

COMPOSICIONES DE GENOTERAPIA Y MÉTODOS DE USO DE LAS MISMAS

Antecedentes

Existe un subconjunto de enfermedades humanas que pueden atribuirse a cambios en el ADN heredados o adquiridos en una fase temprana del desarrollo embrionario. De especial interés para los creadores de genoterapias son las enfermedades causadas por una mutación en un solo gen, conocidas como enfermedades monogénicas. Se cree que existen más de 6000 enfermedades monogénicas. Por lo general, cualquier enfermedad genética particular causada por mutaciones hereditarias es relativamente poco frecuente, pero, en conjunto, el número de enfermedades relacionadas con la genética es elevado. Entre las enfermedades genéticas más conocidas se incluyen la fibrosis quística, la distrofia muscular de Duchenne, la enfermedad de Huntington y la drepanocitosis (o anemia drepanocítica). Otras clases de enfermedades genéticas son los trastornos metabólicos, como las acidemias orgánicas, y las tesarismosis lisosómicas, en las que los genes disfuncionales provocan defectos en los procesos metabólicos y la acumulación de subproductos tóxicos que pueden dar lugar a una grave morbilidad y mortalidad tanto a corto como a largo plazo.

Sumario

Las enfermedades genéticas causadas por genes disfuncionales representan una gran parte de las enfermedades en todo el mundo. La genoterapia se perfila como una prometedora forma de tratamiento destinada a mitigar los efectos de las enfermedades genéticas.

Antes de la presente divulgación, determinadas genoterapias con AAV que utilizaban recombinación homóloga, empleaban composiciones de vector vírico en ausencia de una nucleasa para introducir un transgén de interés en un sitio particular. Entre otras cosas, la presente divulgación reconoce que la combinación de dicho vector con una o más nucleasas (p. ej., proporcionadas en una sola composición o como composiciones distintas) puede mejorar las tasas de y/o eficacia de integración de transgenes en seres humanos.

Alternativa o adicionalmente, la presente divulgación abarca además el reconocimiento de que una combinación de un vector como el descrito en el presente documento con una o más nucleasas, puede proporcionar mejoras sorprendentes e inesperadas en las tasas de integración y/o eficacia de transgenes en seres humanos, por ejemplo, en donde un sitio de corte diana es distal con respecto a un sitio de integración de transgenes. En parte, la presente divulgación abarca el reconocimiento y la observación de que los diseños de vectores optimizados y los sitios de corte preferidos pueden ser diferentes entre especies. Como se demuestra en el presente documento, los diseños óptimos en un sistema humano o humanizado pueden diferir significativamente de los de otra especie o sistema modelo (p. ej., un ratón de tipo silvestre (*wild-type*)). En algunas realizaciones, un sistema humano o humanizado puede permitir una mayor flexibilidad en la selección de un sitio de corte para mejorar la recombinación homóloga y, por tanto, la eficacia de ciertas genoterapias.

En algunas realizaciones, la presente divulgación proporciona composiciones que incluyen (i) una nucleasa o secuencia de polinucleótidos que codifica una nucleasa, y (ii) un casete polinucleotídico que comprende un casete de expresión que comprende una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén, y la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes al integrarse en un sitio de integración diana en el genoma de una célula, una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' de un sitio de integración diana en el genoma de una célula, y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' de un sitio de integración diana en el genoma de la célula, en donde: el casete polinucleotídico no incluye una secuencia promotora, la nucleasa es capaz

de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula, y el sitio de corte es distal con respecto al sitio de integración diana.

En algunas realizaciones, la presente divulgación proporciona métodos de integración de un transgén en el genoma de una célula, comprendiendo los métodos administrar a un sujeto una composición que comprende: (i) una nucleasa o secuencia de polinucleótidos que codifica una nucleasa, y (ii) un casete polinucleotídico que comprende un casete de expresión que comprende una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén, y la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' de la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes al integrarse en un sitio de integración específico en el genoma de una célula, una tercera secuencia de ácido nucleico situada en posición 5' del casete de expresión y que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' de un sitio de integración específico en el genoma de una célula, y una cuarta secuencia de ácido nucleico situada en posición 3' del casete de expresión y que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' de un sitio de integración diana en el genoma de la célula, en donde: el casete polinucleotídico no incluye una secuencia promotora, la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula, y el sitio de corte es distal con respecto al sitio de integración diana, en donde después de administrar la composición, el transgén se integra en el genoma de la población de células. En algunas realizaciones, una célula se edita *in vivo*. En algunas realizaciones, la integración de un transgén se lleva a cabo *ex vivo*.

Aunque se contempla cualquiera de una variedad de sistemas de suministro, en algunas realizaciones, una composición comprende además un vector vírico recombinante. Aunque se contempla cualquiera de una variedad de vectores víricos, en algunas realizaciones, un vector vírico recombinante es un vector de AAV

recombinante. En algunas realizaciones, un vector vírico recombinante es o comprende un polipéptido de la cápside que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 95 % con la secuencia de aminoácidos de sL65, LK03, AAV8, AAV-DJ; AAV-LK03; o AAVNP59.

5 En algunas realizaciones, una composición comprende además secuencias de ITR de AAV2.

De acuerdo con diversas realizaciones, se contempla cualquiera de una diversidad de longitudes y/o relaciones de brazos de homología. En algunas realizaciones, cada una de una tercera y cuarta secuencia de ácido nucleico, tiene una longitud de entre 50 y 1600 nt. En algunas realizaciones, una tercera y una
10 cuarta secuencia de ácido nucleico tienen la misma longitud. En algunas realizaciones, una tercera y cuarta secuencia de ácido nucleico, tienen diferentes longitudes.

En algunas realizaciones, un casete polinucleotídico no comprende una
15 secuencia promotora.

De acuerdo con diversas realizaciones, las composiciones proporcionadas pueden incluir integración en un sitio de integración diana. En algunas realizaciones, después de la integración de un casete polinucleotídico en un sitio de integración diana en el genoma de la célula, el transgén se expresa bajo el control de un
20 promotor endógeno en el sitio de integración diana. En algunas realizaciones, un sitio de integración diana es un locus de albúmina que comprende un promotor de albúmina endógeno y un gen de albúmina endógeno. En algunas realizaciones, un sitio de integración diana es un locus de colágeno que comprende un promotor de colágeno endógeno y un gen de colágeno endógeno. En algunas realizaciones, un
25 sitio de integración diana es un locus de actina que comprende un promotor de actina endógeno y un gen de actina endógeno. En algunas realizaciones, un sitio de integración diana se encuentra dentro de una secuencia codificante del locus de albúmina y adyacente en posición 5' a un codón finalizador. En algunas realizaciones, un sitio de integración diana es adyacente en posición 5' a un codón
30 finalizador en el exón 14 del locus de albúmina.

De acuerdo con diversas realizaciones, puede utilizarse cualquier sitio de corte apropiado para la aplicación. Como ejemplo no limitativo, en algunas realizaciones un sitio de corte se encuentra dentro de una secuencia no codificante del locus de albúmina. En algunas realizaciones, un sitio de corte se encuentra dentro de un intrón, una región no traducida, un potenciador, un promotor, un silenciador o un aislante del locus de albúmina. En algunas realizaciones, un sitio de corte se encuentra dentro del intrón 13 o 14 del locus de albúmina. En algunas realizaciones, un sitio de corte se encuentra a una distancia de entre 1 y 2000 pb con respecto al sitio de integración diana. En algunas realizaciones, un sitio de corte se encuentra a una distancia de hasta 100 pb del sitio de integración diana.

Como se describe en el presente documento, la presente divulgación abarca el reconocimiento de que cualquiera de una variedad de nucleasas puede ser útil en los métodos proporcionados y junto con las composiciones proporcionadas. En algunas realizaciones, una nucleasa se selecciona de una meganucleasa, una TALEN (nucleasa de tipo activador de transcripción, *transcription activator-like effector nuclease*), una nicasa TALE, una ZFN (nucleasa de dedos de zinc, *zinc-finger nuclease*), una nicasa de ZF, una enzima Cas o una variante de las mismas.

De acuerdo con diversas realizaciones, puede usarse una variedad de segundos ácidos nucleicos para promover la producción de dos productos génicos independientes después de la integración en un sitio de integración diana en el genoma de una célula. Por ejemplo, en algunas realizaciones una segunda secuencia de ácido nucleico puede ser o comprender a) una secuencia de ácido nucleico que codifica un péptido 2A, b) una secuencia de ácido nucleico que codifica un sitio interno de entrada al ribosoma (IRES, *internal ribosome entry site*), c) una secuencia de ácido nucleico que codifica una región de corte y empalme aminoterminal de inteína y una región de corte y empalme carboxiterminal de inteína o d) una secuencia de ácido nucleico que codifica un donante de corte y empalme y un aceptor de corte y empalme. En algunas realizaciones, un segundo ácido nucleico es o comprende una secuencia de ácido nucleico que codifica un péptido 2A seleccionado del grupo que consiste en P2A, T2A, E2A y F2A.

La presente divulgación abarca el reconocimiento de que cualquiera de una variedad de transgenes puede usarse con las composiciones proporcionadas y/o en los métodos proporcionados. Por ejemplo, en algunas realizaciones, un transgén puede seleccionarse de CBS, UGT1A1, MUT, FAH, ATP7B, A1AT, ASL, LIPA, Factor IX, o de una variante de los mismos.

Cualquiera de una variedad de tipos celulares puede modificarse mediante la aplicación de una o más composiciones y/o métodos proporcionados. Por ejemplo, en algunas realizaciones, una célula es una célula sanguínea, hepática, muscular o del SNC.

De acuerdo con diversas realizaciones, los métodos y composiciones proporcionados pueden ser útiles en el contexto del trasplante de órganos y/o tejidos. En algunas realizaciones, después de la integración del transgén, se administra una célula en un trasplante autólogo. En algunas realizaciones, después de la integración del transgén, se administra una célula en un trasplante alogénico.

En un aspecto, la invención presenta una composición o un conjunto de composiciones que incluyen: una nucleasa o secuencia de polinucleótidos que codifica una nucleasa, en donde la nucleasa se selecciona de una enzima (Cas) asociada a repeticiones palindrómicas cortas regularmente interespaciadas agrupadas (CRISPR, *clustered regularly interspaced short palindromic repeats*), una nucleasa efectora similar a un activador de la transcripción (TALEN, *transcription activator-like effector (TALE) nuclease*), una nucleasa TALE, una nucleasa de dedos de zinc (ZFN, *zinc finger (ZF) nuclease*), una nucleasa de ZF o una meganucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico está situada en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes al integrarse en un sitio de integración diana en el genoma de una célula humana; una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que incluye una secuencia sustancialmente homóloga a

una secuencia genómica humana situada en posición 5' del sitio de integración diana en el genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana situada en posición 3' del sitio de integración diana en el genoma de la célula humana; en donde el casete polinucleotídico no incluye una secuencia promotora; y la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula humana.

En algunas realizaciones del aspecto anterior, el sitio de integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno.

En otro aspecto, la invención presenta una composición o un conjunto de composiciones que incluyen: una nucleasa o secuencia de polinucleótidos que codifica una nucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes al integrarse en un sitio de integración diana en el genoma de una célula humana, en donde el sitio de integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno; una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' del sitio de integración diana en un genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' del sitio de integración diana en el genoma de la célula humana; en donde el casete polinucleotídico no incluye una secuencia promotora; y la nucleasa es capaz de

inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula humana.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la composición o el conjunto de composiciones incluye además un vector vírico recombinante. En algunas realizaciones, el vector vírico recombinante es un vector de AAV recombinante. En algunas realizaciones, el vector vírico recombinante es o incluye un polipéptido de la cápside que incluye una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 95 % con la secuencia de aminoácidos de sL65, LK03, AAV8, AAV-DJ; AAV-LK03; o AAVNP59. En algunas realizaciones, el vector vírico recombinante incluye el polipéptido de la cápside y la secuencia de polinucleótidos que codifica una nucleasa y/o el casete polinucleotídico se encapsida en el vector vírico recombinante. En algunas realizaciones, el casete polinucleotídico se encapsida en el vector vírico recombinante.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la composición o el conjunto de composiciones incluye además secuencias de repeticiones terminales invertidas (ITR, *inverted terminal repeat*) de AAV2. En algunas realizaciones, las secuencias de ITR de AAV2 flanquean los extremos 5' y 3' de la secuencia de polinucleótidos que codifica la nucleasa y/o el casete polinucleotídico.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, cada una de la tercera y cuarta secuencia de ácido nucleico, tiene una longitud de entre 50 y 1600 nt.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la tercera y cuarta secuencia de ácido nucleico tienen la misma longitud.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la tercera y cuarta secuencia de ácido nucleico, tienen diferentes longitudes.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, después de la integración del casete polinucleotídico en el sitio de integración diana

en el genoma de la célula, el transgén se expresa bajo control de un promotor endógeno en el sitio de integración diana.

5 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de integración diana se encuentra dentro de una secuencia codificante del locus de albúmina y adyacente en posición 5' a un codón finalizador. En algunas realizaciones, el sitio de integración diana es adyacente en posición 5' a un codón finalizador en el exón 14 del locus de albúmina.

10 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de corte se encuentra dentro de una secuencia no codificante del locus de albúmina. En algunas realizaciones, el sitio de corte se encuentra dentro de un intrón, una región no traducida, un potenciador, un promotor, un silenciador o un aislante del locus de albúmina. En algunas realizaciones, el sitio de corte se encuentra dentro del intrón 12, 13 o 14 del locus de albúmina.

15 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa se selecciona de una enzima Cas, una TALEN, una nicasa TALE, una ZFN, una nicasa de ZF o una meganucleasa.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa se selecciona de una enzima Cas o una TALEN.

20 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa es una enzima Cas.

25 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la enzima Cas se selecciona de Cas9 de *Staphylococcus aureus* (saCas9), *Streptococcus pyogenes* (spCas9), Nucleasa de AZ, HF1-Cas9, HF2-Cas9 o HiFi-Cas9. En algunas realizaciones, la composición o el conjunto de composiciones incluye además un ARN guía (ARNg). En algunas realizaciones, el ARNg incluye una secuencia de ácido nucleico de una cualquiera de las SEQ ID NO: 27-45, 71-86 o 93-98, o una secuencia de ácido nucleico que tiene una identidad de al menos 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % con una cualquiera de las SEQ ID NO: 27-45, 71-86 o 93-98.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el ARNg se formulan conjuntamente.

5 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o secuencia de polinucleótidos que codifica la nucleasa y el ARNg se formulan por separado.

10 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la segunda secuencia de ácido nucleico es o incluye: una secuencia de ácido nucleico que codifica un péptido 2A; una secuencia de ácido nucleico que codifica un sitio interno de entrada al ribosoma (IRES); una secuencia de ácido nucleico que codifica una región de corte y empalme aminoterminal de inteína y una región de corte y empalme carboxiterminal de inteína; o una secuencia de ácido nucleico que codifica un donante de corte y empalme y un aceptor de corte y empalme.

15 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la segunda secuencia de ácido nucleico es o incluye una secuencia de ácido nucleico que codifica un péptido 2A.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, el segundo ácido nucleico es o incluye una secuencia de ácido nucleico que codifica un péptido 2A seleccionado del grupo que consiste en P2A, T2A, E2A y F2A.

20 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de corte se encuentra a una distancia de entre 1 y 2000 pb con respecto al sitio de integración diana. En algunas realizaciones, el sitio de corte se encuentra a una distancia de hasta 100 pb del sitio de integración diana.

25 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el transgén se selecciona de CBS, UGT1A1, MUT, FAH, ATP7B, A1AT, ASL, LIPA, PAH, G6PC, Factor IX, o de una variante de los mismos.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la composición o el conjunto de composiciones es un conjunto de composiciones, en donde la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se

formula en una nanopartícula lipídica (LNP) y el casete polinucleotídico se encapsida en un vector de AAV recombinante.

En algunas realizaciones de cualquiera de los aspectos anteriores, el sitio de corte es distal con respecto al sitio de integración diana.

5 En otras realizaciones de cualquiera de los aspectos anteriores, el sitio de corte se solapa con el sitio de integración.

En otro aspecto, la invención presenta un método de integración de un transgén en el genoma de una célula humana, incluyendo dicho método administrar a un sujeto una composición o un conjunto de composiciones que incluyen: una

10 nucleasa o una secuencia de polinucleótidos que codifica una nucleasa, en donde la nucleasa se selecciona de una enzima Cas, una TALEN, una nicasa TALE, una ZFN, una nicasa ZF o una meganucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera secuencia de ácido nucleico y una

15 segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico está situada en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la

20 producción de dos productos génicos independientes al integrarse en un sitio de integración diana en el genoma de una célula humana; una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que incluye una

25 secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' del sitio de integración diana en el genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de

expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' del sitio de integración diana en el

genoma de la célula humana; en donde: el casete polinucleotídico no incluye una secuencia promotora; y la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula; en donde, después de administrar la composición o el conjunto de composiciones, el transgén se integra en el genoma de la célula humana. En algunas realizaciones, un sitio de

integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno.

En otro aspecto, la invención presenta un método de integración de un transgén en el genoma de una célula humana, incluyendo dicho método poner en
5 contacto la célula humana con una composición o un conjunto de composiciones que incluyen: una nucleasa o una secuencia de polinucleótidos que codifica una nucleasa, en donde la nucleasa se selecciona de una enzima Cas, una TALEN, una nicasa TALE, una ZFN, una nicada ZF o una meganucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera
10 secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico está situada en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes al integrarse en un sitio de integración diana en el genoma de una célula humana; una
15 tercera secuencia de ácido nucleico situada en posición 5' del casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' del sitio de integración diana en el genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que incluye una secuencia que es sustancialmente
20 homóloga a una secuencia genómica humana en posición 3' del sitio de integración diana en el genoma de la célula humana; en donde: el casete polinucleotídico no incluye una secuencia promotora; y la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula; en donde, después de poner en contacto la célula humana con la
25 composición o el conjunto de composiciones, el transgén se integra en el genoma de la célula humana. En algunas realizaciones, un sitio de integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno.

En otro aspecto, la invención presenta un método de integración de un transgén en el genoma de una célula humana, incluyendo dicho método administrar
30

a un sujeto una composición o conjunto de composiciones que incluyen: una nucleasa o una secuencia de polinucleótidos que codifica una nucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde

5 la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes después de la integración en un sitio de integración diana en el genoma de una

10 célula humana, en donde el sitio de integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno; una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' de un sitio de integración diana en el genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3'

15 en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' de un sitio de integración diana en el genoma de la célula humana; en donde: el casete polinucleotídico no incluye una secuencia promotora; la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en

20 el genoma de la célula; y el sitio de corte es distal con respecto al sitio de integración diana; en donde, después de administrar la composición o el conjunto de composiciones, el transgén se integra en el genoma de la célula humana.

En otro aspecto, la invención presenta un método de integración de un transgén en el genoma de una célula humana, incluyendo dicho método poner en

25 contacto la célula humana con una composición o conjunto de composiciones que incluyen: una nucleasa o una secuencia de polinucleótidos que codifica una nucleasa; y un casete polinucleotídico que incluye: un casete de expresión que incluye una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y

30 la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' en la primera

secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes después de la integración en un sitio de integración diana en el genoma de una célula humana, en donde el sitio de integración diana es un locus de albúmina que incluye un promotor de albúmina endógeno y un gen de albúmina endógeno; una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 5' de un sitio de integración diana en el genoma de una célula humana; y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que incluye una secuencia que es sustancialmente homóloga a una secuencia genómica humana en posición 3' de un sitio de integración diana en el genoma de la célula humana; en donde: el casete polinucleotídico no incluye una secuencia promotora; la nucleasa es capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte en el genoma de la célula; y el sitio de corte es distal con respecto al sitio de integración diana; en donde, después de poner en contacto la célula humana con la composición o el conjunto de composiciones, el transgén se integra en el genoma de la célula humana.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la composición o el conjunto de composiciones incluye además un vector vírico recombinante. En algunas realizaciones, el vector vírico recombinante es un vector de AAV recombinante. En algunas realizaciones, el vector vírico recombinante es o incluye un polipéptido de la cápside que incluye una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 95 % con la secuencia de aminoácidos de LK03, AAV8, AAV-DJ; AAV-LK03; o AAVNP59. En algunas realizaciones, el vector vírico recombinante incluye el polipéptido de la cápside y la secuencia de polinucleótidos que codifica una nucleasa y/o el casete polinucleotídico se encapsida en el vector vírico recombinante. En algunas realizaciones, el casete polinucleotídico se encapsida en el vector vírico recombinante.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la composición o el conjunto de composiciones incluye además secuencias de ITR de AAV2. En algunas realizaciones, las secuencias de ITR de AAV2 flanquean los extremos 5' y 3' de la secuencia de polinucleótidos que codifica la nucleasa y/o el casete polinucleotídico.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, cada una de la tercera y cuarta secuencia de ácido nucleico, tiene una longitud de entre 50 y 1600 nt.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la tercera y cuarta secuencia de ácido nucleico tienen la misma longitud.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la tercera y cuarta secuencia de ácido nucleico, tienen diferentes longitudes.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, después de la integración del casete polinucleotídico en el sitio de integración diana en el genoma de la célula, el transgén se expresa bajo control de un promotor endógeno en el sitio de integración diana.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de integración diana se encuentra dentro de una secuencia codificante del locus de albúmina y adyacente en posición 5' a un codón finalizador. En algunas realizaciones, el sitio de integración diana es adyacente en posición 5' a un codón finalizador en el exón 14 del locus de albúmina.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de corte se encuentra dentro de una secuencia no codificante del locus de albúmina. En algunas realizaciones, el sitio de corte se encuentra dentro de un intrón, una región no traducida, un potenciador, un promotor, un silenciador o un aislante del locus de albúmina. En algunas realizaciones, el sitio de corte se encuentra dentro del intrón 12, 13 o 14 del locus de albúmina.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa se selecciona de una enzima Cas, una TALEN, una nicasa TALE, una ZFN, una nicasa de ZF o una meganucleasa.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa se selecciona de una enzima Cas o una TALEN.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa es una enzima Cas.

5 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la enzima Cas se selecciona de Cas9 de *Staphylococcus aureus* (saCas9), *Streptococcus pyogenes* (spCas9), Nucleasa de AZ, HF1-Cas9, HF2-Cas9 o HiFi-Cas9.

10 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el método incluye además un ARN guía (ARNg). En algunas realizaciones, el ARNg incluye una secuencia de ácido nucleico de una cualquiera de las SEQ ID NO: 27-45, 71-86 o 93-98, o una secuencia de ácido nucleico que tiene una identidad de al menos 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % con una cualquiera de las SEQ ID NO: 27-45, 71-86 o 93-98.

15 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el ARNg se formulan conjuntamente.

20 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o secuencia de polinucleótidos que codifica la nucleasa y el ARNg se formulan por separado.

25 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la segunda secuencia de ácido nucleico es o incluye: una secuencia de ácido nucleico que codifica un péptido 2A; una secuencia de ácido nucleico que codifica un sitio interno de entrada al ribosoma (IRES); una secuencia de ácido nucleico que codifica una región de corte y empalme aminoterminal de inteína y una región de corte y empalme carboxiterminal de inteína; o una secuencia de ácido nucleico que codifica un donante de corte y empalme y un aceptor de corte y empalme.

30 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la segunda secuencia de ácido nucleico es o incluye una secuencia de ácido nucleico que codifica un péptido 2A. En algunas realizaciones, la segunda secuencia de

ácido nucleico es o incluye una secuencia de ácido nucleico que codifica un péptido 2A seleccionado del grupo que consiste en P2A, T2A, E2A y F2A.

5 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el sitio de corte se encuentra a una distancia de entre 1 y 1000 pb con respecto al sitio de integración diana. En algunas realizaciones, el sitio de corte se encuentra a una distancia de hasta 100 pb del sitio de integración diana.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, el transgén se selecciona de CBS, UGT1A1, MUT, FAH, ATP7B, A1AT, ASL, LIPA, PAH, G6PC, Factor IX, o de una variante de los mismos.

10 En algunas realizaciones de uno cualquiera de los aspectos anteriores, el método, que es un conjunto de composiciones, en donde la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se formula en una nanopartícula lipídica y el casete polinucleotídico se encapsida en un vector de AAV recombinante.

15 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la célula se edita *in vivo*.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, el método de integración de un transgén se lleva a cabo *ex vivo*.

20 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la célula es una célula sanguínea, hepática, muscular o del SNC. En algunas realizaciones, la célula se administra en un trasplante autólogo después de la integración del transgén. En algunas realizaciones, la célula se administra en un trasplante alogénico después de la integración del transgén.

25 En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el casete polinucleotídico se administran al sujeto el mismo día. En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el casete polinucleotídico se administran al sujeto en días diferentes.

30 En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se administra al sujeto entre 1 hora y 3 días después del casete

polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se administra al sujeto entre 1 hora y 24 horas después del casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se administra al sujeto entre 1 hora y 3 días antes que el casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se administra al sujeto entre 1 hora y 4 horas antes que el casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se administra al sujeto 4 horas antes que el casete polinucleotídico.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el casete polinucleotídico se ponen en contacto con la célula humana el mismo día.

En algunas realizaciones de uno cualquiera de los aspectos anteriores, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa y el casete polinucleotídico se ponen en contacto con la célula humana en días diferentes.

En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se ponen en contacto con la célula humana entre 1 hora y 3 días después del casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se ponen en contacto con la célula humana entre 1 hora y 24 horas después del casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se ponen en contacto con la célula humana entre 1 hora y 3 días antes que el casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se ponen en contacto con la célula humana entre 1 y 4 horas antes que el casete polinucleotídico. En algunas realizaciones, la nucleasa o la secuencia de polinucleótidos que codifica la nucleasa se ponen en contacto con la célula humana 4 horas antes que el casete polinucleotídico.

En algunas realizaciones de cualquiera de los aspectos anteriores, el sitio de corte es distal con respecto al sitio de integración diana.

En otras realizaciones de cualquiera de los aspectos anteriores, el sitio de corte se solapa con el sitio de integración.

Como se utilizan en esta solicitud, los términos "alrededor de" y "aproximadamente" se utilizan como equivalentes. Todas las citas de publicaciones, patentes o solicitudes de patentes incluidas en el presente documento se incorporan por referencia en su totalidad. Cualquier número utilizado en esta solicitud con o sin alrededor de/aproximadamente, pretende cubrir cualquier fluctuación normal apreciada por un experto en la materia.

Otras características, objetos y ventajas de la presente invención son evidentes en la descripción detallada más adelante. Sin embargo, debe entenderse que aunque la descripción detallada indica realizaciones de la presente invención, se ofrece únicamente a título ilustrativo, no limitativo. Varios cambios y modificaciones dentro del alcance de la invención serán evidentes para los expertos en la materia a partir de la descripción detallada.

Definiciones

Aproximadamente: El término "aproximadamente", cuando se utiliza en el presente documento en referencia a un valor, se refiere a un valor que es similar, en contexto, al valor referenciado. En general, los expertos en la materia, familiarizados con el contexto, apreciarán el grado pertinente de variación abarcado por el término "aproximadamente" en ese contexto. Por ejemplo, en algunas realizaciones, el término "aproximadamente" puede abarcar un intervalo de valores que se encuentren dentro del 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, o menos del valor referido.

Adulto: Como se utiliza en el presente documento, el término "adulto" se refiere a un ser humano de dieciocho años o mayor. En algunas realizaciones, un adulto humano tiene un peso en el intervalo de aproximadamente 90 libras a aproximadamente 250 libras.

Asociados(as): Dos acontecimientos o entidades están "asociados(as)" entre sí, tal y como se utiliza este término en el presente documento, si la presencia, el nivel y/o la forma de uno está correlacionado con el del otro. Por ejemplo, se

considera que una entidad concreta (*p. ej.*, polipéptido, firma genética, metabolito, microbio y similares) está asociada a una enfermedad, trastorno o afección concretos, si su presencia, nivel y/o forma se correlaciona con la incidencia de y/o la susceptibilidad a la enfermedad, trastorno o afección (*p. ej.*, en una población relevante). En algunas realizaciones, dos o más entidades están físicamente "asociadas" entre sí si interactúan, directa o indirectamente, de modo que están y/o permanecen en proximidad física entre sí. En algunas realizaciones, dos o más entidades que están físicamente asociadas entre sí están unidas de manera covalente entre sí; en algunas realizaciones, dos o más entidades que están físicamente asociadas entre sí no están unidas de manera covalente entre sí, sino que están asociadas de manera no covalente, por ejemplo mediante enlaces de hidrógeno, interacción de van der Waals, interacciones hidrófobas, magnetismo y combinaciones de los mismos.

Muestra biológica: Como se utiliza en el presente documento, la expresión "muestra biológica" se refiere normalmente a una muestra obtenida o derivada de una fuente biológica (*p. ej.*, un tejido u organismo o cultivo celular) de interés, como se describe en el presente documento. En algunas realizaciones, una fuente de interés comprende un organismo, tal como un animal o un ser humano. En algunas realizaciones, una muestra biológica es o comprende tejido o líquido biológico. En algunas realizaciones, una muestra biológica puede ser o comprender médula ósea; sangre; células sanguíneas; ascitis; muestras de tejido o de biopsia con aguja fina; líquidos corporales que contengan células; ácidos nucleicos flotantes libres; esputo; saliva; orina; líquido cefalorraquídeo, líquido peritoneal; líquido pleural; heces; linfa; líquidos ginecológicos; frotis cutáneos; frotis vaginales; frotis orales; frotis nasales; baños o lavados tales como lavados ductales o lavados broncoalveolares; aspirados; raspados; muestras de médula ósea; muestras de biopsia tisular; muestras quirúrgicas; heces, otros líquidos corporales, secreciones y/o excreciones; y/o células de los mismos, y similares. En algunas realizaciones, una muestra biológica es o comprende células obtenidas de un individuo. En algunas realizaciones, las células obtenidas son o incluyen células de un individuo del que

se obtiene la muestra. En algunas realizaciones, una muestra es una "muestra primaria" obtenida directamente de una fuente de interés por cualquier medio apropiado. Por ejemplo, en algunas realizaciones, una muestra biológica primaria se obtiene mediante métodos seleccionados del grupo que consiste en biopsia (p. Ej., aspiración con aguja fina o biopsia tisular), cirugía, recogida de líquido corporal (p. ej., sangre, linfa, heces y similares.), y muestras similares. En algunas realizaciones, como quedará claro por el contexto, el término "muestra" se refiere a una preparación que se obtiene procesando (p. ej., eliminando uno o más componentes de y/o añadiendo uno o más agentes a) una muestra primaria. Por ejemplo, filtrado mediante una membrana semipermeable. Dicha "muestra procesada" puede comprender, por ejemplo, ácidos nucleicos o proteínas extraídos de una muestra u obtenidos sometiendo una muestra primaria a técnicas tales como amplificación o transcripción inversa de ARNm, aislamiento y/o purificación de determinados componentes, y técnicas similares.

Biomarcador. En el presente documento el término "biomarcador" se utiliza en consonancia con su uso en la técnica para referirse a una entidad cuya presencia, nivel o forma se correlaciona con un acontecimiento biológico o estado de interés particular, de modo que se considera un "marcador" de ese acontecimiento o estado. Entre otras cosas, la presente divulgación abarca biomarcadores de genoterapia (p. ej., que son útiles para evaluar uno o más rasgos o características de un tratamiento de genoterapia, tales como, por ejemplo, grado, nivel, y/o persistencia de la expresión de la carga útil). En algunas realizaciones, un biomarcador es un marcador de superficie celular. En algunas realizaciones, un biomarcador es intracelular. En algunas realizaciones, un biomarcador se encuentra fuera de las células (p. ej., se secreta o se genera o está presente de otro modo fuera de las células, p. ej., en un líquido corporal tal como sangre, orina, lágrimas, saliva, líquido cefalorraquídeo, y similares). En ciertas realizaciones, la presente divulgación demuestra la eficacia de los biomarcadores que pueden detectarse en una muestra obtenida de un sujeto que ha recibido genoterapia para su uso en la evaluación de uno o más rasgos o características de esa genoterapia; en algunas

de dichas realizaciones, la muestra es de células, tejido y/o líquido distinto de aquel al que se suministró la genoterapia y/o distinto de aquel en el que la carga útil está activa.

Optimización codónica: Como se utiliza en el presente documento, la expresión "optimización codónica" se refiere a un proceso de cambio de codones de un gen dado de tal manera que la secuencia de polipéptidos codificada por el gen sigue siendo la misma mientras que los codones cambiados mejoran el proceso de expresión de la secuencia de polipéptidos. Por ejemplo, si el polipéptido es de una secuencia de proteína humana y se expresa en *E. coli*, la expresión mejorará a menudo si se realiza una optimización codónica en la secuencia de ADN para cambiar los codones humanos por codones que sean más eficaces para la expresión en *E. coli*.

Fracción detectable: La expresión "fracción detectable", como se utiliza en el presente documento, se refiere a cualquier entidad (p. ej., molécula, complejo o parte o componente de los mismos). En algunas realizaciones, una fracción detectable se proporciona y/o utiliza como una entidad molecular distinta; en algunas realizaciones, forma parte de y/o está asociada a otra entidad molecular. Como ejemplos de fracciones detectables se incluyen, pero sin limitación: varios ligandos, radionúclidos (p. ej., ^3H , ^{14}C , ^{18}F , ^{19}F , ^{32}P , ^{35}S , ^{135}I , ^{125}I , ^{123}I , ^{64}Cu , ^{187}Re , ^{111}In , ^{90}Y , $^{99\text{m}}\text{Tc}$, ^{177}Lu , ^{89}Zr y similares), colorantes fluorescentes (para ejemplos específicos de colorantes fluorescentes, véase más adelante), agentes quimioluminiscentes (tales como, por ejemplo, ésteres acridínicos, dioxetanos estabilizados y similares), agentes bioluminiscentes, nanocristales semiconductores fluorescentes inorgánicos espectralmente resolubles (es decir, puntos cuánticos), nanopartículas metálicas (p. ej., oro, plata, cobre, platino y similares), nanoagrupaciones, iones metálicos paramagnéticos, enzimas (para ejemplos específicos de enzimas, véase más adelante), etiquetas colorimétricas (tales como, por ejemplo, colorantes, oro coloidal y similares), biotina, dioxigenina, haptenos, anticuerpos y/o proteínas para las que se dispone de antiseros o anticuerpos monoclonales.

