de 90 ácidos nucleicos a alrededor de 95 ácidos nucleicos, de alrededor de 95 ácidos nucleicos a

alrededor de 120 ácidos nucleicos, de alrededor de 100 ácidos nucleicos a alrededor de 120 ácidos nucleicos, de alrededor de 105 ácidos nucleicos a alrededor de 120 ácidos nucleicos, de alrededor de 110 ácidos nucleicos a alrededor de 120 ácidos nucleicos o de alrededor de 115 ácidos nucleicos a alrededor de 120 ácidos nucleicos.

5 **V. ARN guía (ARNg)**

[0061] Los ARN guía (ARNg), incluidos los ARNg diana y los ARNg donantes descritos en el presente documento, en general hacen referencia a un ARN dirigido a ADN que contiene (1) una secuencia guía que es complementaria a una secuencia de ácido nucleico afín (por ejemplo, una secuencia diana de CDL y/o secuencia diana genómica) y guía la nucleasa guiada por ARN hacia la secuencia de ácido nucleico afín y (2) una secuencia de andamiaje que interactúa y se une con la nucleasa guiada por ARN. En algunas modalidades de la descripción, un ARNg diana y un ARNg donante tienen la misma secuencia. En algunas modalidades de la descripción, un ARNg diana y un ARNg donante incluyen la misma secuencia, por ejemplo, comparten una secuencia de ácido nucleico capaz de hibridarse tanto con una secuencia diana que se va a modificar como con una secuencia diana de CDL. En otras modalidades de la descripción, un ARNg diana y un ARNg donante tienen secuencias diferentes. En las composiciones y métodos descritos en el presente documento, un ARNg comprende una parte que es complementaria a una secuencia de ácido nucleico afín. Una vez que un ARNg forma un complejo RNP con una nucleasa dirigible (por ejemplo, una primera nucleasa guiada por ARN), el complejo RNP puede guiarse a una secuencia de ácido nucleico afín mediante la complementariedad entre el ARNg y la secuencia de ácido nucleico afín. En algunas modalidades, la nucleasa dirigible es una proteína Cas9. La proteína Cas9 «identifica» una secuencia de ácido nucleico afín identificando primero un motivo adyacente de protoespaciador (PAM) de 3 pares de bases ubicado hacia 3' de la secuencia de ácido nucleico afín. Una vez que se identifica el PAM, un ARNg en el complejo RNP se hibrida con la secuencia de ácido nucleico afín secuencia arriba del PAM. En algunas modalidades, un ARNg incluye una parte de nucleótidos que son complementarias a una parte en la secuencia de ácido nucleico afín que está aproximadamente 20 nucleótidos secuencia arriba de la secuencia PAM. En general, una proteína Cas9 o una variante de esta escinde aproximadamente tres nucleótidos secuencia arriba de la secuencia PAM. Se puede seleccionar un ARNg usando un software. Como ejemplo no taxativo,

las consideraciones para seleccionar un ARNg pueden incluir, por ejemplo, la secuencia PAM para la nucleasa guiada por ARN que se va a usar y estrategias para minimizar modificaciones fuera de la diana. Herramientas, como NUPACK® y CRISPR Design Tool, pueden proporcionar secuencias para preparar el ARNg, evaluar la eficiencia de la modificación de la diana y/o evaluar la escisión en sitios fuera de la diana.

[0062]En algunas modalidades, un ARNg comprende una parte de al menos 17 nucleótidos (por ejemplo, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos) que son complementarios a una secuencia de ácido nucleico afín. En algunas modalidades, un ARNg puede ser completamente complementario o parcialmente complementario a una secuencia de ácido nucleico afín. En algunas modalidades, al menos el 60 % (por ejemplo, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 97 %) de los nucleótidos en una secuencia de ácido nucleico afín pueden interactuar en emparejamiento de bases de Watson-Crick con sus nucleótidos correspondientes en un ARNg.

[0063]Como se describe con más detalle en el presente documento, una secuencia diana de CDL y el PAM en uno o ambos extremos de la plantilla de HDR se pueden diseñar con diferentes configuraciones.

[0064]En algunas modalidades de la descripción, un ARNg diana y un ARNg donante tienen la misma secuencia (por ejemplo, una secuencia diana genómica y una secuencia diana de CDL son las mismas) y cada uno forma un complejo con la misma especie de nucleasa dirigible (por ejemplo, ambos tienen la misma secuencia y forman un complejo RNP con la misma especie de proteína Cas). En este caso, el ARNg puede formar un primer complejo RNP con la nucleasa guiada por ARN. El primer complejo RNP puede unirse al ácido nucleico diana mediante la hibridación entre el ARNg y el ácido nucleico diana. El ARNg también puede formar un segundo complejo RNP con la nucleasa guiada por ARN y la plantilla donante. En este segundo complejo RNP, el ARNg puede unirse a la secuencia diana de la proteína de unión al ADN en la plantilla donante para llevar la plantilla donante a la ubicación intracelular deseada (por ejemplo, el núcleo) para que se produzca la recombinación homóloga en el ácido nucleico diana escindido. En algunas modalidades, el ARNg y la secuencia diana de la proteína de unión al ADN solo tienen complementariedad parcial. En algunas modalidades de la descripción, un ARNg diana y un ARNg donante tienen una secuencia diferente (por ejemplo, una secuencia diana genómica y una

secuencia diana de CDL son distintas). Los ARNg diana y ARNg donantes distintos pueden formar cada uno un complejo con la misma especie de nucleasa dirigible (por ejemplo, cada uno forma un complejo RNP con la misma especie de proteína Cas). Los ARNg diana y ARNg donantes distintos pueden formar cada uno un complejo con una especie distinta de nucleasa dirigible (por ejemplo, cada uno forma un complejo RNP con la misma especie de proteína Cas).

VI. Plantilla donante

[0065] La plantilla de HDR, las secuencias diana de CDL y las secuencias PAM correspondientes pueden tener varias configuraciones diferentes en una plantilla donante de plásmido para mejorar la reparación dirigida por homología entre la plantilla de HDR y un ácido nucleico diana. En general, una plantilla donante incluye (i) una plantilla de reparación dirigida por homología (HDR); (ii) una primera secuencia diana de linealización de coadministración (CDL), en donde la primera secuencia diana de CDL está operativamente unida hacia 5' de la plantilla de HDR; y (iii) una segunda secuencia diana de CDL, en donde la segunda secuencia diana de CDL está operativamente unida hacia 3' de la plantilla de HDR, y en donde cada una de la primera y la segunda secuencias diana de CDL son capaces de escindirse mediante la proteína nucleasa dirigible o un complejo que comprende la proteína nucleasa dirigible. En modalidades donde la nucleasa dirigible es una proteína Cas (por ejemplo, Cas9), cada una de las secuencias diana de CDL primera y segunda están operativamente unidas a un PAM de 3 pares de bases ubicado hacia 3' de las secuencias diana de CDL. Una secuencia diana de CDL y el PAM correspondiente se pueden orientar de manera que el PAM se ubique entre la secuencia diana de CDL y una plantilla de HDR. Una secuencia diana de CDL y el PAM correspondiente se pueden orientar de manera que la secuencia diana de CDL se ubique entre el PAM y una plantilla de HDR. Cada secuencia diana de CDL y el PAM correspondiente en una plantilla donante de plásmido se pueden orientar de forma independiente. Un PAM que corresponde a la primera secuencia diana de CDL puede ubicarse entre la secuencia diana de CDL y una plantilla de HDR. La primera secuencia diana de CDL puede ubicarse entre el PAM y una plantilla de HDR. Un PAM que corresponde a la segunda secuencia diana de CDL puede ubicarse entre la secuencia diana de CDL y una plantilla de HDR. La segunda secuencia diana de CDL puede ubicarse entre el PAM y una plantilla de HDR. En un ejemplo no taxativo, un PAM correspondiente a la primera secuencia diana de CDL se ubica entre la primera

secuencia diana de CDL y una plantilla de HDR, y un PAM correspondiente a la segunda secuencia diana de CDL se ubica entre la segunda secuencia diana de CDL y la plantilla de HDR.

[0066]En algunas modalidades, el tamaño o la longitud de la plantilla donante de plásmido es mayor que alrededor de 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 1 kb, 1,1 kb, 1,2 kb, 1,3 kb, 1,4 kb, 1,5 kb, 1,6 kb, 1,7 kb, 1,8 kb, 1,9 kb, 2,0 kb, 2,1 kb, 2,2 kb, 2,3 kb, 2,4 kb, 2,5 kb, 2,6 kb, 2,7 kb, 2,8 kb, 2,9 kb, 3 kb, 3,1 kb, 3,2 kb, 3,3 kb, 3,4 kb, 3,5 kb, 3,6 kb, 3,7 kb, 3,8 kb, 3,9 kb, 4,0 kb, 4,1 kb, 4,2 kb, 4,3 kb, 4,4 kb, 4,5 kb, 4,6 kb, 4,7 kb, 4,8 kb, 4,9 kb, 5,0 kb, 5,1 kb, 5,2 kb, 5,3 kb, 5,4 kb, 5,5 kb, 5,6 kb, 5,7 kb, 5,8 kb, 5,9 kb, 6,0 kb, 6,1 kb, 6,2 kb, 6,3 kb, 6,4 kb, 6,5 kb, 6,6 kb, 6,7 kb, 6,8 kb, 6,9 kb, 7,0 kb, 7,1 kb, 7,2 kb, 7,3 kb, 7,4 kb, 7,5 kb, 7,6 kb, 7,7 kb, 7,8 kb, 7,9 kb, 8,0 kb, 8,1 kb, 8,2 kb, 8,3 kb, 8,4 kb, 8,5 kb, 8,6 kb, 8,7 kb, 8,8 kb, 8,9 kb, 9,0 kb, 9,1 kb, 9,2 kb, 9,3 kb, 9,4 kb, 9,5 kb, 9,6 kb, 9,7 kb, 9,8 kb, 9,9 kb, 10,0 kb, cualquier tamaño de plantilla entre estos tamaños, o mayor que 10 kb. Por ejemplo, el tamaño de la plantilla puede ser de alrededor de 200 pb a alrededor de 500 pb, de alrededor de 200 pb a alrededor de 750 pb, de alrededor de 200 pb a alrededor de 1 kb, de alrededor de 200 pb a alrededor de 1,5 kb, de alrededor de 200 pb a alrededor de 2,0 kb, de alrededor de 200 pb a alrededor de 2,5 kb, de alrededor de 200 pb a alrededor de 3,0 kb, de alrededor de 200 pb a alrededor de 3,5 kb, de alrededor de 200 pb a alrededor de 4,0 kb, de alrededor de 200 pb a alrededor de 4,5 kb, de alrededor de 200 pb a alrededor de 5,0 kb. En algunos casos, el tamaño de la plantilla es lo suficientemente grande y tiene la cantidad suficiente para ser letal como el ADN no conjugado.

[0067]En algunas modalidades, una plantilla donante codifica una proteína heteróloga o un fragmento de esta. En algunas modalidades, una plantilla donante incluye secuencias reguladoras, por ejemplo, una secuencia promotora y/o una secuencia potenciadora para regular la expresión de la proteína heteróloga o fragmento de esta, por ejemplo, después de la inserción en el genoma de una célula. Una proteína heteróloga puede incluir un receptor de antígeno quimérico (CAR). Una proteína heteróloga puede incluir un receptor de linfocito T (TCR).

[0068]En algunas modalidades, una plantilla donante de plásmido incluye una secuencia exógena tal como una secuencia de nucleótidos exógena. Una secuencia exógena puede incluir una proteína heteróloga codificada o un fragmento de esta. Una secuencia exógena puede incluir un gen o un

fragmento de este. Una secuencia de nucleótidos exógena puede ser una secuencia corta, por ejemplo, de 3-100 nucleótidos de longitud. Una secuencia de nucleótidos exógena de interés puede ser un único nucleótido. Además, una secuencia de nucleótidos exógena de interés puede ser una secuencia larga, por ejemplo, de 500-3000 nucleótidos de longitud. Una secuencia de nucleótidos exógena de interés puede codificar o no codificar una secuencia polipeptídica. Además, se puede insertar una secuencia de nucleótidos exógena de interés en una célula de manera que forme un gen quimérico tras la inserción. Por ejemplo, se puede insertar una parte de receptor exógeno en marco en una secuencia codificante de receptor endógeno para producir una secuencia codificante de receptor quimérico que, después de la edición, incluye la parte de receptor exógeno operativamente unida a una parte intracelular endógena (por ejemplo, para la transducción de señales).

[0069]En algunos ejemplos, un gen o parte puede ser una secuencia de nucleótidos que codifica una proteína (es decir, una secuencia de nucleótidos que codifica una secuencia de polipéptido). En general, puede usarse cualquier nucleótido codificante de proteínas. En algunos ejemplos, una secuencia de nucleótidos que codifica una proteína codifica una proteína útil en terapias con células autólogas (por ejemplo, terapias con linfocitos T autólogos). En algunos ejemplos, una secuencia de nucleótidos codificante de proteína puede incluir, de modo no taxativo, un factor que modula el sistema inmunitario, una citocina, un factor que modula la función de los linfocitos T, un factor que promueve la supervivencia de los linfocitos T, un factor que promueve la función de los linfocitos T o un inhibidor de punto de control inmunitario. Una secuencia de nucleótidos codificante de proteína, particularmente una proteína secretada o proteínas unidas a membrana, puede incluir una secuencia de nucleótidos que codifica un péptido señal. El péptido señal puede ser endógeno a la proteína codificada por la secuencia de nucleótidos codificante de proteína. El péptido señal puede ser exógeno a la proteína codificada por la secuencia de nucleótidos codificante de proteína.

[0070]En algunos ejemplos, un gen o parte de este puede ser una secuencia de nucleótidos no codificante de proteína. En general, puede usarse cualquier nucleótido no codificante de proteína. En algunos casos, una secuencia de nucleótidos no codificante de proteína puede ser una secuencia de nucleótidos útil en terapias con células autólogas (por ejemplo, terapias con linfocitos T autólogos). En algunos casos, una secuencia de nucleótidos no codificante de proteína puede incluir, de modo no taxativo, un ARNhc, un ARNpi, un miARN y un ARNlnc.

[0071]Aunque una secuencia de nucleótidos que codifica al menos una parte de un gen (por ejemplo, un gen exógeno de interés) puede, en general, tener cualquier tamaño, se pueden tener en cuenta consideraciones prácticas, tales como el impacto del tamaño del gen en el tamaño general de la plantilla y en la consiguiente eficiencia general de edición. Por lo tanto, en un aspecto particular, se proporcionan en el presente documento células modificadas que se editan genómicamente, o que son capaces de editarse genómicamente, para expresar un gen exógeno mayor o igual a 100 bases de longitud a tasas de eficiencia de HR mayores que las descritas previamente (por ejemplo, un mayor porcentaje de una población que tiene una secuencia de polinucleótidos integrada), particularmente cuando se utilizan métodos de administración no virales. Las tasas de eficiencia de HR mejoradas se aplican de manera similar a genes de más de 100 bases de longitud, tales como la introducción de secuencias exógenas mayores o iguales a 200 bases de longitud, mayores o iguales a 400 bases de longitud, mayores o iguales a 500 bases de longitud, mayores o iguales a 600 bases de longitud, mayores o iguales a 750 bases de longitud, mayores o iguales a 1000 bases de longitud, mayores o iguales a 1500 bases de longitud, o iguales a 2000 bases de longitud, mayores o iguales a 3000 bases de longitud o mayores o iguales a 4000 bases de longitud. La al menos una parte de un gen puede tener una longitud mayor o igual a 800 bases. La al menos una parte de un gen puede tener una longitud mayor o igual a 1600 bases.

[0072]Las secuencias exógenas pueden tener entre 100-200 bases de longitud, entre 100-300 bases de longitud, entre 100-400 bases de longitud, entre 100-500 bases de longitud, entre 100-600 bases de longitud, entre 100-700 bases de longitud, entre 100-800 bases de longitud, entre 100-900 bases de longitud o entre 100-1000 bases de longitud. Las secuencias exógenas pueden tener entre 100-2000 bases de longitud, entre 100-3000 bases de longitud, entre 100-4000 bases de longitud, entre 100-5000 bases de longitud, entre 100-6000 bases de longitud, entre 100-7000 bases de longitud, entre 100-8000 bases de longitud, entre 100-9000 bases de longitud o entre 100-10.000 bases de longitud. Las secuencias exógenas pueden tener entre 1000-2000 bases de longitud, entre 1000-3000 bases de longitud, entre 1000-4000 bases de longitud, entre 1000-5000 bases de longitud, entre 1000-6000 bases de longitud, entre 1000-7000 bases de longitud, entre 1000-8000 bases de longitud, entre 1000-9000 bases de longitud o entre 1000-10.000 bases de longitud.

[0073]Las secuencias exógenas pueden ser mayores o iguales a 10 bases de longitud, mayores o iguales a 20 bases de longitud, mayores o iguales a 30 bases de longitud, mayores o iguales a 40

bases de longitud, mayores o iguales a 50 bases de longitud, mayores o iguales a 60 bases de longitud, mayores o iguales a 70 bases de longitud, mayores o iguales a 80 bases de longitud, mayores o iguales a 90 bases de longitud o mayores o iguales a 95 bases de longitud. Las secuencias exógenas pueden tener entre 1-100 bases de longitud, entre 1-90 bases de longitud, entre 1-80 bases de longitud, entre 1-70 bases de longitud, entre 1-60 bases de longitud, entre 1-50 bases de longitud, entre 1-40 bases de longitud o entre 1-30 bases de longitud. Las secuencias exógenas pueden tener entre 1-20 bases de longitud, entre 2-20 bases de longitud, entre 3-20 bases de longitud, entre 5-20 bases de longitud, entre 10-20 bases de longitud o entre 15-20 bases de longitud. Las secuencias exógenas pueden tener entre 1-10 bases de longitud, entre 2-10 bases de longitud, entre 3-10 bases de longitud, entre 5-10 bases de longitud, entre 1-5 bases de longitud o entre 1-15 bases de longitud. Las secuencias exógenas pueden tener 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 115, 120, 125, 150, 175, 200, 225 o 250 bases de longitud. Las secuencias exógenas pueden tener 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14 bases de longitud. Las secuencias exógenas pueden ser mayores que alrededor de 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 1 kb, 1,1 kb, 1,2 kb, 1,3 kb, 1,4 kb, 1,5 kb, 1,6 kb, 1,7 kb, 1,8 kb, 1,9 kb, 2,0 kb, 2,1 kb, 2,2 kb, 2,3 kb, 2,4 kb, 2,5 kb, 2,6 kb, 2,7 kb, 2,8 kb, 2,9 kb, 3 kb, 3,1 kb, 3,2 kb, 3,3 kb, 3,4 kb, 3,5 kb, 3,6 kb, 3,7 kb, 3,8 kb, 3,9 kb, 4,0 kb, 4,1 kb, 4,2 kb, 4,3 kb, 4,4 kb, 4,5 kb, 4,6 kb, 4,7 kb, 4,8 kb, 4,9 kb, 5,0 kb, 5,1 kb, 5,2 kb, 5,3 kb, 5,4 kb, 5,5 kb, 5,6 kb, 5,7 kb, 5,8 kb, 5,9 kb, 6,0 kb, 6,1 kb, 6,2 kb, 6,3 kb, 6,4 kb, 6,5 kb, 6,6 kb, 6,7 kb, 6,8 kb, 6,9 kb, 7,0 kb o cualquier tamaño de plantilla entre estos tamaños.

[0074]En ejemplos en los que se introducen múltiples secuencias exógenas, las múltiples secuencias exógenas pueden tener diferentes tamaños, por ejemplo, una primera secuencia exógena puede ser mayor o igual a 100 bases y una segunda secuencia exógena puede ser mayor o igual a 100 bases, o una primera secuencia exógena puede ser mayor o igual a 100 bases y una segunda secuencia exógena puede tener menos de 100 bases (por ejemplo, entre 1-100 bases de longitud).

[0075]En general, una plantilla donante de plásmido es un plásmido de ADN circular. En algunos casos, una plantilla donante de plásmido es un plásmido bicatenario. En algunos casos, una plantilla

donante de plásmido es un plásmido monocatenario. En algunos casos, una plantilla donante de plásmido es un minicírculo. En algunos casos, una plantilla donante de plásmido es un nanoplásmido.

5 **[0076]** Las secuencias diana de CDL y/o componentes de plantilla de HDR (brazos de homología, gen de interés, etc.) pueden introducirse en plantillas de ADNcd de cualquier formato, incluidas secuencias de ADNcd lineales producidas por PCR, digestiones con enzimas de restricción o cualquier otro método de linealización, así como secuencias de ADNcd circulares como los plásmidos. En el caso de un plásmido, las secuencias diana de CDL se pueden clonar en el plásmido fuera de los brazos de homología y las regiones de inserción de ADN, incluso, de modo no taxativo, 10 de forma adyacente a los bordes de los brazos de homología. De manera similar a las plantillas de ADNcd lineal, el complejo de proteína de unión al ADN (por ejemplo, RNP elaborado a partir de Cas9 y ARNg) se puede incubar brevemente con la plantilla de ADN plasmídico para permitir la unión del plásmido de ADN mediante la RNP antes de su introducción en la célula (por ejemplo, a través de electroporación). Véanse, por ejemplo, las Figuras 1, 2 y 10B de la Publicación de 15 Patente Internacional No. WO2018232356 y el párrafo [0100] de la Publicación de Patente Internacional No. WO2019084552.

[0077] Una plantilla donante de plásmido puede contener además una o más secuencias espaciadoras adicionales entre una secuencia diana de CDL y una plantilla de HDR. En algunas modalidades, una secuencia espaciadora puede tener al menos 2 nucleótidos, por ejemplo, entre 2 20 y 24 nucleótidos (por ejemplo, entre 2 y 22, entre 2 y 20, entre 2 y 18, entre 2 y 16, entre 2 y 14, entre 2 y 12, entre 2 y 10, entre 2 y 8, entre 2 y 6, entre 2 y 4, entre 4 y 24, entre 6 y 24, entre 8 y 24, entre 10 y 24, entre 12 y 24, entre 14 y 24, entre 16 y 24, entre 18 y 24, entre 20 y 24 o entre 22 y 24 nucleótidos; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24 nucleótidos).

25 **VII. Nucleasa dirigible**

[0078] Como se describió anteriormente, en algunas modalidades de las composiciones y métodos descritos en el presente documento, una nucleasa dirigible es una nucleasa guiada por ARN (por ejemplo, una proteína Cas). Una nucleasa dirigible puede reconocer una secuencia de una secuencia de ácido nucleico afín (por ejemplo, un gen diana dentro de un genoma y/o una secuencia diana de

CDL), unirse a la secuencia de ácido nucleico afín y modificar la secuencia de ácido nucleico afín. En otras modalidades, la nucleasa dirigible puede ser una proteína de fusión que incluye una proteína que puede unirse a una secuencia de ácido nucleico afín y una proteína que puede modificar una secuencia de ácido nucleico afín (por ejemplo, una nucleasa, un activador o represor de la transcripción).

[0079]En algunas modalidades, una nucleasa dirigible tiene actividad nucleasa. Por ejemplo, la nucleasa dirigible puede modificar una secuencia de ácido nucleico afín mediante escisión de la secuencia de ácido nucleico afín. La secuencia de ácido nucleico afín escindida puede luego experimentar recombinación homóloga (por ejemplo, a través de HMEJ) con una plantilla cercana de reparación dirigida por homología (HDR) (por ejemplo, una plantilla de HDR proporcionada por una plantilla donante de plásmido). Por ejemplo, una nucleasa Cas puede dirigir la escisión de una o ambas cadenas en una ubicación en una secuencia de ácido nucleico afín. Los ejemplos no taxativos de nucleasas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida como Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, Cpf1, homólogos de estas, variantes de estas, mutantes de estas y derivados de estas. Hay tres tipos principales de nucleasas Cas (de tipo I, tipo II y tipo III), y 10 subtipos que incluyen 5 proteínas de tipo I, 3 de tipo II y 2 de tipo III (véase, por ejemplo, Hochstrasser y Doudna, *Trends Biochem Sci*, 2015:40(1):58-66). Las nucleasas Cas de tipo II incluyen Cas1, Cas2, Csn2, Cas9 y Cfp1. Estas nucleasas Cas son conocidas por los expertos en la técnica. Por ejemplo, la secuencia de aminoácidos del polipéptido de Cas9 de tipo salvaje de *Streptococcus pyogenes* se establece, por ejemplo, en el n.º de sec. de ref. de NBCI NP_269215, y la secuencia de aminoácidos del polipéptido Cas9 de tipo salvaje de *Streptococcus thermophilus* se establece, por ejemplo, en el n.º de sec. de ref. de NBCI WP_011681470.