Niño: Como se utiliza en el presente documento, el término "niño" se refiere a un ser humano entre dos y 18 años. El peso corporal puede variar mucho en función de la edad y de cada niño, con un intervalo habitual de entre 30 y 150 libras.

Combiterapia: Como se utiliza en el presente documento, el término "combiterapia" se refiere a aquellas situaciones en las que un sujeto se expone simultáneamente a dos o más pautas posológicas terapéuticas (p. ej., dos o más agentes terapéuticos, por ejemplo, una modalidad terapéutica de genoterapia y una modalidad terapéutica que no es genoterapia). En algunas realizaciones, las dos o más pautas posológicas pueden administrarse simultáneamente; en algunas realizaciones, dichas pautas posológicas pueden administrarse secuencialmente (p. ej., todas las "dosis" de una primera pauta posológica se administran antes de la administración de cualquier dosis de una segunda pauta posológica); en algunas realizaciones, dichos agentes se administran en pautas posológicas que se solapan. En algunas realizaciones, la "administración" de la combiterapia puede implicar la administración de uno o más agente(s) o modalidad(es) a un sujeto que recibe el otro agente(s) o modalidad(es) en la combinación. Para mayor claridad, la combiterapia no requiere que los agentes individuales se administren juntos en una única composición (ni siquiera necesariamente al mismo tiempo).

Composición: Los expertos en la materia apreciarán que el término "composición", como se utiliza en el presente documento, puede utilizarse para referirse a una entidad física distinta que comprende uno o más componentes especificados. En general, a menos que se especifique lo contrario, una composición puede tener cualquier forma - p. ej., gas, gel, líquido o sólido.

Determinar: Muchas de las metodologías descritas en el presente documento incluyen una etapa de "determinación". Los expertos en la materia, al leer la presente memoria descriptiva, comprenderán que dicha "determinación" puede utilizar o llevarse a cabo mediante el uso de cualquiera de las diversas técnicas disponibles para los expertos en la materia, incluidas, por ejemplo, las técnicas específicas a las que se hace referencia explícita en el presente documento. En algunas realizaciones, la determinación implica la manipulación de

una muestra física. En algunas realizaciones, la determinación implica el estudio y/o la manipulación de datos o información, por ejemplo utilizando un ordenador u otra unidad de procesamiento adaptada para realizar un análisis relevante. En algunas realizaciones, la determinación implica recibir información y/o materiales relevantes de una fuente. En algunas realizaciones, la determinación implica comparar una o más características de una muestra o entidad con una referencia comparable.

Distal: Como se utiliza en el presente documento, el término "distal" con respecto a la posición de un sitio de corte en relación con un sitio de integración diana significa que el sitio de corte y el sitio de integración diana no son idénticos y/o no se solapan. En algunas realizaciones, el sitio de corte está a una distancia de aproximadamente 2 kB (p. ej., 100 pb, 200 pb, 300 pb, 400 pb, 500 pb, 600 pb, 700 pb, 800 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 o 2.0 kB) del sitio de integración. En algunas realizaciones, el sitio de corte está dentro de los límites de los brazos de homología. Debe entenderse que, en otras realizaciones, el sitio de corte puede ser idéntico al sitio de integración diana o solaparse con éste.

Gen: Como se utiliza en el presente documento, el término "gen" se refiere a una secuencia de ADN que codifica un producto génico (p. ej., un producto de ARN y/o un producto polipeptídico). En algunas realizaciones, un gen incluye una secuencia codificante (p. ej., una secuencia que codifica un producto génico particular); en algunas realizaciones, un gen incluye una secuencia no codificante. En algunas realizaciones particulares, un gen puede incluir secuencias tanto codificantes (p. ej., exónicas) como no codificantes (p. ej., intrónicas). En algunas realizaciones, un gen puede incluir uno o más elementos reguladores (p. ej., promotores, potenciadores, silenciadores, señales de terminación) que, por ejemplo, pueden controlar o afectar a uno o más aspectos de la expresión génica (p. ej., expresión específica de tipo celular, expresión inducible). En algunas realizaciones, un gen se sitúa o se encuentra (o tiene una secuencia de nucleótidos idéntica a la localizada o encontrada) en un genoma (p. ej., en o sobre un cromosoma u otro ácido nucleico replicable).

Producto génico o producto de expresión: Como se utiliza en el presente documento, la expresión "producto génico" o "producto de expresión" se refiere generalmente a un ARN transcrito a partir del gen (pre y/o pos-procesamiento) o a un polipéptido (pre y/o pos-modificación) codificado por un ARN transcrito a partir del gen.

Identidad: Como se utiliza en el presente documento, el término "identidad" se refiere al parentesco global entre moléculas poliméricas, p.ej., entre moléculas de ácido nucleico (p. ej., moléculas de ADN y/o moléculas de ARN) y/o entre moléculas polipeptídicas. En algunas realizaciones, se considera que las moléculas poliméricas son "sustancialmente idénticas" entre sí si sus secuencias son al menos idénticas en un 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 %. El cálculo del porcentaje de identidad de dos secuencias de ácido nucleico o polipéptido, puede realizarse, por ejemplo, alineando las dos secuencias con la finalidad de realizar una comparación óptima (p. ej., pueden introducirse huecos en una o ambas de una primera y una segunda secuencias para realizar una alineación óptima y las secuencias que no son idénticas pueden ignorarse con fines de comparación). En ciertas realizaciones, la longitud de una secuencia alineada con fines de comparación es al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, al menos el 95 %, o sustancialmente el 100 % de la longitud de una secuencia de referencia. A continuación, se comparan los nucleótidos en las posiciones correspondientes. Cuando una posición de la primera secuencia está ocupada por el mismo resto (p. ej., nucleótido o aminoácido) que la posición correspondiente de la segunda secuencia, entonces las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias depende del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos y la longitud de cada hueco, que es necesario introducir para una alineación óptima de las dos secuencias. La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias pueden lograrse utilizando un algoritmo matemático. Por ejemplo,

el porcentaje de identidad entre dos secuencias de nucleótidos puede determinarse utilizando el algoritmo de Meyers y Miller (CABIOS, 1989, 4: 11-17), que se ha incorporado al programa ALIGN (versión 2.0). En algunas realizaciones ilustrativas, las comparaciones de secuencias de ácidos nucleicos realizadas con el programa

5 ALIGN utilizan una tabla de restos de peso PAM120, una penalización de longitud de hueco de 12 y una penalización de hueco de 4. Como alternativa, el porcentaje de identidad entre dos secuencias de nucleótidos puede determinarse utilizando el programa GAP del paquete informático GCG utilizando una matriz NWSgapdna.CMP.

10 **"Mejorar", "aumentar", "inhibir" o "reducir":** Como se utiliza en el presente documento, los términos "mejorar", "aumentar", "inhibir", "reducir", o sus equivalentes gramaticales, indican valores relativos a un valor inicial u otra medición de referencia. En algunas realizaciones, una medición de referencia apropiada puede ser o comprender una medición en un sistema particular (p. ej., en un único

15 individuo) en condiciones por lo demás comparables sin presencia de (p. ej., antes y/o después de) un agente o tratamiento particular, o en presencia de un agente de referencia comparable apropiado. En algunas realizaciones, una medición de referencia apropiada puede ser o comprender una medición en un sistema comparable que se sabe o se espera que responda de una manera particular, en

20 presencia del agente o tratamiento pertinente.

Lactante: Como se utiliza en el presente documento, el término "lactante" se refiere a un ser humano menor de dos años. El peso corporal normal de un lactante varía entre 3 y 20 libras.

25 **Neonato:** Como se utiliza en el presente documento, el término "neonato" se refiere a un ser humano recién nacido.

Ácido nucleico: Como se utiliza en el presente documento, en su sentido más amplio, se refiere a cualquier compuesto y/o sustancia que se incorpora o puede incorporarse a una cadena de oligonucleótidos. En algunas realizaciones, un ácido nucleico es un compuesto y/o sustancia que se incorpora o puede

30 incorporarse a una cadena de oligonucleótidos mediante un enlace fosfodiéster.

Como quedará claro por el contexto, en algunas realizaciones, "*ácido nucleico*" se refiere a un resto de ácido nucleico individual (p. ej., un nucleótido y/o nucleósido); en algunas realizaciones, "*ácido nucleico*" se refiere a una cadena de oligonucleótidos que comprende restos de ácido nucleico individuales. En algunas realizaciones, un "*ácido nucleico*" es o comprende ARN; en algunas realizaciones, un "*ácido nucleico*" es o comprende ADN. En algunas realizaciones, un ácido nucleico es, comprende o consiste en uno o más restos de ácido nucleico natural. En algunas realizaciones, un ácido nucleico es, comprende o consiste en uno o más análogos de ácido nucleico. En algunas realizaciones, un análogo de ácido nucleico difiere de un ácido nucleico en que no utiliza un esqueleto fosfodiéster. En algunas realizaciones, un ácido nucleico es, comprende o consiste en uno o más nucleósidos naturales (p. ej., adenosina, timidina, guanosina, citidina, uridina, desoxiadenosina, desoxitimidina, desoxiguanosina y desoxicitidina). En algunas realizaciones, un ácido nucleico es, comprende o consiste en uno o más análogos de nucleósido (p. ej., 2-aminoadenosina, 2-tiotimidina, inosina, pirrolo-pirimidina, 3-metil adenosina, 5-metilcitidina, C-5 propinil-citidina, C-5 propinil-uridina, 2-aminoadenosina, C5-bromouridina, C5-fluorouridina, C5-yodouridina, C5-propinil-uridina, C5-propinil-citidina, C5-metilcitidina, 2-aminoadenosina, 7-deazaadenosina, 7-deazaguanosina, 8-oxoadenosina, 8-oxoguanosina, 0(6)-metilguanina, 2-tiocitidina, bases metiladas, bases intercaladas y combinaciones de las mismas). En algunas realizaciones, un ácido nucleico tiene una secuencia de nucleótidos que codifica un producto génico funcional, tal como un ARN o una proteína. En algunas realizaciones, un ácido nucleico incluye uno o más intrones. En algunas realizaciones, los ácidos nucleicos se preparan mediante uno o más de aislamiento de una fuente natural, síntesis enzimática por polimerización basada en un molde complementario (*in vivo* o *in vitro*), reproducción en una célula o sistema recombinante, y síntesis química. En algunas realizaciones, un ácido nucleico tiene una longitud de al menos 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900,

1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 o más restos. En algunas realizaciones, un ácido nucleico es parcial o totalmente monocatenario; en algunas realizaciones, un ácido nucleico es parcial o totalmente bicatenario. En algunas realizaciones, un ácido nucleico tiene una secuencia de nucleótidos que comprende al menos un elemento que codifica, o es el complemento de una secuencia que codifica, un polipéptido. En algunas realizaciones, un ácido nucleico tiene actividad enzimática.

Péptido: Como se utiliza en el presente documento, el término "péptido" o "polipéptido" se refiere a cualquier cadena polimérica de aminoácidos. En algunas realizaciones, un péptido tiene una secuencia de aminoácidos que se da en la naturaleza. En algunas realizaciones, un péptido tiene una secuencia de aminoácidos que no se da en la naturaleza. En algunas realizaciones, un péptido tiene una secuencia de aminoácidos que el hombre ha diseñado y/o producido de manera artificial. En algunas realizaciones, un péptido puede comprender o consistir en aminoácidos naturales, aminoácidos no naturales, o ambos. En algunas realizaciones, un péptido puede comprender o consistir únicamente en aminoácidos naturales o únicamente en aminoácidos no naturales. En algunas realizaciones, un péptido puede comprender D-aminoácidos, L-aminoácidos, o ambos. En algunas realizaciones, un péptido puede comprender únicamente D-aminoácidos. En algunas realizaciones, un péptido puede comprender únicamente L-aminoácidos. En algunas realizaciones, un péptido es lineal. En algunas realizaciones, el término "péptido" puede añadirse al nombre de un péptido, actividad o estructura de referencia; en tales casos, se utiliza en el presente documento para referirse a péptidos que comparten la actividad o estructura relevante y, por tanto, pueden considerarse miembros de la misma clase o familia de péptidos. Para cada una de estas clases, la presente memoria descriptiva proporciona péptidos ilustrativos dentro de la clase cuyas secuencias de aminoácidos y/o funciones se conocen y/o los expertos en la materia son conscientes de ello; en algunas realizaciones, dichos péptidos ilustrativos son péptidos de referencia para la clase o familia de péptidos. En algunas realizaciones, un miembro de una clase o familia de péptidos muestra

homología o identidad de secuencia significativa con, comparte un motivo de secuencia común (p. ej., un elemento de secuencia característico) con, y/o comparte una actividad común (en algunas realizaciones a un nivel comparable o dentro de un intervalo designado) con un péptido de referencia de la clase; en algunas realizaciones con todos los péptidos dentro de la clase). Por ejemplo, en algunas realizaciones, un péptido miembro muestra un grado global de homología o identidad de secuencia con un péptido de referencia que es de al menos aproximadamente el 30-40 %, y a menudo es mayor que aproximadamente el 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más y/o que incluye al menos una región (p. ej., una región conservada que, en algunas realizaciones, puede ser o comprender un elemento de secuencia característico) que muestra una identidad de secuencia muy alta, a menudo mayor del 90 % o incluso del 95 %, 96 %, 97 %, 98 % o 99 %. Dicha región conservada suele abarcar al menos 3-4 y a menudo hasta 20 o más aminoácidos; en algunas realizaciones, una región conservada abarca al menos un tramo de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o más aminoácidos contiguos.

Puerto seguro: Como se usa en el presente documento, la expresión "puerto seguro", "locus de puerto seguro", "gen de puerto seguro" o "sitio de puerto seguro" se refiere a una o más regiones del genoma de un sujeto o célula (p. Ej., genoma de ratón, genoma humano, genoma animal humanizado, genoma animal quimérico) que permiten la expresión estable de transgenes integrados sin afectar negativamente a un sujeto y/o a una célula. En algunas realizaciones, la integración en un puerto seguro permite una función predecible de los transgenes integrados. En algunas realizaciones, la integración en un puerto seguro no causa una alteración no intencionada del genoma de un sujeto o célula que suponga un riesgo para el sujeto o la célula. En algunas realizaciones, los genes de puerto seguro pueden alterarse sin integración de transgenes (p. ej., mediante NHEJ, *non-homologous end joining*, unión de extremos no homólogos) sin que afecte negativamente a un sujeto y/o a una célula (p. ej., sin causar oncogénesis).

Sujeto: Como se utiliza en el presente documento, el término "sujeto" se refiere a un organismo, normalmente un mamífero (p. ej., un ser humano, incluyendo en algunas realizaciones formas humanas prenatales). En algunas realizaciones, un sujeto padece una enfermedad, trastorno o afección relevante. En algunas realizaciones, un sujeto es susceptible a una enfermedad, trastorno o afección. En algunas realizaciones, un sujeto muestra uno o más síntomas o características de una enfermedad, trastorno o afección. En algunas realizaciones, un sujeto no muestra ningún síntoma o característica de una enfermedad, trastorno o afección. En algunas realizaciones, un sujeto es alguien con uno o más rasgos característicos de susceptibilidad o riesgo de una enfermedad, trastorno o afección. En algunas realizaciones, un sujeto es un paciente. En algunas realizaciones, un sujeto es un individuo al que se le administra y/o se le ha administrado un diagnóstico y/o una terapia. En realizaciones preferidas, el sujeto es un ser humano.

Sustancialmente: Como se utiliza en el presente documento, el término "sustancialmente" se refiere a la condición cualitativa de presentar una extensión o grado total o casi total de una característica o propiedad de interés. Un experto familiarizado con el campo de la biología comprenderá que los fenómenos biológicos y químicos rara vez, si es que alguna vez lo hacen, llegan a finalizar y/o prosiguen hasta finalizar o alcanzar o evitar un resultado absoluto. Por lo tanto, en el presente documento, el término "sustancialmente" se utiliza para captar la posible falta de exhaustividad intrínseca a muchos fenómenos biológicos y químicos.

Variante: Como se utiliza en el presente documento en el contexto de moléculas, p. ej., ácidos nucleicos, proteínas o moléculas pequeñas, el término "variante" se refiere a una molécula que muestra identidad estructural significativa con una molécula de referencia pero difiere estructuralmente de la molécula de referencia, p. ej., en la presencia o ausencia o en el nivel de una o más fracciones químicas en comparación con la entidad de referencia. En algunas realizaciones, una variante también difiere funcionalmente de su molécula de referencia. En general, el hecho de que una molécula concreta se considere adecuadamente una "variante" de una molécula de referencia se basa en su grado de identidad

estructural con la molécula de referencia. Como apreciarán los expertos en la materia, cualquier molécula de referencia biológica o química tiene ciertos elementos estructurales característicos. Una variante, por definición, es una molécula distinta que comparte uno o más de esos elementos estructurales

5 característicos pero difiere en al menos un aspecto de la molécula de referencia. Por citar sólo algunos ejemplos, un polipéptido puede tener un elemento de secuencia característico compuesto por una pluralidad de aminoácidos que tienen posiciones designadas relativas entre sí en el espacio lineal o tridimensional y/o que contribuyen a un motivo estructural particular y/o a una función biológica; un ácido

10 nucleico puede tener un elemento de secuencia característico compuesto por una pluralidad de restos de nucleótidos que tienen posiciones designadas relativas entre sí en el espacio lineal o tridimensional. En algunas realizaciones, un polipéptido o ácido nucleico variante puede diferir de un polipéptido o ácido nucleico de referencia como resultado de una o más diferencias en la secuencia de aminoácidos o

15 nucleótidos y/o una o más diferencias en las fracciones químicas (p. ej., hidratos de carbono, lípidos, grupos fosfato) que son componentes del polipéptido o ácido nucleico unidos de manera covalente (p. ej., que están unidos al esqueleto del polipéptido o ácido nucleico). En algunas realizaciones, un polipéptido o ácido nucleico variante muestra una identidad de secuencia global con un polipéptido o

20 ácido nucleico de referencia que es al menos del 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 % o 99 %. En algunas realizaciones, un polipéptido o ácido nucleico variante no comparte al menos un elemento de secuencia característico con un polipéptido o ácido nucleico de referencia. En algunas realizaciones, un polipéptido o ácido nucleico de referencia tiene una o más

25 actividades biológicas. En algunas realizaciones, un polipéptido o ácido nucleico variante comparte una o más de las actividades biológicas del polipéptido o ácido nucleico de referencia. En algunas realizaciones, un polipéptido o ácido nucleico variante carece de una o más de las actividades biológicas del polipéptido o ácido nucleico de referencia. En algunas realizaciones, un polipéptido o ácido nucleico

30 variante muestra un nivel reducido de una o más actividades biológicas en

comparación con el polipéptido o ácido nucleico de referencia. En algunas realizaciones, se considera que un polipéptido o ácido nucleico de interés es una "variante" de un polipéptido o ácido nucleico de referencia si tiene una secuencia de aminoácidos o nucleótidos idéntica a la de la referencia pero con un pequeño número de alteraciones de secuencia en posiciones concretas. Normalmente, menos de aproximadamente el 20 %, aproximadamente el 15 %, aproximadamente el 10 %, aproximadamente el 9 %, aproximadamente el 8 %, aproximadamente el 7 %, aproximadamente el 6 %, aproximadamente el 5 %, aproximadamente el 4 %, aproximadamente el 3 % o aproximadamente el 2 % de los restos de una variante están sustituidos, insertados o delecionados, en comparación con la referencia. En algunas realizaciones, un polipéptido o ácido nucleico variante comprende aproximadamente 10, aproximadamente 9, aproximadamente 8, aproximadamente 7, aproximadamente 6, aproximadamente 5, aproximadamente 4, aproximadamente 3, aproximadamente 2 o aproximadamente 1 resto sustituido en comparación con una referencia apropiada. A menudo, un polipéptido o ácido nucleico variante comprende un número muy pequeño (p. ej., menos de aproximadamente 5, aproximadamente 4, aproximadamente 3, aproximadamente 2 o aproximadamente 1) de restos funcionales sustituidos, insertados o delecionados (es decir, restos que participan en una actividad biológica particular) en relación con la referencia. En algunas realizaciones, un polipéptido o ácido nucleico variante comprende no más de aproximadamente 5, aproximadamente 4, aproximadamente 3, aproximadamente 2 o aproximadamente 1 adición o deleción, y, en algunas realizaciones, no comprende adiciones o deleciones, en comparación con la referencia. En algunas realizaciones, un polipéptido o ácido nucleico variante comprende menos de aproximadamente 25, aproximadamente 20, aproximadamente 19, aproximadamente 18, aproximadamente 17, aproximadamente 16, aproximadamente 15, aproximadamente 14, aproximadamente 13, aproximadamente 10, aproximadamente 9, aproximadamente 8, aproximadamente 7, aproximadamente 6, y comúnmente menos de aproximadamente 5, aproximadamente 4, aproximadamente 3 o

aproximadamente 2 adiciones o deleciones en comparación con la referencia. En algunas realizaciones, un polipéptido o ácido nucleico de referencia es uno que se encuentra en la naturaleza. En algunas realizaciones, un polipéptido o ácido nucleico de referencia es un polipéptido o ácido nucleico humano.

5 **Breve descripción de los dibujos**

La FIG. 1A muestra un diagrama ilustrativo de posibles sitios de corte, por ejemplo, mediante una o más nucleasas, en el gen de la albúmina (ALB) humana. La FIG. 1B muestra la secuencia de la ALB humana alrededor del codón FINALIZADOR en el exón 14, indicando la secuencia subrayada una región
10 ilustrativa del corte diana.

La FIG. 2A muestra que una rotura bicatenaria (DSB, *double-strand break*) inducida en el intrón 14 del gen de la ALB humana mejoró la eficacia de integración. La FIG. 2B muestra que una DSB inducida en el intrón 13 del gen de la ALB humana también mejoró la eficacia de integración. La FIG. 2C muestra el grado de mejora
15 de la eficacia de integración en relación con la ubicación de la DSB introducida en varios lugares del intrón 13 o intrón 14 del gen de la ALB. La FIG. 2D muestra una imagen de fluorescencia representativa que muestra que la proteína Cas9 junto con el ARN guía puede aumentar las células positivas a GFP (*green fluorescent protein*, proteína verde fluorescente). La FIG. 2E muestra que una DSB inducida en el intrón
20 13 y/o 14 de la ALB humana mejoró la eficacia de integración (medida como porcentaje de células positivas a GFP). La FIG. 2F muestra que una DSB inducida en el intrón 13 y/o 14 de la ALB humana mejoró la eficacia de integración (medida como niveles ARNm fusionado).

La FIG. 3A muestra imágenes de fluorescencia representativas que muestran
25 que el uso adecuado de Cas9 y de ARN guía puede aumentar la eficacia de integración. La FIG. 3B muestra diferentes proporciones entre ARNm de Cas9 y ARNg y la eficacia de integración (medida como porcentaje de células positivas a GFP).

La FIG. 4A muestra un esquema de una construcción GENERIDE™ antes de
30 la integración (AAV) y después de la integración mediada por HR en el genoma en

un locus diana (p. ej., Albúmina). La FIG. 4B muestra que las alineaciones de homología más largas pueden predecir una mayor eficacia de integración (medida como nivel de ARNm fusionado). La FIG. 4C muestra que las alineaciones de homología más largas pueden predecir una mayor eficacia de integración (medida como porcentaje de células positivas a GFP).

La FIG. 5A muestra que una variedad de Cas9 (p. ej., HF1-Cas9, HF2-Cas9) puede mejorar la integración específica. La FIG. 5B muestra que una versión de alta fidelidad (*HiFi*, *high-fidelity*) de Cas9 (p. ej., HiFi-Cas9), puede mejorar la integración específica.

La FIG. 6 muestra que Cas9 de *Staphylococcus aureus* (saCas9) puede mejorar la integración del vector GENERIDE™ humano.

La FIG. 7 muestra un plásmido ilustrativo diseñado para expresar saCas9 y un ARNg (panel superior). Cuando el plásmido se empaqueta en un AAV puede mejorar la integración después de la transfección (panel inferior).

La FIG. 8 muestra imágenes de fluorescencia representativas que muestran que el uso de Cas9 y de ARN guía adecuados puede aumentar la eficacia de integración.

La FIG. 9A muestra que una DSB inducida en el intrón 13 y/o 14 de la ALB humana mejoró la eficacia de integración (medida como porcentaje de células positivas a GFP). La FIG. 9B muestra que una DSB inducida en el intrón 13 y/o 14 de la ALB humana mejoró la eficacia de integración (medida como niveles de ARNm fusionado).

La FIG. 10 muestra la evaluación de biomarcadores de GENERIDE™ 3 semanas después de la administración *in-vivo* de GENERIDE™, Cas9 y ARN guía apropiados. D, día de la administración de LNP en relación con la administración de AAV.

La FIG. 11 muestra imágenes que representan que la administración *in vivo* de Cas9 y de ARN guía apropiados puede aumentar la eficacia de integración.

La FIG. 12 muestra los resultados de un experimento en el que se administraron construcciones de GENERIDE™ (GR-hATP7B o GR-GFP) solas o

junto con Cas9, con un control sin tratamiento previo. La construcción GR-hATP7b incluía brazos de homología de distintos tamaños (0.4/0.8 kB; 0.8/0.4 kB o 0.6/0.6 kB), mientras que la GR-GFP tenía brazos de homología de 1.0/1.0 kB. GOI, gen de interés (*gene of interest*). La FIG. 12 muestra una serie de gráficos que miden los niveles de ARNm fusionado (copias/20 mg) (panel izquierdo) y ARNm fusionado normalizado (norm. a 0.6/0.6 GR-hATP7B) (panel derecho). GR-GFP tenía un valor inicial más alto, lo que, sin querer ligarse a ninguna teoría en particular, puede deberse a brazos de homología más largos o a un GOI más pequeño.

La FIG. 13 muestra los resultados de un experimento en el que el factor de aumento del ARNm fusionado entre los GOI se midió como ARNm fusionado normalizado (norm. a GR solamente) en las construcciones indicadas. El nivel de ARNm fusionado es una medida de la eficacia de integración.

La FIG. 14 muestra los resultados de un experimento en el que se administró una construcción de GENERIDE™ (GR-hAAT) a células HepG2 humanas sola o junto con ARNm de spCas9, con un control sin tratamiento previo y H2O. La FIG. 14 muestra un gráfico que mide los niveles de ARNm fusionado (copias/100 ng). El ARNm de spCas9 aumentó la eficacia de integración de GR-hAAT.

Las FIGS. 15A y 15B muestran los resultados de un experimento en el que se midió la eficacia de integración de plásmidos de tipo silvestre (WT, *wild type*) frente a plásmidos mutantes, en formato circular o lineal con o sin Cas9. Los plásmidos tenían sitios de corte de enzimas de restricción para linealizar el plásmido, y las mutaciones de PAM pueden mejorar aún más la integración. La FIG. 15A muestra un gráfico que mide los niveles de ARNm fusionado (copias/100 ng) en presencia o ausencia de Cas9 de plásmidos de WT y mutantes, en formato circular y lineal, con un control sin tratamiento previo (panel izquierdo). La FIG. 15B muestra un gráfico que mide los niveles de ARNm fusionado en presencia de Cas9 (norm. a C-WT+Cas9) de plásmidos de WT y mutantes, tanto en formato circular como lineal (panel derecho).

Las FIGS. 16A y 16B muestran los resultados de un experimento en el que se midió la eficacia de edición (*Editing _efficiency*), el porcentaje de lecturas con

integración de donante mediante HDR (*homology-directed repair*, reparación dirigida por homología) (KI_HDR) y el porcentaje de lecturas con integración de donante mediante NHEJ en orientación directa (*forward*, *Fw*) o inversa (*reverse*, *Rv*) (NHEJ_Fw+Rv) globales, después de la transfección con ARNm de Nucleasa de AZ (FIG. 16A) o SpCas9 (FIG. 16B) de diferentes proporciones entre donante y ARNg respectivamente. Las flechas indican solapamiento de los ARNg para SpCas9 y Nucleasa de AZ; los asteriscos indican los ARNg con el mayor KI-HDR para la Nucleasa de AZ.

Las FIGS. 17A y 17B muestran los resultados de un experimento en el que se midió la eficacia de integración del transgén para diferentes ARNg humanos en células HEPG2. La FIG. 17A muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de spCas9 en células HEPG2, con un control de GR solo (PD128). La FIG. 17B muestra un gráfico que mide los niveles de ARNm fusionado normalizados (norm. a vt solamente) en presencia de spCas9 en células HEPG2, con un control de GR solo (PD128).

Las FIGS. 18A y 18B muestran los resultados de un experimento en el que se evaluó el efecto de la ubicación del corte (es decir, del corte dentro del brazo de homología (HA, *homology arm*) de GR, en el que los segmentos del HA donante en 3' y 5' del sitio de corte participan en la HDR; en otras palabras, el sitio de DSB se encuentra dentro de los límites del brazo de homología, frente al corte fuera del brazo de homología de GR, en el que solo un extremo puede experimentar HDR) para spCas9 en células humanas. La FIG. 18A muestra un gráfico en donde se comparan los niveles de ARNm fusionado (norm.) cuando el corte se realiza fuera del HA (1 brazo) frente a cuando el corte se realiza dentro del HA (2 brazos); el corte dentro del HA dio como resultado una eficacia de integración significativamente más alta para spCas9. La FIG. 18B muestra un gráfico que mide los niveles de ARNm fusionado (norm.) en función de la longitud alineada (intrón 12 y 13) cuando el corte se realiza fuera del HA y dentro del HA.

Las FIGS. 19A-19C muestran los resultados de un experimento en el que se midió la eficacia de integración del transgén de GENERIDE™ con la Nucleasa de

AZ Cas9 para diferentes ARNg humanos en células HEPG2. La FIG. 19A muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-mut y Nucleasa de AZ Cas9 en células HEPG2, con un control PD128. La FIG. 19B muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-GFP y Nucleasa de AZ Cas9 en células HEPG2, con un control sin tratamiento previo y un control de GR solo (vt-0290). La FIG. 19C muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-ATP7B y Nucleasa de AZ Cas9 en células HEPG2, con un control sin tratamiento previo y un control de GR solo (vt-235).

Las FIGS. 20A y 20B muestran los resultados de un experimento que mide la eficacia de la integración de GENERIDE™ con Nucleasa de AZ Cas9 al cortar dentro del brazo de homología, HA, frente a fuera del HA en células HEPG2. La FIG. 20A muestra un gráfico en donde se comparan los niveles de ARNm fusionado (norm. a vt únicamente) cuando el corte se realiza fuera del HA (1 brazo) frente a cuando el corte se realiza dentro del HA (2 brazos) (panel izquierdo). La FIG. 20B muestra un gráfico que mide los niveles de ARNm fusionado (norm.) en función de la longitud alineada (intrón 12 y 13) cuando el corte se realiza fuera del HA y dentro del HA (panel derecho).

La FIG. 21 muestra los resultados de un experimento en el que se midió la eficacia de edición (*Editing_efficiency*), el porcentaje de lecturas con integración de donante mediante HDR (KI_HDR) y el porcentaje de lecturas con integración de donante mediante NHEJ en orientación directa o inversa (NHEJ_Fw+Rv) globales, después de la transfección con ARNm de Nucleasa de AZ de diferentes proporciones entre donante y ARNg respectivamente. La FIG. 21 muestra una serie de gráficos en los que se midió la eficacia de edición (%) (panel superior), la KI_HDR (%) (panel central) y la NHEJ_Fw (*forward*, directa) +Rv (*reverse*, inversa) (%) (panel inferior), después de la transfección con ARNm de Nucleasa de AZ de los ARNg y donantes indicados (Sin donante, ADNmc o ADNbc) en células HEPG2. La media de KI_HDR (%) fue mayor para el ADNmc (monocatenario) que para el ADNbc (bicatenario) y viceversa en el caso de NHEJ_Fw+Rv (%). Las flechas

indican solapamiento de los ARNg para SpCas9 y Nucleasa de AZ; los asteriscos indican los ARNg con el mayor KI-HDR para la Nucleasa de AZ.

Las FIGS. 22A-22E muestran resultados de experimentos realizados para optimizar una estrategia de dosificación de GENERIDE™ y de la nanopartícula lipídica (LNP) GenVoy-ILM (*Ionizable Lipid Mix*, mezcla lipídica ionizable) en células hepáticas, para confirmar el suministro de Cas9 por la LNP GenVoy-ILM, y determinar la cronología de los componentes para la integración máxima. A diferentes grupos de ratones se les administraron diferentes combinaciones de GR (3e13 vg/kg) solo, GR (3e13 vg/kg) junto con Cas9 (1 mg/kg) después de 1, 3 o 7 días, o solo LNP que contenía Cas9 + ARNg (1 mg/kg). El vector de la herramienta GENERIDE™ que se utilizó fue DJ-GFP (LB-Vt-0298-001). La FIG. 22A muestra imágenes de inmunohistoquímica (IHQ) de células hepáticas que detectan Cas9-Flag (Cas9 con etiqueta Flag), que confirman la expresión de Cas9 después del suministro de la LNP (nanopartícula lipídica) que contenía Cas9, con un control de tipo silvestre (WT) sin tratamiento previo. La FIG. 22B muestra una serie de gráficos que miden los niveles de ALB-2A (µg/ml) hasta 6 semanas después de la dosificación de LNP (panel izquierdo) y que miden los niveles de ALB-2A (µg/ml) a las 3 semanas después de la dosificación de LNP (panel derecho) para GR dosificado con LNP en diferentes puntos temporales (D1, D3 y D7), con un control de GR únicamente. La FIG. 22C muestra una imagen IHQ de hepatocitos que confirma una mejor integración en GR2.0 (Cas9) después del suministro de GR (3e13) + LNP (1 mg/kg) que contenía Cas9. Para este experimento, ratones FvB/NJ recibieron dosis de 3e13 vk de GR-GFP (1.0/1.0) y después, el Día 1, de LNP-Cas9 GenVoy-ILM a 1 mg/kg. La FIG. 22D muestra un gráfico de barras que mide los niveles de ALB-2A (µg/ml) 1 semana después de la administración para confirmar la potencia de la LNP almacenada cuando la LNP se ha almacenado durante 1 o 2 meses. La FIG. 22E muestra una serie de gráficos que cuantifican la frecuencia de indel (inserción/delección) mediante un ensayo TIDE de D4-1 (panel superior) y D4-2 (panel inferior).