[0080]Las nucleasas Cas, por ejemplo, nucleasas Cas9, pueden derivarse de una variedad de especies bacterianas que incluyen, de modo no taxativo, *Veillonella atypical*, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus*

kitaharae, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Finnegoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subesp. *Torquens*, *Ilyobacter polytropus*,
5 *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus cellulolyticus*, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratifactor salsuginis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subesp. *Succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum*
10 *rubrum*, *Candidatus Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subesp. *Jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subesp. *Multocida*, *Sutterella*
15 *wadsworthensis*, *proteobacterium*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes* y *Francisella novicida*.

[0081]La proteína Cas9 hace referencia a una proteína nickasa o una proteína nucleasa de unión a ADN bicatenario guiada por ARN. La nucleasa Cas9 en estado natural tiene dos dominios funcionales, por ejemplo, RuvC y HNH, que cortan diferentes cadenas de ADN. Cas9 puede
20 inducir las roturas bicatenarias en ADN genómico (ADN diana) cuando ambos dominios funcionales están activos. La enzima Cas9 puede comprender uno o más dominios catalíticos de una proteína Cas9 derivada de bacterias pertenecientes al grupo que consiste en *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*,
25 *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor* y *Campylobacter*. En algunas modalidades, una proteína Cas9 puede ser una proteína de fusión, por ejemplo, los dos dominios catalíticos derivan de especies diferentes de bacterias.

[0082]En algunas modalidades, una proteína Cas puede ser una variante de la proteína Cas. Por ejemplo, las variantes útiles de la nucleasa Cas9 pueden incluir un dominio catalítico inactivo
30 simple, tal como una enzima RuvC⁻ or HNH⁻ o una nickasa. Una nickasa Cas9 tiene solamente un

dominio funcional activo y puede cortar solamente una cadena de una secuencia de ácido nucleico afín, lo que crea de este modo un corte o una rotura de cadena simple. En algunas modalidades, una nucleasa Cas9 puede ser una nucleasa Cas9 mutante que tiene una o más mutaciones de aminoácidos. Por ejemplo, una Cas9 mutante que tiene al menos una mutación D10A es una nickasa Cas9. En otras modalidades, una nucleasa Cas9 mutante que tiene al menos una mutación H840A es una nickasa Cas9. Otros ejemplos de mutaciones presentes en una nickasa Cas9 incluyen, de modo no taxativo, N854A y N863A. Se puede introducir una rotura de cadena doble utilizando una nickasa Cas9 si se utilizan al menos dos ARN dirigidos a ADN que se dirigen a cadenas de ADN opuestas. Una rotura de cadena doble inducida por doble corte puede repararse mediante NHEJ o HDR (Ran *et al.*, 2013, *Cell*, 154:1380-1389). Ejemplos no limitantes de nucleasas o nickasas Cas9 se describen, por ejemplo, en las patentes estadounidenses n.º 8,895,308; 8,889,418; y 8,865,406 y las publicaciones de solicitudes estadounidenses n.º 2014/0356959, 2014/0273226 y 2014/0186919. La nickasa o nucleasa Cas9 se puede someter a una optimización de codones para la célula diana o el organismo diana.

[0083] En algunas modalidades, una variante de la proteína Cas carece de actividad de escisión (por ejemplo, escisión completa o nickasa). Una variante de la proteína Cas puede contener una o más mutaciones puntuales que eliminan la actividad nickasa de la proteína. En algunas modalidades, las variantes de la proteína Cas pueden fusionarse con otras proteínas y servir como dominios de dirección para dirigir las otras proteínas al ácido nucleico diana. Por ejemplo, las variantes de la proteína Cas sin actividad de escisión pueden fusionarse con dominios de activación o represión transcripcional para controlar la expresión génica (Ma *et al.*, *Protein and Cell*, 2(11):879-888, 2011; Maeder *et al.*, *Nature Methods*, 10:977-979, 2013; y Konermann *et al.*, *Nature*, 517:583-588, 2014). Se puede usar una variante de la proteína Cas que carece de actividad de escisión para la dirección hacia regiones genómicas, lo que da como resultado un control transcripcional dirigido por ARN. En algunas modalidades, se puede usar una variante de la proteína Cas sin ninguna actividad de escisión para dirigir una proteína exógena al ácido nucleico diana. Se puede fusionar una proteína exógena a la variante de la proteína Cas. Una proteína exógena puede ser un dominio de proteína efectora. Una proteína exógena puede ser un activador o represor de la transcripción. Otros ejemplos de proteínas exógenas incluyen, de modo no taxativo, VP64-p65-Rta (VPR), VP64, P65, Krab, metilcitosina dioxigenasa de translocación diez-once (TET) y ADN metiltransferasa

(DNMT). A continuación, también se describen variantes específicas de la proteína Cas que carecen de actividad de escisión (por ejemplo, nickasa).

[0084]En algunas modalidades, una nucleasa Cas puede ser una variante del polipéptido Cas9 de alta fidelidad o especificidad mejorada con efectos fuera del objetivo reducidos y una escisión robusta dentro del objetivo. Los ejemplos no taxativos de variantes del polipéptido Cas9 con especificidad mejorada dentro del objetivo incluyen las variantes SpCas9 (K855A), SpCas9 (K810A/K1003A/R1060A) (también conocida como eSpCas9(1.0)) y SpCas9 (K848A/K1003A/R1060A) (también conocida como eSpCas9(1.1)) descritas en Slaymaker *et al.*, *Science*, 351(6268):84-8 (2016), y las variantes SpCas9 descritas en Kleinstiver *et al.*, *Nature*, 529(7587):490-5 (2016) que contiene una, dos, tres o cuatro de las siguientes mutaciones: N497A, R661A, Q695A y Q926A (por ejemplo, SpCas9-HF1 contiene las cuatro mutaciones).

[0085]En algunas modalidades, una nucleasa dirigible puede ser también una proteína de fusión que contiene una proteína que puede unirse a una secuencia de ácido nucleico afín y una proteína que puede escindir la secuencia de ácido nucleico afín. Por ejemplo, una proteína que puede reconocer y unirse a una secuencia de ácido nucleico afín puede ser una variante de la proteína Cas sin ninguna actividad de escisión. Una variante de la proteína Cas sin ninguna actividad de escisión puede ser un polipéptido Cas9 que contiene dos mutaciones silenciadoras de los dominios de nucleasa RuvC1 y HNH (D10A y H840A), también conocida como dCas9 (Jinek *et al.*, *Science*, 2012, 337:816-821; Qi *et al.*, *Cell*, 152(5):1173-1183). En una modalidad, el polipéptido dCas9 de *Streptococcus pyogenes* comprende al menos una mutación en la posición D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, A987 o cualquier combinación de estas. Las descripciones de dichos polipéptidos dCas9 y variantes de estos se proporcionan, por ejemplo, en la publicación de patente internacional n.º WO 2013/176772. La enzima dCas9 puede contener una mutación en D10, E762, H983 o D986, así como una mutación en H840 o N863. En algunos casos, la enzima dCas9 puede contener una mutación D10A o D10N. Además, la enzima dCas9 puede contener una H840A, H840Y o H840N. En algunas modalidades, la enzima dCas9 puede contener D10A y H840A; D10A y H840Y; D10A y H840N; D10N y H840A; D10N y H840Y; o sustituciones de D10N y H840N. Las sustituciones pueden ser sustituciones conservadoras o no conservadoras para hacer que el polipéptido Cas9 sea catalíticamente inactivo mientras aún es capaz de unirse a una secuencia de ácido nucleico afín.

[0086]En otras modalidades, una proteína que puede reconocer y unirse a una secuencia de ácido nucleico afín puede ser una proteína efectora de unión al ADN de tipo activador de la transcripción (TAL) o una proteína de unión a ADN con dedos de cinc. La proteína efectora de unión al ADN TAL tiene un dominio central de repeticiones en tándem de unión al ADN que generalmente contienen 33-35 aminoácidos de longitud y dos residuos de aminoácidos hipervariables en las posiciones 12 y 13 que pueden reconocer uno o más pares de bases de ADN específicos. La proteína de unión al ADN con dedos de cinc tiene un motivo de unión al ADN que a menudo se caracteriza por la ausencia o presencia de uno o más iones de cinc para coordinar y estabilizar el pliegue del motivo. La proteína de unión al ADN con dedos de cinc contiene múltiples protuberancias en forma de dedos que establecen contactos en tándem con su molécula diana. Algunas proteínas de unión al ADN con dedos de cinc también forman puentes salinos para estabilizar los pliegues en forma de dedos. Se identificaron por primera vez como un motivo de unión al ADN en el factor de transcripción TFIIA de *Xenopus laevis* (rana de uñas africana); sin embargo, ahora se reconoce que se unen a sustratos de ADN, ARN, proteínas y/o lípidos.

[0087]En algunas modalidades, una nucleasa dirigible en las composiciones y métodos descritos en el presente documento puede ser una proteína de fusión que contiene una proteína efectora de unión al ADN TAL y una proteína que puede escindir una secuencia de ácido nucleico afín (también denominada «nucleasas efectoras de tipo activador de la transcripción (TALEN)»). En otras modalidades, una nucleasa dirigible en las composiciones y métodos descritos en el presente documento puede ser una proteína de fusión que contiene una proteína de unión al ADN con dedos de cinc y una proteína que puede escindir una secuencia de ácido nucleico afín. Por ejemplo, una proteína que puede escindir una secuencia de ácido nucleico afín puede ser una endonucleasa FokI de tipo salvaje o mutada o el dominio catalítico de FokI. Se encuentran descripciones detalladas de TALEN y sus usos para la edición de genes, por ejemplo, en las patentes estadounidenses n.º 8,440,431; 8,440,432; 8,450,471; 8,586,363; y 8,697,853; Scharenberg *et al.*, *Curr Gene Ther*, 2013, 13(4):291-303; Gaj *et al.*, *Nat Methods*, 2012, 9(8):805-7; Beurdeley *et al.*, *Nat Commun*, 2013, 4:1762; y Joung y Sander, *Nat Rev Mol Cell Biol*, 2013, 14(1):49-55. En la técnica se describen ejemplos de una proteína de unión al ADN con dedos de cinc fusionada a una proteína que puede escindir el ácido nucleico diana e incluyen, de modo no taxativo, los descritos en Urnov *et al.*, *Nature Reviews Genetics*, 2010, 11:636-646; Gaj *et al.*, *Nat Methods*, 2012, 9(8):805-7; las

patentes estadounidenses n.º 6,534,261; 6,607,882; 6,746,838; 6,794,136; 6,824,978; 6,866,997; 6,933,113; 6,979,539; 7,013,219; 7,030,215; 7,220,719; 7,241,573; 7,241,574; 7,585,849; 7,595,376; 6,903,185; 6,479,626; y las publicaciones de solicitudes estadounidenses n.º. 2003/0232410 y 2009/0203140.

- 5 **[0088]** En algunas modalidades, una nucleasa dirigible no tiene actividad nucleasa. Por ejemplo, la nucleasa dirigible (por ejemplo, una nucleasa dirigible sin ninguna actividad nucleasa) puede regular la expresión de una secuencia de ácido nucleico afín. En algunas modalidades, la nucleasa dirigible puede ser una proteína de fusión que incluye una proteína que puede unirse a una secuencia de ácido nucleico afín, tal como una variante de la proteína Cas sin ninguna actividad de
- 10 escisión (por ejemplo, una dCas9), una proteína efectora de unión al ADN TAL y una proteína de unión al ADN con dedos de cinc como se describe anteriormente, y una proteína que puede modificar una secuencia de ácido nucleico afín, tal como un activador o represor de la transcripción.

- [0089]** También se puede fusionar una nucleasa dirigible con un péptido o proteína de localización.
- 15 Por ejemplo, una nucleasa dirigible puede fusionarse con una o más secuencias de señal de localización nuclear (NLS), que pueden dirigir una nucleasa dirigible, y los complejos RNP que forma, hacia el núcleo para modificar una secuencia de ácido nucleico afín. En la técnica se conocen ejemplos de secuencias de NLS, por ejemplo, como se describe en Lange et al., *J Biol Chem.* 282(8):5101-5, 2007, y también incluyen, de modo no taxativo,
- 20 AVKRPAATKKAGQAKKKKLD, MSRRRKANPTKLS ENAKKLAKEVEN, PAAKRVKLD, KLKIKRPVK y PKKKRKV. Ejemplos de otros péptidos o proteínas que pueden usarse para una nucleasa dirigible, tales como péptidos que penetran en las células y péptidos que se dirigen a las células, están disponibles en la técnica y se describen, por ejemplo, en Vivès et al., *Biochim Biophys Acta.* 1786(2):126-38, 2008.

- 25 **[0090]** Una nucleasa dirigible que modifica un ácido nucleico diana (por ejemplo, una secuencia diana genómica, tal como en un genoma de linfocito T) puede ser la misma que una nucleasa dirigible capaz de escindir una secuencia diana de CDL. En un ejemplo ilustrativo no taxativo, una nucleasa dirigible que modifica un ácido nucleico diana y una nucleasa dirigible capaz de escindir una secuencia diana de CDL son ambas una proteína Cas9. Una nucleasa dirigible que modifica un

ácido nucleico diana puede ser distinta de una nucleasa dirigible capaz de escindir una secuencia diana de CDL. Por ejemplo, una nucleasa dirigible que modifica un ácido nucleico diana puede ser una proteína Cas9 y una nucleasa dirigible capaz de escindir una secuencia diana de CDL puede ser una proteína TALEN, ZFN, Cas9. En otro ejemplo, una nucleasa dirigible que modifica un
5 ácido nucleico diana puede ser una proteína TALEN, ZFN o no Cas9 y una nucleasa dirigible capaz de escindir una secuencia diana de CDL puede ser una proteína Cas9.

[0091]Una nucleasa dirigible capaz de escindir una primera secuencia diana de CDL puede ser la misma que una nucleasa dirigible capaz de escindir una segunda secuencia diana de CDL. En un ejemplo ilustrativo no taxativo, una nucleasa dirigible capaz de escindir una primera secuencia
10 diana de CDL y una nucleasa dirigible capaz de escindir una segunda secuencia diana de CDL son ambas una proteína Cas9. Una nucleasa dirigible capaz de escindir una primera secuencia diana de CDL puede ser distinta de una nucleasa dirigible capaz de escindir una segunda secuencia diana de segund CDL. Por ejemplo, una nucleasa dirigible capaz de escindir una primera secuencia diana de CDL puede ser una proteína Cas9 y una nucleasa dirigible capaz de escindir una segunda
15 secuencia diana de CDL puede ser una proteína TALEN, ZFN o no Cas9. En otro ejemplo, una nucleasa dirigible capaz de escindir una primera secuencia diana de CDL puede ser una proteína TALEN, ZFN o no Cas9 y una nucleasa dirigible capaz de escindir una segunda secuencia diana de CDL puede ser una proteína Cas9.

VIII. Secuencia diana de CDL

[0092]Una secuencia diana de linealización de coadministración (CDL) es una secuencia de nucleótidos que es reconocida y unida por una nucleasa dirigible capaz de escindir la secuencia diana de CDL. En las composiciones y métodos descritos en el presente documento, las secuencias diana de CDL flanquean una plantilla de HDR, de modo que la plantilla de HDR se pueda linealizar o extirpar de una plantilla donante de plásmido. Sin desear ceñirse a ninguna teoría, una plantilla
20 de HDR linealizada/extirpada puede promover la unión de extremos mediada por homología (HMEJ), mientras que la administración como un plásmido puede reducir la citotoxicidad. Por lo tanto, una secuencia diana de CDL puede ayudar a mejorar el rendimiento celular de activación, por ejemplo, mediante una eficiencia de reparación dirigida por homología mejorada y/o una citotoxicidad reducida.

- [0093] En algunas modalidades, una secuencia diana de CDL puede reconocerse y unirse directamente mediante una proteína de unión a ADN, por ejemplo, una proteína efectora de unión a ADN TAL o proteínas de unión a ADN con dedos de cinc. En otras modalidades, una secuencia diana de CDL puede reconocerse y unirse directamente mediante una proteína de unión a ADN, por ejemplo, una nucleasa guiada por ARN, a través de un ARNg donante. En algunas modalidades, al menos el 60 % (por ejemplo, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 97 %) de los nucleótidos en una secuencia diana de CDL pueden interactuar en emparejamiento de bases de Watson-Crick con sus nucleótidos correspondientes en el ARNg donante. En algunas modalidades, una secuencia diana de CDL puede tener al menos un (por ejemplo, uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve o diez) nucleótido que no coincide con su nucleótido correspondiente en un ARNg donante cuando la secuencia diana de CDL y un ARNg donante se hibridan. Los ejemplos de bases no coincidentes incluyen un emparejamiento de guanina y uracilo, guanina y timina, y adenina y citosina. En algunas modalidades, una secuencia diana de CDL es una parte de un ácido nucleico diana (por ejemplo, una diana genómica de una célula).
- [0094] En general, las secuencias diana de CDL están presentes en ambos extremos de una plantilla de HDR en una plantilla donante de plásmido como se describe anteriormente. Una secuencia diana de CDL y, cuando es un sistema Cas, un PAM correspondiente en una plantilla donante de plásmido pueden tener diferentes configuraciones como se describió anteriormente. En algunas modalidades, una secuencia diana de CDL puede ser complementaria a una parte de igual longitud de la secuencia de un ARNg donante. En algunas modalidades, una secuencia diana de CDL tiene al menos 17 nucleótidos, por ejemplo, entre 17 y 20 nucleótidos (por ejemplo, entre 17 y 19, entre 17 y 18, entre 18 y 20, entre 18 y 19, o 17, 18, 19 o 20 nucleótidos). En algunas modalidades, una secuencia diana de CDL es parcialmente complementaria (es decir, comprende faltas de coincidencia de nucleótidos, en comparación con una parte de igual longitud de la secuencia de un ARNg donante. Por ejemplo, una secuencia diana de proteína de unión al ADN que tiene 20 nucleótidos puede tener entre 1 y 6 faltas de coincidencia (por ejemplo, entre 1 y 5, entre 1 y 4, entre 1 y 3, entre 1 y 2 faltas de coincidencia de nucleótidos; 1, 2, 3, 4, 5 o 6 faltas de coincidencia de nucleótidos) en comparación con una parte de 20 nucleótidos de la secuencia de un ARNg donante.

IX. Gen dirigido a ácidos nucleicos en las células

[0095] Las composiciones descritas en el presente documento se pueden usar en métodos para modificar un ácido nucleico diana en una célula, por ejemplo, células eucariotas, células procarióticas, células animales, células vegetales, células fúngicas y similares. Opcionalmente, la célula es una célula de mamífero, por ejemplo, una célula humana. La célula puede encontrarse *in vitro*, *ex vivo* o *in vivo*. La célula también puede ser una célula primaria, una célula germinal, una célula madre o una célula precursora. La célula precursora puede ser, por ejemplo, una célula madre pluripotente o una célula madre hematopoyética. En algunas modalidades, la célula es una célula hematopoyética primaria, una célula madre hematopoyética primaria o un linfocito T primario. En algunas modalidades, la célula hematopoyética primaria es una célula inmunitaria. En algunas modalidades, la célula inmunitaria es un linfocito T. En algunas modalidades, el linfocito T es un linfocito T regulador, un linfocito T efector o un linfocito T indiferenciado. En algunas modalidades, el linfocito T es un linfocito T CD4⁺. En algunas modalidades, el linfocito T es un linfocito T CD8⁺. En algunas modalidades, el linfocito T es un linfocito T CD4⁺CD8⁺. En algunas modalidades, el linfocito T es un linfocito T CD4⁺CD8⁻. En algunas modalidades, el linfocito T es un linfocito T αβ, en algunas modalidades el linfocito T es un linfocito T γδ. También se proporcionan poblaciones de cualquiera de las células modificadas por cualquiera de los métodos descritos en el presente documento. En algunas modalidades, los métodos comprenden además expandir la población de células modificadas.

[0096] En un aspecto particular, se proporciona una población de células (por ejemplo, una población de linfocitos T). La población de células puede comprender cualquiera de las células modificadas descritas en el presente documento. La célula modificada puede estar dentro de una población heterogénea de células y/o una población heterogénea de diferentes tipos de células. La población de células puede ser heterogénea con respecto al porcentaje de células que se editan genómicamente. Una población de células puede tener más del 10 %, más del 20 %, más del 30 %, más del 40 %, más del 50 %, más del 60 %), más del 70 %, más del 80 % o más del 90 % de la población que comprende una secuencia de nucleótidos integrada. En un determinado aspecto, una población de células comprende una secuencia de nucleótidos integrada, en donde la secuencia de nucleótidos integrada comprende al menos una parte de un gen, la secuencia de nucleótidos integrada está integrada en un locus diana genómico endógeno y la secuencia de nucleótidos integrada está orientada de tal manera que la al menos una parte del gen sea capaz de expresarse,

en donde la población de células está sustancialmente libre de componentes de administración mediados por virus, y en donde más del 10 %), más del 20 %, más del 30 %, más del 40 %, más del 50 %, más del 60 %), más del 70 %, más del 80 % o más del 90 % de las células de la población comprenden la secuencia de nucleótidos integrada.

5 **[0097]**Una población de células puede tener más del 91 %, más del 92 %, más del 93 %), más del 94 %, más del 95 %, más del 96 % o más del 97 %, más del 98 %), más del 99 %, más del 99,5 % o más del 99,9 % de la población que comprende una secuencia de nucleótidos integrada. Una población de células puede tener más del 20 % de la población que comprende una secuencia de nucleótidos integrada. Una población de células puede tener más del 30 % de la población que
10 comprende una secuencia de nucleótidos integrada. Una población de células puede tener más del 60 % de la población que comprende una secuencia de nucleótidos integrada. Una población de células puede tener más del 70 % de la población que comprende una secuencia de nucleótidos integrada.

[0098]Una célula puede incluir una célula que comprende una secuencia insertada de forma no viral tal como una secuencia exógena. Una célula puede estar libre de virus. Una célula puede estar
15 sustancialmente libre de virus. Una célula puede incluir al menos una secuencia de ácido nucleico (por ejemplo, que comprende al menos un gen heterólogo) insertada de forma no viral en al menos una región diana. En determinados aspectos, una célula no comprende un vector viral, por ejemplo, para introducir al menos una secuencia de ácido nucleico tal como una plantilla donante.

20 **[0099]**Una célula puede incluir una o más células primarias que incluyen una secuencia exógena insertada de forma no viral que tiene un tamaño de al menos 200 pares de bases. Las células primarias pueden ser células hematopoyéticas primarias o células madre hematopoyéticas primarias. Las células primarias pueden ser células hematopoyéticas primarias y las células hematopoyéticas primarias pueden ser células inmunitarias. Las células inmunitarias pueden ser
25 linfocitos T. Las células primarias pueden ser células humanas. En algunos aspectos, las células primarias no comprenden un vector viral. El tamaño de la secuencia exógena puede ser mayor que una longitud seleccionada del grupo que consiste en: 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 1 kb, 1,1 kb, 1,2 kb, 1,3 kb, 1,4 kb, 1,5 kb, 1,6 kb, 1,7 kb, 1,8 kb, 1,9 kb, 2,0 kb, 2,1 kb, 2,2 kb, 2,3 kb, 2,4 kb, 2,5 kb,

2,6 kb, 2,7 kb, 2,8 kb, 2,9 kb, 3 kb, 3,1 kb, 3,2 kb, 3,3 kb, 3,4 kb, 3,5 kb, 3,6 kb, 3,7 kb, 3,8 kb, 3,9 kb, 4,0 kb, 4,1 kb, 4,2 kb, 4,3 kb, 4,4 kb, 4,5 kb, 4,6 kb, 4,7 kb, 4,8 kb, 4,9 kb y 5,0 kb. El tamaño de la secuencia exógena puede ser superior a 1,5 kb. El tamaño de la secuencia exógena puede ser de alrededor de 200 pb a alrededor de 500 pb, de alrededor de 200 pb a alrededor de 750 pb, de
5 alrededor de 200 pb a alrededor de 1 kb, de alrededor de 200 pb a alrededor de 1,5 kb, de alrededor de 200 pb a alrededor de 2,0 kb, de alrededor de 200 pb a alrededor de 2,5 kb, de alrededor de 200 pb a alrededor de 3,0 kb, de alrededor de 200 pb a alrededor de 3,5 kb, de alrededor de 200 pb a alrededor de 4,0 kb, de alrededor de 200 pb a alrededor de 4,5 kb, de alrededor de 200 pb a alrededor de 5,0 kb. El tamaño de la secuencia exógena puede ser superior a alrededor de 1 kb. La secuencia
10 exógena puede incluir una secuencia reguladora, opcionalmente en donde la secuencia reguladora comprende una secuencia promotora y/o una secuencia potenciadora.