Las FIGS. 23A-23F muestran los resultados de experimentos realizados para evaluar el ARN de *Integrated DNA Technologies* (IDT) para garantizar la viabilidad con la LNP de AstraZeneca (AZ). El donante GR-HDR utilizado fue DJ-mHA (1.0/1.0) GFP-LB-Vt-0298-001 y la LNP utilizada fue la LNP AZ. Se administraron 5 combinaciones diferentes a distintos grupos de ratones de WT: Únicamente GR WT, GR WT + LNP N:P=3 (Co-formulación) escalonado el D1, GR WT + LNP N:P=6 (Co-formulación) escalonado el D1, GR WT + LNP N:P=6 (Formulación por separado (Sep.)) escalonado el D1 y GR WT + LNP N:P=6 (Co-formulación) escalonado en H4. La FIG. 23A muestra los resultados de un experimento realizado para medir los niveles de ALB-2A ($\mu\text{g/ml}$) hasta 3 semanas después de la dosificación de LNP en las 5 combinaciones diferentes enumeradas anteriormente, con el GR únicamente como control. La FIG. 23B es un gráfico de barras que mide las copias de ARNm fusionado/100 ng de ARN de las 5 combinaciones diferentes administradas, con el GR únicamente como control. La FIG. 23C es un conjunto de imágenes de IHQ, realizándose la IHQ para etiquetar y medir células GFP+ después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23D es un gráfico de barras que mide el % de células positivas a GFP después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23E es un gráfico de barras que mide los niveles de ALB-2A ($\mu\text{g/ml}$) después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23F muestra los niveles de ALB-2A ($\mu\text{g/ml}$) en función del % de células positivas a GFP después de la administración de las 5 combinaciones enumeradas.

Las FIGS. 24A y 24B muestran resultados de experimentos realizados para optimizar un calendario de tratamiento, en los que las células se trataron de acuerdo con 2 protocolos diferentes para determinar cuál era mejor para mejorar la integración. La FIG. 24A muestra los resultados de experimentos en los que las células se trataron con GR el mismo día que la transfección de ARNm (protocolo original). La siembra se realizó el D0 seguido de la transfección de ARNm de la nucleasa junto con el tratamiento GR-GFP el D1 y la recogida se realizó el D4 (panel superior). La FIG. 24A muestra un gráfico de barras que mide el % de células HA+

(positivas a hemaglutinina) de las 4 combinaciones diferentes administradas: sin tratamiento previo (control), GR-GFP, GR-GFP + Cas9, y sólo ARNm (panel inferior izquierdo) y otro gráfico de barras que mide el % de células HA+ (norm.) de las mismas 4 combinaciones administradas (panel inferior derecho). La FIG. 24B muestra los resultados de experimentos en los que las células se trataron con GR el día después de la transfección de ARNm. La siembra se realizó el D0, seguido de la transfección del ARNm de la nucleasa el D1, seguido del tratamiento con GR-GFP el D2 y, por último, la recogida se realizó el D5 (panel superior). La FIG. 24B muestra un gráfico de barras que mide el % de células HA+ de las 4 combinaciones diferentes administradas: sin tratamiento previo (control), GR-GFP, GR-GFP + Cas9, y sólo ARNm (panel inferior izquierdo) y otro gráfico de barras que mide el % de células HA+ (norm.) de las mismas 4 combinaciones administradas (panel inferior derecho).

La FIG. 25 muestra los resultados de experimentos realizados para comparar la eficacia de transfección de diferentes reactivos mediante inmunofluorescencia. La FIG. 25 muestra 4 imágenes de inmunofluorescencia (panel superior izquierdo) que prueban la eficacia de transfección de Lipofectamine™ 3000 para la transfección de mCherry y eGFP, con un canal DAPI y una imagen fusionada de eGFP-mCherry y 4 imágenes de inmunofluorescencia (panel superior derecho) que prueban la eficacia de transfección de MessengerMax™ para la transfección de mCherry y eGFP, con un canal DAPI y una imagen fusionada de eGFP-mCherry. La FIG. 25 también muestra gráficos de barras que miden el % de eficacia de transfección para eGFP y mCherry tanto para Lipofectamine™ 3000 como para MessengerMax™, paneles inferior izquierdo e inferior derecho, respectivamente.

La FIG. 26 muestra los resultados de experimentos realizados para medir la cantidad de expresión de proteína (% de células HA+) al utilizar ARNm de Cas9 junto con MessengerMax™ como reactivo de transfección. La siembra se realizó el D0, seguido de la transfección del ARNm de la nucleasa el D1, seguido del tratamiento con GR-GFP el D2 y, por último, la recogida se realizó el D5 (panel superior). La FIG. 26 muestra un gráfico de barras que mide el % de células HA+

de las 8 combinaciones diferentes administradas: sin tratamiento previo (control), ARNm solo, Vehículo + GR-GFP + ARNm de Cas9, Vehículo + ODNbc + GR-GFP + ARNm de Cas9, ARNg n.º 2 + GR-GFP + ARNm de Cas9, ARNg n.º 2 + ODNbc + GR-GFP + ARNm de Cas9, ARNg n.º 10 + GR-GFP + ARNm de Cas9, y ARNg n.º 10 + ODNbc + GR-GFP + ARNm de Cas9 (panel inferior izquierdo), un gráfico de barras que mide el % de células HA+ (norm.) de las mismas 8 combinaciones administradas (panel inferior central), y otro gráfico de barras que mide el recuento de células/imagen de las mismas 8 combinaciones administradas (panel inferior derecho).

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

Genoterapia

Las genoterapias alteran el perfil de expresión génica de las células de un paciente mediante transferencia génica, un proceso de suministro de un gen terapéutico exógeno, denominado transgén. Se conocen varios vehículos de suministro que se utilizan como vectores para transportar transgenes al núcleo de una célula para alterar o aumentar las capacidades de una célula (p. ej., proteoma, funcionalidad y similares). Los inventores han logrado grandes avances en la introducción de genes en células de tejidos como el hígado, la retina del ojo y las células hematopoyéticas de la médula ósea utilizando diversos vectores. En algunos casos, estas estrategias han dado lugar a terapias aprobadas y, en otros, han mostrado resultados muy prometedores en ensayos clínicos.

Existen múltiples estrategias de genoterapia. En la genoterapia convencional con AAV (también descrita en el presente documento como genoterapia tradicional o convencional), se introduce un transgén en el núcleo de una célula hospedadora pero no se pretende que se integre en el ADN cromosómico. El transgén se expresa a partir de un elemento genético no integrado denominado episoma que existe en el interior del núcleo. Un segundo tipo de genoterapia emplea el uso de un tipo de virus diferente, tal como un lentivirus, que se por sí solo se inserta, junto con el transgén, en el ADN cromosómico pero en sitios arbitrarios.

La expresión episomal de un gen debe dirigirla un promotor exógeno, lo que conduce a la producción de una proteína que corrige o mejora el estado de la enfermedad.

Limitaciones de las genoterapias conocidas hasta ahora

5 En el caso de la genoterapia basada en la expresión episomal, cuando las células se dividen durante el proceso de crecimiento o regeneración tisular, los beneficios de la terapia suelen disminuir porque los transgenes no estaban destinados a integrarse en un cromosoma hospedador y no se replican durante la división celular. Por tanto, cada nueva generación de células reduce aún más la
10 proporción de células que expresan el transgén en el tejido diana, lo que conduce a la reducción o eliminación del beneficio terapéutico con el tiempo.

Aunque el uso de algunas genoterapias que utilizan la inserción mediada por virus tiene el potencial de proporcionar beneficios a largo plazo porque el gen se inserta en el cromosoma hospedador, no hay capacidad de controlar dónde se
15 inserta el gen, lo que presenta el riesgo de alterar un gen esencial o de insertarlo en un lugar que pueda promover efectos no deseados tales como la formación de tumores. Por este motivo, estas estrategias de genoterapia integradora se limitan principalmente a estrategias *ex vivo*, donde las células se tratan fuera del organismo, se comprueba si la integración es satisfactoria y, a continuación, se
20 trasplantan a un sujeto.

Una característica común de muchas estrategias de genoterapia utilizadas anteriormente es que se introduce un transgén en las células junto con un promotor exógeno. Los promotores son necesarios para iniciar la transcripción y amplificación del ADN a ARN mensajero, o ARNm, que finalmente se traducirá en proteínas. La
25 expresión de altos niveles de proteínas terapéuticas a partir de un transgén mediante métodos de genoterapia anteriores requería promotores fuertes y diseñados. Aunque estos promotores eran un componente necesario de los anteriores sistemas de genoterapia para producir la expresión de proteínas, los estudios realizados en modelos animales han demostrado que la integración
30 inespecífica de vectores de genoterapia puede dar lugar a un aumento significativo

del desarrollo de tumores. La fuerza del promotor puede desempeñar un papel crucial en el aumento del desarrollo de estos tumores utilizando métodos de genoterapia anteriores. Por tanto, los intentos de dirigir altos niveles de expresión con promotores fuertes y exógenos pueden tener consecuencias nocivas a largo plazo.

A. Reescritura genómica

La reescritura genómica se utiliza normalmente para referirse a la delección, alteración o aumento de genes mediante la introducción de roturas en el ADN de las células utilizando mecanismos de reescritura genómica suministrados exógenamente. La eficacia de muchas de las estrategias actuales de reescritura genómica se ha visto limitada por las elevadas tasas de modificaciones inespecíficas no deseadas y por baja eficacia de la corrección génica, que en parte se debe a que la célula intenta reparar rápidamente las roturas de ADN introducidas. Muchas técnicas de reescritura genómica se centran en desactivar un gen disfuncional o en corregir u omitir mutaciones nocivas individuales dentro de un gen. Estas estrategias anteriores a menudo no tratan o previenen suficientemente las enfermedades genéticas debido a la cantidad de mutaciones posibles en un gen, en comparación con un método que comprende la inserción de un gen corrector completo.

A diferencia de las estrategias de genoterapia que recurren a la expresión episomal, la reescritura genómica permite que una región genética reparada se propague a nuevas generaciones de células a través de la división celular normal. Además, una proteína deseada puede expresarse utilizando la propia maquinaria reguladora de una célula. Las estrategias tradicionales de reescritura genómica se basan en nucleasas y emplean enzimas nucleasas para introducir una rotura mono- o bicatenaria en el ADN en un sitio de corte específico para provocar una delección, realizar una alteración o aplicar una secuencia correctora. Estos métodos están diseñados para proporcionar reescritura genómica (p. ej., delección, alteración, inserción de secuencia correctora, y similares) en la ubicación de un sitio de corte.

Una vez que las nucleasas han cortado el ADN, las técnicas tradicionales de reescritura genómica modifican el ADN utilizando dos vías: la reparación dirigida por homología (HDR) y la unión de extremos no homólogos (NHEJ). La HDR en estos métodos anteriores implica una incorporación muy exacta de secuencias de ADN complementarias a un sitio de integración, que se encuentra en la misma ubicación que el sitio de corte. La HDR tiene ventajas clave, ya que puede reparar el ADN con alta fidelidad y evita la introducción de mutaciones no deseadas en un sitio de integración. La NHEJ es un proceso menos selectivo y más propenso a errores que une rápidamente los extremos del ADN roto, lo que da lugar a una alta frecuencia de inserciones o deleciones en un sitio de corte.

1. Reescritura genómica tradicional basada en nucleasas

La reescritura genómica basada en nucleasas utiliza nucleasas, que son enzimas endógenas o diseñadas que producen roturas mono- o bicatenarias en el ADN en un sitio de corte. La reescritura genómica basada en nucleasas es un proceso de dos etapas. En primer lugar, una nucleasa exógena, que es capaz de cortar una o ambas cadenas de ADN bicatenario, se dirige a un sitio deseado (p. ej., mediante el uso de un ARN guía, restos específicos del sitio dentro de la nucleasa, y similares) y realiza un corte específico en un sitio de corte. A continuación, la maquinaria celular endógena de reparación del ADN se activa y finaliza el proceso de reescritura mediante NHEJ o, con menor frecuencia, mediante HDR.

La NHEJ puede producirse sin molde de ADN para que una célula copie mientras repara un corte de ADN. Aunque las tasas de NHEJ en comparación con otros mecanismos de reparación (p. ej., HDR) dependen de la especie, en ausencia de compuestos exógenos (p. ej., secuencia de molde de reparación, enzimas y similares) las células suelen emplear NHEJ preferentemente para reparar roturas bicatenarias. La reescritura genómica mediante NHEJ puede utilizarse para introducir pequeñas inserciones o deleciones, conocidas como indeles, que pueden dar lugar a una desactivación o reducción de la función génica. La NHEJ crea inserciones y deleciones en el ADN debido a su modo de reparación y también

puede dar lugar a la introducción de mutaciones inespecíficas, no deseadas, incluyendo aberraciones cromosómicas.

La HDR mediada por nucleasas se produce suministrando una nucleasa y un molde de ADN con complementariedad parcial o total con respecto a un sitio de integración diana. En métodos anteriores, el sitio de integración y el sitio de corte coincidían o se solapaban. Las células emplean el molde de ADN para construir ADN reparador, lo que da lugar a la inserción de una secuencia genética corregida y/o alternativa.

La reescritura genómica tradicional (también descrita en el presente documento como reescritura genómica convencional) ha empleado frecuentemente tres clases diferentes de nucleasas para las estrategias basadas en nucleasas: Nucleasas efectoras similares a activadores de la transcripción (TALEN), nucleasas asociadas a repeticiones palindrómicas cortas regularmente interespaciadas agrupadas (CRISPR/Cas) y nucleasas de dedos de zinc (ZFN).

2. Limitaciones de la reescritura genómica tradicional basada en nucleasas

Las anteriores estrategias de reescritura genómica basadas en nucleasas estaban limitadas al uso de enzimas nucleasas para producir el solapamiento de sitios de corte y de sitios de integración. Además, estos métodos tradicionales dependían de promotores exógenos para la expresión del transgén.

Las nucleasas pueden producir mutaciones específicas e inespecíficas. Una de las principales preocupaciones a la hora de implementar estrategias de reescritura genómica basadas en nucleasas, es la frecuencia relativamente alta de efectos inespecíficos. Las tecnologías de reescritura genómica anteriores podían provocar genotoxicidad, incluidas alteraciones cromosómicas, debido al proceso de NHEJ, propenso a errores, o a la interrupción de la expresión de genes naturales al dirigirse a los exones.

Plataforma tecnológica GENERIDE™

3. Características GENERIDE™

GENERIDE™ es una novedosa tecnología de reescritura genómica (*genome editing*) basada en AAV que inserta con precisión un transgén terapéutico en un genoma celular mediante recombinación homóloga. GENERIDE™ proporciona una expresión transgénica duradera independientemente de la proliferación celular y del crecimiento tisular, y las células corregidas con GENERIDE™ muestran una ampliación selectiva dentro de los tejidos afectados por una enfermedad (p. ej., ampliación selectiva de hepatocitos en un hígado de un sujeto enfermo). Sin querer estar limitado por ninguna teoría en particular, se contempla que GENERIDE™ es una tecnología de reescritura genómica que aprovecha la recombinación homóloga (HR, *homologous recombination*), que es un proceso de HDR (*homology-directed repair*, reparación dirigida por homología) natural que conserva la fidelidad del genoma. En algunas realizaciones, GENERIDE™ permite la inserción de transgenes en localizaciones genómicas diana específicas (p. ej., mediante HR) sin nucleasas exógenas. En algunas realizaciones, GENERIDE™ permite la inserción de transgenes en localizaciones genómicas diana específicas (p. ej., mediante HR) junto con una o más nucleasas exógenas. La integración de transgenes dirigida por GENERIDE™ está diseñada para aprovechar promotores endógenos en un sitio de integración diana para conducir altos niveles de expresión de genes específicos de tejido, sin producir problemas perjudiciales asociados al uso de promotores exógenos en la genoterapia tradicional.

4. Ventajas de la Tecnología GENERIDE™

La tecnología GENERIDE™ está diseñada para integrar con precisión genes correctores en el genoma de un paciente para proporcionar un efecto terapéutico estable y duradero. Entre otras cosas, GENERIDE™ se puede aplicar al direccionamiento de trastornos poco frecuentes (p. ej., trastornos del hígado, SNC, músculo, sangre y similares) en pacientes pediátricos donde es crítico proporcionar tratamiento temprano en la vida de un paciente antes de que pueda producirse una patología irreversible de la enfermedad.

En algunas realizaciones, la tecnología GENERIDE™ proporciona mejoras sobre determinadas limitaciones clave tanto de la genoterapia tradicional como de las estrategias de reescritura genómica convencionales de una manera que está bien posicionada para tratar enfermedades genéticas (p. ej., enfermedades genéticas en pacientes pediátricos). En algunas realizaciones, GENERIDE™ utiliza un vector de AAV para suministrar un transgén en el núcleo de una célula, seguido de la inserción del transgén en un sitio de integración diana dentro del genoma celular, donde la expresión del transgén está regulada por uno o más promotores endógenos. En algunas realizaciones, GENERIDE™ puede permitir la producción de proteínas de por vida, incluso cuando el organismo crece y cambia con el tiempo, lo cual no es factible con la genoterapia de AAV convencional.

Entre otras cosas, la tecnología GENERIDE™ también puede proporcionar mejoras sorprendentes e inesperadas cuando se combina con nucleasas diana. En algunas realizaciones, GENERIDE™ se combina con una o más nucleasas dirigidas a un sitio de corte particular en el genoma celular, en donde el sitio de corte es distal con respecto al sitio de integración diana (p. ej., el sitio de corte y el sitio de integración diana no son idénticos y/o no se solapan). Como se demuestra en el presente documento, el uso de dichas nucleasas diana junto con componentes GENERIDE™ adicionales puede producir tasas de integración y/o eficacia de genes transgénicos mejoradas (p. ej., niveles mejorados de ADN modificado, ARNm modificado, y/o expresión de proteínas). En particular, pueden diseñarse nucleasas diana (p. ej., en solitario o junto con una molécula de direccionamiento, incluyendo pero sin limitación a un ARN guía) para introducir una rotura monocatenaria o bicatenaria en un sitio de corte en una secuencia de ADN, en donde el sitio de corte es distal con respecto al sitio de integración diana. Aunque se ha demostrado que la introducción de un sitio de corte distal con respecto a un sitio diana para la integración de transgenes puede proporcionar mejoras en ratones (véase, Caneva *et al*, JCI Insight, 2019), se demostró que estas mejoras dependen en gran medida de la ubicación exacta de un sitio de corte. Por ejemplo, Caneva demostró que el corte dentro del intrón 14 flanqueante de la albúmina de ratón producía mejoras

notables en la eficacia de corte y HDR en comparación con los niveles limitados de integración cuando el sitio de corte se introducía en el intrón 13 de la albúmina de ratón (Figura S1 de Caneva). Aunque Caneva demostró que determinados sitios de corte de "puntos calientes" distales con respecto a un sitio de integración podrían aumentar posiblemente la eficacia de la HDR y la integración del transgén, no se demostró que estas mejoras condujeran a la integración específica de sitio y a la expresión de un transgén de interés. Además, Caneva limitó todos los experimentos a células de ratón que comprendían albúmina de ratón de tipo silvestre. Como comprendería un experto en la materia, se necesitarían pruebas e investigaciones considerables para determinar si un sistema de este tipo podría alterar su funcionalidad en una especie diferente. Por ejemplo, se sabe que el alto grado de variabilidad de secuencia en genes particulares (p. ej., albúmina) a través de especies tales como ratón y ser humano, no permiten a un experto utilizar los resultados de una especie para determinar si, y dónde, pueden encontrarse sitios de corte o integración apropiados en la otra especie. Además, en la técnica se sabe que los genes de ratón y de ser humano (p. ej., albúmina) pueden comprender diferencias significativas de secuencia, epigenéticas y estructurales que pueden alterar la eficacia del corte por nucleasas y/o la recombinación homóloga en sitios particulares.

La presente divulgación proporciona la sorprendente e inesperada idea de que la combinación de nucleasas diana con componentes GENERIDE™, en donde un sitio de corte de nucleasa es distal con respecto a un sitio de integración diana (p. ej., el sitio de corte y el sitio de integración no son idénticos y/o no se solapan), puede proporcionar mejoras inesperadas en la tasa de integración y/o eficacia de expresión de transgenes en células y tejidos humanos. En algunas realizaciones, dichas mejoras pueden observarse en múltiples sitios de corte, p. ej., en diferentes secuencias no codificantes en una secuencia de ADN humano de interés (p. ej., tanto el intrón 13 como el 14 de la albúmina humana).

Las estrategias modulares desveladas en el presente documento pueden aplicarse para permitir que GENERIDE™ proporcione una fuerte expresión génica

específica del tejido, que sea reproducible en todas las diferentes opciones terapéuticas suministradas a un tejido (p. ej., hígado, músculo, SNC, sangre).

En el documento WO 2013/158309, incorporado al presente documento por referencia, se describen trabajos anteriores sobre el direccionamiento génico sin alteración. En el documento WO 2015/143177, incorporado al presente documento por referencia, se describen trabajos anteriores sobre la reescritura genómica sin nucleasas. En el documento WO 2020/032986, incorporado al presente documento por referencia, se describen trabajos anteriores sobre genoterapia sin alteración para el tratamiento de la MMA (acidemia metilmalónica). En el documento WO/2020/214582, incorporado al presente documento por referencia, se describen trabajos anteriores sobre la supervisión de la genoterapia.

B. Componentes de GENERIDE™

Vehículos de suministro

En la técnica se conocen múltiples estrategias de genoterapia. Como tal, existen múltiples mecanismos de suministro conocidos en la técnica. En algunas realizaciones, se proporciona un transgén utilizando un vehículo de suministro. En algunas realizaciones, las composiciones de la presente divulgación comprenden un vehículo de suministro. En algunas realizaciones, un vehículo de suministro es o comprende una partícula vírica (p. ej., un vector vírico). En algunas realizaciones, un vehículo de suministro es o comprende una partícula no vírica y/o una carga útil no particulada (p. ej., una nucleasa). En algunas realizaciones, un vehículo de suministro es una partícula lipídica (p. ej., una nanopartícula lipídica). En la técnica se conocen diversas nanopartículas lipídicas para el suministro de ácidos nucleicos, por ejemplo, en las descritas en los documentos WO2015184256; WO2013149140; WO2014089486A1; WO2009127060; WO2011071860; WO2020219941 incorporando el contenido de cada uno de ellos en el presente documento por referencia.

En algunas realizaciones, un vehículo de suministro es o comprende un exosoma. Un experto en la materia reconocerá varios métodos de producción y uso

de exosomas. En Luan et al., *Acta Pharmacologica Sinica* volumen 38, páginas 754-763 (2017), se describen ejemplos de dichos métodos y usos.

En algunas realizaciones, puede suministrarse una combinación de una o más cargas útiles diferentes con uno o más sistemas de suministro descritos en el presente documento (p. ej., vector vírico, nanopartícula lipídica, y similares). En algunas realizaciones, un sistema de suministro puede suministrar una carga útil (p. ej., una o más nucleasas (p. ej., proteínas Cas, endonucleasas, TALEN, ZFN)) junto con una segunda carga útil distinta (p. ej., una secuencia de polinucleótidos que comprenda un transgén). En algunas realizaciones, un primer sistema de suministro (p. ej., una nanopartícula lipídica (LNP) puede suministrar una carga útil (p. ej., una o más nucleasas (p. ej., proteínas Cas, endonucleasas, TALEN, ZFN))) junto con un segundo sistema de suministro distinto (p. ej., un vector vírico tal como un vector de AAV recombinante), que suministre una segunda carga útil distinta (p. ej., una secuencia de polinucleótidos que comprenda un transgén).

Vectores víricos

En algunas realizaciones, un vehículo de suministro es o comprende un vector vírico. Los vectores víricos comprenden virus o material cromosómico vírico, dentro del cual puede insertarse una secuencia de ácido nucleico heterólogo para su transferencia a una secuencia diana de interés (p. ej., para su transferencia al ADN genómico dentro de una célula). Como vectores víricos, pueden utilizarse diversos virus incluyendo, p. ej., virus de ADN monocatenario (ADNmc), virus de ADN bicatenario (ADNbc) y/o virus de ARN con una fase de ADN en su ciclo de vida. En algunas realizaciones, un vector vírico es o comprende un dependoparvovirus (anteriormente virus adenoasociado, AAV, *Adeno-Associated Virus*) o una variante de AAV.

En algunas realizaciones, una partícula vectorial es una única unidad de virus que comprende una cápside que encapsida un polinucleótido basado en virus (p. ej., un genoma vírico de tipo silvestre o un vector vírico recombinante). En algunas realizaciones, una partícula vectorial es o comprende una partícula vectorial de AAV. En algunas realizaciones, una partícula vectorial de AAV se refiere a una

partícula vectorial compuesta por al menos una proteína de cápside de AAV y un vector de AAV encapsidado. En algunas realizaciones, una partícula vectorial (también denominada vector vírico) comprende al menos una proteína de cápside de AAV y un vector de AAV encapsidado, en donde el vector comprende además una o más secuencias de polinucleótidos heterólogas.

Proteínas de cápside

En algunas realizaciones, una construcción de expresión comprende secuencias de polinucleótidos que codifican proteínas de cápside de uno o más subtipos de AAV, incluyendo AAV naturales y recombinantes. En algunas realizaciones, una construcción de expresión comprende secuencias de polinucleótidos que codifican proteínas de cápside de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAVC11.01, AAVC11.02, AAVC11.03, AAVC11.04, AAVC11.05, AAVC11.06, AAVC11.07, AAVC11.08, AAVC11.09, AAVC11.10, AAVC11.11 (denominadas indistintamente en el presente documento sL65), AAVC11.12, AAVC11.13, AAVC11.14, AAVC11.15, AAVC11.16, AAVC11.17, AAVC11.18, AAVC11.19, AAV-DJ, AAV-LK03, AAV-LK19, AAVrh.74, AAVrh.10, AAVhu.37, AAVrh.K, AAVrh.39, AAV12, AAV 13, AAVrh.8, AAV aviar, AAV bovino, AAV canino, AAV equino, AAV de primate, AAV no de primate, AAV ovino, un AAV híbrido (p. ej., un AAV que comprende una o más secuencias de un subtipo de AAV y una o más secuencias de un segundo subtipo), y/o un AAV que comprende una proteína de cápside de AAV mutante o una cápside de AAV quimérica (p. ej., una cápside con secuencias de polinucleótidos derivadas de dos o más serotipos diferentes de AAV), o variantes de los mismos.

En algunas realizaciones, los vectores víricos se empaquetan dentro de proteínas de cápside (p. ej., proteínas de cápside de uno o más subtipos de AAV). En algunas realizaciones, las proteínas de cápside proporcionan un aumento o una mejora de la transducción de células (p. ej., células humanas o murinas) en relación con una proteína de cápside de referencia. En algunas realizaciones, las proteínas de cápside proporcionan un aumento o una mejora de la transducción de determinadas células o tipos tisulares (p. ej., tropismo hepático, tropismo muscular,

tropismo del SNC, tropismo pulmonar) en relación con una proteína de cápside de referencia. En algunas realizaciones, las proteínas de cápside aumentan o mejoran la transducción de células o tejidos (p. ej., hígado, músculo y/o SNC) en al menos aproximadamente 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, o más en relación con una proteína de cápside de referencia. En algunas realizaciones, las proteínas de cápside aumentan o mejoran la transducción de células o tejidos (p. ej., hígado, músculo, pulmón y/o SNC) en al menos aproximadamente 1.2x, 1.5x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 11x, 12x, 13x, 14x, 15x, 16x, 17x, 18x, 19x, 20x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x o más en relación con una proteína de cápside de referencia.

En algunas realizaciones, una secuencia que codifica una proteína de cápside puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una proteína de cápside de tipo silvestre correspondiente. En algunas realizaciones, una secuencia que codifica una proteína de cápside puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o 100 % con una secuencia correspondiente que codifica una proteína de cápside diseñada (p. ej., una proteína de cápside quimérica, una proteína de cápside con optimización de codones y similares). En algunas realizaciones, una secuencia que codifica una proteína de cápside puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o 100 % con una secuencia ilustrativa de la Tabla 1 siguiente.