[0100]La secuencia exógena puede codificar una proteína heteróloga o un fragmento de esta. La secuencia exógena puede codificar un receptor de antígeno quimérico (CAR). La secuencia exógena puede codificar un receptor de linfocito T (TCR).

15 **[0101]**Una célula puede incluir un linfocito T humano primario que comprende una plantilla de ADN insertada de forma no viral que tiene un tamaño superior a 1 kb.

[0102]Una célula puede incluir un linfocito T humano primario que comprende: al menos una secuencia de ácido nucleico que comprende al menos un gen heterólogo insertado de forma no viral en al menos una región diana de uno o ambos de: gen constante de la subunidad alfa del receptor
20 de linfocitos T endógeno (TRAC) y gen constante de la subunidad beta del receptor de linfocitos T endógeno (TRBC), el al menos un gen heterólogo comprende al menos uno de: (1) una región variable de un gen de cadena alfa del receptor de linfocitos T heterólogo (TCR- α) y (2) una región variable de un gen de cadena beta del receptor de linfocitos T heterólogo (TCR- β). En algunos aspectos, el linfocito T no comprende un vector viral para introducir la al menos una secuencia de
25 ácido nucleico en el linfocito T. En algunos aspectos, la al menos una secuencia de ácido nucleico tiene un tamaño de al menos 1,5 kb. En algunos aspectos, la al menos una secuencia de ácido nucleico tiene un tamaño de al menos 500 pb. En algunos aspectos, la región diana está en el exón 1, 2 o 3 de TRAC. En algunos aspectos, la región diana está en el exón 1, 2 o 3 de TRBC. En algunos aspectos, el linfocito T es un linfocito T CD8⁺ o un linfocito T CD4⁺. En algunos aspectos,

el al menos un gen heterólogo comprende al menos uno de: (1) a) región variable o b) región variable y región constante del gen de cadena alfa del receptor de linfocitos T heterólogo (TCR- α) y (2) a) región variable o b) región variable y región constante del gen de cadena beta del receptor de linfocitos T heterólogo (TCR- β). En algunos aspectos, el al menos un gen heterólogo comprende

5 cada uno de: (1) la a) región variable o b) región variable y región constante del gen de cadena alfa del receptor de linfocitos T heterólogo (TCR- α) y (2) la a) región variable o b) región variable y región constante del gen de cadena beta del receptor de linfocitos T heterólogo (TCR- β). En algunos aspectos, el linfocito T comprende cada uno de (1) la a) región variable o b) región variable y región constante del gen de cadena TCR- α heterólogo y (2) la a) región variable o b) región

10 variable y región constante del gen de cadena TCR- β heterólogo. En algunos aspectos, los genes heterólogos forman un receptor de linfocitos T (TCR) específico de antígeno tras la expresión. En algunos aspectos, el gen de cadena TCR- α heterólogo y el gen de cadena TCR- β heterólogo están unidos operativamente mediante una secuencia conectora, opcionalmente la secuencia conectora es una secuencia conectora escindible o un elemento multicistónico. En algunos aspectos, el gen

15 de cadena TCR- α heterólogo y el gen de cadena TCR- β heterólogo se insertan en TRAC. En algunos aspectos, la expresión del al menos un gen heterólogo está impulsada por un promotor endógeno. En algunos aspectos, la expresión de uno o ambos de TRAC y TRBC se reduce en la célula con respecto a un linfocito T de control, en donde el linfocito T de control es un linfocito T humano primario que carece de la inserción no viral.

20 **[0103]**Una célula puede incluir un linfocito T humano primario que comprende: al menos una secuencia de ácido nucleico que comprende al menos un gen heterólogo insertado en al menos una región diana de uno o ambos de: gen constante de la subunidad alfa del receptor de linfocitos T endógeno (TRAC) y gen constante de la subunidad beta del receptor de linfocitos T endógeno (TRBC), el al menos un gen heterólogo comprende al menos uno de: (1) una región variable de un

25 gen de cadena alfa del receptor de linfocitos T heterólogo (TCR- α) y (2) una región variable de un gen de cadena beta del receptor de linfocitos T heterólogo (TCR- β), y en donde el linfocito T no comprende un vector viral para introducir la al menos una secuencia de ácido nucleico al linfocito T.

[0104]Una célula puede incluir una célula primaria que comprende: al menos una secuencia de

ácido nucleico que comprende al menos un gen heterólogo insertada de forma no viral en al menos una región diana del genoma de la célula. En algunos aspectos el al menos un gen heterólogo codifica un CAR u otro receptor quimérico. En algunos aspectos el al menos un gen heterólogo comprende al menos uno o ambos de: (1) una región variable de un gen de cadena alfa del receptor de linfocitos T heterólogo (TCR- α) y (2) una región variable de un gen de cadena beta del receptor de linfocitos T heterólogo (TCR- β). En algunos aspectos la célula primaria es un linfocito T. En algunos aspectos, la célula no comprende un vector viral para introducir la al menos una secuencia de ácido nucleico en la célula. En algunos aspectos, la al menos una secuencia de ácido nucleico tiene un tamaño de al menos 1,5 kb. En algunos aspectos, la al menos una secuencia de ácido nucleico tiene un tamaño de al menos 500 pb. En algunos aspectos, la región diana está en TRAC, por ejemplo, el exón 1, 2 o 3 de TRAC. En algunos aspectos, la región diana está en TRBC, por ejemplo, el exón 1, 2 o 3 de TRBC. En algunos aspectos, el linfocito T es un linfocito T CD8⁺ o un linfocito T CD4⁺. En algunos aspectos, la expresión del al menos un gen heterólogo está impulsada por un promotor endógeno. En algunos aspectos, la expresión de uno o ambos de TRAC y TRBC se reduce en el linfocito T con respecto a un linfocito T de control, en donde el linfocito T de control es un linfocito T humano primario que carece de la inserción no viral.

[0105]En algunos casos, se hace referencia a los CAR como CAR de primera, segunda y/o tercera generación. En algunos aspectos, un CAR de primera generación es aquel que únicamente proporciona una señal inducida por la cadena CD3 tras la unión al antígeno; en algunos aspectos, un CAR de segunda generación es uno que proporciona tal señal y señal coestimuladora, tal como una que incluye un dominio de señalización intracelular de un receptor coestimulador tal como CD28 o CD137; en algunos aspectos, un CAR de tercera generación es uno que incluye múltiples dominios coestimuladores de diferentes receptores coestimuladores. En algunas modalidades, el receptor de antígeno quimérico incluye una parte extracelular que contiene un anticuerpo o fragmento descrito en el presente documento. En algunos aspectos, el receptor de antígeno quimérico incluye una parte extracelular que contiene un anticuerpo o fragmento descrito en el presente documento y un dominio de señalización intracelular. En algunas modalidades, un anticuerpo o fragmento incluye un scFv o un anticuerpo VH de dominio único y el dominio intracelular contiene un ITAM. En algunos aspectos, el dominio de señalización intracelular incluye un dominio de señalización de una cadena ζ de un CD3 (cadena CD3 ζ). En algunas

modalidades, el receptor de antígeno quimérico incluye un dominio transmembrana que une el dominio extracelular y el dominio de señalización intracelular. En algunos aspectos, el dominio transmembrana contiene una parte transmembrana de CD28. El dominio extracelular y el transmembrana pueden estar unidos directa o indirectamente. En algunas modalidades, el dominio extracelular y el transmembrana están unidos mediante un espaciador, como cualquiera de los descritos en el presente documento. En algunas modalidades, el receptor de antígeno quimérico contiene un dominio intracelular de una molécula coestimuladora de linfocitos T, por ejemplo, entre el dominio transmembrana y el dominio de señalización intracelular. En algunos aspectos, la molécula coestimuladora de linfocitos T es CD28 o 41BB. En algunas modalidades, el CAR contiene un anticuerpo, por ejemplo, un fragmento de anticuerpo, un dominio transmembrana que es o contiene una parte transmembrana de CD28 o una variante funcional de este, y un dominio de señalización intracelular que contiene una parte de señalización de CD28 o una variante funcional de este y una parte de señalización de CD3 zeta o variante funcional de esta. En algunas modalidades, el CAR contiene un anticuerpo, por ejemplo, fragmento de anticuerpo, un dominio transmembrana que es o contiene una parte transmembrana de CD28 o una variante funcional de este, y un dominio de señalización intracelular que contiene una parte de señalización de un 4-1BB o una variante funcional de este y una parte de señalización de CD3 zeta o variante funcional de esta. En algunas de tales modalidades, el receptor incluye además un espaciador que contiene una parte de una molécula de Ig, tal como una molécula de Ig humana, tal como una bisagra de Ig, por ejemplo, una bisagra de IgG4, tal como un espaciador únicamente bisagra. En algunas modalidades, el dominio transmembrana del receptor, por ejemplo, el CAR, es un dominio transmembrana de CD28 humano o una variante de este, por ejemplo, un dominio transmembrana de 27 aminoácidos de un CD28 humano (n.º de acceso: P10747.1). En algunas modalidades, el receptor de antígeno quimérico contiene un dominio intracelular de una molécula coestimuladora de linfocitos T. En algunos aspectos, la molécula coestimuladora de linfocitos T es CD28 o 41BB. En algunas modalidades, el dominio de señalización intracelular comprende un dominio de señalización coestimulador intracelular de CD28 humano o parte o variante funcional de este, tal como un dominio de 41 aminoácidos de este y/o dicho dominio con una sustitución de LL por GG en las posiciones 186-187 de una proteína CD28 nativa. En algunas modalidades, el dominio intracelular comprende un dominio de señalización coestimulador intracelular de 41BB o parte o

variante funcional de este, tal como un dominio citoplasmático de 42 aminoácidos de un 4-1BB humano (n.º de acceso Q07011.1) o parte o variante funcional de este. En algunas modalidades, el dominio de señalización intracelular comprende un dominio de señalización estimulador de CD3 zeta humano o una variante funcional de este, tal como un dominio citoplasmático de 112 AA de la isoforma 3 de CD3ζ humana (n.º de acceso: P20963.2) o un dominio de señalización de CD3ζ como se describe en la patente estadounidense n.º 7,446,190 o la patente estadounidense n.º 8,911,993.

[0106] Los métodos para modificar un ácido nucleico diana en una célula que se describen en el presente documento comprenden introducir en la célula una composición descrita en el presente documento, en donde la plantilla de HDR se integra en el ácido nucleico diana. Como se demuestra en los Ejemplos, la preincubación de una nucleasa dirigible (por ejemplo, una nucleasa guiada por ARN) y un complejo RNP de ARNg donante con una plantilla donante (que comprende una plantilla de HDR modificada con secuencia diana de CDL) antes de introducir la composición en la célula mejora el rendimiento celular de activación. En algunas modalidades, una composición descrita en el presente documento se introduce en la célula mediante electroporación.

[0107] En algunos casos, las células se eliminan de un sujeto, se modifican usando cualquiera de los métodos descritos en el presente documento y se administran al sujeto. En otros casos, una composición descrita en el presente documento se puede administrar al sujeto *in vivo*. Véase, por ejemplo, la patente estadounidense n.º 9737604 y Zhang et al. «Lipid nanoparticle-mediated efficient delivery of CRISPR/Cas9 for tumor therapy», *NPG Asia Materials* Volumen 9, página e441 (2017).