Tabla 1: Secuencias ilustrativas que codifican la(s) proteína(s) de la cápside

Descripción	Secuencia	SEQ ID NO
AAV LK03 (nucleótido)	atggctgctgacgggtatctccagattggctcgaggacaaccttctgaag gcattcgagagtgggtggcgctgcaacctggagcccctaaaccaaggc aaatcaacaacatcaggacaacgctcgggggtcttgcttccgggtacaa	1

5	atacctcggacccggcaacggactcgacaagggggaacccgtcaacg cagcggacgcggcagccctcgagcacgacaaggcctacgaccagcag ctcaaggccggtgacaacccctacctcaagtacaaccacgccgacgcc gagttccaggagcggctcaaagaagatacgtctttgggggcaacctcgg gcgagcagtcttcaggccaaaaagaggcttctgaacctcttggtctggt gaggaagcggctaagacggctcctggaaagaagaggcctgtagatcag tctcctcaggaaccggactcatctggtgttggaatcgggcaaacag cctgccagaaaaagactaaatttcggtcagactggcgactcagagtcagt cccagaccctcaacctctcggagaaccaccagcagccccacaaagtttg ggatctaataacaatggcttcaggcgggtggcgaccaatggcagacaata acgaggggtccgatggagtggttaattcctcaggaaattggcattgagatt cccaatggctgggacagagtcaccaccagcaccagaacctggg ccctgccacttacaacaacctctctacaagcaaatctccagccaatca ggagcttcaaacgacaaccactactttggctacagcacccttgggggtat tttgactttaacagattccactgccacttctcaccacgtgactggcagcgact cattaacaacaactggggattccggcccaagaaactcagcttcaagctctt caacatccaagttaaagaggtcacgcagaacgatggcacgacgactatt gccaataaccttaccagcacggttcaagtgtttacggactcggagtatcag ctcccgtacgtgctcgggtcggcgaccaaggctgtctccgcccgtttcca gcgagcgtcttcatggctccctcagtatggatacctcaccctgaacaacgga agtcaagcgggtgggacgctcatcctttactgcctggagtacttcccttcga gatgctaaggactggaaataaactccaattcagctataccttcgaggatgta ccttttcacagcagctacgctcacagccagagttggatcgcttgatgaatc ctcttattgatcagtatctgtactacctgaacagaacgcaaggaacaacctc tgaacaaccaaccaatcacggctgcttttagccaggctgggcctcagtc tatgtctttgcaggccagaaattggctacctgggcccgtgctaccgggaaca gagactttcaaagactgctaacgacaacaacaacagtaactttccttgga agcggccagcaaatatcatctcaatggccgcgactcgctggtgaatccag gaccagctatggccagtcacaaggacgatgaagaaaaattttccctatg cacggcaatctaattttggcaaagaagggacaacggcaagtaacgca	
10		
15		
20		
25		
30		

5		gaattagataatgtaatgattacggatgaagaagagattcgtaccaccaat cctgtggcaacagagcagtatggaactgtggcaaataacttcagagctc aaatacagctcccacgactagaactgtcaatgatcagggggccttacctg gcatggtgtggcaagatcgtgacgtgtacctcaaggacctatctgggcaa agattcctcacacggatggacactttcatccttctcctctgatgggaggcttg gactgaaacatccgcctcctcaaatacatgatcaaaaaatactccggtaccg gcaaatactccgacgactttcagcccggccaagtttgcttcatttatcactca gtactccactggacaggtcagcgtggaaattgagtgggagctacagaaa gaaaacagcaaacgttggaatccagagattcagtacactccaactaca acaagtctgttaatgtggactttactgtagacactaatgggtgttatagtgaac ctcgccccattggcaccggttaccttaccggtcccctgtaa	
15	AAV-LK03 (proteína)	MAADGYLPDWLEDNLSEGIREWVALQPGAPKPKAN QQHQDNARGLVLPGYKYLGPNGLDKGEPVNAADA AALEHDKAYDQQLKAGDNPYLYKNHADADEFQERLKE DTSFGGNLGRAVFQAKRLLLEPLGLVEEAAKTAPGK KRPVDQSPQEPDSSSGVGKSGKQPARKRLNFGQTG DSESVDPDQPLGEPPAAPTSLGSNTMASGGGAPMA DNNEGADGVGNSSGNWHCDSQWLGDREVITTSTRT WALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGY FDFNRFHCHFSPRDWQRLINNNWGFRPKKLSFKLFN IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL GSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGR SSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYA HSQSLDRLMNPLIDQYLYLNRTQGTTSGTTNQSRL FSQAGPQMSLQARNWLPGPCYRQQRLSKTANDN NNSNFPWTAASKYHLNGRDSLVPNGPAMASHKDDE EKFFPMHGNIIFGKEGTTASNAELDNVMITDEEEIRTT NPVATEQYGTVANLQSSNTAPTTRTVNDQGALPG	2

5		MVWQDRDVYLQGPIWAKIPHTDGHFHPSPMLMGGFG LKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTG QVSVEIEWELQKENSkrWNPEIQYTSNynKSVNVDF TVDTNGVYSEPRPIGTRYLTRPL	
10	AAV8 (nucleótido)	atggctgccgatgggtatctccagattggctcgaggacaacctctctgagg gcattcgcgagtgggtgggcgctgaaacctggagccccgaagcccaaag ccaaccagcaaaagcaggacgacggccgggtctagtgttctctggcta caagtacctcgaccctcaacggactcgacaagggggagcccgtaa cgcgccggacgcagcgccctcgagcacgacaaggctacgaccagc agctgcaggcgggtgacaatccgtacctgcggtataaccacgccgacgc cgagttcaggagcgtctgcaagaagatacgtctttggggcaacctcgg gcgagcagtcctccaggccaagaagcgggtctcgaacctctcggtctggt tgaggaaggcgctaagacggctcctggaaagaagagaccggtagagc catcaccagcggttctccagactcctctacgggcatcggcaagaaaggc caacagcccgccagaaaaagactcaatttggtcagactggcgactcag agtcagttccagaccctcaacctctcggagaacctccagcagcgccctct ggtgtgggacctaatacaatggctgcaggcgggtggcgaccaatggcag acaataacgaaggcgccgacggagtggttagttcctcgggaaattggca 20 ttgcgattccacatggctgggcgacagagtcaccaccagcaccgga acctgggccctgccacctacaacaaccacctctacaagcaaatctcca acgggacatcgggaggagccaccaacgacaacacctacttcggctaca gcacccctgggggtatttgactttaacagattccactgccactttcaccac gtgactggcagcgactcatcaacaacaactggggattccggcccaagag 25 actcagcttcaagctcttcaacatccaggtaaggaggtcacgcagaatg aaggcaccaagaccatcgccaataacctcaccagcaccatccagggtgtt acggactcggagtaccagctgccgtacgttctcggtctgccaccaggg ctgcctgcctccgttccggcgacgtgttcattgattccccagtacggctac ctaactcaacaacggtagtcaggccgtgggacgctcctccttactgc	3

5		ctggaataactttccttcgcagatgctgagaaccggcaacaacttccagttta cttacaccttcgaggacgtgcctttccacagcagctacgcccacagccag agcttggaaccggctgatgaatcctctgattgaccagtacctgtactacttgtct cggactcaaacaacaggaggcacggcaaatacgcagactctgggcttc agccaaggtgggcctaatacaatggccaatcaggcaaagaactggctg ccaggaccctgttaccgccaacaacgcgtctcaacgacaaccggggcaa aacaacaatagcaactttgcctggactgctgggaccaaataccatctgaat ggaagaaattcattggctaatacctggcatcgctatggcaacacacaaaga cgacgaggagcgttttttcccagtaacgggatcctgattttggcaaaciaa 10 aatgtgccagagacaatgcggattacagcgatgtcatgtcaccagcga ggaagaaatcaaaaccactaaccctgtggctacagaggaatacgggtatc gtggcagataacttgcagcagcaaaacacggctcctcaaattggaactgt caacagccagggggccttaccgggtatggtctggcagaaccggggacgtg 15 tacctgcaggggtcccatctgggccaagattcctcacacggacggcaacttc caccgtctccgtgatgggaggctttggcctgaaacatcctccgcctcag atcctgatcaagaacacgcctgtacctgcggatcctccgaccacctcaac cagtaaaagctgaactctttcatcacgcaatacagcaccggacaggtcag cgtggaaattgaatgggagctgcagaaggaaaacagcaagcgctgga 20 accccgagatccagtacacctccaactactacaaatctacaagtgtggac tttgctgtaatacagaaggcgtgtactctgaaccccgccccattggcaccc gttacctcaccgtaatctgtaa	
25	AAV8 (proteína)	MAADGYLPDWLEDNLSEGIREWALKPGAPKPKAN QQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA AALEHDKAYDQQLQAGDNPYLRYNHADADEFQERLQ EDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPG KKRPVEPSPQRSPDSSTGIGKKGQQPARKRLNFGQT GDSESVDPDQPLGEPPAAPSGVGPNTMAAGGGAPM ADNNEGADGVGSSSGNWHCDSTWLGDRVITTSTRT	4

5		WALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPW GYFDFNRHFHCHFSPRDWQRLINNNWGFRPKRLSFK LFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPY VLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVG RSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSY AHSQSLDRLMNPLIDQYLYLSRTQTTGGTANTQTL GFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQ NNNSNFAWTAGTKYHLNGRNSLANPGIAMATHKDD EERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIK TTNPVATEEYGIVADNLQQQNTAPQIGTVNSQGALP GMVWQNRDVYLQGPIWAKIPHTDGNFHPSPLMGGF GLKHPPPQILIKNTPVPADPPTTFNQSKLNSFITQYST GQVSVEIEWELQKENSkrWNPEIQYTSNYYKSTSVD FAVNTEGVYSEPRPIGTRYLTRNL	
15	AAV-DJ (nucleótido)	atggctgccgatgggtatctccagattggctcgaggacactctctctgaagg aataagacagtgggtgaagctcaaacctggcccaccaccaccaaagcc cgcagagcggcataaggacgacagcaggggtcttgcttctcgggtac aagtacctcgaccctcaacggactcgacaagggagagccgggtcaac gaggcagacgccgcccctcgagcacgacaaagcctacgaccggca gctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgc cgagttccaggagcggctcaaagaagatacgtctttgggggcaacctcg ggcgagcagtcctccaggccaaaaagaggcttcttgaaaccttggctcgtg tgaggaagcggctaagacggctcctggaaagaagaggcctgtagagca ctctcctgtggagccagactcctcctcggaaccggaaaggcgggcccag cagcctgcaagaaaaagattgaatttggtcagactggagacgcagactc agtcccagaccctcaaccaatcgagaaacctcccgagccccctcaggt gtgggatctcttacaatggctgcaggcgggtggcgaccaatggcagaca ataacgagggcgccgacggagtgggtaattcctcggaattggcattgc gattccacatggatgggcgacagagtcaccaccagcaccggaacct	5

5	gggccctgccacctaacaaccacctctacaagcaaattccaacag cacatctggaggatcttcaaatgacaacgcctacttcgggtacagcacccc ctgggggtatTTtgactttaacagattccactgccacttttcaccacgtgactg gcagcgactcatcaacaacaactggggattccggcccaagagactcag cttcaagctcttcaacatccagggtcaaggagggtcacgcagaatgaaggca ccaagaccatcgccaataacctcaccagcaccatccagggtttacggac tcggagtaccagctgccgtacgttctcggctctgccaccagggtgctgctg cctccgttcccgcgaggacgtgtcatgattccccagtagggctacctaacac tcaacaacggtagtcaggccgtgggacgctcctccttactgcctggaat actttccttcgcagatgctgagaaccggcaacaacttcagttacttacac cttcgaggacgtgcctttccacagcagctacgcccacagccagagcttgg accggctgatgaatcctctgattgaccagtacctgtactactgtctcggact caaacaacaggaggcacgacaaaatacgcagactctgggcttcagccaa ggtgggcctaatacaatggccaatcaggcaaagaactgggtgccagga ccctgttaccgccagcagcgagtatcaaagacatctgcggataacaaca acagtgaatactcgtggactggagctaccaagtaccacctaattggcaga gactctctggtgaatccgggcccggccatggcaagccacaaggacgatg aagaaaagTTTTtctcagagcgggggttctcatctttggaagcaaggctc agagaaaacaaatgtggacattgaaaagggtcatgattacagacgaaga ggaaatcaggacaaccaatcccgtggctacggagcagtaggttctgtat ctaccaacctccagagaggcaacagacaagcagctaccgcagatgtca acacacaaggcgttcttcaggcatggtctggcaggacagagatgtgtac cttcagggggcccatctgggcaaagattccacacacggacggacatttca cccctctcccctcatgggtggattcggacttaaacacctccgcctcagatc ctgatcaagaacacgcctgtacctgcggatcctccgaccaccttcaacca gtcaaagctgaactctttcatcaccagatttctactggccaagtcagcgtg gagatcgagtgggagctgcagaaggaaaacagcaagcgttgaaccc cgagatccagtagacctccaactactacaaatctacaagtgtggacttgc gtttaatacagaaggcgtgtactctgaaccccgccccattggcaccggttac ctcacccgtaattctgtaa	
10		
15		
20		
25		
30		

AAV-DJ (proteína)	MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPA ERHKDDSRGLVLPGYKYLGPFNGLDKGEPVNEADAA ALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKED TSFGGNLGRAVFQAKKRILLEPLGLVEEAAKTAPGKK RPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGD ADSVPDPQPIGEPPAAPSGVGSLTMAAGGGAPMAD NNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTW ALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGY FDFNRFHCHFSPRDWQRLINNNWGFRPKRLSFKLFN IQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLG SAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSS FYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAH QSLDRLMNPLIDQYLYLSRTQTTGGTTNTQTLGFSQ GGPNTMANQAKNWLPGPCYRQQRVSKTSADNNNS EYSWTGATKYHLNGRDSLVPNGPAMASHKDDEEKF FPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPV ATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVW QDRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHP PPQILIKNTPVPADPPTTFNQSKLNSFITQYSTGQVSV EIEWELQKENS KRWNPEIQYTSNYYKSTSVDFAVNT EGVYSEPRPIGTRYLTRNL	6
sL65 (nucleótido)	atggctgccgatgggtatctccagattggctcgaggacactctctgaagg cattcgcgagtggtggcgctgaaacctggagctccacaaccaaggcc aaccaacagcatcaggacaacggcaggggtctgtgcttctgggtacaa gtacctcggaccctcaacggactcgacaagggagagccggtcaacga ggcagacgccgcggccctcgagcacgacaaggcctacgacaagcagc tcgagcagggggacaacccgtacctcaagtacaaccacgccgacgccg agtttcaggagcgtctgcaagaagatacgtctttgggggcaacctcgggc	7

5	gagcagtctccaggccaagaagcgggtctcgaacctctcggtctggtg aggaaggcgctaagacggctcctggaaagaagagaccggtagagccg tcacctcagcgttccccgactcctccacgggcatcggcaagaaaggcc agcagcccgcagaaagagactcaatttcggtcagactggcgactcaga gtcagtccccgacctcaacctctcggagaacctccagcagcgccctcta gtgtgggatctggtacagtggctgcaggcgggtggcgaccaaattggcaga caataacgaaggtgccgacggagtgggtaatgcctcaggaaattggcatt gcgattccacatggctgggacagagtcattaccaccagcaccggaac ctgggcccctgcccacctacaacaaccacctctacaagcaaatctccagcc aatcaggagcttcaaacgacaaccactactttggctacagcacccttg gggtattttgactttaacagattccactgccacttctaccacgtgactggca gagactcattaacaacaactggggattccggccaagagactcaactca agctcttcaacatccaagtcaaggaggtcacgacgaatgatggcgtcacg accatcgctaataaccttaccagcacgggtcaagtcttctcggactcggagt accagttgccgtacgtcctcggctctgcgcaccagggtgcctccctccgtt cccggcggacgtgttcattgattccccagtcaggctacctaactcaacaa cggtagtcaggccgtgggacgctcctcctttactgcctggaatatttccat cgcagatgctgagaacgggcaataactttgagttcagctacagcttcgag gacgtgcctttccacagcagctacgcacacagccagagcttgaccgact gatgaatcctctcattgaccagctacgtgtactacttatccagaactcagtcca caggaggaactcaaggtagccagcaattgttattttctcaagctgggcctgc aaacatgtcggctcaggccaagaactggctgcctggacctgtaccggc agcagcgagtctccacgacactgtcgaaaacaacaacagcaactttgc ttggactgggtgccaccaaatacacctgaacggcagaaactcgttgggttaa tcccggtcgtcgccatggcaactcacaaggacgacgaggaccgcttttcc catccagcggagtcctgattttggaaaaactggagcaactaacaaaact acattggaaaatgtgttaatgacaaatgaagaagaaattcgtcctactaat cctgtagccacggaagaatacgggatagtcagcagcaacttacaagcg gctaatactgcagcccagacacaagttgtcaacaaccaggggagccttac ctggcatggtctggcagaaccgggacgtgtacctgcagggtcccatttggg	
---	---	--

5		ccaaaattcctcacacagatggacactttcacccgtctcctcttatgggcgg ctttggactcaagaacccgcctcctcagatcctcatcaaaaacacgcctgtt cctgcgaatcctccggcggagtttcagctacaaagtttgcttcattcatcac ccagtattccacaggacaagtgagcgtggagattgaatgggagctgcag aaagaaaacagcaaacgctggaatcccgaagtgcagtatacatctaact atgcaaaatctgccaacgttgattcactgtggacaacaatggactttatact gagcctcgccccattggcacccgttaccttaccgcgtcccctgtaa	
10	sL65 (proteína)	MAADGYLPDWLEDLTLSEGIREWALKPGAPQPKAN QQHQDNNGRGLVLPGYKYLGPFNGLDKGEPVNEADA AALEHDKAYDKQLEQGDNPYLKYNHADADEFQERLQE DTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGK KRPVEPSPQRSPTSSTGIGKKGQQPARKRLNFGQT GDSESVDPDPQPLGEPPAAPSSVSGSTVAAGGGAPM ADNNEGADGVGNASGNWHCDSTWLGDREVITTSTRT WALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGY FDFNRHFHCHFSRPDWQRLINNNWGFRPKRLNFKLFN IQVKEVTTNDGVTTIANNLSTVQVFSDSEYQLPYVL GSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRS SFYCLEYFPSQMLRTGNNEFEFSYSFEDVPFHSSYAH SQSLDRLMNPLIDQYLYLSRTQSTGGTQGTQQLLF SQAGPANMSAQAKNWLPGPCYRQQRVSTTLSQNN NSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDDED RFFPSSGVLFIFGKTGATNKTTLENVLMTNEEEIRPTNP VATEEYGIVSSNLQAANTAAQTQVVNNQGALPGMV WQNRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLK NPPPQILIKNTPVPANPPAEFSATKFASFITQYSTGQV SVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTV DNNGLYTEPRPIGTRYLTRPL	8

5	AAV-NP59 (nucleótido)	atggctgccgatgggtatcttccagattggctcgaggacactctctgaagg aataagacagtgggtgaagctcaaaccctggcccaccaccaccaaagcc cgcagagcggcataaggacgacagcaggggtcttgtgcttctgggtac aagtacctcggacccttcaacggactcgacaagggagagccgggtcaac gaggcagacgccgcgccctcgagcacgacaaagcctacgaccggca gctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgc ggagtttcaggagcgccttaaagaagatacgtctttgggggcaacctcg acgagcagtcttccaggcgaaaaagaggggttctgaacctctggcctgg ttgaggaacctgttaagacggctccgggaaaaaagaggccggtagagc actctctgtggagccagactcctcctcggaaccggcaagacaggcca gcagcccgtataaaaagagactcaattttggtcagactggcgactcagagt cagtcccagaccctcaacctctcggagaaccaccagcagccccctctggt ctgggaactaatac gatggctacaggcagtgggcgaccaatggcagac aataacgaggggtccgatggagtgggtaattcctcaggaaattggcattgc gattcccaatggctgggacagagtcacaccaccagcaccggaacct gggccctgccacctaacaaccatctctacaagcaaatctccagcca atcaggagcttcaaacgacaaccactactttggctacagcacccttggg ggatctttgactttaacagattccactgccactctcaccacgtgactggcag cgactcattaacaacaactggggattccggccaagaaactcagcttaa gctcttaacatccaagttaaagaggtcacgcagaaacgatggcacgacg actattgccaaataaccttaccagcacgggtcaagtgttactgactcggagta ccagctcccgtacgtcctcggctcggcgcatcaaggatgcctcccgcggtt cccagcagacgtcttcatggtgccacagtatggatacctcaccctgaaca acgggagtcaggcagtaggacgctcttcattttactgctggagtagcttctt ctcagatgctgcgtaccggaaacaactttacctcagctacacttttgagga cgttcctttccacagcagctacgctcacagccagagtctggaccgtctcatg aatcctctcatcgaccagtacctgtattacttgagcagaacaaactcca agtggaaccaccacgcagtcagggtcagttttctcaggccggagcgag tgacattcgggaccagtctaggaactggcttctggaccctgttaccgcca gcagcgagtatcaaagacatctcggataacaacaacagtgaatactcg	9
10			
15			
20			
25			
30			

5		tgggctggagctaccaagtaccacctcaatggcagagactctctggtgaat ccgggcccggccatggcaagccacaaggacgatgaagaaaagtttttc ctgagagcgggggtctcatctttgggaagcaaggctcagagaaaacaaat gtggacattgaaaaggatcatgattacagacgaagaggaaatcaggaca accaatcccgtggctacggagcagtatggttctgtatctaccaacctccag agaggcaacagacaagcagctaccgcagatgtcgacacacaaggcgtt cttccaggcatggtatggcaggacagagatgtgtaccttcagggacctc tgggcaaagattccacacacggacggacattttcacccctctcccctcatg ggtggattcggacttaaacaccctcctccacagattctcatcaagaacacc ccggtacctgcgaatccttcgaccaccttcagtgcggcaaagtttgcttcctt catcacacagtactcaacgggacaggtcagcgtggagatcgagtggga gctgcagaaggaaaacagcaaacgctggaatcccgaaattcagtacac ttccaactacaacaagtctgttaatgtggactttactgtggacactaatggcg tgtattcagagcctcgccccattggcaccagataacctgactcgtaatctgta a	
10			
15			
20	AAV-NP59 (proteína)	MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPA ERHKDDSRGLVLPGYKYLGPFNGLDKGEPVNEADAA ALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKED TSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKK RPVEHSPVEPDSSSGTGKTGQQPAKKRLNFGQTGD SESVPDPQPLGEPPAAPSGLTNTMATGSGAPMAD NNEGADGVGNSSGNWHCDSQWLGDREVITTSTRTW ALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFD FNRFHCHFSPRDWQRLINNNWGFRPKKLSFKLFNIQ VKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGS AHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSS FYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHS QSLDRLMNPLIDQYLYLSRTNTPSGTTTQSRLQFSQ	10
25			
30			

5		AGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSE YSWAGATKYHLNGRDSLVPNPAMASHKDDEEKFF PQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVA TEQYGSVSTNLQRGNRQAATADVDTQGVLPGMVW QDRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHP PPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSV EIEWELQKENSkrWNPEIQYTSNynKSVNVDFTVDT NGVYSEPRPIGTRYLTRNL	
10	AAV2 (nucleótido)	atggctgccgatgggtatcttcagattggctcgaggacactctctgaagg aataagacagtgggtgaagctcaaacctggcccaccaccaccaaagcc cgcagagcggcataaggacgacagcaggggtcttgcttctgggtac aagtacctcggaccctcaacggactcgacaagggagagccggtcaac gaggcagacgccgcgccctcgagcacgacaaagcctacgaccggca gctcgacagcggagacaacccgtacctcaagtacaaccacgccgacgc 15 ggagtttcaggagcgccttaaagaagatacgtctttgggggcaacctcgg acgagcagtcttcaggcgaaaaagaggggtcttgaaacctctgggcctgg ttgaggaacctgttaagacggctccgggaaaaaagaggccggtagagc actctctgtggagccagactcctcctcggaaccggaaaggcgggcca gcagcctgcaagaaaaagattgaattttgtcagactggagacgcagact 20 cagtacctgacccccagcctctcgacagccaccagcagccccctctggt ctgggaactaatac gatggctacaggcagtgggcgaccaatggcagac aataacgagggcgccgacggagtggttaattcctcggaattggcattg cgattccacatggatgggacagagtcacaccaccagcaccgaacc tgggccctgcccacctacaacaaccacctctacaaacaaatttcagcca 25 atcaggagcctcgaacgacaatcactactttggctacagcacccttggg ggatatttgactcaacagattccactgccactttcaccacgtgactggcaa agactcatcaacaacaactggggattccgacccaagagactcaactca agctcttaacattcaagtcaaagaggtcacgcagaatgacggtagcagc 30 acgattgccaaataaccttaccagcacgggtcaggtgtttactgactcggagt	11

5

10

15

20

25

30

	accagctcccgtagctcctcggctcggcgcatcaaggatgcctcccgccg tcccagcagacgtcttcatgggtgccacagtatggatacctcaccctgaac aacgggagtcaggcagtaggacgctcttcattttactgcctggagtactttcc ttctcagatgctgctaccggaacaactttaccttcagctacacttttgagg acgttcctttccacagcagctacgctcacagccagagtctggaccgtctcat gaatcctctcatcgaccagtacctgtattacttgagcagaacaaacactcc aagtggaaccaccacgcagtcaggcttcagttttctcaggccggagcgga gtgacattcgggaccagtctaggaactggcttcttgaccctgttaccgcc agcagcgagtatcaaagacatctgcggataacaacaacagtgaatactc gtggactggagctaccaagtaccacctaattggcagagactctctggtga atccgggcccggccatggcaagccacaaggacgatgaagaaaagtttt tctcagagcggggttctcatctttggaagcaaggctcagagaaaacaa atgtggacattgaaaaggctatgattacagacgaagaggaaatcaggac aaccaatcccgtggctacggagcagtatgggtctgtatctaccaacctcca gagaggcaacagacaagcagctaccgcagatgtcaacacacaaggcg ttcttcaggcatgggtctggcaggacagagatgtgtaccttcaggggccc atctgggcaaagattccacacacggacggacattttacccctctcccctcat gggtggattcggacttaaacaccctcctccacagattctcatcaagaacac cccggtacctgcgaatcctcgaccaccttcagtgcggcaaagttgcttct tcatcacacagtactccacgggacaggtcagcgtggagatcgagtggga gctgcagaaggaaaacagcaaacgctggaatcccgaaattcagtacac ttccaactacaacaagtctgttaatgtggactttactgtggacactaatggcg tgtattcagagcctcgccccattggcaccagatacctgactcgtaatctgta a	
AAV2 (proteína)	MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPA ERHKDDSRGLVLPGYKYLGPFNGLDKGEPVNEADAA ALEHDKAYDRQLDSDNPYLKYNHADA EFQERLKED TSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTPGKK RPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGD	12

5		ADSVPDPQPLGQPPAAPSGGLGTNTMATGSGAPMAD NNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTW ALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFD FNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQ VKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGS AHQGCCLPPFPADVFMVPQYGYLTLNNGSQAVGRSS FYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHS QSLDRLMNPLIDQYLYLSRTNTPSGTTTQSRLQFSQ AGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSE YSWTGATKYHLNGRDSLVPNPGPAMASHKDDEEKFF PQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVA TEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVW QDRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHP PPQILIKNTPVPANPSTTFSAKFASFITQYSTGQVSV EIEWELQKENSKRWNPEIQYTSNYNKSVMVDFTVDT NGVYSEPRPIGTRYLTRNL	
10			
15			
20	AAV5 (nucleótido)	ATGTCTTTTGTTGATCACCTCCAGATTGGTTGGAA GAAGTTGGTGAAGGTCTTCGCGAGTTTTTGGGCCT TGAAGCGGGCCCCACCGAAACCAAAACCCAATCAG CAGCATCAAGATCAAGCCCGTGGTCTTGTGCTGCC TGGTTATAACTATCTCGGACCCGGAACGGTCTCG ATCGAGGAGAGCCTGTCAACAGGGCAGACGAGGT CGCGCGAGAGCACGACATCTCGTACAACGAGCAG CTTGAGGCGGGAGACAACCCCTACCTCAAGTACAA CCACGCGGACGCCGAGTTTCAGGAGAAGCTCGCC GACGACACATCCTTCGGGGGAAACCTCGGAAAGG CAGTCTTTCAGGCCAAGAAAAGGGTTCTCGAACCT TTTGGCCTGGTTGAAGAGGGTGCTAAGACGGCCC CTACCGGAAAGCGGATAGACGACCACTTTCCAAAA	13
25			
30			

5	AGAAAGAAGGCTCGGACCGAAGAGGACTCCAAGC CTTCCACCTCGTCAGACGCCGAAGCTGGACCCAG CGGATCCCAGCAGCTGCAAATCCCAGCCCAACCA GCCTCAAGTTTGGGAGCTGATACAATGTCTGCGGG AGGTGGCGGCCCATTTGGGCGACAATAACCAAGGT GCCGATGGAGTGGGCAATGCCTCGGGAGATTGGC ATTGCGATTCCACGTGGATGGGGGACAGAGTCGT CACCAAGTCCACCCGAACCTGGGTGCTGCCCAGC TACAACAACCACCAGTACCGAGAGATCAAAAGCGG	
10	CTCCGTCGACGGAAGCAACGCCAACGCCTACTTT GGATACAGCACCCCCTGGGGGTACTTTGACTTTAA CCGCTTCCACAGCCACTGGAGCCCCCGAGACTGG CAAAGACTCATCAACAACCTACTGGGGCTTCAGACC CCGGTCCCTCAGAGTCAAAATCTTCAACATTCAAG TCAAAGAGGTCACGGTGCAGGACTCCACCACCAC CATCGCCAACAACCTCACCTCCACCGTCCAAGTGT TTACGGACGACGACTACCAGCTGCCCTACGTCGTC GGCAACGGGACCGAGGGATGCCTGCCGGCCTTCC CTCCGCAGGTCTTTACGCTGCCGCAGTACGGTTAC	
15	GCGACGCTGAACCGCGACAACACAGAAAATCCCA CCGAGAGGAGCAGCTTCTTCTGCCTAGAGTACTTT CCCAGCAAGATGCTGAGAACGGGCAACAACCTTTGA GTTTACCTACAACCTTTGAGGAGGTGCCCTTCCACT CCAGCTTCGCTCCCAGTCAGAACCTGTTCAAGCTG GCCAACCCGCTGGTGGACCAGTACTTGTACCGCTT CGTGAGCACAAATAAACTGGCGGAGTCCAGTTCA ACAAGAACCTGGCCGGGAGATACGCCAACACCTA CAAAAACCTGGTTCCCGGGGGCCCATGGGCCGAACC CAGGGCTGGAACCTGGGCTCCGGGGTCAACCGC	
20	GCCAGTGTGAGCGCCTTCGCCACGACCAATAGGA	
25		
30		

5		TGGAGCTCGAGGGCGCGAGTTACCAGGTGCCCCC GCAGCCGAACGGCATGACCAACAACCTCCAGGGC AGCAACACCTATGCCCTGGAGAACACTATGATCTT CAACAGCCAGCCGGCGAACCCGGGCACCACCGC CACGTACCTCGAGGGCAACATGCTCATCACCAGC GAGAGCGAGACGCAGCCGGTGAACCGCGTGGCG TACAACGTCGGCGGGCAGATGGCCACCAACAACC AGAGCTCCACCACTGCCCCCGCGACCGGCACGTA CAACCTCCAGGAAATCGTGCCCGGCAGCGTGTGG 10 ATGGAGAGGGACGTGTACCTCCAAGGACCCATCT GGGCAAGATCCCAGAGACGGGGGCGCACTTTCA CCCCTCTCCGGCCATGGGCGGATTCGGACTCAA CACCCACCGCCCATGATGCTCATCAAGAACACGCC TGTGCCCCGAAATATCACCAGCTTCTCGGACGTGC 15 CCGTCAGCAGCTTCATCACCCAGTACAGCACCGG GCAGGTCACCGTGGAGATGGAGTGGGAGCTCAAG AAGGAAAACCTCCAAGAGGTGGAACCCAGAGATCC AGTACACAAACAACTACAACGACCCCCAGTTTGTG GACTTTGCCCCGGACAGCACCGGGGAATACAGAA 20 CCACCAGACCTATCGGAACCCGATACCTTACCCGA CCCCTTTAA	
25	AAV5 (proteína)	MSFVDHPPDWLEEVGEGLEAGPPKPKPNQ QHQQDQARGLVLPGYNYLGPNGNLDREGPVNRADEV AREHDISYNEQLEAGDNPYLKYNHADAQEFQEKLADD TSFGGNLKGAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQL QIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGN ASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQY REIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHW	14
30			

5		SPRDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQD STTTIANNLSTVQVFTDDDYQLPYVVGNGTEGCLPA FPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFP SKMLRTGNNFETYNFEEVPPFHSSFAPSQNLFLAN PLVDQYLYRFVSTNNTGGVQFNKNLAGRYANTYKN WFPGPMGRTQGWNLGSGVNRASVSAFATTNRMEL EGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQ PANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQ MATNNQSSTTAPATGTYNLQEIVPGSVWMERDVYLQ GPIWAKIPETGAHFHPSPAMGGFGLKHPPPMMLIKN TPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKK ENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTR PIGTRYLTRPL	
15	AAV6 (nucleótido)	atggctgccgatgggtatctccagattggctcgaggacaacctctctgagg gcattcgcgagtgggtgggacttgaaacctggagccccgaaacccaaagc caaccagcaaaagcaggacgacggccgggtctggtgcttcttggttac aagtacctcgagaccttcaacggactcgacaagggggagcccgtaac gcggcggtatgcagcgccctcgagcacgacaaggcctacgaccagca gctcaaagcgggtgacaatccgtacctgcggtataaccacgccgacgcc gagtttcaggagcgtctgcaagaagatacgtctttgggggcaacctcggg cgagcagtcctccaggccaagaagaggggtctcgaaccttttggtctggtg aggaagggtgctaagacggctcctggaaagaaacgtccggtagagcagt cgccacaagagccagactcctcctcgggcattggcaagacaggccagc agcccgctaaaaagagactcaattttggtcagactggcgactcagagtca gtccccgaccacaacctctcggagaacctccagcaacccccgctgctgt gggacctactacaatggcttcaggcgggtggcgaccaatggcagacaat aacgaaggcgccgacggagtgggtaatgcctcaggaaattggcattgcg attccacatggctgggcgacagagtcaccaccagcaccgaacatg ggccttgccacacctataacaaccacctctacaagcaaattctccagtgttc	15
20			
25			
30			

5

10

15

20

25

aacgggggcccagcaacgacaaccactacttcgggtacagcaccctg
 ggggtatgttgattcaacagattccactgccatttctaccacgtgactggca
 gcgactcatcaacaacaattggggattccggcccaagagactcaactca
 agctcttcaacatccaagtcaaggaggtcacgacgaatgatggcgtcacg
 accatcgctaataacctaccagcacgggtcaagtcttctcgactcggagt
 accagttgccgtacgtcctcggtctgcgaccagggctgcctccctcgtt
 cccggcggacgtgttcatgattccgcagtacgggtacctaacgctcaaca
 atggcagccaggcagtgggacgggtcatcctttactgcctggaatattccc
 atcgcagatgctgagaacgggcaataactttaccttcagctacacctcga
 ggacgtgcctttccacagcagctacgcgacagccagagcctggaccgg
 ctgatgaatcctctcatcgaccagtacctgtattacctgaacagaactcaga
 atcagtcgggaagtgccaaaacaaggacttgctgttagccgggggtctc
 cagctggcatgtctgttcagccaaaaactggctacctggaccctgttacc
 ggcagcagcgcgtttctaaaacaaaaacagacaacaacagcaact
 ttacctggactggtgcttcaaaatataaccttaatgggcgtgaatctataatc
 aacctggcactgctatggcctcacacaaagacgacgaagacaagttctt
 tccatgagcgggtgtcatgattttggaaaggagagcgccggagctcaaa
 cactgcattggacaatgtcatgatcacagacgaagaggaaatcaaagcc
 actaaccctggccaccgaaagatttgggactgtggcagtcaatctcca
 gagcagcagcacagaccctgcgaccggagatgtgcatgttatgggagcc
 ttacctggaatggtgtggcaagacagagacgtataacctgcagggctctattt
 gggccaaaattcctcacacggatggacactttcaccgtctcctctcatggg
 cggctttggacttaagcaccgcctcctcagatcctcatcaaaaacacgcc
 tgttctgcgaatcctccggcagagtttccggctacaaagttgttcttcatc
 acccagtattccacaggacaagtgagcgtggagattgaatgggagctgc
 agaaagaaaacagcaaacgctggaatcccgaagtgcagtatacatcta
 actatgcaaaatctgccaacgttgatttactgtggacaacaatggactttat
 actgagcctcgccccattggcaccggtacctcaccgtccctgtaa

30

AAV6 (proteína)	MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKAN QQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQ EDTSFGGNLGRAVFQAKKRVLEPFGLVEEGAKTAPG KKRPEQSPQEPDSSSGIGKTGQQPAKKRLNFGQT GDSESVDPDPQLGEPPATPAAVGPTTMASGGGAPM ADNNEGADGVGNASGNWHCDSTWLGDREVITTSTRT WALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGY FDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFN IQVKEVTTNDGVTTIANNNLTSTVQVFSDSEYQLPYVL GSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRS SFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAH SQSLDRLMNPLIDQYLYLNRTQNNQSGSAQNKDLLF SRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN NSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEDK FFPMSGVMIFGKESAGASNTALDNVMITDEEEIKATN PVATERFGTVAVNLQSSSTD PATGDVHVMGALPGM VWQDRDVYLQGGPIWAKIPHTDGHFHPSPLMGGFGL KHPPPQILIKNTPVPANPPAEFSATKFASFITQYSTGQ VSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFT VDNNGLYTEPRPIGTRYLTRPL	16
AAV9 (nucleótido)	atggctgccgatgggtatcttccagattggctcgaggacaaccttagtgaag gaattcgcgagtggtgggcttgaaacctggagcccctcaacccaaggca aatcaacaacatcaagacaacgctcgaggctctgtgctccgggttataaaa taccttggaacccggcaacggactcgacaagggggagccggtcaacgca gcagacgcggcgccctcgagcacgacaaggcctacgaccagcagct caaggccggagacaacccgtacctcaagtacaaccacgccgacgccc agttccaggagcggctcaaagaagatacgtctttgggggcaacctcggg cgagcagctctccaggccaaaaagaggcttctgaacctcttggtctggttg	17

5	aggaagcggctaagacggctcctggaaagaagaggcctgtagagcagt ctctcaggaaccggactcctccgcggtattggcaaatcgggtgcacag cccgctaaaaagagactcaatttcgggtcagactggcgacacagagtcagt cccagaccctcaaccaatcggagaacctcccgagccccctcaggtgtg ggatctcttacaatggcttcaggtggtggcgcaccagtggtgagacaataa cgaagggtgccgatggagtggttagttcctcgggaaattggcattgcgattc ccaatggctgggggacagagtcaccaccagcaccgaacctgggc cctgcccacctacaacaatcacctctacaagcaaattccaacagcacat ctggaggatcttcaatgacaacgcctacttcggctacagcacccttg gggatatttgactcaacagattccactgccactctcaccacgtgactggca gcgactcatcaacaacaactggggattccggcctaagcgactcaactca agctcttcaacattcaggtcaaagaggttacggacaacaatggagtcaag accatcgccaataaccttaccagcacgggtccaggtcttcacggactcaga ctatcagctcccgtagctgctcgggtcgggtcacgagggtgcctcccgcc gtcccagcggacgttttcatgattcctcagtagcgggtatctgacgcttaatga tggagccaggccgtgggtcgttcgtcctttactgcctggaatatttccgctc gcaaatgctaagaacgggtaacaacttcagttcagctacgagtttgaga acgtacctttcatagcagctacgctcacagccaaagcctggaccgacta atgaatccactcatcgaccaatactgtactatctctcaaagactattaacgg ttctggacagaatcaacaaacgctaaaattcagtggtggccggaccagca acatggctgtccagggaagaaactacatacctggaccagctaccgaca acaacgtgtctcaaccactgtgactcaaaacaacaacagcgaatttgcttg gcctggagcttcttgggctctcaatggacgtaatagcttgatgaatcctg gacctgctatggccagccacaaagaaggagaggaccgtttcttcttctg tggatctttaatttttgcaacaaggaactggaagagacaacgtggatgc ggacaaagtcataataaccaacgaagaagaaattaaaactactaacc ggtagcaacggagtcctatggacaagtggccacaaaccaccagagtg ccaagcacaggcgagaccggctgggttcaaaaccaaggaatacttcc gggatggtttggcaggacagagatgtgtacctgcaaggaccatttgggc caaaattcctcacacggacggcaactttacccttctccgctgatgggagg	
10		
15		
20		
25		
30		

5		gtttgaatgaagcaccgcctcctcagatcctcatcaaaaacacacctgt acctgcggatcctccaacggccttcaacaaggacaagctgaactcttcat caccagatttctactggccaagtcagcgtggagatcgagtgggagctgc agaaggaaaacagcaagcgctggaacccggagatccagtacacttcca actattacaagtctaataatgttgaattgctgttaatactgaagggtatatag tgaaccccgccccattggcaccagatacctgactcgtaatctgtaa	
10	AAV9 (proteína)	MAADGYLPDWLEDNLSEGIREWALKPGAPQPKAN QQHQDNARGLVLPGYKYLPGNGLDKGEPVNAADA AALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKE DTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGK KRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTG DTESVPDPQPIGEPPAAPSGVGS LT MASGGGAPVAD NNEGADGVGSSSGNWHCDSQWLGD RVITTSTRTW ALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGY FDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFN IQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVL GSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGN NFQFSYEFENVPFHSSYAH SQSLDRLMNPLIDQYLYLSKTINGSGQNQQTLKFSV AGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSE FAWPGASSWALNGRNSLMNPGPAMASHKEGEDRF FPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNP VATESYGQVATNHQSAQAQAQTGWVQNQGILPGMV WQDRDVYLGQPIWAKIPHTDGNFHPSPLMGGFGMK HPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQV SVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAV NTEGVYSEPRPIGTRYLTRNL	18
25	AAV-rh74 (nucleótido)	atggctgccgatgggtatctccagattggctcgaggacaacctctctgagg gcattcgcgagtgggtgggacctgaaacctggagccccgaaacccaaag	19

5	ccaaccagcaaaagcaggacaacggccggggtctggtgcttctggcta caagtacctcggacccttcaacggactcgacaagggggagcccgtcaa cgcggcggacgcagcgccctcgagcacgacaaggcctacgaccagc agctccaagcgggtgacaatccgtacctgcggtataatcacgccgacgc cgagtttcaggagcgtctgcaagaagatacgtctttgggggcaacctcgg gcgcgagctctccaggccaaaaagcgggttctgaacctctgggcctgg ttgaatcgccggttaagacggctcctggaaagaagagaccggtagagcc atcaccccagcgctctccagactcctctacgggcatcggaagaaaggc cagcagcccgcaaaaaagagactcaattttgggcagactggcgactcag 10 agtcagtccccgaccctcaaccaatcggaagaaccaccagcaggcccct ctggtctgggatctggtacaatggctgcaggcggtaggctccaatggca gacaataacgaaggcgccgacggagtggttagttcctcaggaaattggc attgcgattccacatggctgggcgacagagtcaccaccagcaccgcg acctgggccctgcccacctacaacaaccacctctacaagcaaatctcca 15 acgggacctcgggaggaagcaccaacgacaacacctacttcggctaca gcacccctgggggtattttgacttcaacagattccactgccactttcacca cgtgactggcagcgactcatcaacaacaactggggattccggcccaaga ggctcaacttcaagctcttcaacatccaagtcaaggaggtcacgcagaat gaaggcaccaagaccatcgccaataaccttaccagcacgattcaggtctt 20 tacggactcggaataccagctcccgtagctgctcggctcggcgccaccagg gctgcctgcctccgttcccgcgagcttctcatgattcctcagtagcggtac ctgactctgaacaatggcagtcaggctgtgggcccgtcgtccttctactgcc tggagtactttccttctcaaatgctgagaacgggcaacaactttgaattcag ctacaacttcgaggacgtgcccttcacagcagctacgcgcacagccag 25 agcctggaccggctgatgaaccctctcatcgaccagtacttgactacctgt cccgactcaaagcacgggcggtactgcaggaactcagcagttgctatttt ctcaggccgggcctaacaacatgtcggctcaggccaagaactggctacc cggctccctgctaccggcagcaacgcgtctccacgacactgtcgagaac aacaacagcaactttgcctggacgggtgccaccaagtatcatctgaatgg 30 cagagactctctggtgaatcctggcggttgccatggctaccacaaggacg	
---	--	--

5		acgaagagcgatttttccatccagcggagtcttaatgtttgggaaacaggg agctggaaaagacaacgtggactatagcagcgtgatgctaaccagcga ggaagaaataaagaccaccaacccagtgccacagaacagtacggcg tggaggccgataacctgcaacagcaaaacgccgctcctatttaggggcc gtcaatagtcaaggagccttacctggcatggtgtggcagaaccgggacgt gtacctgcagggtcccatctgggccaagattcctcatacggacggcaactt tcatccctcgccgctgatgggaggctttggactgaagcatccgcctcctcag atcctgattaaaaacacacctgttcccgcggatcctccgaccaccttcaatc aggccaagctggcttcttcatcacgcagtacagtaccggccagggtcagc gtggagatcgagtgggagctgcagaaggagaacagcaaacgctggaa cccagagattcagtacactccaactactacaaatctacaaatgtggacttt gctgtcaatactgagggtacttattccgagcctcgccccattggcaccggt acctcaccgtaatctgtaa	
15	AAV-rh74 (proteína)	MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKAN QKQDNGRGLVLPGYKYLGPFNGLDKGEPVNAADA AALEHDKAYDQQLQAGDNPYLRYNHADADEFQERLQ EDTSFGGNLGRAVFQAKKRVLEPLGLVESPVKTAPG KKRPEPSPQRSPDSSTGIGKKGQQPAKKRLNFGQT GDSESVDPDQPIGEPPAGPSGLGSGTMAAGGGAPM ADNNEGADGVGSSSGNWHCDSTWLGDRVITTSTRT WALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPW GYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFK LFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPY VLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVG RSSFYCLEYFPSQMLRTGNNFEFSYNFEDVPFHSSY AHSQSLDRLMNPLIDQYLYLSRTQSTGGTAGTQQL LFSQAGPNNMSAQAKNWLPGPCYRQQRVSTTLSQN NNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDE ERFFPSSGVLMFGKQGAGKDNVDYSSVMLTSEEEIK	20

5

	TTNPVATEQYGVVADNLQQQNAAPIVGAVNSQGALP GMVWQNRDVYLQGPIWAKIPHTDGNFHPSPLMGGF GLKHPPPQILIKNTPVPADPPTTFNQAKLASFITQYST GQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTNVD FAVNTEGTYSEPRPIGTRYLTRNL	
--	--	--

Estructura y función del AAV

10

15

20

El dependoparvovirus (AAV) es un parvovirus compuesto por una cápside de proteína icosaédrica y un genoma de ADN monocatenario. La cápside vírica del AAV comprende tres subunidades, VP1, VP2 y VP3, y dos regiones de repeticiones terminales invertidas (ITR, *inverted terminal repeat*), que se encuentran en los extremos de la secuencia genómica. Las ITR actúan como orígenes de replicación y desempeñan un papel en el empaquetamiento vírico. El genoma vírico también comprende los genes *rep* y *cap*, asociados a la replicación y al empaquetamiento de la cápside, respectivamente. En la mayoría de los AAV de tipo silvestre, el gen *rep* codifica cuatro proteínas necesarias para la replicación vírica, Rep 78, Rep68, Rep52 y Rep40. El gen *cap* codifica las subunidades de la cápside, así como la proteína activadora del ensamblaje (AAP), que promueve el ensamblaje de las partículas víricas. Los AAV son generalmente deficientes en replicación, requiriendo la presencia de un virus auxiliar o la funciones de virus auxiliares (p. ej., virus del herpes simple (HSV) y/o adenovirus (AdV)) para replicarse dentro de una célula infectada. Por ejemplo, en algunas realizaciones, los AAV requieren genes de ARN adenovírico E1A, E2A, E4 y VA para replicarse dentro de una célula hospedadora.

AAV recombinante

25

30

En general, los vectores de AAV recombinantes (rAAV) pueden comprender muchos de los mismos elementos que se encuentran en los AAV de tipo silvestre, incluyendo secuencias y estructuras de cápside similares, así como secuencias de polinucleótidos que no son de origen AAV (p. ej., un polinucleótido heterólogo al AAV). En algunas realizaciones, los rAAV sustituirán las secuencias de AAV originarias, de tipo silvestre, por secuencias de polinucleótidos que codifican una carga útil. Por ejemplo, en algunas realizaciones, un rAAV comprenderá secuencias

de polinucleótidos que codifican uno o más genes destinados a fines terapéuticos (p. ej., para genoterapia). Los rAAV pueden modificarse para eliminar una o más secuencias codificantes víricas de tipo silvestre. Por ejemplo, los rAAV pueden diseñarse para que comprendan sólo una ITR, y/o uno o más genes necesarios para el empaquetamiento (p. ej., *genes rep* y *cap*) que los que se encontrarían en un AAV de tipo silvestre. La expresión génica con los rAAV suele limitarse a uno o más genes que totalizan 5 kb o menos, ya que las secuencias de mayor tamaño no se empaquetan eficazmente dentro de la cápside vírica. En algunas realizaciones, se pueden utilizar dos o más rAAV para proporcionar partes de una carga útil más grande, por ejemplo, para proporcionar una secuencia codificante completa para un gen que normalmente sería demasiado grande para caber en un solo AAV.

Entre otras cosas, la presente divulgación proporciona vectores víricos que comprenden uno o más polipéptidos descritos en el presente documento. En algunas realizaciones, los rAAV pueden comprender una o más proteínas de cápside (p. ej., una o más proteínas de cápside de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAVC11.01, AAVC11.02, AAVC11.03, AAVC11.04, AAVC11.05, AAVC11.06, AAVC11.07, AAVC11.08, AAVC11.09, AAVC11.10, AAVC11.11 (denominadas indistintamente en el presente documento como sL65), AAVC11.12, AAVC11.13, AAVC11.14, AAVC11.15, AAVC11.16, AAVC11.17, AAVC11.18, AAVC11.19, AAV-DJ, AAV-LK03, AAV-LK19, AAVrh.74, AAVrh.10, AAVhu.37, AAVrh.K, AAVrh.39, AAV12, AAV 13, AAVrh.8,), AAV aviar, AAV bovino, AAV canino, AAV equino, AAV primate, AAV no primate, AAV ovino, un AAV híbrido (p. ej., un AAV que comprenda una o más secuencias de un subtipo de AAV y una o más secuencias de un segundo subtipo), y/o un AAV que comprenda una proteína de cápside de AAV mutante o una cápside de AAV quimérica (p. ej., una cápside con secuencias de polinucleótidos procedentes de dos o más serotipos diferentes de AAV). En algunas realizaciones, los rAAV pueden comprender una o más secuencias de polinucleótidos que codifican un gen o ácido nucleico de interés (p. ej., un gen para el tratamiento de una enfermedad/trastorno genético y/o una secuencia de ácido nucleico inhibidora).

En algunas realizaciones, un vector de AAV recombinante puede comprender al menos una ITR. En algunas realizaciones, un vector de AAV recombinante comprende dos ITR. En algunas realizaciones, un vector de AAV recombinante comprende una ITR en 5'. En algunas realizaciones, un vector de AAV recombinante comprende una ITR en 3'. En algunas realizaciones, un vector de AAV recombinante comprende una ITR de AAV2. En algunas realizaciones, un vector de AAV recombinante comprende una parte de una ITR de AAV2. En algunas realizaciones, un vector de AAV recombinante comprende una ITR que tiene una identidad de secuencia de al menos 80 %, 85 %, 90 %, 95 %, 99 % o 100 % con una ITR de AAV2. En algunas realizaciones, un vector de AAV recombinante comprende una ITR que tiene una identidad de secuencia del 90 %, 95 %, 99 % o 100 % con una de las secuencias ilustrativas de la Tabla 2 siguiente.

Tabla 2: Secuencias de ITR ilustrativas

Descripción de la ITR	Secuencia (5' a 3')	Tamaño (pb)	SEQ ID NO
ITR de tipo silvestre, orientación flip	AGGAACCCCTAGTGATGGAGTTGG CCACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGCCCCGGGCA AAGCCCGGGCGTCGGGCGACCTT TGGTCGCCCCGGCCTCAGTGAGCG AGCGAGCGCGCAGAGAGGGAGTG GCCAA	145	21
ITR de tipo silvestre, orientación flop	AGGAACCCCTAGTGATGGAGTTGG CCACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGGGCGACCA AAGGTCGCCCCGACGCCCGGGCTT TGCCCGGGCGGCCTCAGTGAGCG	145	22

		AGCGAGCGCGCAGAGAGGGAGTG GCCAA		
5	ITR con bucle B delecionado	AGGAACCCCTAGTGATGGAGTTGG CCTACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGCCCCGGGCT TTGCCCCGGGCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGT GGCCAA	123	23
10	ITR con bucle C delecionado	AGGAACCCCTAGTGATGGAGTTGG CCTACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGGGCGACCA AAGGTCGCCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGT GGCCAA	123	24
15	ITR con región A delecionada	AGGAACCCCTAGTGATGGAGTTGG CCTACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGGGCGACCA AAGGTCGCCCCGACGCCCGGGCTT TGCCCCGGGCGGCCTCAGTGAGCG AGCGAGCGCGCAG	130	25
20				
25	ITR con región D delecionada	CTGCGCGCTCGCTCGCTCACTGAG GCCGCCCGGGCAAAGCCCCGGGCG TCGGGCGACCTTTGGTCGCCCGG CCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTG	105	26

Los vectores de VAA pueden ser capaces de replicarse en una célula hospedadora infectada (replicación competente) o incapaces de replicarse en una célula hospedadora infectada (replicación incompetente). Un AAV con replicación competente (*rcAAV*, *replication competent AAV*) requiere la presencia de uno o más genes de empaquetamiento de AAV funcionales. Los vectores de AAV

recombinantes se diseñan generalmente para que sean incompetentes en cuanto a su replicación en células de mamífero, con el fin de reducir la posibilidad de que se generen rcAAV por recombinación con secuencias que codifican genes de empaquetamiento de AAV. En algunas realizaciones, las preparaciones de vectores de rAAV descritas en el presente documento están diseñadas para que comprendan pocos vectores de rcAAV, o ninguno. En algunas realizaciones, las preparaciones de vectores de rAAV comprenden menos de aproximadamente 1 rcAAV por 10^2 vectores de rAAV. En algunas realizaciones, las preparaciones de vectores de rAAV comprenden menos de aproximadamente 1 rcAAV por 10^4 vectores rAAV. En algunas realizaciones, las preparaciones de vectores de rAAV comprenden menos de aproximadamente 1 rcAAV por 10^8 vectores rAAV. En algunas realizaciones, las preparaciones de vectores de rAAV comprenden menos de aproximadamente 1 rcAAV por 10^{12} vectores de rAAV. En algunas realizaciones, las preparaciones de vectores rAAV no comprenden vectores de rcAAV.

Nanopartículas lipídicas (LNP)

Las nanopartículas lipídicas (LNP) son sistemas de suministro que, entre otras cosas, pueden lograr el suministro intracelular de ácidos nucleicos en forma intacta, lo que permite cambios biológicos que incluyen efectos terapéuticos. En un aspecto, las nanopartículas lipídicas son composiciones lipídicas que comprenden al menos un lípido. En algunas realizaciones, las nanopartículas lipídicas pueden comprender además al menos un ácido nucleico (p. ej., ADN, ARN y similares). En algunas realizaciones, las nanopartículas lipídicas pueden comprender un ácido nucleico terapéutico (p. ej., ADN, ARN y similares) encapsulado en una parte lipídica de la nanopartícula.

La presente divulgación proporciona composiciones que comprenden nanopartículas lipídicas. En algunas realizaciones, las nanopartículas lipídicas comprenden uno o más componentes. En algunas realizaciones, las nanopartículas lipídicas comprenden uno o más componentes tales como lípidos ionizables, esteroides, lípidos conjugados con enlazadores y fosfolípidos. En algunas realizaciones, las nanopartículas lipídicas pueden comprender uno o más

compuestos descritos en el presente documento. Entre otras cosas, la presente divulgación describe que la selección y combinación de uno o más de los componentes descritos en el presente documento puede influir en las características de las nanopartículas lipídicas, como el diámetro, el pKa, la estabilización y la capacidad de ionización.

Entre otras cosas, la presente divulgación describe que la selección y combinación de uno o más de los componentes descritos en el presente documento influye en la actividad funcional de las nanopartículas lipídicas como el tropismo, la estabilización y la eficacia de suministro. Por ejemplo, la presente divulgación describe que una combinación de componentes puede adecuarse mejor al suministro de una carga útil particular (p. ej., ADN (p. ej., ADN plasmídico, ADN linealizado (p. ej., ADN linealizado y dimerizado (CELID, *close-ended linear duplex*, dúplex lineal cerrado))) en comparación con el de una referencia apropiada (p. ej., secuencias alternativas de ácido nucleico, proteínas, y similares).

En algunas realizaciones, las nanopartículas lipídicas pueden ser o comprender una estructura producida por un método disponible (Véase, Jayaraman *et al.*, *Angewandte Chemie (International Ed. In English)*, 2012; Li *et al.* *PLOS One*, 2013; Kulkarni *et al.* *Nanomedicine: Nanotechnology, Biology, and Medicine*, 2017; US 2019/0240345; WO 2019/089828; WO/2021/102411; WO/2019/046809; WO/2018089540, cada uno de los cuales se incorpora en el presente documento por referencia en su totalidad). A la luz de las enseñanzas expuestas en el presente documento, los expertos en la materia serían conscientes de la posibilidad de sustituir las nanopartículas lipídicas descritas en este documento por otras alternativas disponibles.

Lípidos ionizables

En algunas realizaciones, las nanopartículas lipídicas pueden comprender uno o más lípidos ionizables, como se describe en el presente documento. En algunas realizaciones, un lípido ionizable puede incluir un grupo que contiene amina en el grupo de cabeza. En algunas realizaciones, un lípido ionizable es o comprende un compuesto descrito en el presente documento (p. ej., un compuesto de Fórmula

II, Fórmula III o Fórmula IV). En algunas realizaciones, un lípido ionizable está presente en una preparación de nanopartículas lipídicas (LNP) de aproximadamente 30 por ciento molar a aproximadamente 70 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica. En algunas realizaciones, el lípido ionizable está presente de aproximadamente 33 por ciento molar a aproximadamente 60 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica. En algunas realizaciones, el lípido ionizable está presente de aproximadamente 34 por ciento molar a aproximadamente 55 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica. En algunas realizaciones, el lípido ionizable está presente de aproximadamente 33 por ciento molar a aproximadamente 51 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica. En algunas realizaciones, el lípido ionizable está presente a aproximadamente 34.7 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica. En algunas realizaciones, el lípido ionizable está presente a aproximadamente 50 por ciento molar, basado en moles totales de componentes de la nanopartícula lipídica.

Esteroles

En algunas realizaciones, las nanopartículas lipídicas comprenden uno o más esteroides como se describe en el presente documento. En algunas realizaciones, un esteroide es un colesterol, o una variante o derivado del mismo. En algunas realizaciones, se modifica un colesterol. En algunas realizaciones, un colesterol es un colesterol oxidado. En algunas realizaciones, el colesterol es colesterol esterificado. Las enzimas pueden actuar sobre el colesterol no modificado para formar variantes con cadena lateral oxidada o anillo oxidado. En algunas realizaciones, un colesterol puede oxidarse en la estructura de anillo beta o en la estructura de cola de hidrocarburo. En algunas realizaciones, un esteroide es un fitosteroide. Como esteroides ilustrativos que se consideran para su uso en las nanopartículas lipídicas desveladas se incluyen, pero sin limitación, 25-hidroxicolesterol (25-OH), 20 α -hidroxicolesterol (20 α -OH), 27-hidroxicolesterol, 6-ceto-5 α -hidroxicolesterol, 7-cetocolesterol, 7 β -hidroxicolesterol, 7 α -

hidroxicolesterol, 7 β -25-dihidroxicoolesterol, beta-sitosterol, estigmasterol, brasicasterol, campesterol, o combinaciones de los mismos. En algunas realizaciones, un colesterol con cadena lateral oxidada puede mejorar el suministro de carga en relación con otras variantes de colesterol. En algunas realizaciones, un

5

colesterol es un colesterol no modificado.

En algunas realizaciones, la composición de LNP comprende de aproximadamente 20 por ciento molar a aproximadamente 50 por ciento molar de esterol. En algunas realizaciones, la composición de LNP comprende aproximadamente 38 por ciento molar de esterol. En algunas realizaciones, la

10

Lípidos conjugados con enlazadores

En algunas realizaciones, las nanopartículas lipídicas comprenden uno o más

lípidos conjugados con enlazadores, como se describe en el presente documento. En algunas realizaciones, un lípido conjugado con enlazador es o comprende un lípido con polietilenglicol (PEG) o un lípido modificado con PEG. En algunas realizaciones, los lípidos con PEG o modificados con PEG, pueden denominarse

15

20

25

30

PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPE, PEG-DSG o un lípido PEG-DSPE.

En algunas realizaciones, el lípido conjugado con enlazador comprende un lípido de polietilenglicol. En algunas realizaciones, el lípido conjugado con enlazador comprende dimiristoil glicerol (DMG), 1,2-dipalmitoil-rac-glicerol, metoxipolietilén glicol (DPG-PEG) o 1,2-distearoil-rac-glicero-3-metilpolioxietileno (DSG - PEG). En algunas realizaciones, el lípido conjugado con enlazador tiene una masa molecular promedio de aproximadamente 500 Da a aproximadamente 5000 Da. En algunas realizaciones, el lípido conjugado con enlazador tiene una masa molecular promedio de aproximadamente 2000 Da. En algunas realizaciones, la composición de LNP comprende de aproximadamente 0 por ciento molar a aproximadamente 5 por ciento molar de lípido conjugado con enlazador. En algunas realizaciones, la composición de LNP comprende aproximadamente 1.5 por ciento molar de lípido conjugado con enlazador. En algunas realizaciones, la composición de LNP comprende aproximadamente 3 por ciento molar de lípido conjugado con enlazador.

Fosfolípidos

En algunas realizaciones, las nanopartículas lipídicas comprenden uno o más fosfolípidos, como se describe en el presente documento. En algunas realizaciones, uno o más fosfolípidos pueden ensamblarse en una o más bicapas lipídicas. En algunas realizaciones, uno o más fosfolípidos pueden incluir una fracción fosfolipídica. En algunas realizaciones, uno o más fosfolípidos pueden incluir una o más fracciones fosfolipídica. En algunas realizaciones, uno o más fosfolípidos pueden incluir una fracción fosfolipídica y una o más fracciones de ácido graso. En algunas realizaciones, un fosfolípido es o comprende un compuesto descrito en el presente documento (p. ej., un compuesto de Fórmula I). En algunas realizaciones, una fracción fosfolipídica incluye, pero sin limitación, fosfatidilcolina, fosfatidiletanolamina, fosfatidilglicerol, fosfatidilserina, ácido fosfatídico, 2-lisofosfatidilcolina y esfingomieline. En algunas realizaciones, una fracción de ácido graso incluye, pero sin limitación, ácido láurico, ácido mirístico, ácido miristoleico, ácido palmítico, ácido palmitoleico, ácido esteárico, ácido oleico, ácido linoleico,

ácido alfa-linolénico, ácido erúcico, ácido fitánico, ácido araquídico, ácido araquidónico, ácido eicosapentaenoico, ácido behénico, ácido docosapentaenoico y ácido docosahexaenoico. También se contemplan especies no naturales incluyendo especies naturales con modificaciones y sustituciones incluyendo ramificación, oxidación, ciclación y alquinos. Por ejemplo, un fosfolípido puede funcionalizarse con uno o más alquinos, o reticularse con ellos, (p. ej., un grupo alquenilo en el que uno o más dobles enlaces se sustituye(n) por un triple enlace). En condiciones de reacción apropiadas, un grupo alquino puede experimentar una cicloadición catalizada por cobre al exponerse a una azida. Dichas reacciones pueden ser útiles en la funcionalización de una bicapa lipídica de una composición de nanopartícula para facilitar la infiltración en la membrana o el reconocimiento celular o en la conjugación de una composición de nanopartículas a un componente útil tal como una fracción de direccionamiento o de obtención de imágenes (p. ej., un colorante).

Como fosfolípidos ilustrativos se incluyen, pero sin limitación, 1,2-distearoil-*sn*-glicero-3-fosfocolina (DSPC), 1,2-dioleoil-*sn*-glicero-3-fosfoetanolamina (DOPE), 1,2-dilinoil-*sn*-glicero-3-fosfocolina (DLPC), 1,2-dimiristoil-*sn*-glicero-3-fosfocolina (DMPC), 1,2-dioleoil-*sn*-glicero-3-fosfocolina (DOPC), 1,2-dipalmitoil-*sn*-glicero-3-fosfocolina (DPPC), 1,2-diundecanoil-*sn*-glicero-3-fosfocolina (DUPC), 1-palmitoil-2-oleoil-*sn*-glicero-3-fosfocolina (POPC), 1,2-di-*n*-octadecenil-*sn*-glicero-3-fosfocolina (18:0 Diéter PC), 1-oleoil-2-colesterilhemisuccinoyl-*sn*-glicero-3-fosfocolina (OChemPC), 1-hexadecil-*sn*-glicero-3-fosfocolina (C16 Liso PC), 1,2-dilinoil-*sn*-glicero-3-fosfocolina, 1,2-diarachidonoyl-*sn*-glicero-3-fosfocolina, 1,2-didocosahexaenoil-*sn*-glicero-3-fosfocolina, 1,2-difitanoil-*sn*-glicero-3-fosfoetanolamina (ME 16.0 PE), 1,2-distearoil-*sn*-glicero-3-fosfoetanolamina, 1,2-dilinoil-*sn*-glicero-3-fosfoetanolamina, 1,2-dilinoil-*sn*-glicero-3-fosfoetanolamina, 1,2-diarachidonoyl-*sn*-glicero-3-fosfoetanolamina, 1,2-didocosahexaenoil-*sn*-glicero-3-fosfoetanolamina, sal sódica de 1,2-dioleoil-*sn*-glicero-3-fosfo-rac-(1-glicerol) (DOPG), dipalmitoilfosfatidilglicerol (DPPG), palmitoilfosfatidiletanolamina (POPE), distearoil-fosfatidiletanolamina (DSPE),

dipalmitoilfosfatidiletanolamina (DPPE), dimiristoilfosfoetanolamina (DMPE), 1-estearoil-2-oleoil-fosfatidiletanolamina (SOPE), 1-estearoil-2-oleoilfosfatidilcolina (SOPC), esfingomielina, fosfatidilcolina, fosfatidiletanolamina, fosfatidilserina, fosfatidilinositol, ácido fosfatídico, palmitoilooilfosfatidilcolina, lisofosfatidilcolina, lisofosfatidiletanolamina (LPE), o combinaciones de los mismos. En algunas realizaciones, un fosfolípido es DSPC. En algunas realizaciones, un fosfolípido es DMPC.

En algunas realizaciones, el fosfolípido comprende 1,2-dioleoil-*sn*-glicero-3-fosfoetanolamina-N-(succinil) (succinil PE), 1,2-distearoil-*sn*-glicero-3-fosfocolina (DSPC), colesterol, 1,2-distearoil-*sn*-glicero-3-fosfoetanolamina (DSPE), 1,2-dipalmitoil-*sn*-glicero-3-fosfoetanolamina-N-(succinil) (succinil-DPPE), 1,2-dioleoil-*sn*-glicero-3-fosfoetanolamina (DOPE), 1,2-dimiristoil-*sn*-glicero-3-fosfocolina (DMPC), 1,2-dipalmitoil-*sn*-glicero-3-fosfocolina (DPPC), o una combinación de los mismos.

Diámetro

Entre otras cosas, la presente divulgación describe composiciones que tienen un diámetro hidrodinámico promedio de aproximadamente 30 a aproximadamente 220 nm. En algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento pueden tener un diámetro hidrodinámico promedio de aproximadamente 30 a aproximadamente 220 nm. En algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento pueden tener un diámetro hidrodinámico promedio de aproximadamente 30 nm, 35 nm, 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, 150 nm, 155 nm, 160 nm, 165 nm, 170 nm, 175 nm, 180 nm, 185 nm, 190 nm, 195 nm, 200 nm, 205 nm, 210 nm, 215 nm, 220 nm, o cualquier intervalo cuyos extremos estén definidos por cualquiera de dos de los valores mencionados anteriormente. Por ejemplo, en algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento pueden tener un diámetro hidrodinámico promedio de aproximadamente 50 nm a 200 nm.

pKa

Entre otras cosas, la presente divulgación describe composiciones, preparaciones, nanopartículas, y/o nanomateriales que tienen un pKa de aproximadamente 5 a aproximadamente 9.

5 En algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento tienen un pKa de aproximadamente 5 a aproximadamente 9. En algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento tienen un pKa que es de aproximadamente 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, o cualquier intervalo cuyos extremos estén definidos por cualquiera de dos
10 de los valores mencionados anteriormente. En algunas realizaciones, las nanopartículas lipídicas descritas en el presente documento tienen un pKa que es de aproximadamente 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9.0, o cualquier intervalo cuyos extremos estén definidos por cualquiera de dos de los valores
15 mencionados anteriormente.

Composiciones ilustrativas de LNP

La presente invención proporciona composiciones que comprenden nanopartículas lipídicas. En algunas realizaciones, una composición de nanopartículas lipídicas comprende de aproximadamente 30 por ciento molar a
20 aproximadamente 70 por ciento molar de lípidos ionizables, de aproximadamente 5 por ciento molar a aproximadamente 25 por ciento molar de fosfolípidos, de aproximadamente 25 por ciento molar a aproximadamente 45 por ciento molar de colesterol y de aproximadamente 0 por ciento molar a aproximadamente 5 por ciento molar de lípido conjugado con enlazador. En algunas realizaciones, una
25 composición de nanopartículas lipídicas comprende aproximadamente 50 por ciento molar de lípidos ionizables, aproximadamente 20 por ciento molar de fosfolípidos, aproximadamente 39 por ciento molar de colesterol y aproximadamente 1 por ciento molar de lípido conjugado con enlazador.

En algunas realizaciones, una composición de nanopartículas lipídicas
30 comprende de aproximadamente 30 por ciento molar a aproximadamente 70 por

ciento molar de lípidos ionizables de Fórmula II, Fórmula III o Fórmula IV, de aproximadamente 5 por ciento molar a aproximadamente 25 por ciento molar de DSPC, de aproximadamente 25 por ciento molar a aproximadamente 45 por ciento molar de colesterol y de aproximadamente 0 por ciento molar a aproximadamente 5 por ciento molar de PEG lipídico, basándose en los moles totales de estos ingredientes.

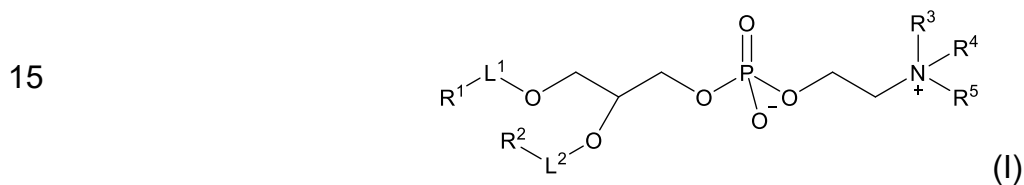
En algunas realizaciones, una preparación de nanopartículas lipídicas (LNP) comprende una relación de masa de (el total de uno o más lípidos ionizables, esteroides, lípidos conjugados con enlazador y fosfolípidos):ADN de aproximadamente 2:1 y 50:1. En algunas realizaciones, una preparación de LNP comprende una relación de masa de (el total de uno o más lípidos ionizables, esteroides, lípidos conjugados con enlazador y fosfolípidos):ADN de aproximadamente 2:1, aproximadamente 3:1, aproximadamente 4:1, aproximadamente 5:1, aproximadamente 6:1, aproximadamente 7:1, aproximadamente 8:1, aproximadamente 9:1, aproximadamente 10:1, aproximadamente 11:1, aproximadamente 12:1, aproximadamente 13:1, aproximadamente 14:1, aproximadamente 15:1, aproximadamente 16:1, aproximadamente 17:1, aproximadamente 18:1, aproximadamente 19:1, aproximadamente 20:1, aproximadamente 21:1, aproximadamente 22:1, aproximadamente 23:1, aproximadamente 24:1, aproximadamente 25:1, aproximadamente 26:1, aproximadamente 27:1, aproximadamente 28:1, aproximadamente 29:1, aproximadamente 30:1, aproximadamente 31:1, aproximadamente 32:1, aproximadamente 33:1, aproximadamente 34:1, aproximadamente 35:1, aproximadamente 36:1, aproximadamente 37:1, aproximadamente 38:1, aproximadamente 39:1, aproximadamente 40:1, aproximadamente 41:1, aproximadamente 42:1, aproximadamente 43:1, aproximadamente 44:1, aproximadamente 45:1, aproximadamente 46:1, aproximadamente 47:1, aproximadamente 48:1, aproximadamente 49:1, aproximadamente 50:1.