[0108] En modalidades particulares, las composiciones descritas en el presente documento se pueden usar en métodos para modificar un ácido nucleico diana en una célula primaria. Las composiciones descritas en el presente documento se pueden usar en métodos para inducir una modificación génica estable de un ácido nucleico diana en una célula primaria. En algunas modalidades, el método incluye introducir en la célula primaria una composición que comprende una proteína Cas (por ejemplo, una proteína Cas9), uno o más ARN guía simple (ARNgs) y un polímero aniónico. El ARNg puede comprender una primera secuencia de nucleótidos que es complementaria al ácido nucleico diana y una segunda secuencia de nucleótidos que interactúa con

la proteína Cas (por ejemplo, proteína Cas9). En algunas modalidades, se pueden incubar juntos una proteína Cas (por ejemplo, una proteína Cas9) y un ARNs para formar un complejo RNP antes de introducirlo en la célula primaria. Se puede electroporar en la célula primaria una composición que comprende una proteína Cas (por ejemplo, una proteína Cas9), uno o más ARN guía simple (ARNgs) y una plantilla donante de plásmido. En algunas modalidades, la célula primaria se selecciona del grupo que consiste en una célula inmunitaria (por ejemplo, un linfocito T primario), una célula sanguínea, una célula progenitora o madre de esta, una célula mesenquimatosa y una combinación de estas. En algunos casos, la célula inmunitaria se selecciona del grupo que consiste en un linfocito T, un linfocito B, una célula dendrítica, un linfocito citolítico natural, un macrófago, un neutrófilo, un eosinófilo, un basófilo, un mastocito, un precursor de estos, y una combinación de estos. El progenitor o célula madre se puede seleccionar del grupo que consiste en una célula progenitora hematopoyética, una célula madre hematopoyética y una combinación de estas. En algunos casos, la célula sanguínea es una célula madre sanguínea. En algunos casos, la célula mesenquimatosa se selecciona del grupo que consiste en una célula madre mesenquimatosa, una célula progenitora mesenquimatosa, una célula precursora mesenquimatosa, una célula mesenquimatosa diferenciada y una combinación de estas. La célula mesenquimatosa diferenciada se puede seleccionar del grupo que consiste en una célula ósea, una célula de cartílago, una célula muscular, una célula adiposa, una célula estromal, un fibroblasto, una célula dérmica y una combinación de estas. En algunas modalidades, la célula primaria puede comprender una población de células primarias. En algunas modalidades, la población de células primarias comprende una población heterogénea de células primarias. En otros casos, la población de células primarias comprende una población homogénea de células de primarias.

[0109]En algunas modalidades, la célula primaria se aísla de un mamífero antes de introducir una composición descrita en el presente documento en la célula primaria. Por ejemplo, la célula primaria se puede recolectar de un sujeto humano. En algunos casos, la célula primaria o una descendencia de esta se devuelve al mamífero después de introducir la composición descrita en el presente documento en la célula primaria. En otras palabras, la célula primaria modificada genéticamente se somete a un autotrasplante. En otros casos, la célula primaria modificada genéticamente se somete a un alotrasplante. Por ejemplo, una célula primaria que no ha experimentado una modificación genética estable se aísla de un sujeto donante y luego la célula

primaria modificada genéticamente se trasplanta a un sujeto receptor que es diferente del sujeto donante.

[0110]Una composición descrita en el presente documento se puede introducir en una célula (por ejemplo, una célula primaria) usando métodos y técnicas disponibles en la técnica. Ejemplos no taxativos de métodos adecuados incluyen electroporación, tecnología de pistola de partículas y microinyección directa. En algunas modalidades, la etapa de introducir la composición descrita en el presente documento en la célula comprende electroporar la composición en la célula.

[0111]En algunas modalidades, la modificación genética estable del ácido nucleico diana se induce en más de alrededor del 5 % de la población de células (por ejemplo, la población de células primarias), por ejemplo, alrededor del 6 %, alrededor del 7 %, alrededor del 8 %, alrededor del 9 %, alrededor del 10 %, alrededor del 12 %, alrededor del 14 %, alrededor del 16 %, alrededor del 18 %, alrededor del 20 %, alrededor del 22 %, alrededor del 24 %, alrededor del 26 %, alrededor del 28 % o alrededor del 30 % de la población de células. En algunas modalidades, la modificación genética estable del ácido nucleico diana se induce en más de alrededor del 50 % de la población de células (por ejemplo, la población de células primarias), por ejemplo, alrededor del 51 %, alrededor del 52 %, alrededor del 53 %, alrededor del 54 %, alrededor del 55 %, alrededor del 56 %, alrededor del 57 %, alrededor del 58 %, alrededor del 59 %, alrededor del 60 %, alrededor del 61 %, alrededor del 62 %, alrededor del 63 %, alrededor del 64 %, alrededor del 65 %, alrededor del 66 %, alrededor del 67 %, alrededor del 68 %, alrededor del 69 %, alrededor del 70 %, alrededor del 71 %, alrededor del 72 %, alrededor del 73 %, alrededor del 74 %, alrededor del 75 %, alrededor del 76 %, alrededor del 77 %, alrededor del 78 %, alrededor del 79 %, alrededor del 80 %, alrededor del 81 %, alrededor del 82 %, alrededor del 83 %, alrededor del 84 %, alrededor del 85 %, alrededor del 86 %, alrededor del 87 %, alrededor del 88 %, alrededor del 89 %, alrededor del 90 %, alrededor del 91 %, alrededor del 92 %, alrededor del 93 %, alrededor del 94 %, alrededor del 95 %, alrededor del 96 %, alrededor del 97 %, alrededor del 98 %, alrededor del 99 % o alrededor del 100 % de la población de células. En otras modalidades, la modificación genética estable del ácido nucleico diana se induce en más de alrededor del 70 % de la población de células, por ejemplo, alrededor del 71 %, alrededor del 72 %, alrededor del 73 %, alrededor del 74 %, alrededor del 75 %, alrededor del 76 %, alrededor del 77 %, alrededor del 78 %, alrededor del 79 %, alrededor del 80 %, alrededor del 81 %, alrededor del 82 %, alrededor del 83 %, alrededor del 84 %, alrededor del 85 %, alrededor

del 86 %, alrededor del 87 %, alrededor del 88 %, alrededor del 89 %, alrededor del 90 %, alrededor del 91 %, alrededor del 92 %, alrededor del 93 %, alrededor del 94 %, alrededor del 95 %, alrededor del 96 %, alrededor del 97 %, alrededor del 98 %, alrededor del 99 % o alrededor del 100 % de la población de células. En aún otras modalidades, la modificación genética estable del ácido nucleico diana se induce en más de alrededor del 90 % de la población de células, por ejemplo, alrededor del 90 %, alrededor del 91 %, alrededor del 92 %, alrededor del 93 %, alrededor del 94 %, alrededor del 95 %, alrededor del 96 %, alrededor del 97 %, alrededor del 98 %, alrededor del 99 % o alrededor del 100 % de la población de células.

[0112]En otras modalidades, la modificación genética estable del ácido nucleico diana comprende la sustitución de una mutación genética en el ácido nucleico diana (por ejemplo, para corregir una mutación puntual o un polimorfismo de un solo nucleótido (SNP, por sus siglas en inglés) en el ácido nucleico diana que está asociado con una enfermedad) o la inserción de un marco de lectura abierto (ORF, por sus siglas en inglés) que comprende una copia normal del ácido nucleico diana (por ejemplo, para activar un ADNc de tipo salvaje del ácido nucleico diana que está asociado con una enfermedad).

[0113]En algunas modalidades, cualquiera de los métodos descritos en el presente documento también puede incluir purificar la célula (por ejemplo, una célula primaria) que tiene la modificación genética estable del ácido nucleico diana. En algunos casos, la composición aislada mediante la etapa de purificación incluye al menos alrededor del 80 % de células que tienen la modificación genética estable del ácido nucleico diana, por ejemplo, alrededor del 80 %, alrededor del 81 %, alrededor del 82 %, alrededor del 83 %, alrededor del 84 %, alrededor del 85 %, alrededor del 86 %, alrededor del 87 %, alrededor del 88 %, alrededor del 89 %, alrededor del 90 %, alrededor del 91 %, alrededor del 92 %, alrededor del 93 %, alrededor del 94 %, alrededor del 95 %, alrededor del 96 %, alrededor del 97 %, alrededor del 98 %, alrededor del 99 % o más células que tienen la modificación genética estable del ácido nucleico diana.

[0114]Se describen materiales, composiciones y componentes que pueden emplearse para los métodos y las composiciones descritos, que pueden utilizarse junto con estos, que pueden utilizarse en su preparación o que son productos de estos. Estos y otros materiales se describen en el presente documento, y se entiende que cuando se describen combinaciones, subconjuntos, interacciones,

- grupos, etc. de estos materiales, si bien la referencia específica de cada una de las diversas combinaciones y permutaciones individuales y colectivas de estos compuestos puede no describirse explícitamente, cada uno se contempla y describe específicamente en el presente documento. Por ejemplo, si se describe y analiza un método y se analizan una serie de modificaciones que se pueden
- 5 hacer a una o más moléculas, incluso en el método, todas y cada una de las combinaciones y permutaciones del método y las modificaciones que son posibles se contemplan específicamente a menos que se indique específicamente lo contrario. Asimismo, también se contempla y se describe específicamente cualquier subconjunto o combinación de estas. Este concepto se aplica a todos los aspectos de la presente descripción, lo cual incluye, de modo no taxativo, las etapas en los métodos
- 10 que usan las composiciones descritas. Por lo tanto, si existe una variedad de etapas adicionales que pueden realizarse, se entiende que cada una de estas etapas adicionales puede realizarse con cualquier etapa de método o combinación de etapas de método específicas de los métodos que se describen, y que cada combinación o subconjunto de combinaciones de ese tipo se encuentra específicamente contemplado y debe considerarse descrito.
- 15 **[0115]** Las publicaciones citadas en la presente y el material para el que se citan se incorporan a la presente específicamente mediante referencia en su totalidad.

EJEMPLOS

- [0116]** Los siguientes ejemplos se proporcionan a modo únicamente ilustrativo y no taxativo. Los expertos en la técnica reconocerán fácilmente una variedad de parámetros no críticos que pueden
- 20 cambiarse o modificarse para producir resultados esencialmente iguales o similares.

Ejemplo 1 – Métodos experimentales

Cultivo de linfocitos T

- [0117]** Se enriquecieron linfocitos T a partir de células mononucleares de sangre periférica (PBMC, por sus siglas en inglés) preparadas usando Lymphoprep (STEMCELL Technologies) de
- 25 Leukopaks de donante normal (STEMCELL Technologies) y usando el kit de aislamiento de linfocitos T humanos EasySep (STEMCELL Technologies). Posteriormente, los linfocitos T se activaron con CD3/CD28 Dynabeads en una relación de perlas con respecto a células de 1:1 (ThermoFisher, 40203D) en medio TexMACS (Miltenyi 130-197-196) complementado con suero

AB humano al 3 % (Gemini Bio) y 12,5 ng/ ml de IL-7 e IL-15 humanas (grado premium de Miltenyi) y se cultivaron a 37 °C, 5 % de CO₂ durante 48 horas antes de la electroporación.

Preparación de RNP y electroporación de linfocitos T

5 [0118]Se prepararon RNP CRISPR combinando 120 µM de ARNGs (Synthego) dirigido a la secuencia de ADN AAGTCTCTCAGCTGGTACA (SEQ ID NO:1), proteína sNLS-SpCas9-sNLS 62,5 µM (Aldevron) y solución amortiguadora P3 (Lonza) en una relación de 5:1:3:6. Se mezcló un rango de masas (la concentración final osciló entre 0 - 100 mg/l) de ADN plasmídico (plásmidos que tienen plantillas de HDR [una inserción de ~5,7 kb que codifica un transgén que tiene una etiqueta de epítipo de Myc y brazos de homología flanqueantes de 450 pb] con o sin secuencias

10 de CDL que flanquean los brazos de homología de HDR) con 3,5 µl de RNP. Se contaron los linfocitos T, se centrifugaron a 90 X G durante 10 minutos y se resuspendieron a 10⁶ células/14,5 µl de P3 con complemento añadido (Lonza). Se añadieron 14,5 µl de suspensión de linfocitos T a la mezcla de ADN/RNP, se transfirieron a un pocillo en una placa Nucleocuvette de 384 pocillos de Lonza y se pulsaron en un sistema HT Nucleofector de Lonza con código EH-115. Se dejó que

15 las células reposaran durante 15 minutos a temperatura ambiente antes de transferirlas a placas de 96 pocillos (Sarstedt) en medio TexMACS complementado con 12,5 ng/ml de IL-7 e IL-15 humanas (grado premium de Miltenyi).

Análisis de citometría de flujo

[0119]La expresión de transgén (gen de interés con una etiqueta de epítipo de Myc) se detectó

20 mediante tinción con anticuerpo anti-Myc (clon 9B11 de Cell Signaling Technology) y análisis en un citómetro de flujo Attune NxT. Otros anticuerpos utilizados fueron el anticuerpo TCRalfa/beta (clon IP26 de BioLegend), el anticuerpo CD4 (clon RPA-T4 de BioLegend), el anticuerpo CD8 (clon SK1 de BioLegend).