Compuestos ilustrativos

Entre otras cosas, la presente divulgación describe composiciones, y/o nanopartículas que comprenden uno o más compuestos como los descritos en el presente documento.

5 En algunas realizaciones, la presente divulgación proporciona nanopartículas lipídicas que comprenden compuestos que comprenden uno o más grupos lipídicos parcialmente insaturados. En algunas realizaciones, las nanopartículas lipídicas proporcionadas comprenden compuestos que comprenden uno o más grupos amina trimetilados. En algunas realizaciones, las nanopartículas lipídicas
10 proporcionadas comprenden compuestos que comprenden uno o más grupos amina dimetilados.

En algunas realizaciones, la presente divulgación proporciona un compuesto de Fórmula I:



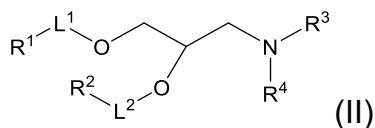
o una sal farmacéuticamente aceptable del mismo, en donde:

20 cada uno de L^1 y L^2 es independientemente un enlace covalente, $-C(O)-$, $-OC(O)-$, o una cadena hidrocarbonada C_1-C_{12} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida;

25 cada uno de R^1 y R^2 es independientemente un grupo opcionalmente sustituido seleccionado de una cadena hidrocarbonada C_1-C_{20} saturada o insaturada, recta o ramificada en donde 1-3 unidades de metileno están opcionalmente sustituidas por $-O-$; y

cada uno de R^3 , R^4 y R^5 es independientemente un hidrógeno o una cadena hidrocarbonada C_1-C_{10} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida.

30 En algunas realizaciones, la presente divulgación proporciona un compuesto de Fórmula II:



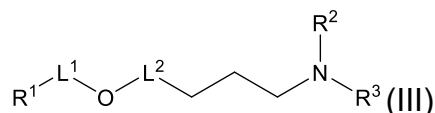
o una sal farmacéuticamente aceptable del mismo, en donde:

5 cada uno de L^1 y L^2 es independientemente un enlace covalente, $-C(O)-$, $-OC(O)-$, o una cadena hidrocarbonada C_1-C_{12} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida;

10 cada uno de R^1 y R^2 es independientemente un grupo opcionalmente sustituido seleccionado de una cadena hidrocarbonada C_1-C_{20} saturada o insaturada, recta o ramificada en donde 1-3 unidades de metileno están opcionalmente sustituidas por $-O-$; y

cada uno de R^3 y R^4 es independientemente un hidrógeno o una cadena hidrocarbonada C_1-C_{10} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida.

15 En algunas realizaciones, la presente divulgación proporciona un compuesto de Fórmula III:



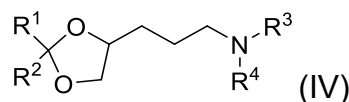
o una sal farmacéuticamente aceptable del mismo, en donde:

20 cada uno de L^1 y L^2 es independientemente un enlace covalente, $-C(O)-$, $-OC(O)-$, o una cadena hidrocarbonada C_1-C_{12} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida;

25 R^1 es independientemente un hidrógeno, o un grupo opcionalmente sustituido seleccionado de una cadena hidrocarbonada C_1-C_{20} saturada o insaturada, recta o ramificada en donde 1-3 unidades de metileno están opcionalmente sustituidas por $-O-$; y

cada uno de R^2 y R^3 es independientemente un hidrógeno o una cadena hidrocarbonada C_1-C_{10} bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida.

En algunas realizaciones, la presente divulgación proporciona un compuesto de Fórmula IV:



5 o una sal farmacéuticamente aceptable del mismo, en donde:

cada uno de R¹ y R² es independientemente un grupo opcionalmente sustituido seleccionado de una cadena hidrocarbonada C₁-C₂₀ saturada o insaturada, recta o ramificada en donde 1-3 unidades de metileno están opcionalmente sustituidas por -O-; y

10 cada uno de R³ y R⁴ es independientemente un hidrógeno o una cadena hidrocarbonada C₁-C₁₀ bivalente, saturada o insaturada, recta o ramificada, opcionalmente sustituida.

Nucleasas diana

15 En la técnica se conocen múltiples nucleasas diana dirigidas al ADN. En algunas realizaciones, las nucleasas descritas en el presente documento pueden comprender cualquier secuencia de polipéptidos o de proteínas capaz de inducir una rotura una rotura monocatenaria o bicatenaria de ADN. En algunas realizaciones, las nucleasas descritas en el presente documento pueden ser proteínas de tipo silvestre, proteínas de fusión, proteínas genomodificadas o

20 variantes de las mismas. En algunas realizaciones, las nucleasas pueden seleccionarse de una clase de enzimas que comprende: TALEN, ZFN, Meganucleasas, Nicasas TALE, Nicasas de dedos de zinc y nucleasas Cas (p. ej., Cas9 de *Staphylococcus aureus* (saCas9), *Streptococcus pyogenes* (spCas9), Nucleasa de AZ (de *ArcticZymes*), HF1-Cas9, HF2-Cas9 o HiFi-Cas9). En algunas

25 realizaciones, las nucleasas se combinan con otros componentes de GENERIDE™ para producir la integración dirigida de un transgén en un sitio de integración diana. En algunas realizaciones, las nucleasas se combinan con otros componentes de GENERIDE™ para producir la integración dirigida de un transgén en un sitio de integración diana que es distal con respecto un sitio de corte de nucleasa.

30 *TALEN (nucleasas efectoras similares a un activador de la transcripción)*

Las TALEN son enzimas de restricción que comprenden un dominio de unión al ADN (dominio efector similar a un activador de la transcripción (TAL, *transcription activator-like*)) fusionado directa o indirectamente a una nucleasa. Los efectores TAL (TALE) comprenden un conjunto altamente conservado de repeticiones de secuencias de 33-34 aminoácidos, que pueden diseñarse para unirse a secuencias específicas de ADN dentro de un genoma celular. La fusión de TALE modificadas a un dominio de escisión de nucleasa (p. ej., la nucleasa FokI) puede producir nucleasas TALE (TALEN) capaces de dirigirse a una secuencia de ADN específica para producir roturas en el ADN. Pueden utilizarse combinaciones de TALEN (p. ej., pares de TALEN) para producir roturas bicatenarias (DSB) en una secuencia de ADN.

En algunas realizaciones, las TALEN pueden diseñarse para dirigirse a secuencias de ADN específicas e inducir DSB. En algunas realizaciones, una o más TALEN (p. ej., pares de TALEN) pueden combinarse para inducir DSB a través de la unión a secuencias de ADN diana en cadenas de ADN genómico. En algunas realizaciones, determinadas características de las TALEN (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para mejorar la eficacia de unión y/o corte del ADN. En algunas realizaciones, determinadas características de las TALEN (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas) pueden diseñarse u optimizarse para reducir la unión inespecífica.

Nicasa TALE

Los TALE (*transcription activator-like effector*, efectores similares a un activador de la transcripción) comprenden un conjunto altamente conservado de repeticiones de secuencias de 33-34 aminoácidos, que pueden diseñarse para unirse a secuencias de ADN específicas dentro de un genoma celular. Pueden utilizarse combinaciones de TALEN (p. ej., pares de TALEN) para producir roturas de ADN diana en una secuencia de ADN. Una Nicasa TALE puede diseñarse introduciendo mutaciones (p. ej., D450A) en un dominio de escisión (p. ej., la nucleasa FokI) de un monómero TALEN en las TALEN diseñadas (p. ej., pares de

TALEN). Las Nicasas TALE son capaces de reconocer regiones específicas de una secuencia de ADN diana y generar roturas monocatenarias (SSB, *single strand breaks*) diana en una secuencia de ADN.

En algunas realizaciones, las Nicasas TALE pueden diseñarse para dirigirse a secuencias de ADN específicas e inducir DSB. En algunas realizaciones, determinadas características de las Nicasas TALE (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para mejorar la eficacia de unión y/o corte del ADN. En algunas realizaciones, determinadas características de las Nicasas TALE (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas) pueden diseñarse u optimizarse para reducir la unión inespecífica.

ZFN (Zinc Finger Nucleases, nucleasas de dedos de zinc)

Las ZFN son enzimas de restricción que comprenden un dominio de unión al ADN (proteína de dedos de zinc (ZFP, *Zinc Finger Protein*)) fusionado directa o indirectamente a una nucleasa. Las ZFP comprenden de 3 a 6 repeticiones individuales de dedos de zinc, que son capaces de reconocer entre 9 pb y 18 pb de una secuencia de ADN diana. La fusión de ZFP diseñadas a un dominio de escisión de nucleasa (p. ej., la nucleasa FokI) puede producir ZFN capaces de dirigirse a una secuencia específica de ADN para producir roturas en el ADN. Pueden utilizarse combinaciones de ZFN (p. ej., pares de ZFN) para producir roturas bicatenarias (DSB) diana en una secuencia de ADN.

En algunas realizaciones, las ZFN pueden diseñarse para dirigirse a secuencias de ADN específicas e inducir roturas bicatenarias, DSB. En algunas realizaciones, una o más ZFN (p. ej., pares de ZFN) pueden combinarse para inducir DSB a través de la unión a secuencias de ADN diana en cadenas en ADN genómico. En algunas realizaciones, determinadas características de las ZFN (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para mejorar la eficacia de unión y/o corte del ADN. En algunas realizaciones, determinadas características de las ZFN (p. ej., regiones de

unión al ADN, regiones enlazadoras, región de fusión de nucleasas) pueden diseñarse u optimizarse para reducir la unión inespecífica.

Nicasa de dedos de Zinc

Las ZFP comprenden 3-6 repeticiones individuales de dedos de zinc, que son capaces de reconocer entre 9 pb y 18 pb de una secuencia de ADN diana. Para producir roturas dirigidas en el ADN en una secuencia de ADN, pueden utilizarse combinaciones de ZFN (p. ej., pares de ZFN). Introduciendo mutaciones (p. ej., D450A) en un dominio de escisión (p. ej., nucleasa FokI) de un monómero de ZFN en una ZFN diseñada, puede diseñarse una Nicasa de ZF. Las Nicasas de ZF son capaces de reconocer regiones específicas de una secuencia de ADN diana y generar roturas monocatenarias (SSB, *single strand breaks*) dirigidas en una secuencia de ADN.

En algunas realizaciones, las nicasas de ZF pueden diseñarse para dirigirse a secuencias de ADN específicas e inducir SSB. En algunas realizaciones, determinadas características de la Nicasa de ZF (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para mejorar la eficacia de unión y/o corte del ADN. En algunas realizaciones, determinadas características de la Nicasa de ZF (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para reducir la unión inespecífica.

Meganucleasas

Las meganucleasas (también denominadas endonucleasas de asentamiento (*homing*)) son endonucleasas específicas de secuencia capaces de reconocer entre 12 pb y 40 pb de una secuencia de ADN diana. La clase más numerosa de endonucleasas de asentamiento es la familia LAGLIDADG, que incluye, entre otras, las enzimas I-CreI e I-SceI, bien caracterizadas y de uso común. El rediseño de estas endonucleasas de asentamiento puede producir endonucleasas de asentamiento capaces de dirigirse a una secuencia de ADN específica para producir roturas del ADN. Se han diseñado proteínas quiméricas que comprenden fusiones de meganucleasas, ZFP y TAL para generar nuevas enzimas que aprovechen la

afinidad de unión de las ZF y las TALE y la especificidad de escisión de las meganucleasas.

En algunas realizaciones, pueden diseñarse meganucleasas que se dirijan a secuencias de ADN específicas e induzcan DSB (roturas bicatenarias). En algunas realizaciones, una o más meganucleasas pueden combinarse para inducir DSB a través de la unión a secuencias diana de ADN en cadenas de ADN genómico. En algunas realizaciones, una o más meganucleasas pueden combinarse con ZFP y/o TAL para inducir DSB a través de la unión a secuencias diana de ADN en cadenas de ADN genómico. En algunas realizaciones, determinadas características de las meganucleasas (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para mejorar la eficacia de unión y/o corte del ADN. En algunas realizaciones, determinadas características de las meganucleasas (p. ej., regiones de unión al ADN, regiones enlazadoras, región de fusión de nucleasas), pueden diseñarse u optimizarse para reducir la unión inespecífica.

Sistemas asociados a CRISPR

Los sistemas asociados a CRISPR comprenden normalmente una nucleasa Cas (p. ej., Cas9, Cas13, Cas12a, Cas12b, Cas9 y similares) o una variante de la misma y una secuencia de ARN guía (ARNg) diseñada. Las secuencias de ARNg están diseñadas para tener complementariedad parcial con una secuencia de ADN genómico diana de interés dentro de una determinada distancia de un motivo adyacente de protoespaciador (PAM, *protospacer adjacent motif*). Las secuencias PAM y la distancia a los sitios de corte de la nucleasa Cas suelen ser específicas de determinados tipos de enzimas (p. ej., Cas9, Cas13, Cas12a, Cas12b, Cas9). Determinadas nucleasas Cas, o sus variantes, son capaces de producir DSB en una secuencia de ADN diana.

Cas9

Cas9 es una enzima endonucleasa de ADN guiada por ARN que actúa a través de la complementariedad de pares de bases entre los 17-20 primeros nucleótidos de un ARNg diseñado y una cadena complementaria de ADN genómico

diana. El ADN genómico diana debe presentar una secuencia PAM apropiada (p. ej., NGG o NAG) adyacente a la región de complementariedad del ARNg. Una vez unida a una secuencia diana apropiada, Cas9 induce una DSB en una secuencia de ADN diana. Las nucleasas Cas9 pueden encontrarse en diversas especies diferentes, incluyendo, p. ej., *S. pyogenes* (SpCas9), *S. aureus* y *N. meningitidis*, entre otras. Las nucleasas Cas9 diseñadas pueden mejorar la actividad de corte o reducir los efectos de corte inespecíficos.

Cas12a / Cpf1

Cas12a (Cpf1) es una enzima endonucleasa de ADN guiada por ARN que actúa a través de la complementariedad de pares de bases entre los 20 primeros nucleótidos de un ARNg diseñado y una cadena complementaria de ADN genómico diana. El ADN genómico diana debe presentar una secuencia PAM apropiada (p. ej., TTN/TTTN/TTTV) adyacente a la región de complementariedad del ARNg. Una vez unida a una secuencia diana apropiada, Cas12a induce un corte bicatenario escalonado en una secuencia de ADN diana.

Nicasas Cas9

Las mutaciones (p. ej., D10A y/o H840A) en uno de los dos dominios de nucleasa Cas9 de tipo silvestre dan lugar a una variante de Cas9 (nicasa Cas9). Al igual que las Cas 9 de tipo silvestre, las nicasas actúan a través de la complementariedad de pares de bases entre los nucleótidos de un ARNg diseñado y una cadena complementaria de ADN genómico diana. A diferencia de Cas9 de tipo silvestre, una vez unida a una secuencia diana apropiada, la nicasa Cas9 corta sólo una cadena del ADN generando una SSB que puede repararse sin inducir indeles (inserciones/deleciones).

ARNg (ARN guía)

Una molécula de ARN guía puede ser o comprender un ácido nucleico que promueva el direccionamiento o asentamiento específico de un complejo de ARNg/Cas a una diana. En algunas realizaciones, un ARNg incorporará funciones y estructuras de un ARN crispr (ARNcr) y/o de un ARN transactivante crispr (ARNtracr). En algunas realizaciones, un ARNg puede ser quimérico y comprender

rasgos tanto de ARNcr como de ARNtracr en una única secuencia de ácido nucleico (p. ej., ARN guía único, o ARNgu). En algunas realizaciones, la molécula de ARNg comprende dominios múltiples. En algunas realizaciones, una molécula de ARNg comprende un dominio de direccionamiento (complementario a un ácido nucleico diana), un primer dominio de complementariedad, un dominio de enlace, un segundo dominio de complementariedad (complementario al primer dominio de complementariedad), un dominio proximal, y/o un dominio de cola.

En algunas realizaciones, una molécula de ARNg puede comprender un dominio de direccionamiento que comprenda una secuencia de ácido nucleico que sea complementaria (p. ej., al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 % o 100 % complementaria) a una secuencia de ADN diana. En algunas realizaciones, una molécula de ARNg puede comprender un dominio de direccionamiento que comprenda una secuencia de ácido nucleico que sea complementaria (p. ej., al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 % o 100 % complementaria) a una cadena sencilla de una secuencia de ADN diana. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 5-50 nucleótidos (p. ej., 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50 y similares). En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 16 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 16 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 17 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 18 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 19 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 20 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 21 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 22 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 23 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 24 nucleótidos. En algunas realizaciones, el

dominio de direccionamiento tiene una longitud de 25 nucleótidos. En algunas realizaciones, el dominio de direccionamiento tiene una longitud de 26 nucleótidos.

En algunas realizaciones, una molécula de ARNg puede diseñarse para reducir el direccionamiento inespecífico dentro de un genoma celular. En algunas realizaciones, una molécula de ARNg puede diseñarse utilizando un programa informático para una secuencia de ADN diana específica con una similitud de secuencia limitada (p. ej., homología, identidad y similares) con otra región del genoma.

En algunas realizaciones, una molécula de ARNg se diseña para dirigirse a una secuencia de ADN no codificante (p. ej., un intrón, una región no traducida, un potenciador, un promotor, un silenciador o secuencia aislante). En algunas realizaciones, una molécula de ARNg se diseña para dirigirse a una secuencia no codificante en un gen humano. En algunas realizaciones, una molécula de ARNg se diseña para dirigirse a una secuencia no codificante en un gen de puerto seguro humano (p. ej., albúmina, colágeno, actina, CCR5, y similares). En algunas realizaciones, una molécula de ARNg se diseña para dirigirse a una secuencia no codificante en un gen de albúmina humana. En algunas realizaciones, una molécula de ARNg se diseña para dirigirse al intrón 13 o al intrón 14 de un gen de albúmina humana.

En algunas realizaciones, una molécula de ARNg puede diseñarse para dirigirse a una secuencia y/o incluir una secuencia seleccionada de una secuencia de la Tablas 3 y 4 mostradas más adelante. En algunos ejemplos, la secuencia de la Tabla 3 o de la Tabla 4 puede ser una secuencia espaciadora para un ARNg.

Tabla 3: Secuencias ilustrativas de ARNg para spCas9.

Locus diana	Cadena	Secuencia	SEQ ID NO:
Albúmina humana Intrón 14_1	-	GTGGTCGGTGCTGGTCTATA	27
Albúmina humana	-	TTAGAATAGTGGTCGGTGC	28

	Intrón 14_2			
	Albúmina humana Intrón 14_3	-	CGCATCATATAAGTCAGAAC	29
5	Albúmina humana Intrón 14_4	-	TGAGCCATCACTTACTTGGA	30
	Albúmina humana Intrón 14_5	-	CTGCTGAGCCATCACTTACT	31
	Albúmina humana Intrón 14_6	-	TTAAGCCTGTCCTAGACTTA	32
10	Albúmina humana Intrón 13_1	+	TTGCATGTTTGGTTAGGCTA	33
	Albúmina humana Intrón 13_2	-	ACAGTATGGCACAATAGAGC	34
15	Albúmina humana Intrón 13_3	-	AAGTTACTGTCAGGCCATAA	35
	Albúmina humana Intrón 13_4	-	AATCTCTGGACGGAAGCTCA	36

Tabla 4: Secuencias ilustrativas de ARNg para saCas9.

	Locus diana	Cadena	Secuencia	SEQ ID NO:
20	Albúmina humana Intrón 13_1	+	ATGTTTGGTTAGGCTAGGGCTT	37
25	Albúmina humana Intrón 13_2	+	GGAGGCTTTGTACATGTGGGAC	38
30	Albúmina humana Intrón 13_3	+	TTACAAACAATTGTCTTACAAA	39

5	Albúmina humana Intrón 14_1	-	CACTGAACTTTGGTTTCAATT	40
	Albúmina humana Intrón 14_2	-	CCACTGCTGAGCCATCACTTAC	41
	Albúmina humana Intrón 14_3	+	CAAGTAAGTGATGGCTCAGCAG	42
10	Albúmina humana Intrón 14_4	+	GGCTCAGCAGTGGGAATACTCT	43
15	Albúmina humana Intrón 14_5	+	TTCAACTCATCCTTTCCATTG	44
	Albúmina humana Intrón 14_6	+	TGACTTATATGATGCGGTACA	45

Ácidos nucleicos heterólogos

Cargas útiles

En algunas realizaciones, uno o más vectores o construcciones descritos en el presente documento pueden comprender una secuencia de polinucleótidos que codifique una o más cargas útiles. De acuerdo con diversos aspectos, puede utilizarse cualquiera de una variedad de cargas útiles (p. ej., las que tienen un fin diagnóstico y/o terapéutico), solas o combinadas. En algunas realizaciones, una carga útil puede ser o comprender una secuencia de polinucleótidos que codifique un péptido o polipéptido. En algunas realizaciones, una carga útil es un péptido que tiene actividad celular intrínseca o extrínseca, que promueve un proceso biológico para tratar una dolencia. En algunas realizaciones, una carga útil puede ser o comprender un transgén (también denominado en el presente documento gen de

interés (GOI, *gene of interest*)). En algunas realizaciones, una carga útil puede ser o comprender una o más secuencias de repeticiones terminales invertidas (ITR) (p. ej., una o más ITR de AAV). En algunas realizaciones, una carga útil puede ser o comprender uno o más transgenes con secuencias de ITR flanqueantes. En algunas realizaciones, una carga útil puede ser o comprender una o más secuencias de ácido nucleico heterólogas que codifiquen un gen indicador (p. ej., un indicador fluorescente o luminiscente). En algunas realizaciones, una carga útil puede comprender uno o más biomarcadores (p. ej., uno indirecto para la expresión de la carga útil). En algunas realizaciones, una carga útil puede comprender una secuencia para la expresión policistónica (incluyendo, p. ej., un péptido 2A, o secuencia intrónica, un sitio interno de entrada al ribosoma). En algunas realizaciones, los péptidos 2A son secuencias peptídicas pequeñas (p. ej., de aproximadamente 18-22 aminoácidos) que permiten la coexpresión de dos o más productos de proteína distintos dentro de una sola secuencia codificante. En algunas realizaciones, los péptidos 2A permiten la coexpresión de dos o más productos de proteína distintos, independientemente de la disposición de las secuencias codificantes de proteína. En algunas realizaciones, los péptidos 2A son o comprenden un motivo consenso (p. ej., DVEXNPGP). En algunas realizaciones, los péptidos 2A promueven la escisión de proteínas. En algunas realizaciones, los péptidos 2A son o comprenden secuencias víricas (p. ej., virus de la fiebre aftosa (F2A), virus de la rinitis equina A, teschovirus porcino-1 (P2A) o virus Thosa asigna (T2A)).

En algunas realizaciones, una carga útil puede ser o comprender una secuencia de polinucleótidos, que comprenda un casete de expresión. En algunas realizaciones, un casete de expresión comprende una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico, en donde la primera secuencia de ácido nucleico codifica un transgén; y la segunda secuencia de ácido nucleico se sitúa en posición 5' o 3' en la primera secuencia de ácido nucleico y promueve la producción de dos productos génicos independientes (p. ej., una secuencia codificante de un péptido 2A).

En algunas realizaciones, los biomarcadores son o comprenden un péptido 2A (p. ej., P2A, T2A, E2A y/o F2A). En algunas realizaciones, los biomarcadores son o comprenden un motivo de escisión de Furina (véase, Tian et al., FurinDB: A Database of 20-Residue Furin Cleavage Site Motifs, Substrates and Their Associated Drugs, (2011), Int. J. Mol. Sci., vol. 12: 1060-1065). En algunas realizaciones, los biomarcadores son o comprenden una etiqueta (p. ej., una etiqueta inmunológica). En algunas realizaciones, una carga útil puede comprender uno o más ácidos nucleicos funcionales (p. ej., uno o más ARNpi (ARN pequeño de interferencia) o miARN (microARN)). En algunas realizaciones, una carga útil puede comprender uno o más ácidos nucleicos inhibidores (incluyendo, p. ej., ribozima, miARN, ARNpi o ARNh_c, entre otros). En algunas realizaciones, una carga útil puede comprender una o más nucleasas (p. ej., proteínas Cas, endonucleasas, TALEN, ZFN).

En algunas realizaciones, una secuencia que codifica un péptido 2A puede tener una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej., una secuencia P2A de tipo silvestre). En algunas realizaciones, una secuencia que codifica un péptido P2A puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una parte de una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej., una secuencia génica de tipo silvestre). En algunas realizaciones, una secuencia que codifica un péptido P2A puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una secuencia ilustrativa de la Tabla 5 mostrada a continuación.

Tabla 5: Secuencias ilustrativas de péptido 2A

Descripción	Secuencia	SEQ ID NO
Secuencia de nucleótidos de P2A	GGAAGCGGCGCCACCAATTTTCAGCCTGCTGAAAC AGGCCGGCGACGTGGAAGAGAACCCTGGCCCT	46
Secuencia de aminoácidos de P2A	GSGATNFSLLKQAGDVEENPGP	47
Secuencia de nucleótidos de P2A versión 2	GGCAGCGGCGCCACCAACTTCAGCCTGCTGAAAC AGGCCGGCGACGTGGAAGAGAACCCTGGCCCT	48

Transgenes

En algunas realizaciones, un transgén es un gen corrector elegido para mejorar uno o más signos y/o síntomas de una enfermedad, trastorno o afección. En algunas realizaciones, un transgén puede integrarse en el genoma de una célula hospedadora utilizando un vector o vectores incluidos en la presente divulgación. En algunas realizaciones, los transgenes son versiones funcionales de genes asociados a enfermedades (es decir, una o más isoformas de genes que están asociadas a la manifestación o empeoramiento de una enfermedad, trastorno o afección) que se encuentran en una célula hospedadora. En algunas realizaciones, los transgenes son una versión optimizada de genes asociados a enfermedades que se encuentran en una célula hospedadora (p. ej., variantes con optimización de codones o expresión). En algunas realizaciones, los transgenes son variantes de genes asociados a enfermedades que se encuentran en una célula hospedadora (p. ej., un fragmento génico funcional o una variante del mismo). En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que normalmente se expresa en uno o más tejidos sanos. En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que normalmente se

expresa en células hepáticas. En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que normalmente se expresa en células de músculo. En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que normalmente se expresa en células del sistema nervioso central.

5 En algunas realizaciones, un transgén puede ser o comprender un gen que causa la expresión de un péptido que no se expresa normalmente en uno o más tejidos sanos (p. ej., un péptido expresado de manera ectópica). En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que se expresa de manera ectópica en uno o más tejidos sanos (p. ej., hígado, músculo, sistema nervioso central (SNC), pulmón). En algunas realizaciones, un transgén es un gen que causa la expresión de un péptido que se expresa de manera ectópica en uno o más tejidos sanos y que normalmente se expresa en uno o más tejidos sanos (p. ej., hígado, músculo, sistema nervioso central (SNC), pulmón).

10 En algunas realizaciones, un transgén puede ser o comprender un gen que codifique un ácido nucleico funcional. En algunas realizaciones, un agente terapéutico es o comprende un agente que tiene un efecto terapéutico sobre una célula hospedadora o sobre un sujeto hospedador (incluyendo, p. ej., ribozima, ARN guía (ARNg), oligonucleótido antisentido (ASO, *antisense oligonucleotide*), miARN, ARNpi y/o ARNhc). Por ejemplo, en algunas realizaciones, un agente terapéutico promueve un proceso biológico para tratar una dolencia, p. ej., al menos un síntoma de una enfermedad, trastorno o afección.

15 En algunas realizaciones, la expresión del transgén en un sujeto es sustancialmente el resultado de la integración en un sitio de integración diana. En algunas realizaciones, un 75 % o más (p. ej., 80 % o más, 85 % o más, 90 % o más, 95 % o más, 99 % o más, 99.5 % o más) de la expresión total del transgén en un sujeto procede de la integración del transgén en un sitio de integración diana. En algunas realizaciones, un 25 % o menos (p. ej., 20 % o menos, 15 % o menos, 10 % o menos, 5 % o menos, 1 % o menos, 0.5 % o menos, 0.1 % o menos) de la expresión total del transgén en un sujeto procede de una fuente distinta de la

integración del transgén en un sitio de integración diana (p. ej., expresión episomal, integración en un sitio de integración no diana).

5 En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares). En algunas realizaciones, un 75 % o más (p. ej., 80 % o más, 85 % o más, 90 % o más, 95 % o más, 99 % o más, 99.5 % o más) de la expresión total del transgén en un sujeto procede de expresión transitoria. En algunas realizaciones, el 25 % o menos (p. ej., 20 % o menos, 15 % o menos, 10 % o menos, 5 % o menos, 1 % o menos, 0.5 % o menos, 0.1 % o menos) de la expresión total del transgén en
10 un sujeto procede de una fuente distinta de la expresión transitoria (p. ej., integración en un sitio de integración no diana). En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante una o más semanas después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante uno o más meses después del tratamiento.

20 En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) una o más semanas después del tratamiento a un nivel comparable al observado en uno o más días después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) uno o más meses después del tratamiento a un nivel comparable al observado en uno o más días después del tratamiento.

25 En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) una o más semanas después del tratamiento a un nivel que se reduce con respecto al observado en uno o más días después del tratamiento. En algunas
30 realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej.,

expresión episomal de plásmidos, ADN minicirculares, virus y similares) uno o más meses después del tratamiento a un nivel que se reduce con respecto al observado en uno o más días después del tratamiento.

5 En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de un mes después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de dos meses después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de tres meses después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de cuatro meses después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de cinco meses después del tratamiento. En algunas realizaciones, los transgenes se expresan de manera transitoria en un sujeto (p. ej., expresión episomal de plásmidos, ADN minicirculares, virus y similares) durante no más de seis meses después del tratamiento.

25 En algunas realizaciones, el tamaño combinado de los transgenes y de los brazos de homología puede optimizarse para aumentar la probabilidad de que estos transgenes tengan una longitud de secuencia adecuada para empaquetarse de manera eficaz en un vehículo de suministro, lo que puede aumentar la probabilidad de que los transgenes se suministren en definitiva de forma adecuada al paciente.

30 En algunas realizaciones, una secuencia de nucleótidos que codifica un transgén presenta optimización de codones. En algunas realizaciones, una secuencia de nucleótidos que codifica un transgén presenta optimización de codones para un determinado tipo celular (p. ej., mamífero, insecto, bacteria, hongo, y similares). En algunas realizaciones, una secuencia de nucleótidos que codifica

un transgén presenta optimización de codones para una célula humana. En algunas realizaciones, una secuencia de nucleótidos que codifica un transgén presenta optimización de codones para una célula humana de un tipo de tejido particular (p. ej., hígado, músculo, SNC, pulmón).

5 En determinadas realizaciones, una secuencia de nucleótidos que codifica un transgén puede optimizarse con codones para que tenga una homología de nucleótidos con una secuencia de nucleótidos de referencia (p. ej., una secuencia génica de tipo silvestre) de menos del 100 %. En determinadas realizaciones, la homología de nucleótidos entre una secuencia de nucleótidos con optimización de
10 codones que codifica un transgén y una secuencia de nucleótidos de referencia es inferior al 100 %, inferior al 99%, inferior al 98%, inferior al 97%, inferior al 96%, inferior al 95%, inferior al 94%, inferior al 93%, inferior al 92%, inferior al 91%, inferior al 90%, inferior al 89%, inferior al 88%, inferior al 87%, inferior al 86%, inferior al 85%, inferior al 84%, inferior al 83%, inferior al 82%, inferior al 81%, inferior al 80%, inferior al 78%, inferior al 76%, inferior al 74%, inferior al 72%, inferior al 70%, inferior al 68%, inferior al 66%, inferior al 64%, inferior al 62%, inferior al 60%, inferior al 55%, inferior al 50% e inferior al 40%.

En algunas realizaciones, un transgén puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o
20 del 100 % con una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej., una secuencia génica de tipo silvestre). En algunas realizaciones, un transgén puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una parte de una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej.,
25 una secuencia génica de tipo silvestre). En algunas realizaciones, un transgén puede ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o 100 % con una secuencia ilustrativa de la Tabla 6 mostrada a continuación.