Secuencias

25 [0120]Protoespaciador CRISPR (por ejemplo, secuencia diana de CDL):
GAGCCATGCTTGGCTTACGA

[0121]Sitio CRISPR completo, protoespaciador y PAM de ejemplo:

GAGCCATGCTTGGCTTACGAGGG

[0122]Secuenciade cadena principal donante de plásmido estándar utilizada (X indica dónde estaba nuestra inserción de transgén de 5696 pb de longitud; la mayúscula estándar indica brazos de homología):

5 CGACCAACCCATCAAACCTCCCCGCCCCCAGCACTTTTATTTCTCCTCTTTAGGAAGTA
CACTTCAGTATCTTTGGCACAGTGCATGAGCACGACTAAAGTAAAACATCGCAGAAA
ACATAGCTTTAGTCTACCCTTCGTGTCCTAAAAGGAAAACCAGTAGCTTCCCAGGCC
ACCGGAAGGGCAACACATGTCCTCTGCAGTTTCTGCACACGGGAAGGTAAAGACAG
AGAGAGGACCTACTCCTCAACACAGAAACATTTCAAAATCTTTCTCGCCTGCAACC
10 CAAGCTGAAGTCATTCTCCCCAGAAATAACAAAAGTTGGAAGAGAAGCCGGAGACA
GGATAGGTGCAGGAAGCCCACACTTTGAGGGCAGCACTCAGACACCCTCTCCTGTGT
GCAGGACGTGCCGAATGTTTCAGGTGCAATGAGAATGAGCCATGCTTGGCTTA-X-
CGAGGGCAATCTGGCCCATCAAGTGGCCTTCGCCTCTGGGAGTAACAAAAATGCACT
TCAAAATAGCTTCTGTAATCAAGCTGCATGGGTGGAGTACTCCCCAGCTGACTCCAG
15 GAAGTTCTCTATCCAAAGCTATTTCATTAGGCCAGAGCTGTGCAAATAATTAGTCACC
CACTTGCTCCATAACCCTCCATGACAGCCCAGGCATTGAGTCCAGGTGGGACCATCA
AGCCATGCTCTGGTGGCTCATGCATTATCATAGAAATGGGAGGCTTTATTTATTTTAC
TAAAAAGAACAAAAACAACAGACTGCTGTCCTTTAGACAATAGGATCACGTCATCTG
AGCCCTCTGTGCCCCAGGTGACAAGCCCAGCCCCAAGTTCTCTTTCTCAGCCTCCCC
20 ACACATGTTCTGGAGGAGATGGGCCCAGCAGGCTGCTCTGAGGCCTGGCatcccaatggcg
cgccgagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgtatccgctcacaattccacacaacatacgagccggaagcataaa
gtgtaaagcctgggggtgcctaagtgtgagtaactcacattaattgctgtgcgctcactgcccgtttccagtcgggaaacctgtcgtgcc
gctgcattaatgaatcgccaacgcgcggggagaggcggttgcgtattgggcgctgttccgcttcctcgtcactgactcgtgcgctcgg
tcgttcggctgcggcgagcggtatcagctcactcaaaggcggtatccaggaatcaggggataacgcaggaaagaacatgtg
25 agcaaaaggccagcaaaaggccaggaaccgtaaaaggccggttgcgtggcggtttccataggctccgccccctgacgagcatcaca
aaatcgacgctcaagttagaggtggcgaaacccgacaggactataaagataccaggcggtttccccctggaagctccctcgtgcgctcct
gttccgacctgcccgttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttc
gggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggttaactatcgtcttagtc
caaccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagtt
30 cttgaagtgggtggcctaactacggctacactagaagaacagtatgtggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggt

agctcttgatccggcaaacaccaccgctggtagcgggtgggtttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaa
 gatcctttgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgtaagggattttggcatgagattatcaaaaaggatcttcacc
 tagatccttttaaattaaaaatgaagttttaaataatctaaagtatatatgagtaaacttgggtctgacagttaccaatgcttaatcagtgaggcacc
 tatctcagcgatctgtctatttcgtcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggccttaccatctggccccag
 5 tgcctgcaatgataccgcgagaaccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtg
 gtcctgcaactttatccgctccatccagcttattaattgttgcgggaagctagagtaagtagttcgcagtaatagtttgcgcaacgttggtgc
 cattgctacaggcatcgtgggtgcacgctcgtcgtttggtagtggttcattcagctccgggtcccaacgatcaaggcgagttacatgatcccca
 tgtgtgcaaaaaagcgggttagctccttcggctccgatcgtgtcagaagtaagttggccgcagtggtatcactcatggttatggcagcactg
 cataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtagtactcaaccaagtcattctgagaatagtgtagcggcgaccg
 10 agttgctcttgcggcgctcaatacgggataataccgcgccacatagcagaactttaaagtgtctcatcattggaaaacgttcttcggggcga
 aaactctcaaggatcttaccgctgttgagatccagttcgatgtaaccactcgtgcaccaactgatcttcagcatcttttactttcaccagcgttt
 ctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaagggaataagggcgacacggaaatgtgaatactcatactcttccttttcaa
 tattattgaagcatttatcaggggtattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaataggggttccgcgcacatttcc
 ccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgatcacgaggcccttttgtctcgcgcgtt
 15 tcgggtgatgacggtgaaaacctctgacacatgcagctcccggagactgtcacagcttgtctgtaagcggatgccgggagcagacaagccc
 gtcagggcgcgctcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgccacatagcgc
 gtgtgaaataccgcacagatgcgtaaggagaaaaataccgcacagggcgccattcgccattcaggctgcgcaactgttgggaagggcgatc
 ggtgcggggcctcttcgctattacgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccaggggtttcccagtcac
 gacgttgtaaacgacggccagtggaattgacgcgtattgggat

20 **[0123]**Secuencia de cadena principal de donante de plásmido de CDL utilizada (X indica dónde
 estaba nuestra inserción de transgén de 5696 pb de longitud; la negrita indica protoespaciador
 (secuencia diana de CDL); la minúscula cursiva indica PAM; la mayúscula estándar indica brazos
 de homología):

GAGCCATGCTTGGCTTACGA*ggg*CGACCAACCCATCAAACCTCCCCGCCCCCAGCACT
 25 TTTATTTCTCCTCTTTAGGAAGTACACTTCAGTATCTTTGGCACAGTGCATGAGCACG
 ACTAAAGTAAAACATCGCAGAAAACATAGCTTTAGTCTACCCTTCGTGTCCTAAAAG
 GAAAACCAGTAGCTTCCCAGGCCACCGGAAGGGCAACACATGTCCTCTGCAGTTTCT
 GCACACGGGAAGGTAAAGACAGAGAGAGGACCTACTCCTCAACACAGAAACATTTC
 AAAATCTTTCCTCGCCTGCAACCCAAGCTGAAGTCATTCTCCCCAGAAATAACAAAA
 30 GTTGGAAGAGAAGCCGGAGACAGGATAGGTGCAGGAAGCCCACACTTTGAGGGCAG

CACTCAGACACCCTCTCCTGTGTGCAGGACGTGCCGAATGTTTCAGGTGCAATGAGAA
TGAGCCATGCTTGGCTTA-X-

CGAGGGCAATCTGGCCCATCAAGTGGCCTTCGCCTCTGGGAGTAACAAAAATGCACT
TCAAAATAGCTTCTGTAATCAAGCTGCATGGGTGGAGTACTCCCCAGCTGACTCCAG
5 GAAGTTCTCTATCCAAAGCTATTCATTAGGCCAGAGCTGTGCAAATAATTAGTCACC
CACTTGCTCCATAACCCTCCATGACAGCCCAGGCATTGAGTCCAGGTGGGACCATCA
AGCCATGCTCTGGTGGCTCATGCATTATCATAGAAATGGGAGGCTTTATTTATTTTAC
TAAAAAGAACAAAAACAACAGACTGCTGTCCTTTAGACAATAGGATCACGTCATCTG
AGCCCTCTGTGCCCCAGGTGACAAGCCCAGCCCCAAGTTCTCTTTCTCAGCCTCCCC

10 ACACATGTTCTGGAGGAGATGGGCCCAGCAGGCTGCTCTGAGGCCTGGC~~ccc~~**TCGTA**
AGCCAAGCATGGCTCatcccaatggcgcgccgagcttggcgtaatcatggtcatagctgttctgtgtgaaattgttatccgc
tcacaattccacacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaatagtgagtaactcacattaattgcgttgcgt
cactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgccaacgcgcggggagagggcggttgcgtattggcg
ctgttccgcttctcgtcactgactcgtcgcgtcggctcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatcagggtatc
15 cacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgcgtggc
gttttccataggtccgccccctgacgagcatcacaataatcgacgctcaagttagggtggcgaaacccgacaggactataagatac
caggcggttccccctggaagctccctcgtcgcgtctcctgttccgacctgccgttaccggatacctgtccgcctttctcccttcgggaagcg
tggcgctttctcatagctcacgctgtaggtatctcagttcgggttaggtcgttcgctccaagctgggctgtgtgcagaaacccccgttcagcc
cgaccgctgcgccttatccggttaactatcgtcttgagccaacccggttaagacacgacttatcgccactggcagcagccactggtaacagga
20 ttagcagagcgaggtatgtaggcggtgtacagagttctgaagtgggtggcctaactacggctacactagaagaacagtatttgggtatctgcg
ctctgctgaagccagttaccttcggaaaaagagttggttagctcttgatccggcaaaacaaaccaccgctggtagcggtgtttttgtttgcaag
cagcagattacgcgcagaaaaaaaggatctcaagaagatccttgatctttctacggggctgacgctcagtggaacgaaaactcacgttaa
gggattttgggtcatgagattatcaaaaaggatctcacctagatcctttaaataaaaaatgaagttttaaatcaatctaaagtatatatgagtaaact
tggctgacagttaccaatgcttaacagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgcctgactccccgtcgtgtaga
25 taactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagaaccacgctcaccggctccagatttatcagcaata
aaccagccagccggaagggccgagcgcagaagtggctctgcaactttatccgcctccatccagtctattaattgttccgggaagctagagt
aagtagttcgccagttaatagtttgcgcaacggtgttgccattgctacagggcatcgtgggtgcacgctcgtcgttgggtatggcttcattcagctcc
gggtcccaacgatcaaggcgagttacatgatccccatgtgtgtcaaaaaagcggttagctccttcggctcctccgatcgtgtcagaagtaagt
tggccgagtggttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactc
30 aaccaagtcattctgagaatagtgtatgcggcgaccgagttgtcttggccggcgtaatacgggataataccgcgccacatagcagaacttt

aaaagtgtcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgaacccactcgtgca
 cccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggggaataagggc
 gacacggaaatgttgaatactcatactcttcttttcaatattattgaagcatttatcaggggtattgtctcatgagcggatacatatttgaatgtattt
 agaaaaataaacaatataggggtccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataa
 5 aaataggcgtatcacgaggcccttttctcgcgcgtttcgggtgatgacggtgaaaacctctgacacatgcagctcccggagactgtcacag
 cttgtctgtaagcgggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcgg
 catcagagcagattgtactgagagtgcacatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgccattc
 gccattcaggctgcgcaactgttgggaagggcgatcgggtcgggcctcttcgctattacgccagctggcgaaagggggatgtgctgcaag
 gcgattaagtgggtaacgccaggggtttccagtcacgacgttgtaaacgacgcccagtgattgacgcgtattgggat

10

Ejemplo 2 – Diseños de linealización de coadministración (CDL) para la edición genética mediada por CRISPR de linfocitos T

[0124] Se exploró la modificación genética de linfocitos T mediante el uso de electroporación de RNP comparando una plantilla donante de plásmido con (1) un casete de plantilla de reparación dirigida por homología (HDRT, por sus siglas en inglés) que presenta un diseño de HDRT estándar de un transgén simplemente flanqueado por brazos de homología (Figura 1 – panel izquierdo) o
 15 (2) un diseño de linealización de coadministración (CDL) donde la HDRT estaba flanqueada adicionalmente con secuencias diana de CDL (sitios diana de CRISPR con una coincidencia para el protoespaciador del ARNg CRISPR usada para dirigirse al genoma de linfocito T) y motivo adyacente a protoespaciador (PAM) (Figura 1 – panel derecho).
 20

[0125] Como se muestra en la Figura 2, los linfocitos T se electroporaron con complejos RNP con cantidades variables de plantillas donantes de plásmido estándar o CDL y luego se evaluó la integración de transgén mediante citometría de flujo (Figura 2A, cuantificada en la Tabla 1). El diseño de CDL demostró mayor KI% a dosis menores de ADN de HDRT que los plásmidos de HDRT estándar sin las modificaciones de secuencia de CDL (Figura 2B, cuantificada en la Tabla
 25 2). Además, el rendimiento de linfocitos T general aumentó a estas dosis de ADN bajas en comparación con los plásmidos de HDRT estándar (Figura 2C, cuantificada en la Tabla 3). Por

consiguiente, el diseño de plásmido donante de CDL demuestra rendimientos significativamente ($P < 0,0001$) aumentados de linfocitos T modificados genéticamente en comparación con los diseños de plásmido estándar. Adicionalmente, la relación de linfocitos T CD8 a CD4 fue menos de $\sim 0,5$ cuando se usó la plantilla HDRT de CDL para todas las concentraciones probadas, en

5 contraste con el plásmido de HDRT estándar (Figura 3, cuantificada en la Tabla 4). La variabilidad del rendimiento celular de activación en linfocitos T derivados de 4 donantes individuales también se evaluó mediante coeficiente de variación porcentual (%CV) y relación máxima a mínima (veces Max/Min), usando plásmidos estándar o plásmidos CDL en diferentes dosis. Cabe destacar que la diferencia entre rendimientos fue claramente menor con los diseños de CDL en comparación con

10 los diseños de plásmido estándar a dosis $\leq 0,5$ (Tabla 5). Los datos demostraron que el diseño de plásmido donante de CDL generó mayor rendimiento y mayor uniformidad de rendimiento entre donantes para la edición de linfocitos T de Cas9 mediada por electroporación y que es un beneficio importante en el proceso de modificación genética celular.