Tabla 6: Secuencias de transgenes ilustrativas

Descripción del Transgén	Secuencia	SEQ ID NO
FAH humano	tccttcatcccggtggccgaggattccgacttcccatccacaacctgcct acggcgtcttctcgaccagaggcgaccaagaccgaggatagggtgtggc cattggcgaccagatcctggacctcagcatcatcaagcacctcttactggt cctgtcctctccaaacaccaggatgtcttcaatcagcctacactcaacagct tcatgggcctgggtcaggctgcctggaaggaggcgagagtgttctgcag aacttgctgtctgtgagccaagccaggctcagagatgacaccgaacttcg gaagtgtgattcatctcccaggctctgccacgatgcacctccagccacc ataggagactacacagacttctattcctctcggcagcatgctaccaacgtc ggaatcatgttcagggaacaaggagaatgcgttgatgcaaattggctgca cttaccagtgggctaccatggccgtgcctcctctgtcgtggtgtctggcacc caatccgaaggcccatgggacagatgaaacctgatgactctaagcctcc cgtatatggtgcctgcaagctcttgacatggagctggaaatggctttttgt aggccttgaaacagattgggagagccgatccccattccaaggcccat gagcacattttggaatggctcttatgaacgactggagtgcacgagacattc agaagtgggagtatgtccctctcgggccattccttggaagagtttgggac cactgtctctccgtgggtggtgcccattgatgctctcatgcccttgctgtgcc caacccgaagcaggaccccaggccccctgccgtatctgtgcatgacgag ccctacacatttgacatcaacctctctgttaacctgaaaggagaaggaaatg agccaggcggctaccatattgaagtccaattttaagtacatgtactggacg atgctgcagcagctcactcaccactctgtcaacggctgcaacctgcggcc gggggacctcctggcttctgggaccatcagcgggcccggagccagaaaa cttcggctccatgttgaactgtcgtggaagggaacgaagcccatagacct ggggaatggtcagaccaggaagtttctgtgacggggatgaagtcatca taacagggtactgccagggggatggttaccgcatcggccttggccagtgtg ctggaaaagtgtgcctgtctcctgccatca	49

FAH humano Variante 3 con optimización de codones	tccttcattcctgtggccgaggattctgactttccattcacaacctgccctatg gcgtctttcaaccaggggagatcccaggcccagaatcggggtggccatt ggagatcagatcctggacctgtcaatcatcaagcacctgtttacaggccct gtgctgagcaagcaccaggatgtgttcaaccagccaactctgaacagctt catggggctggggcaggctgcctggaaggaggcaaggggtttctgcag aatctgctgagcgtgtcacaggctaggctgagggatgataccgagctgag gaagtgtgcatttatctcacaggcatcagccacaatgcatctgccagctac aatcggcgactataccgacttttactcaagcaggcagcatgccaccaatgt gggcatcatgttcagggaagagaatgccctgatgccaaattggctgc acctgcctgtgggctatcatggcagagccagctcagtgggtggtgcagg acaccaattaggaggcctatgggccagatgaagcctgatgacagcaaa ccacctgtctacggcgctgcaagctcctggatatggaactggagatggct ttttgtggggccaggcaatagactgggagagcccattccaatctcaaag gccatgaacatatctttggcatggtcctgatgaatgactggtcagccaga gacattcagaagtgggagtacgtgccactgggaccatttctgggaaagag ctttggcaccacagtgtcaccatgggtgggtgccatggatgccctgatgcc ctttgctgtgccaaatccaaaacaagaccccaggccccctcccctatctctgt catgatgaaccatatacttttgacattaacctgagcgtgaacctgaaagga gaaggcatgagccaagccgccactatctgtaagagcaactcaaatatat gtactggacaatgctgcagcagctcaccacccatagcgtgaatgggtgta acctgaggccaggcgacctgctggcatcaggcactatttcagggcctgag cctgaaaattttgatcaatgctggagctgtcatggaagggaactaaaccc atcgacctggggaatggccagaccagaaagtttctcctggatggcgatga ggatgacattacaggctactgtcagggggatggctatagaattggattggc caatgcgccggcaaagtccctccctgccctgctgcccagc	50
FAH humano	tccttcacccagtggtgaagattctgacttccattcacaacctcccttat ggagtcttagcacaagaggagaccaagacctagaattggcgtggctat	51

5	Variante 2 con optimización de codones	tggggatcagatcctggacctgagcattatcaaacatctctttacaggccca gtccctcagcaagcaccaagatgtgttcaatcaaccacactgaactcattt atggggctgggccaagctgcctggaaggaggcaagagtgtttctccaga acctcctgtcagtgagccaggccaggctgagagatgacacagagctgag gaagtgtgcctttatctcacaagccagcgccacaatgcacctgccagcaa ccatcggggactacacagacttttacagcagcaggcagcacgccactaa tgtgggcatcatgtttagagacaaagagaatgccctcatgcctaactggct gcatctgccagtgggctaccatggcagagccagcagcgtgggtgtctg gcaccctatcaggaggcccatgggcccagatgaagcctgatgacagca agccacctgtgtatggggcatgtaagctgctggacatggagctggaaatg gctttcttgtggggcctgggaacaggctgggcgaaccaattcccatcagc aaagctcatgaacacattttgggatggtgctcatgaatgactggtctgcca gagacatccagaagtgggagtatgtccccctgggaccattcctgggcaag agctttgggaccactgtgtcaccatgggtggtgccatggatgccctgatgc catttctgtgcctaatcccaaacaagatccaaggccactgccctaccttg ccacgatgagccctacacatttgatatcaacctctctgtgaatctgaaggga gaaggaatgagccaagctgctaccatttgcaaaagcaactttaagtatatg tattggacatgctgcagcaactcaccaccacagcgtgaatggctgtaa cctgaggcctggcgacctgctggcctctggcaccatcagcgccctgaac ctgagaactttggcagcatgctggagctgagctggaagggaaccaagcc cattgacctgggcaatggacagaccaggaaattcctcctggacggggac gaggtgatcatcacaggctattgccagggggatgggtataggattggcttt ggccaatgtgccggcaaagtgtgcctgcctgctccctagc	
25	FAH humano Variante 1 con optimización de codones	tcctttattcctgtggccgaagatagcgacttcctatccataacctccatat ggcgctcttctcaaccagggggcgaccccaggcccaggattggggtggcaa ttggagaccagatcctggacctcagcatcattaagcacctctttacaggac ctgtgctgagcaagcaccaggatgtgttcaaccagcctacctgaatagct ttatgggactgggcccaggcagcatggaaagaagccagagtgttcctcca	52

5		gaatctgctgagcgtgagccaggccaggctcagggatgacacagaact gaggaatgtgccttcattcacaggcatcagccacaatgcacctccctgc cacaattggggactacactgacttctacagcagcagacagcatgccacta atgtggggattatgtttagagacaaggaaaatgccctcatgccaaattggct gcacctgcctgtgggctaccacggcagagcctcatcagtgggtgtcag gcacaccaattaggaggccaatgggacagatgaagcctgacgactcaa agccaccagtctatggcgctgcaaactgctggacatggagctggagatg gctttcttgtggggcctggcaacaggctgggagaaccaatccctatttcaa agggccacgagcacattttggcatgggtctcatgaatgactgggtctgccag agatatccagaagtgggaatacgtgccctggggccctttctgggcaaga gctttggcaccacagtgtcaccttgggtgggtcccaatggatgccctgatgcc ctttggcgtgccaacccaagcaggatcccagaccactgccctacctgt gccacgatgagccctatacctttgacatcaacctgtcagtcaacctgaagg gggagggcatgagccaggccgactatttgaagagcaactttaaatat atgtactggactatgtccaacagctcactcaccactcagtgaatggctgc aatctgaggcctggcgacctcctggctagcggcactatcttggccctgaa cctgagaactttgggagcatgctggagctgtcatggaaagggacaaagc ctattgacctggggaatggccagacaaggaagtttctcctggatggcgatg aggtcatcatcactgggtactgccagggcgatggctacaggattggatttg gacagtgtgctggcaaagtgtcccagctctgctgccctca	
25	ATP7B truncado humano	CCTGAGCAGGAGAGACAGATCACAGCCAGAGAAG GGGCCAGTCGGAAAATCTTATCTAAGCTTTCTTTG CCTACCCGTGCCTGGGAACCAGCAATGAAGAAGA GTTTTGCTTTTGACAATGTTGGCTATGAAGGTGGT CTGGATGGCCTGGGCCCTTCTTCTCAGCCGCAGA AGTGCTTCTTACAGATCAAAGGCATGACCTGTGCA TCCTGTGTGTCTAACATAGAAAGGAATCTGCAGAA AGAAGCTGGTGTCTCTCCGTGTTGGTTGCCTTGA	53

5	TGGCAGGAAAGGCAGAGATCAAGTATGACCCAGA GGTCATCCAGCCCCTCGAGATAGCTCAGTTCATCC AGGACCTGGGTTTTGAGGCAGCAGTCATGGAGGA CTACGCAGGCTCCGATGGCAACATTGAGCTGACAA TCACAGGGATGACCTGCGCGTCCTGTGTCCACAA CATAGAGTCCAAACTCACGAGGACAAATGGCATCA CTTATGCCTCCGTTGCCCTTGCCACCAGCAAAGCC CTTGTTAAGTTTGACCCGGAAATTATCGGTCCACG GGATATTATCAAAATTATTGAGGAAATTGGCTTTCA	
10	TGCTTCCCTGGCCCAGAGAAACCCCAACGCTCATC ACTTGGACCACAAGATGGAAATAAAGCAGTGGAAG AAGTCTTTCCTGTGCAGCCTGGTGTTTGGCATCCC TGTCATGGCCTTAATGATCTATATGCTGATACCCAG CAACGAGCCCCACCAGTCCATGGTCCTGGACCAC	
15	AACATCATTCCAGGACTGTCCATTCTAAATCTCATC TTCTTTATCTTGTGTACCTTTGTCCAGCTCCTCGGT GGGTGGTACTTCTACGTTCAAGGCCTACAAATCTCT GAGACACAGGTCAGCCAACATGGACGTGCTCATC GTCCTGGCCACAAGCATTGCTTATGTTTATTCTCTG	
20	GTCATCCTGGTGGTTGCTGTGGCTGAGAAGGCGG AGAGGAGCCCTGTGACATTCTTCGACACGCCCCC CATGCTCTTTGTGTTCAATTGCCCTGGGCGGTGGC TGGAACACTTGGCAAAGAGCAAAACCTCAGAAGCC CTGGCTAAACTCATGTCTCTCCAAGCCACAGAAGC	
25	CACCGTTGTGACCCTTGGTGAGGACAATTTAATCA TCAGGGAGGAGCAAGTCCCCATGGAGCTGGTGCA GCGGGGCGATATCGTCAAGGTGGTCCCTGGGGGA AAGTTTCCAGTGGATGGGAAAGTCCTGGAAGGCAA TACCATGGCTGATGAGTCCCTCATCACAGGAGAAG	
30	CCATGCCAGTCACTAAGAAACCCGGAAGCACTGTA	

5	ATTGCGGGGTCTATAAATGCACATGGCTCTGTGCT CATTAAAGCTACCCACGTGGGCAATGACACCACTT TGGCTCAGATTGTGAAACTGGTGGAAGAGGCTCA GATGTCAAAGGCACCCATTTCAGCAGCTGGCTGAC CGGTTTAGTGGATATTTTGTCCCATTTATCATCATC ATGTCAACTTTGACGTTGGTGGTATGGATTGTAATC GGTTTTATCGATTTTGGTGTGTTTCAGAGATACTTT CCTAACCCCAACAAGCACATCTCCCAGACAGAGGT GATCATCCGGTTTGCTTTCCAGACGTCCATCACGG	
10	TGCTGTGCATTGCCTGCCCTGCTCCCTGGGGCT GGCCACGCCCACGGCTGTCATGGTGGGCACCGG GGTGGCCGCGCAGAACGGCATCCTCATCAAGGGA GGCAAGCCCCTGGAGATGGCGCACAAGATAAAGA CTGTGATGTTTGACAAGACTGGCACCATTAACCAT	
15	GGCGTCCCCAGGGTCATGCGGGTGCTCCTGCTGG GGGATGTGGCCACACTGCCCCTCAGGAAGGTTCT GGCTGTGGTGGGGACTGCGGAGGCCAGCAGTGA ACACCCCTTGGGCGTGGCAGTCACCAAATACTGTA AAGAGGAACTTGGAACAGAGACCTTGGGATACTGC	
20	ACGGACTTCCAGGCAGTGCCAGGCTGTGGAATTG GGTGCAAAGTCAGCAACGTGGAAGGCATCCTGGC CCACAGTGAGCGCCCTTTGAGTGCACCGGCCAGT CACCTGAATGAGGCTGGCAGCCTTCCCGCAGAAA AAGATGCAGTCCCCCAGACCTTCTCTGTGCTGATT	
25	GGAAACCGTGAGTGGCTGAGGCGCAACGGTTTAA CCATTTCTAGCGATGTCAGTGACGCTATGACAGAC CACGAGATGAAAGGACAGACAGCCATCCTGGTGG CTATTGACGGTGTGCTCTGTGGGATGATCGCAATC GCAGACGCTGTCAAGCAGGAGGCTGCCCTGGCTG	
30	TGCACACGCTGCAGAGCATGGGTGTGGACGTGGT	

5		TCTGATCACGGGGGACAACCGGAAGACAGCCAGA GCTATTGCCACCCAGGTTGGCATCAACAAAGTCTT TGCAGAGGTGCTGCCTTCGCACAAGGTGGCCAAG GTCCAGGAGCTCCAGAATAAAGGGAAGAAAGTCG CCATGGTGGGGGATGGGGTCAATGACTCCCCGGC CTTGGCCCAGGCAGACATGGGTGTGGCCATTGGC ACCGGCACGGATGTGGCCATCGAGGCAGCCGACG TCGTCCTTATCAGAAATGATTTGCTGGATGTGGTG GCTAGCATTACCTTTCCAAGAGGACTGTCCGAAG 10 GATACGCATCAACCTGGTCCTGGCACTGATTTATA ACCTGGTTGGGATACCCATTGCAGCAGGTGTCTTC ATGCCCATCGGCATTGTGCTGCAGCCCTGGATGG GCTCAGCGGCCATGGCAGCCTCCTCTGTGTCTGT GGTGCTCTCATCCCTGCAGCTCAAGTGCTATAAGA 15 AGCCTGACCTGGAGAGGTATGAGGCACAGGCGCA TGGCCACATGAAGCCCCTGACGGCATCCCAGGTC AGTGTGCACATAGGCATGGATGACAGGTGGCGGG ACTCCCCCAGGGCCACACCATGGGACCAGGTCAG CTATGTCAGCCAGGTGTGCTGCTCCTCCCTGACGT 20 CCGACAAGCCATCTCGGCACAGCGCTGCAGCAGA CGATGATGGGGACAAGTGGTCTCTGCTCCTGAATG GCAGGGATGAGGAGCAGTACATC	
25	LDLR humano	gggccctggggctggaaattgcgctggaccgtcgccctgctcctcgccg gcggggactgcagtgggacagatgcgaaagaaacgagttccagtgc caagacgggaaatgcatctctacaagtgggtctgcatggcagcgctg agtgccaggatggctctgatgagtcaggagacgtgctgtgtcacctg caaatccggggacttcagctgtgggggcccgtgtcaaccgctgcattcctca 30 gttctggaggtgcatggccaagtggactgcgacaacggctcagacgag	54

5	caaggctgtcccccaagacgtgctcccaggacgagtttcgctgccacga tgggaagtgcattctcggcagttcgtctgtgactcagaccgggactgcttg gacggctcagacgaggcctcctgcccgggtgctcacctgtggtcccgccag cttcagtgcaacagctccacctgcatccccagctgtgggcctgcgaca acgacccccgactgcaagatggctcggatgagtggccgcagcgctgtag gggtctttacgtgttccaaggggacagtagcccctgctcggccttcgagttc cactgcctaagtggcgagtgcatccactccagctggcgctgtgatggggc cccgactgcaaggacaaatctgacgaggaaaactgcgctgtggccacct gtcgccctgacgaattccagtgctctgatggaaactgcatccatggcagcc	
10	ggcagtgtagccgggaatatgactgcaaggacatgagcgatgaagttgg ctgcgttaatgtgacactctgcgagggacccaacaagttaagtgtcacag cggcgaatgcatcaccctggacaaagtctgcaacatggctagagactgc cgggactggtcagatgaacccatcaaagagtgcgggaccaacgaatgc ttggacaacaacggcgggtgttcccacgtctgcaatgaccttaagatcggc	
15	tacgagtgcctgtgccccgacggcttcagctgggtggcccagcgaagatg cgaagatatcgatgagtgtaggatcccgacacctgcagccagctctgcg tgaacctggaggggtggctacaagtgccagtgtgaggaaggcttcagctg gacccccacacgaaggcctgcaaggctgtgggctccatgcctacctctt cttcaccaaccggcacgaggtcaggaagatgacgctggaccggagcga	
20	gtacaccagcctcatccccaacctgaggaacgtggctcgtctggacacgg aggtggccagcaatagaatctactggtctgacctgtcccagagaatgatct gcagcaccagcttgacagagcccacggcgtctcttctatgacaccgtc atcagcagagacatccaggcccccgacgggctggctgtggactggatcc	
25	acagcaacatctactggaccgactctgtcctgggcactgtctctgttgcgga taccaagggcgtgaagaggaaaacgttattcagggaacggctccaa gccaagggccatcggtggatcctgttcatggcttcatgtactggactgact ggggaactcccgccaagatcaagaaagggggcctgaatggtgtggaca	
30	tctactcgtggtgactgaaaacattcagtgggcccaatggcatcacctag atctcctcagtgggcgctctactgggttgactccaaacttactccatctca agcatcgatgtcaacgggggcaaccggaagaccatcttgaggatgaa	

5	aagaggctggcccccccttctccttgccgtctttgaggacaaagtatttg gacagatatcatcaacgaagccattttcagtgccaaccgcctcacagggtc cgatgtcaacttggtggctgaaaacctactgtccccagaggatatggttctctt ccacaacctcaccagccaagaggagtgaactggtgtgagaggaccac cctgagcaatggcggctgccagtatctgtgcctccctgccccgagatcaa ccccactcgcccaagttacctgcgctgcccgacggcatgtgtctggc cagggacatgaggagctgcctcacagaggctgaggctgcagtgggccac ccaggagacatccaccgtcaggctaaaggctcagctccacagccgtaag gacacagcacacaaccacccgacctgttcccgacacctcccggtgcct 10 ggggccacccctgggctcaccacgggtggagatagtacaatgtctcacc aagctctgggcgacgttgctggcagaggaaatgagaagaagcccagta gcgtgagggtctgtccattgtcctccccatcgtgctcctcgtcttcttgct gggggtcttcttctatggaagaactggcggcttaagaacatcaacagcat 15 caacttgacaaccccgctctatcagaagaccacagaggatgagggtccac atttgccacaaccaggacggctacagctacccctcgagacagatgggtca gtctggaggatgacgtggcg	
20	HAP humano tccactgcggctctggaaaaccaggcttgggcaggaaactctctgacttt ggacaggaaacaagctatattgaagacaactgcaatcaaatggtgcc tatcactgatcttctcactcaaagaagaagtgggtgcattggccaaagtattg cgcttattgaggagaatgatgtaaacctgaccacattgaatctagacctt ctcgtttaaagaaagatgagtatgaattttcaccatttgataaacgtagc ctgcctgctctgacaaacatcatcaagatcttgaggcatgacattggtgcc ctgtccatgagctttcacgagataagaagaaagacacagtgcctggttcc 25 caagaaccattcaagagctggacagatttgccaatcagattctcagctatg gagcggaactggatgctgaccaccctggtttaaagatcctgtgtaccgtgc aagacggaagcagtttgctgacattgcctacaactaccgcatgggcagc ccatccctcgagtgaatacatggaggaagaaaagaaaacatggggca cagtggtcaagactctgaagtccttgataaaacccatgcttgctatgagtac 30 aatcacattttccacttctgaaaagtactgtggcttcatgaagataacatt	55

5		cccagctggaagacgtttctcagttcctgcagactgcactggtttccgcct ccgacctgtggctggcctgcttctcctcctcgggatttcttgggtggcctggccttc cgagtcttccactgcacacagtacatcagacatggatccaagcccatgtat acccccgaacctgacatctgcatgagctgttgggacatgtgccctgttttc agatcgcagctttgccagttttccaggaaattggcctgcctctctgggtg cacctgatgaatacattgaaaagctcgccacaatttactggtttactgtgga gtttgggctctgcaaacaaggagactccataaaggcatatggtgctgggct cctgtcatcctttggtgaattacagtactgcttatcagagaagccaaagcttct cccctggagctggagaagacagccatccaaaattacactgtcacggag ttccagcccctctattacgtggcagagagtttaatgatgccaaggagaaa gtaaggaactttgctgccacaatacctcggcccttctcagttcgctacgacc catacacccaaaggattgaggtcttgacaatacccagcagcttaagattt tggctgattccattaacagtgaattggaatccttgcagtgccctccagaa aataaag	
15	ASL humano	gcctcggagagtgggaagctttgggggtggccggtttgtgggtgcagtggac cccatcatggagaagttcaacgcgtccattgcctacgaccggcacctttgg gaggtggatgttcaaggcagcaaagcctacagcaggggacctggagaag gcagggctcctcaccaaggccgagatggaccagatactccatggcctag acaaggtggctgaggagtgggcccagggcaccttcaaactgaactcaa tgatgaggacatccacacagccaatgagcgcgcctgaaggagctcatt ggtgcaacggcaggggaagctgcacacgggacggagccggaatgacc aggtggtcacagacctcaggctgtggatgcggcagacctgtccacgctc tcgggcctcctctgggagctcattaggacctggtggatcgggcagaggc ggaacgtgatgttcttcccggggtacaccatttgcagagggcccagcc catccgctggagccactggattctgagccacgccgtggcactgacctgag actctgagcggctgctggaggtgcggaagcggatcaatgtcctgcccctg gggagtggggccattgcaggcaatcccctgggtgtggaccgagagctgc tccgagcagaactcaacttggggccatcactctcaacagcatggatgcc actagtgagcgggactttgtggccgagttcctgttctgggcttcgctgtcat	56
20			
25			
30			

5		gacccatctcagcaggatggccgaggacctcatcctctactgcaccaagg aattcagcttcgtgcagctctcagatgcctacagcacgggaagcagcctg atgccccagaagaaaaaccccgacagtttgagctgatccggagcaag gctgggcgtgtgttggcggtgtgccgggctcctgatgacctcaaggga cttcccagcacctacaacaagacttacaggaggacaaggaagctgtgtt tgaagtgtcagacactatgagtgccgtgctccaggtggccactggcgtcat ctctacgctgcagattaccaagagaacatgggacaggctctcagcccc gacatgctggccactgacctgcctattacctgggtccgcaaagggatgcca ttccgccaggccccacgaggcctccgggaaagctgtgttcatggccgaga ccaaggggggtcgccctcaaccagctgtcactgcaggagctgcagacat cagccccctgttctcgggcgacgtgatctgctgtgggactacgggcaca gtgtggagcagtatggtgccctggcggcactgcgcgctccagcgtcgac tggcagatccgccagggtgcgggcgtactgcaggcacagcaggcc	
10			
15	G6PC humano	gaggaaggaatgaatgttctccatgactttgggatccagtcaacacattac ctccaggtgaattaccaagactcccaggactgggtcatcttgggtgccgtgat cgcagacctcaggaatgccttctacgtcctcttccccatctggttccatctca ggaagctgtgggcattaaactccttgggtagctgtgattggagactggctc aacctcgtctttaagtggattctcttggacagcgtccatactgggtgggttttg atactgactactacagcaacactccgtgcccctgataaagcagttccctgt aacctgtgagactggaccaggaggccctctggccatgcatgggcaca gcaggtgtatactacgtgatggtcacatctactcttccatcttcagggaaa gataaagccgacctacagattcgggtgcttgaatgtcattttgtgttgggatt ctgggctgtgcagctgaatgtctgtctgtcacgaatctacctgtgctcatttt cctcatcaagttgttgctggagtccgtgcaggcattgctgttcagaaaacttc agccacatccacagcatctataatgccagcctcaagaaatattttctatta ccttcttctgttcagcttcgccatcggattttatctgctgctcaagggactggg ttagacctcctgtggactctggagaaagcccagaggtggtgcgagcag ccagaatgggtccacattgacaccacaccccttgcagcctcctcaagaa cctgggcacgctcttggcctggggctggctctcaactccagcatgtacag	57
20			
25			
30			

5		ggagagctgcaaggggaaactcagcaagtggtcccattccgcctcagc tctattgtagcctccctcgctcctctgcacgtctttgactccttgaaaccccat cccaagtcgagctggctcttacgtcttgctcttgcaagagtgcggtagtg cccctggcatccgtcagtgatcccctactgcctcgcccaggctctgggcc agccgcacaagaagtcgttg	
10	ABCB11 humano	tctgactcagtaattcttcgaagtataaagaaattggagaggagaatgatg gtttgagtcagataaatcatataataatgataagaaatcaaggttacaaga tgagaagaaaggatgagcggttagagtggctctttcaattgttcggtttctt catcaactgacatttggtgatgtttgtggaagttgtgtgcatttccatgg aatagcccagccaggcgtgctactcattttggcacaatgacagatgtttta ttgactacgacgttgagttacaagaactccagattccaggaaaagcatgtg tgaataacaccattgtatggactaacagttccctcaaccagaacatgacaa atggaacacggttggttgctgaacatcgagagcgaaatgatcaaatttg ccagttactatgctggaattgctgtcgcagtacttatcacaggatatattcaa 15 atatgcttttgggtcattgccgcagctcgtcagatacagaaaatgagaaaat tttactttaggagaataatgagaatggaaatagggtggttgactgcaattca gtgggggagctgaatacaagattctctgatgatattaataaaatcaatgatg ccatagctgaccaaattggccctttcattcagcgcagacctcgaccatctgt 20 ggtttctgttgggattttcaggggttgaaactgaccttggttattttctgtc agccctctcattgggattggagcagccaccattggtctgagtggtccaagtt tacggactatgagctgaaggcctatgcaaagcaggggtggtggctgatg aagtcatttcatcaatgagaacagtggtctgtttggtggtgagaaaagag aggttgaaaggatgagaaaaatcttggttcgcccagcgtggggaatta gaaaaggaatagtgatgggattcttactggattcgtgtggtgtctcatctttt 25 gtgttatgcactggccttctggtacggctccacactgtcctggatgaaggag aatatacaccaggaaccctgtccagattttcctcagtgatagtaggagct ttaaattcttgcaatgcctctcctgttggaagcctttgcaactggacgtgca gcagccaccagcattttgagacaatagacaggaaacccatcattgactg catgtcagaagatggttacaagttggatcgaatcaagggtgaaattgaatt 30	58

5	ccataatgtgaccttccattatccttccagaccagaggtgaagattctaaatg acctcaacatggtcattaaaccaggggaaatgacagctctggtaggaccc agtggagctggaaaaagtacagcactgcaactcattcagcgattctatga cccctgtgaaggaatggtgaccgtggatggccatgacattcgctctcttaac attcagtggttagagatcagattgggtagtgagcaagagccagttctg ttctctaccaccattgcagaaaatattcgctatggcagagaagatgcaaca atggaagacatagtccaagctgccaaggaggccaatgcctacaactcat catggacctgccacagcaatttgacaccctgttgagaaggaggaggcc agatgagtggtagccagaaacaaaggtagctatgccagagccctcat 10ccgaaatcccaagattctgcttttgacatggccacctcagctctggacaat gagagtgaagccatggtgcaagaagtgtgagtaagattcagcatgggc acacaatcatttcagttgctcatcgctgtctacggtcagagctgcagatacc atcattggttttgaacatggcactgcagtggaagagggacccatgaaga attactggaaaggaaagggtgttacttcactctagtactttgcaaagccag 15ggaaatcaagctcttaataagaggacataaaggatgcaactgaagatg acatgcttgcgaggaccttagcagagggagctaccaggatagttaagg gcttccatccggcaacgctccaagtctcagcttcttacctggtgcacgaac ctccattagctgttagatcataagtctacctatgaagaagatagaaagga caaggacattcctgtgcaggaagaagtgaacctgccccagttaggagg 20attctgaaattcagtgctccagaatggccctacatgctggtagggctgtg gtgcagctgtgaacgggacagtcacaccctgtatgccttttattcagccag attcttgggacttttcaattcctgataaagaggaacaaagggtcacagatca atgggtgtgcctactttttagcaatgggctgtgtatctctttcacccaatttc tacagggatatgccttgcataatctggggagctcctaacaaaaaggctac 25gtaaatttggttcagggaatgctggggcaagatattgcctggttgatgac ctcagaaatagccctggagcattgacaacaagacttgctacagatgcttcc caagttcaaggggctgccggctctcagatcgggatgatagcaattccttca ctaacgtcactgtggcatgatcattgccttctccttttagctggaagctgagc ctggatcatctgtgcttctcccttcttggttatcaggagccacacagacca 30ggatgttgacaggattgcctctcgagataagcaggccctggagatggtgg	
---	--	--

5		gacagattacaaatgaagccctcagtaacatccgcactgttgctggaattg gaaaggagaggcggttcattgaagcacttgagactgagctggagaagcc cttcaagacagccattcagaaagccaatattacggattctgctttgccttgc ccagtgcacatgtttattgccaattctgcttctacagatatggaggttactta atctccaatgaggggctccatttcagctatgtgttcagggtgatctctgcagtt gtactgagtgaacagctcttggaagagccttctctacaccccaagttatg caaaagctaaaatatcagctgcacgctttttcaactgctggaccgacaac ccccaatcagtgatacaatactgcaggtgaaaaatgggacaacttcag gggaagattgattttgttgattgtaaattacatatccttctcgacctgactcgc 10 aagttctgaatggctctcagtgctgattagtcagggcagacactggcgttt gttgggagcagtggaatgtggcaaaagcactagcattcagctgttgaacgt ttctatgatcctgatcaaggaaggtgatgatagatggcatgacagcaaa aaagtaaatgtccagttcctccgctcaaacattggaattgttcccaggaac cagtggttgttgctgtagcataatggacaatatcaagtatggagacaacac 15 caaagaaattcccatggaaagagtcatagcagctgcaaaacagggtca gctgcatgattttgtcatgtcactcccagagaaatatgaaactaacgttgggt cccaggggtctcaactctctagaggggagaaacaacgcattgctattgctc gggccattgtacgagatcctaaaatcttgctactagatgaagccacttctgc cttagacacagaaagtgaagacgggtgcaggttgctctagacaaagcc 20 agagagggtcggacctgcattgtcattgcccacgcttgtccaccatccag aacgcggatatcattgctgtcatggcacaggggggtggtgattgaaaagg gacccatgaagaactgatggcccaaaaaggagcctactacaaactagtc accactggatcccccacagt	
25	ABC4 humano	gatcttgaggcggaagaacggaacagcctggcgccccacgagcgc ggagggcgactttgaactgggcatcagcagcaaaacaaaaaggaaaa aaacgaagacagtgaagaatgattggagtattaacattgttgcatactccg attggcaggataaattgtttatgtcgtgggtaccatcatggccatagctcac ggatcaggtctccccctcatgatgatagatttggagagatgactgacaaat ttgtgatactgcaggaaactctccttccagtgaaactttcctgtcgtgcta	59

5	aatccaggcaaaattctggaagaagaaatgactagatatgcatattactac tcaggattgggtgctggagttctgttgctgcctatatacaagtttcatttggac tttggcagctggtcgacagatcaggaaaattaggcagaagttttcatgcta ttctacgacaggaaataggatggttgacatcaacgacaccactgaactc aatacgcggttaacagatgacatctccaaaatcagtgaaggaattggtga caagggtggaatgttcttcaagcagtagccacgtttttgcaggattcatagt gggattcatcagaggatggaagctcacccttgataatggccatcagccc tattctaggactctctgcagccgttgggcaaagatactctcggcatttagtg acaaagaactagctgcttatgcaaaagcaggcgccgtggcagaagagg 10 ctctggggggccatcaggactgtgatagctttcggggggccagaacaaaga gctggaaggtatcagaaacatttagaaaatgccaaagagattggaatta aaaaagctatttcagcaaacattccatgggtattgccttctgttaatatg catcatatgcactggccttctggtatggatccactctagtcatatcaaaagaa tatactattggaaatgcaatgacagttttttcaatcctaattggagcttccagt 15 gttggccaggctgccccatgtattgatgcttttgccaatgcaagaggagca gcatatgtgatctttgatattattgataataatcctaaaattgacagttttcaga gagaggacacaaaccagacagcatcaaaggggaattggagttcaatgat gttcacttttctacccttctcgagctaacgtcaagatcttgaagggcctcaac ctgaagggtcagagtgggcagacggtggccctggttgaagtagtggtg 20 tgggaagagcacaacggtccagctgatacagaggctctatgaccctgat gagggcacaattaacattgatgggcaggatattaggaactttaatgtaaac tatctgagggaaatcattggtgtggtgagtcaggagccggtgctgtttccac cacaattgctgaaaaatattgttatggccgtggaaatgaaccatggatgag ataaagaaagctgtcaaagaggccaacgcctatgagtttatcatgaaatta 25 ccacagaaatttgacaccctggttggagagagagggggccagctgagtg gtgggcagaagcagaggatcgccattgcacgtgccctggtcgaaccc caagatccttctgctggatgaggccacgtcagcattggacacagaaagtg aagctgaggtacaggcagctctggataaggccagagaaggccggacc accattgtgatagcacaccgactgtctacgggtccgaaatgcagatgtcatc 30 gctgggttgaggatggagtaattgtggagcaaggaagccacagcgaac	
---	--	--