Tabla 1 – Cuantificación de imágenes de flujo representativas de porcentaje de activación (Etiqueta Myc %)

KI%	Dosis (ug)	0,125	0,250	0,500	1,000	2,000
estándar		2,68	4,63	8,86	17,22	32,70
CDL		21,98	29,57	33,87	35,99	42,62

Tabla 2 - Cuantificación de porcentaje de activación (Etiqueta Myc %)

KI%	Dosis (ug)	0,125	0,250	0,500	1,000	2,000
estándar	mediana	2,95	4,77	8,45	19,86	37,35
CDL	mediana	18,29	26,80	37,15	41,58	46,15

Tabla 3 - Cuantificación de rendimiento celular de activación

rend. KI /1e6células	Dosis (ug)	0,125	0,250	0,500	1,000	2,000
estándar	mediana	61829,3	92116,2	133162,1	174020,3	183254,9
CDL	mediana	453577,2	604795,1	377600,0	50655,0	14933,3

Tabla 4 - Cuantificación de relación CD8/CD4

relación CD8/CD4	Dosis (ug)	0,000	0,125	0,250	0,500	1,000	2,000
estándar	mediana	0,27	0,56	0,79	0,96	1,11	0,87
CDL	mediana	0,27	0,36	0,49	0,55	0,30	0,31

Tabla 5 - Variabilidad de rendimiento celular de activación en múltiples donantes

	Estándar		CDL	
Dosis (ug)	%CV	veces Max/Min	%CV	veces Max/Min
0,125	80,36	5,69	41,58	2,21
0,250	83,75	8,02	52,84	3,62
0,500	92,34	13,40	89,28	22,56
1,000	99,75	33,37	119,03	144,78
2,000	101,37	45,00	147,25	162,58

5

Ejemplo 3 - El desempeño en una plataforma a escala clínica e inclusión de la secuencia de CDL en el casete de HDR aumenta el porcentaje de células con integración de transgén tras la transfección a escala clínica

Materiales y métodos

- 10 [0126] Aislamiento y activación de linfocitos T humanos primarios. Las poblaciones de células frescas recogidas por aféresis de donantes sanos se obtuvieron de HemaCare Corporation. Se aislaron linfocitos T CD4 y CD8 de la población celular del donante usando el kit de aislamiento de aislamiento positivo de CD4 CD8 de Miltenyi y el AutoMACS Pro Separator de Miltenyi. Los linfocitos T CD4/8 aislados se activaron después del aislamiento mediante el uso de Dynabeads

Human T-Activator CD3/CD28 de ThermoFisher. Las células se activaron en medio TexMACS de Miltenyi complementado con suero AB humano al 3 % (Gemini Bio), e IL7 e IL15 (Miltenyi Biotech). Las células se incubaron durante 48 horas a 37 °C después de la adición del medio de activación.

- 5 **[0127]** Preparación de linfocitos T humanos primarios para electroporación: Después de la activación, los linfocitos T se eliminaron del recipiente de activación y se transfirieron a tubos cónicos de 50 mL. Se siguió el protocolo proporcionado por Miltenyi Biotech para la eliminación de perlas para obtener una población celular activada y libre de perlas.

- 10 **[0128]** Preparación de complejo de ribonucleoproteína para electroporación: Se obtuvo ARN guía simple dirigido al locus GS94 de Synthego Corporation y se resuspendió a una concentración de trabajo en TE/agua. La proteína caspasa spCas9 se obtuvo de Aldevron. Se formaron complejos de ARN guía y proteína caspasa a temperatura ambiente para formar complejo de ribonucleoproteína (RNP).

- 15 **[0129]** ADN plasmídico. Se diseñó internamente ADN plasmídico que contenía las secuencias de CDL y fue preparado por Elim Bio.

- 20 **[0130]** Transfección mediada por electroporación de linfocitos T primarios. Se contaron los linfocitos T activados preparados como se describe anteriormente y se transfirieron 5e7 células por reacción de electroporación a tubos cónicos de 50 mL. Las células se sedimentaron a 300 x g durante 5 minutos, luego se lavaron con DPBS y se sedimentaron a 300 x g durante 5 minutos. A continuación, las células se suspendieron en la solución amortiguadora P3 de Lonza. El volumen de la solución amortiguadora se calculó de modo que el volumen total de suspensión celular después de la adición de RNP y ADN fuera de 1 mL. Para evaluar el rendimiento de múltiples construcciones de ADN, se seleccionaron para la prueba tres construcciones CDL y tres construcciones no CDL que comprendían el transgén «X». Para cada construcción, se usaron 20
25 ug de ADN por transfección. El ADN se mezcló con la solución de RNP y luego se mezcló cuidadosamente con la suspensión celular. Luego se transfirieron las células, RNP y la suspensión de ADN al cartucho LV de Lonza. Las células se electroporaron en el dispositivo LV de Lonza y se dejaron recuperar durante 10 minutos a temperatura ambiente después de la electroporación.

Luego se transfirió la mezcla celular a un G-rex 100 (Wilson Wolf Manufacturing) que contenía 350 ml de medio y se incubó a 37 °C.

[0131] Análisis de resultados mediante citometría de flujo. Los resultados se recogieron mediante citometría de flujo el día 6 después de la transfección o el día 8 después de la activación.

5 Para preparar las células para el análisis, los pocillos de control se mezclaron para obtener una solución homogénea y se contaron utilizando el contador de células Nexcelom K2 con tinción AOPI para diferenciar las células vivas y muertas. Utilizando el valor obtenido, se retiraron 2e5 células de cada pocillo y se transfirieron a una placa con fondo en V de 96 pocillos. Las células se sedimentaron, luego se lavaron una vez en Stain Buffer (BD Biosciences) y se sedimentaron a 400
10 x g durante 5 minutos. Durante la centrifugación se preparó la solución de tinción. El panel de flujo incluye lo siguiente: Myc PE (BioLegend) y tinte Zombie L/D (BioLegend). Las células se resuspendieron en la solución de tinción y se incubaron a 4 °C durante 30 minutos. Después de la incubación, se añadió Stain Buffer a los pocillos y las células se sedimentaron a 400 x g durante 5 minutos. Luego, las células se lavaron una vez con Stain Buffer y se resuspendieron en Stain Buffer
15 complementado con perlas de recuento (ThermoFisher). El análisis de flujo se realizó en el citómetro Attune de ThermoFisher y los resultados se analizaron utilizando el software FlowJo.

Resultados

[0132] Rendimiento en una plataforma a escala clínica. Se evaluó el efecto de modificar la secuencia de ácido nucleico de un plásmido que contiene una plantilla de reparación dirigida por
20 homología para incluir la secuencia de CDL inmediatamente secuencia arriba del brazo de HDR 5' y secuencia abajo del brazo de HDR 3' en el contexto del sistema CRISPR-Cas9 descrito anteriormente en una plataforma y escala pertinente para la elaboración clínica de un producto de terapia celular. El propósito de este estudio fue demostrar la utilidad de la secuencia de CDL en un contexto clínico y resaltar la aplicabilidad de la tecnología para la elaboración de productos de
25 terapia celular.

[0133] La inclusión de la secuencia de CDL en el casete de HDR aumenta el porcentaje de células con integración de transgén tras la transfección a escala clínica. Se evaluaron los linfocitos

T electroporados con plásmidos que contienen la secuencia de CDL

GAGCCATGCTTGGCTTACGA y una secuencia que codifica un transgén o plásmidos de control que comprenden secuencias que codifican un transgén sin la secuencia de CDL para determinar tanto la cantidad de células como la expresión del transgén el día 6 después de la electroporación. La longitud de la secuencia de polinucleótidos que codifica el transgén fue: 6533 pares de bases de longitud (para el plásmido pS3798), 7042 pares de bases de longitud (para el plásmido pS3631) y 6236 pares de bases de longitud (para el plásmido pS3797). Los resultados mostrados en la Figura 4 indican un aumento significativo ($p < 0,05$) en la expresión de transgén en las condiciones que reciben el plásmido que contiene la secuencia de CDL en comparación con las condiciones que reciben el control, plásmidos que no contienen el sitio de CDL. La diferencia es evidente en el porcentaje de células que albergan un transgén integrado. En promedio, la expresión de transgén evaluada mediante citometría de flujo aumentó del 3,35 % para las condiciones sin CDL al 9,76 % para las condiciones de CDL con 20 ug de ADN. En general, estos datos demuestran el efecto mejorado que tiene la inclusión de la secuencia de CDL en el proceso de edición de genes.

REFERENCIAS

- 1 Yin, H., Xue, W. & Anderson, D. G. CRISPR-Cas: a tool for cancer research and therapeutics. *Nat Rev Clin Oncol*, doi:10.1038/s41571-019-0166-8 (2019).
- 2 Dunbar, C. E. *et al.* Gene therapy comes of age. *Science* **359**, doi:10.1126/science.aan4672 (2018).
- 3 Cornu, T. I., Mussolino, C. & Cathomen, T. Refining strategies to translate genome editing to the clinic. *Nat Med* **23**, 415-423, doi:10.1038/nm.4313 (2017).
- 4 David, R. M. & Doherty, A. T. Viral Vectors: The Road to Reducing Genotoxicity. *Toxicol Sci* **155**, 315-325, doi:10.1093/toxsci/kfw220 (2017).
- 5 Roth, T. L. *et al.* Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405-409, doi:10.1038/s41586-018-0326-5 (2018).
- 6 Vo, L. T. *et al.* Regulation of embryonic haematopoietic multipotency by EZH1. *Nature* **553**, 506-510, doi:10.1038/nature25435 (2018).
- 7 Pouton, C. W., Wagstaff, K. M., Roth, D. M., Moseley, G. W. & Jans, D. A. Targeted delivery to the nucleus. *Adv Drug Deliv Rev* **59**, 698-717, doi:10.1016/j.addr.2007.06.010 (2007).

- 8 Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, doi:10.1126/science.1258096 (2014).
- 10 Jiang, F. & Doudna, J. A. CRISPR-Cas9 Structures and Mechanisms. *Annu Rev Biophys* **46**, 505-529, doi:10.1146/annurev-biophys-062215-010822 (2017).
- 5 11 Luecke, S. *et al.* cGAS is activated by DNA in a length-dependent manner. *EMBO Rep* **18**, 1707-1715, doi:10.15252/embr.201744017 (2017).
- 12 Richardson, C. D., Ray, G. J., Bray, N. L. & Corn, J. E. Non-homologous DNA increases gene disruption efficiency by altering DNA repair outcomes. *Nat Commun* **7**, 12463, doi:10.1038/ncomms12463 (2016).
- 10 13 Bernkop-Schnurch, A. Strategies to overcome the polycation dilemma in drug delivery. *Adv Drug Deliv Rev* **136-137**, 62-72, doi:10.1016/j.addr.2018.07.017 (2018).
- 14 Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med* **24**, 1216-1224, doi:10.1038/s41591-018-0137-0 (2018).
- 15 15 Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389, doi:10.1016/j.cell.2013.08.021 (2013).