5	tgatgaagaaggaaggggtgacttcaaactgtcaacatgcagacatca ggaagccagatccagtcagaagaatttgaactaaatgatgaaaaggctg ccactagaatggcccaaatggctggaaatctgcctatttaggcattctac tcagaaaaaccttaaaaattcacaatgtgtcagaagagccttgatgtgga aaccgatggactgaagcaaattgtccaccagtgctcttctgaaggctctg aaactgaataaaacagaatggccctactttgtcgtgggaacagtatgtgcc attgccaatggggggcttcagccggcattttcagtcatttctcagagatcat agcgatttttgaccaggcgatgatgcagtgaagcagcagaagtgaac atattctcttgattttctatttctgggaattatttcttttactttctccttcagggtt tcacgtttgggaaagctggcgagatcctaccagaagactgcggtcaatg gcttttaaagcaatgctaagacaggacatgagctggttgatgaccataaa aacagtactggtgcactttctacaagactgccacagatgctgccaagtc caaggagccacaggaaccagggtggctttaattgcacagaatatagctaa ccttggaaactggtattatcatatcattatctacggttggcagtaaacctattg ctattagcagttgtccaattattgtgtgtcaggaattgtgaaatgaaattgtt ggctggaaatgccaaaagagataaaaaagaactggaagctgctggaa agattgcaacagaggcaatagaaaatattaggacagttgtgtctttgaccc aggaaagaaaattgaaatcaatgtatgtgaaaaattgtatggaccttacag gaattctgtgcagaaggcacacatctatggaattacttttagtatctcacaag catttatgtatttttctatgccggtgtttcgatttggtgcatactcattgtgaat ggacatatgcgcttcagagatgttattctggtgtttctgcaattgtatttggtgc agtggctctaggacatgccagttcattgtctccagactatgctaaagctaag ctgtctgcagcccacttattcatgctgtttgaaagacaacctctgattgacag ctacagtgaagaggggctgaagcctgataaattgaaggaaatataacat ttaatgaagtcgtgttcaactatcccacccgagcaaacgtgccagtgtctca ggggctgagcctggagggtgaagaaaggccagacactagccctggtggg cagcagtggtgtgggaagagcacggtggtccagctcctggagcggttct acgaccccttggcggggacagtgcttctcgatggtcaagaagcaaagaa actcaatgtccagtggtcagagctcaactcggaatcggtgtctcaggagcc tatcctatttgactgcagcattgccgagaatattgcctatggagacaacagc	
10		
15		
20		
25		
30		

5		cgggttgatcacaggatgaaattgtgagtgcagccaaagctgccaacat acatcctttcatcgagacgttaccacacaaatatgaaacaagagtgggag ataaggggactcagctctcaggagggtcaaaaacagaggattgctattgcc cgagccctcatcagacaacctcaaatcctcctgttgatgaagctacatca gctctggatactgaaagtgaaaagggtgtccaagaagccctggacaaag ccagagaaggccgcacctgcattgtgattgctaccgcctgtccaccatcc agaatgcagacttaatagtgggtgttcagaatgggagagtcaaggagcat ggcacgcatcagcagctgctggcacagaaaggcatctattttcaatggtc agtgtccaggctgggacacagaactta	
10	CBS humano	ccttctgagacccccaggcagaagtggggccacaggctgccccac cgctcaggggccacactcggcgaaggggagcctggagaaggggtcccc agaggataaggaagccaaggagcccctgttgatccggcccgatgtcc gagcagggtcacctggcagctggggccggcctgcctccgagtccccacat caccacactgccccggcaaaatctcaaaaatcttgccagatatctgaa gaaaatcggggacacccctatggtcagaatcaacaagattggaagaa gttcggcctgaagtgtgagctcttgccaagtgtgagttctcaacgcgggc gggagcgtgaaggaccgcatcagcctgcggatgattgaggatgctgagc gcgacggggacgtgaagcccggggacacgattatcgagccgacatccg ggaacaccgggatcgggctggccctggctgcggcagtgaggggctatcg ctgcatcatcgtgatgccagagaagatgagctccgagaagggtggacgtg ctgcgggcactgggggctgagattgtgaggacgcccaccaatgccaggtt cgactccccggagtcacacgtgggggtggcctggcggtgaagaacga aatccccaattctcacatcctagaccagtaccgcaacgccagcaaccccc tggctcactacgacaccaccgctgatgagatcctgcagcagtgatggg aagctggacatgctggtggcttcagtgggcacgggacggcaccatcacgg gcattgccaggaagctgaaggagaagtgctcctggatgcaggatcattggg gtggatcccgaagggtccatcctcgagagccggaggagctgaaccag acggagcagacaacctacgaggtggaagggatcggctacgacttcatcc ccacggtgctggacaggacgggtgggtggacaagtgggtcaagagcaacg	60
15			
20			
25			
30			

5		atgaggaggcggtcacctttgcccgcagtgctgatcgcgcaagaggggctg ctgtgcggtggcagtgctggcagcacgggtggcggtggccgtgaaggccg cgcaggagctgcaggagggccagcgctgctggtcattctgcccgactc agtgcggaactacatgaccaagttcctgagcgacaggtggatgctgcag aagggttttctgaaggaggaggacctcacggagaagaagccctgggtgt ggcacctccgtgttcaggagctgggcctgtcagccccgctgacctgtctc cgacctcacctgtgggcacacctcgagatcctccgggagaagggtt cgaccaggcgcccggtggatgaggcggggtaatcctgggaatggtg acgcttgggaacatgctctcgtccctgcttgcgggaagggtgcagccgtca gaccaagttggcaaagtcatctacaagcagttcaaacagatccgcctcac ggacacgctgggcaggctctcgacatcctggagatggaccacttcgccc tgggtggtgcacgagcagatccagtaccacagcaccgggaagtccagtca gcggcagatggtgttcgggggtgtcaccgccattgacttctgaacttcgtg gccgcccaggagcgggaccagaag	
15	hMMUT	ctgagagccaaaaaccagctgttctgctgagccccactatctgagaca ggtcaaagaaagttccgggagtagactgatccagcagagactgctgcac cagcagcagccactgcatcctgagtgggccgctctggccaagaaacag ctgaaggggcaaaaaccagaagacctgatctggcacactccagagggg atttcaatcaagcccctgtacagcaaaagggaactatggatctgccaga ggaactgccaggagtgaagcctttcacccgcgaccttaccgaactatgt ataccttcgacctggacaattcggcagtagccggcttcagtactgtgga ggaatcaacaagttttataaggacaacatcaaggctggacagcagggc ctgagtgtggcattcgatctggccacacatcgcggtatgactcagataatc ccagagtcaggggggacgtgggaatggcaggagtgcctatcgacacag tggagatactaagattctgttcgatggaatccctctggagaaaatgtctgtg agtatgacaatgaacggcgctgtcattcccgtgctggcaaaacttcacgtca ctggcgaggaacagggggtgcctaaggaaaaactgaccggcacaattc agaacgacatcctgaaggagttcatggtgcggaatacttacattttccccct gaacctcatgaaaatcattgccgatatcttcgagtacaccgctaagcac	61
20			
25			
30			

5

10

15

20

25

30

	atgcccaagttcaactcaattagcatctccgggtatcatatgcaggaagca ggagccgacgctattctggagctggcttacacctggcagatggcctgga atattctgaaccggactgcaggcaggcctgacaatcgacgagttcgctc ctagactgagtttctttggggaattggcatgaactttacatggagatcgcca agatgagggctggccggagactgtgggcacacctgatcgagaagatgtt ccagcctaagaactctaagagtctgctgctgcggggccattgccagacat ccggctggtctctgactgaacaggacccatataacaatattgtcagaaccg caatcgaggcaatggcagccgtgttcggaggaacccagagcctgcaca caaactcctttgatgaggccctggggctgcctaccgtgaagtctgtaggat tgcacgcaatacacagatcattatccaggaggaatccggaatcccaaag gtggccgatccctggggaggctcttacaatgatggagtgcctgacaaacga cgtgtatgatgctgcactgaagctgattaatgaaatcgaggaaatggggg gaatggcaaaggccgtggctgagggcattccaaaactgaggatcgagg aatgtgcagctaggcgccaggcacgaattgactcaggaagcgaagtgat cgtcgggtgaataagtaccagctggagaaagaagacgcagtcgaagt gctggccatcgataacacaagcgtgcaatcgacagattgagaagctg aagaaaatcaaaagctcccgcatcaggcactggccgaacgatgcctg gcagccctgactgagtgtgctgcaagcggggacggaaacattctggctct ggcagtcgatgcctcccggttagatgcactgtgggggaaatcacccgac gccctgaagaaagtcttcggagagcacaaggccaatgatcggatggtga gcggcgcttatagacaggagtccgggaatctaaagagattaccagtgcc atcaagaggggtgcacaagttcatggagagagaagggcgacggcccag gctgctggtggcaaagatgggacaggacggacatgatcgcgagcaaaa agtcattgccaccgggttcgctgacctgggattgacgtggatatcggccct ctgttccagacaccacgagaggtcgcacagcaggcagtcgacgctgatg tgcacgcagtcggagtgtccactctggcagctggccataagaccctggtg cctgaactgatcaaagagctgaactctctgggcagaccagacatcctggt catgtgcggcggtgatccacccaggattacgaattcctgtttgaggtc ggggtgagcaacgtgttcggaccaggaaccaggatccctaaggccgca	
--	--	--

	gtgcaggtcctggatgatattgaaaagtgctggaagaaacagcagtc agtg	
--	---	--

Brazos de homología

5 En algunas realizaciones, los vectores víricos descritos en el presente documento comprenden una o más secuencias de polinucleótidos flanqueantes con homología de secuencia significativa con un sitio de integración diana (p. ej., brazos de homología). En algunas realizaciones, los brazos de homología flanquean una secuencia de polinucleótidos que codifica una carga útil (p. ej., un brazo de homología en posición 5' con respecto a una carga útil (también denominado en el
10 presente documento brazo de homología en 5') y un brazo de homología en posición 3' con respecto a una carga útil (también denominado en el presente documento brazo de homología en 3')). En algunas realizaciones, los brazos de homología dirigen la integración específica de sitio de una carga útil.

15 En algunas realizaciones, los brazos de homología tienen la misma longitud (también denominados en el presente documento brazos de homología equilibrados o brazos de homología uniformes). En algunas realizaciones, los vectores víricos comprenden brazos de homología de la misma longitud, en donde los brazos de homología que tienen al menos una determinada longitud, proporcionan efectos mejorados (p. ej., tasa mejorada de integración de diana). En algunas realizaciones,
20 los brazos de homología tienen una longitud de entre 50 nt y 1600 nt. En algunas realizaciones, los brazos de homología tienen una longitud de entre 100 nt y 1000 nt. En algunas realizaciones, los brazos de homología tienen una longitud de entre 200 nt y 1000 nt. En algunas realizaciones, los brazos de homología tienen una longitud de entre 500 nt y 1500 nt. En algunas realizaciones, los brazos de
25 homología tienen una longitud de entre 1000 nt y 2000 nt. En algunas realizaciones, los brazos de homología tienen una longitud superior a 2000 nt. En algunas realizaciones, cada brazo de homología tiene una longitud de al menos 50 nt. En algunas realizaciones, cada brazo de homología tiene una longitud de al menos 750 nt. En algunas realizaciones, cada brazo de homología tiene una longitud de al
30 menos 1000 nt. En algunas realizaciones, cada brazo de homología tiene una

longitud de al menos 1250 nt. En algunas realizaciones, los brazos de homología tienen una longitud inferior a 1000 nt. En algunas realizaciones, los brazos de homología contienen al menos un 70 % de homología con un sitio de integración diana. En algunas realizaciones, los brazos de homología contienen al menos un 80 % de homología con un sitio de integración diana. En algunas realizaciones, los brazos de homología contienen al menos un 90 % de homología con un sitio de integración diana. En algunas realizaciones, los brazos de homología contienen una homología de al menos 95 % con un sitio de integración diana. En algunas realizaciones, los brazos de homología contienen una homología de al menos 99 % con un sitio de integración diana. En algunas realizaciones, los brazos de homología contienen una homología del 100 % con un sitio de integración diana.

En algunas realizaciones, los brazos de homología tienen diferentes longitudes (también denominados en el presente documento brazos de homología no equilibrados o brazos de homología desiguales). En algunas realizaciones, los vectores víricos que comprenden brazos de homología no equilibrados de diferentes longitudes, proporcionan efectos mejorados (p. ej., aumento de la tasa de integración del sitio diana) en comparación con una secuencia de referencia apropiada. En algunas realizaciones, los vectores víricos que comprenden brazos de homología de diferentes longitudes, en donde cada brazo de homología tiene al menos una determinada longitud, proporcionan efectos mejorados (p. ej., aumento de la tasa de integración del sitio diana) en comparación con una secuencia de referencia apropiada (p. ej., un vector vírico que comprende brazos de homología de la misma longitud o un vector vírico que comprende uno o más brazos de homología de menos de 1000 nt de longitud).

En algunas realizaciones, cada brazo de homología tiene una longitud superior a 50 nt. En algunas realizaciones, cada brazo de homología tiene una longitud superior a 100 nt. En algunas realizaciones, cada brazo de homología tiene una longitud superior a 400 nt. En algunas realizaciones, cada brazo de homología tiene una longitud superior a 750 nt. En algunas realizaciones, cada brazo de homología tiene una longitud de al menos 1000 nt. En algunas realizaciones, un

[illegible]

- brazo de homología tiene una longitud de al menos 1500 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1000 nt y otro brazo de homología tiene una longitud de al menos 1600 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1000 nt y otro
- 5 brazo de homología tiene una longitud de al menos 1700 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1000 nt y otro brazo de homología tiene una longitud de al menos 1800 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1000 nt y otro
- 10 brazo de homología tiene una longitud de al menos 1900 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1000 nt y otro brazo de homología tiene una longitud de al menos 2000 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1300 nt y otro
- 15 brazo de homología tiene una longitud de al menos 1400 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1600 nt y otro brazo de homología tiene una longitud de al menos 1000 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 1250 nt y otro
- 20 brazo de homología tiene una longitud de al menos 1250 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 400 nt y otro brazo de homología tiene una longitud de al menos 800 nt. En algunas realizaciones, un brazo de homología tiene una longitud de al menos 600 nt y otro
- brazo de homología tiene una longitud de al menos 600 nt.

- En algunas realizaciones, un brazo de homología en 5' es más largo que un brazo de homología en 3'. En algunas realizaciones, un brazo de homología en 3' es más largo que un brazo de homología en 5'. Por ejemplo, en algunas
- 25 realizaciones, un brazo de homología en 5' tiene una longitud de aproximadamente 1600 nt y un brazo de homología en 3' tiene una longitud de aproximadamente 1000 nt. En algunas realizaciones, un brazo de homología en 5' tiene una longitud de aproximadamente 1000 nt y un brazo de homología en 3' tiene una longitud de aproximadamente 1600 nt. En algunas realizaciones, los vectores víricos que
- 30 comprenden brazos de homología proporcionan efectos mejorados (p. ej., aumento

de la tasa de integración del sitio diana) en comparación con una secuencia de referencia apropiada (p. ej., vectores víricos que carecen de brazos de homología). En algunas realizaciones, los vectores víricos que comprenden brazos de homología proporcionan tasas de integración en el sitio diana del 0.01 % o más (p. ej., 0.05 % o más, 0.1 % o más, 0.2 % o más, 0.3 % o más, 0.4 % o más, 0.5 % o más, 0.6 % o más, 0.7 % o más, 0.8 % o más, 0.9 % o más, 1 % o más, 1.5 % o más, 2 % o más, 5 % o más, 10 % o más, 20 % o más, 30 % o más). En algunas realizaciones, los vectores víricos que comprenden brazos de homología proporcionan mayores tasas de integración del sitio diana a lo largo del tiempo. En algunas realizaciones, las tasas de integración del sitio diana aumentan a lo largo del tiempo en relación con una medición inicial de la integración del sitio diana. En algunas realizaciones, las tasas de integración del sitio diana a lo largo del tiempo son al menos 1.5 veces superiores a la de una medida inicial de integración del sitio diana (p. ej., 1.5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 20 veces, 30 veces, 40 veces, 50 veces, 60 veces, 70 veces, 80 veces, 90 veces, 100 veces, 200 veces). En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más días. En algunas realizaciones, las tasas de integración del sitio diana se miden después de una o más semanas. En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más meses. En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más años.

En algunas realizaciones, los vectores víricos que comprenden brazos de homología de diferentes longitudes proporcionan efectos mejorados (p. ej., aumento de la tasa de integración del sitio diana) en relación con una secuencia de referencia (p. ej., vectores víricos con brazos de homología de la misma longitud, vectores víricos con al menos un brazo de homología inferior a 500 nt). En algunas realizaciones, los vectores víricos que comprenden brazos de homología de diferentes longitudes proporcionan una actividad de edición mejorada de al menos 1.1 veces, al menos 1.2 veces, al menos 1.3 veces, al menos 1.4 veces, al menos 1.5 veces, al menos 1.6 veces, al menos 1.7 veces, al menos 1.8 veces, al menos

1.9 veces, al menos 2.0 veces, al menos 2.5 veces, al menos 3.0 veces, al menos 3.5 veces o de al menos 4.0 veces en relación con una composición de referencia (p. ej., vectores víricos con brazos de homología de la misma longitud, vectores víricos con al menos un brazo de homología inferior a 500 nt).

5 En algunas realizaciones, los vectores víricos que comprenden brazos de homología de diferentes longitudes proporcionan tasas de integración del sitio diana del 0.01 % o más (p. ej., del 0.05 % o más, 0.1 % o más, 0.2 % o más, 0.3 % o más, 0.4 % o más, 0.5 % o más, 0.6 % o más, 0.7 % o más, 0.8 % o más, 0.9 % o más, 1 % o más, 1.5 % o más, 2 % o más, 5 % o más, 10 % o más, 20 % o más, 30 % o más).

10 En algunas realizaciones, los vectores víricos que comprenden brazos de homología de diferentes longitudes proporcionan mayores tasas de integración del sitio diana a lo largo del tiempo. En algunas realizaciones, las tasas de integración del sitio diana aumentan a lo largo del tiempo en relación con una medición inicial de la integración del sitio diana. En algunas realizaciones, las tasas de integración

15 del sitio diana a lo largo del tiempo son al menos 1.5 veces superiores a la de una medida inicial de integración del sitio diana (p. ej., 1.5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 20 veces, 30 veces, 40 veces, 50 veces, 60 veces, 70 veces, 80 veces, 90 veces, 100 veces, 200 veces).

20 En algunas realizaciones, los vectores víricos que comprenden brazos de homología de longitudes diferentes pueden ofrecer una mejora de la reescritura genómica en una especie o en un sistema modelo particular para una especie particular (p. ej., un ratón, un ser humano, o modelos de los mismos). En algunas realizaciones, los vectores víricos pueden comprender diferentes combinaciones de longitudes de brazos de homología cuando se optimizan para la expresión en una

25 especie o en un sistema modelo particular para una especie particular (p. ej., un ratón, un ser humano, o modelos de los mismos). En algunas realizaciones, los vectores víricos que comprenden combinaciones de longitudes de brazos de homología específicas pueden ofrecer una mejora de la reescritura genómica en una especie o en un sistema modelo de una especie (p. ej., un ser humano, modelo

30 de ratón humanizado) en comparación con una segunda especie o con un sistema

modelo de una segunda especie (p. ej., un ratón, modelo de ratón puro). En algunas realizaciones, los vectores víricos que comprenden combinaciones de longitudes de brazos de homología específicas pueden optimizarse con respecto a niveles altos de reescritura genómica en una especie o en un modelo de una especie (p. ej., un ser humano, modelo de ratón humanizado) en comparación con una segunda especie o con un sistema modelo de una segunda especie (p. ej., un ratón, modelo de ratón puro).

En algunas realizaciones, los brazos de homología dirigen la integración de un transgén inmediatamente detrás de un gen endógeno altamente expresado. En algunas realizaciones, los brazos de homología dirigen la integración de un transgén sin que se altere la expresión del gen endógeno (integración no alterada).

En algunas realizaciones, una o más secuencias de brazo de homología pueden tener una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej., una secuencia genómica de tipo silvestre). En algunas realizaciones, una o más secuencias de brazo de homología pueden ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una parte de una secuencia de nucleótidos de referencia de tipo silvestre correspondiente (p. ej., una secuencia genómica de tipo silvestre). En algunas realizaciones, una o más secuencias del brazo de homología pueden ser o comprender una secuencia que tenga una identidad de al menos 80 %, 85 %, 90 %, 95 %, 99 % o del 100 % con una secuencia ilustrativa de la Tabla 7 mostrada a continuación. En algunas realizaciones, las composiciones descritas en el presente documento pueden comprender una o más secuencias de brazos de homología seleccionadas de la Tabla 7 mostrada a continuación.

Tabla 7: Secuencias de brazos de homología ilustrativas

Descripción del HA	Secuencia	SEQ ID NO
Brazo de homología en 5' de 0.4 kb de albúmina humana	GTGATGCTTATGAATATTAATAGGAATATTTGTAAG GCCTGAAATATTTTGATCATGAAATCAAAACATTAA TTTATTTAAACATTTACTTGAAATGTGGTGGTTTGT GATTTAGTTGATTTTATAGGCTAGTGGGAGAATTTA CATTCAAATGTCTAAATCACTTAAAATTGCCCTTTAT GGCCTGACAGTAACTTTTTTTTATTCATTTGGGGAC AACTATGTCCGTGAGCTTCCGTCCAGAGATTATAG TAGTAAATTGTAATTAAAGGATATGATGCACGTGAA ATCACTTTGCAATCATCAATAGCTTCATAAATGTTA ATTTTGTATCCTAATAGTAATGCTAATATTTTCCTAA CATCTGTCATGTCTTTGTGTTTCAGGGTAAAAAACTT GTTGCTGCAAGTCAAGCTGCCTTAGGCTTA	62
Brazo de homología en 5' de 0.6 kb de albúmina humana	GCATGTTTGGTTAGGCTAGGGCTTAGGGATTTATA TATCAAAGGAGGCTTTGTACATGTGGGACAGGGAT CTTATTTTACAAACAATTGTCTTACAAAATGAATAAA ACAGCACTTTGTTTTTATCTCCTGCTCTATTGTGCC AACTGTAAATGTTTATAATGCCTGTTCTGTTTCC AAATTTGTGATGCTTATGAATATTAATAGGAATATTT GTAAGGCCTGAAATATTTTGATCATGAAATCAAAAC ATTAATTTATTTAAACATTTACTTGAAATGTGGTGGT TTGTGATTTAGTTGATTTTATAGGCTAGTGGGAGAA TTTACATTCAAATGTCTAAATCACTTAAAATTGCCCT TTATGGCCTGACAGTAACTTTTTTTTATTCATTTGG GGACAACATATGTCCGTGAGCTTCCGTCCAGAGATT ATAGTAGTAAATTGTAATTAAAGGATATGATGCACG	63

5		TGAAATCACTTTGCAATCATCAATAGCTTCATAAAT GTAAATTTTGTATCCTAATAGTAATGCTAATATTTTC CTAACATCTGTCATGTCTTTGTGTTTCAGGGTAAAAA ACTTGTTGCTGCAAGTCAAGCTGCCTTAGGCTTA	
10	Brazo de homología en 5' de 0.8 kb de albúmina humana	TTCAAACCTCAGTGCACTTGTTGAGCTTGTGAAACA CAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCT GTTATGGATGATTTTCGCAGCTTTTGTAGAGAAGTG CTGCAAGGCTGACGATAAGGAGACCTGCTTTGCC GAGGAGGTACTACAGTTCTCTTCATTTTAATATGTC CAGTATTCATTTTTGCATGTTTGGTTAGGCTAGGGC TTAGGGATTTATATATCAAAGGAGGCTTTGTACATG TGGGACAGGGATCTTATTTTACAAACAATTGTCTTA CAAAATGAATAAAACAGCACTTTGTTTTTATCTCCT GCTCTATTGTGCCATACTGTAAATGTTTATAATGC CTGTTCTGTTTCCAAATTTGTGATGCTTATGAATAT TAATAGGAATATTTGTAAGGCCTGAAATATTTTGAT CATGAAATCAAAACATTAATTTATTTAAACATTTACT TGAAATGTGGTGGTTTGTGATTTAGTTGATTTTATA GGCTAGTGGGAGAATTTACATTCAAATGTCTAAATC ACTTAAAATTGCCCTTTATGGCCTGACAGTAACTTT TTTTTATTCATTTGGGGACAACCTATGTCCGTGAGCT TCCGTCCAGAGATTATAGTAGTAAATTGTAATTAAA GGATATGATGCACGTGAAATCACTTTGCAATCATC AATAGCTTCATAAATGTTAATTTTGTATCCTAATAGT AATGCTAATATTTTCCTAACATCTGTCATGTCTTTGT GTTTCAGGGTAAAAAACTTGTTGCTGCAAGTCAAGC TGCCTTAGGCTTA	64
20			
25			
30			

5	Brazo de homología en 3' de 0.4 kb de albúmina humana	TAACATCACATTTTAAAAGCATCTCAGGTAAGTATAT TTTGAATTTTTTAAAAAAGTAACTATAATAGTTATTA TTAAAATAGCAAAGATTGACCATTTCCAAGAGCCAT ATAGACCAGCACCGACCACTATTCTAAACTATTTAT GTATGTAAATATTAGCTTTTAAAATTCTCAAAATAGT TGCTGAGTTGGGAACCACTATTATTTCTATTTTGTA GATGAGAAAATGAAGATAAACATCAAAGCATAGATT AAGTAATTTTCCAAGGGTCAAAATTCAAAATTGAA ACCAAAGTTTCAGTGTTGCCCATTTGTCCTGTTCTGA CTTATATGATGCGGTACACAGAGCCATCCAAGTAA GTGATGGCTCAGCAGTGGAATACTCTGGGAATTAG GCTGAACCACATGAAAGAGTGCTTTATA	65
15	Brazo de homología en 3' de 0.6 kb de albúmina humana	TAACATCACATTTTAAAAGCATCTCAGGTAAGTATAT TTTGAATTTTTTAAAAAAGTAACTATAATAGTTATTA TTAAAATAGCAAAGATTGACCATTTCCAAGAGCCAT ATAGACCAGCACCGACCACTATTCTAAACTATTTAT GTATGTAAATATTAGCTTTTAAAATTCTCAAAATAGT TGCTGAGTTGGGAACCACTATTATTTCTATTTTGTA GATGAGAAAATGAAGATAAACATCAAAGCATAGATT AAGTAATTTTCCAAGGGTCAAAATTCAAAATTGAA ACCAAAGTTTCAGTGTTGCCCATTTGTCCTGTTCTGA CTTATATGATGCGGTACACAGAGCCATCCAAGTAA GTGATGGCTCAGCAGTGGAATACTCTGGGAATTAG GCTGAACCACATGAAAGAGTGCTTTATAGGGCAAA AACAGTTGAATATCAGTGATTTACATGGTTCAACC TAATAGTTCAACTCATCCTTTCCATTGGAGAATATG ATGGATCTACCTTCTGTGAACTTTATAGTGAAGAAT	66

		CTGCTATTACATTTCCAATTTGTCAACATGCTGAGC TTTAATAGGACTTATCTTCTTATGACAACATTTATTG	
5	Brazo de homología en 3' de 0.8 kb de albúmina humana	TAACATCACATTTAAAAGCATCTCAGGTA ACTATAT TTTGAATTTTTTAAAAAAGTAACTATAATAGTTATTA TTAAAATAGCAAAGATTGACCATTTCCAAGAGCCAT ATAGACCAGCACCGACCACTATTCTAAACTATTTAT GTATGTAAATATTAGCTTTTAAAATTCTCAAAATAGT TGCTGAGTTGGGAACCACTATTATTTCTATTTTGTA GATGAGAAAATGAAGATAAACATCAAAGCATAGATT AAGTAATTTTCCAAAGGGTCAAAATTCAAATTGAA ACCAAAGTTTCAGTGTTGCCCATTTGTCCTGTTCTGA CTTATATGATGCGGTACACAGAGCCATCCAAGTAA GTGATGGCTCAGCAGTGGAATACTCTGGGAATTAG GCTGAACCACATGAAAGAGTGCTTTATAGGGCAAA AACAGTTGAATATCAGTGATTTACATGGTTCAACC TAATAGTTCAACTCATCCTTTCCATTGGAGAATATG ATGGATCTACCTTCTGTGAACTTTATAGTGAAGAAT CTGCTATTACATTTCCAATTTGTCAACATGCTGAGC TTTAATAGGACTTATCTTCTTATGACAACATTTATTG GTGTGTCCCCTTGCCTAGCCCAACAGAAGAATTCA GCAGCCGTAAGTCTAGGACAGGCTTAAATTGTTTT CACTGGTGTAATTGCAGAAAGATGATCTAAGTAAT TTGGCATTATTTTAAATAGGTTTGAAAAACACATGC CATTTTACAAATAAGACTTATATTTGTCCTTTTGTTT TTCAGCCTACCATGAG	67
30	Brazo de homología en	tttctagatgtaaataattatTTAagtttgccctatggtggccccacacatgag acaaaccccccaagatgtgacttttgagaatgagacttgataaaaaaacat	68

5'	de 1.6 kb	gtgaaatgcaagccctgaagctcaactccctattgctatcacaggggttat
	de albúmina	aattgcataaaatttagctatagaaagttgctgtcatctcttgtgggctgtaatc
	humana	atcgtctaggctaagagtaatatgtcaaaacctgtcatgcccacacaaaatc
5		tctccctggcattgttgtctttgcagatgtcagtgaaagagaaccagcagctc
		ccatgagtttgatagccttattttctatagcctccccactattagctttgaagg
		gagcaaagttaagaaccaaataaaagtttctcatctttatagatgagaaa
		aattttaaataaagtccaagataattaaattttaaggatcatttttagctcttta
		atagcaataaaaactcaatatgacataatatggcacttccaaaaactgaata
10		atatataattgcaatgacatacttctttcagagatttactgaaaagaaattgt
		tgacactacataacgtgatgagtggttatactgattgtttcagttggctctccc
		accaactccatgaaagtggattttattatcctcatcatgcagatgagaatatt
		gagacttatagcggatgcctgagcccaaagtactcagagttgcctggct
		ccaagatttataatcttaaatgatgggactaccatccttactctctccattttct
15		atacgtgagtaatgtttttctgtttttttttcttttccattcaaactcagtcactt
		gttgagctcgtgaaacacaagcccaaggcaacaaaagagcaactgaa
		agctgttatggatgatttcgcagctttttagagaagtgtgcaaggctgacg
		ataaggagacctgctttgccgaggaggtagctacagttctcttcattttaatatg
		tccagtattcattttgcatgtttggttaggctagggcttagggatttatatca
20		aaggaggctttgtacatgtgggacagggatcttattttacaacaattgtctt
		acaaaatgaataaaacagcactttgttttatctcctgctctattgtgccatact
		gttaaatgtttataatgcctgttctgtttccaaattgtgatgcttatgaatattaat
		aggaatatttgaaggcctgaaatattttgatcatgaaatcaaacattaattt
		atttaaacatttacttgaaatgtggtgggttgatttagttgattttataggctagt
25		gggagaatttacattcaaagtgtctaaatcacttaaaattgccctttatggcctg
		acagtaactttttttattcatttggggacaactatgtccgtgagctccgtcca
		gagattatagtagtaaattgtaattaaaggatatgatgcacgtgaaatcactt
		tgcaatcatcaatagcttcataaatgttaattttgtatcctaataagtaatgcta
		attttcctaacatctgtcatgtctttgtgttcagggtaaaaaactgttgctgcaa
30		gtcaagctgccttaggctta

Brazo de homología en 3' de 1 kb de albúmina humana	Taacatcacatttaaagcatctcaggtaaactatatttgaatttttaaaaaa gtaactataatagttattattaaaatagcaaagattgaccatttccaagagcc atatagaccagcaccgaccactatttctaaactatttatgtatgtaaattatg cttttaaaattctcaaaaatagttgctgagttgggaaccactattatttctatttgt agatgagaaaaatgaagataaacatcaaagcatagattaagtaatttcca aagggtcaaaattcaaaatgaaaccaaagtttcagtggtgccattgtcct gttctgacttatatgatgcggtacacagagccatccaagtaagtgtggctc agcagtggaaatactctgggaattaggctgaaccacatgaaagagtgttta tagggcaaaaacagttgaatatcagtgatttcacatggttcaacctaatagtt caactcatcctttccattggagaatatgatggatctacctctgtgaactttata gtgaagaatctgctattacatttccaatttgcacatgctgagctttaatagg acttatcttctatgacaacatttattggtgtgtcccctgcctagcccaacaga agaattcagcagccgtaagctaggacaggcttaaattgtttcactgggtga aattgcagaaagatgatctaagtaatttggcatttatttaataggttgaaaa acacatgccattttacaaataagacttatattgtcctttgttttcagcctacca tgagaataagagaaaagaaaatgaagatcaaaagcttattcatctgttttctt tttcgttgggtgaaagccaacaccctgtctaaaaaacataaatttcttaatca tttgcctcttttctctgtgcttcaattaataaaaaatggaaagaatctaataga gtggtacagcactgttattttcaaagatgtgttg	69
Brazo de homología en 5' de 1 kb de albúmina humana	Actccatgaaagtggattttattatcctcatcatgcagatgagaatattgaga cttatagcggatgcctgagcccaaaagtactcagagttgcctgggtccaa gatttataatcttaaagatgggactaccatccttactctctcattttctatacg tgagtaatgtttttctgttttttttttcttttccattcaaactcagtgacttgtga gctcgtgaaacacaagcccaaggcaacaaaagagcaactgaaagctg ttatggatgatttcgcagctttttagagaagtgctgcaaggctgacgataa ggagacctgcttgcgaggagggtactacagttctcttatttaatatgtcca gtattcattttgcatgtttggtaggctagggttagggatttatatacaagg aggctttgtacatgtgggacagggtatcttattttacaacaattgtcttcaaaa	70

5	atgaataaaacagcactttgttttatctcctgctctattgtgccatactgttaaa tgtttataatgcctgttctgtttccaaatttgatgcttatgaatattaataggaa tatttgtaaggcctgaaatatttgatcatgaaatcaaacattaatttatttaa acatttacttgaaatgtggtggtttgtgatttagttgatttataggctagtggga gaatttacattcaaagtctaaatcacttaaaattgccctttatggcctgacag taactttttttattcatttggggacaactatgtccgtgagcttccgtccagagat tatagtagtaaattgtaattaaaggatatgatgcacgtgaaatcactttgcaa tcatcaatagcttcataaatgttaattttgtatcctaataagtaatgctaattttc ctaacatctgtcatgtctttgtgttcagggtaaaaaactgttgctgcaagtca 10 agctgccttaggctta	
---	---	--

En algunas realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender un brazo de homología en 5' y un brazo de homología en 3' diseñados para dirigirse a un locus de albúmina. En algunas realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender una secuencia de brazo de homología en 5' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 62 y una secuencia de brazo de homología en 3' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 67. En algunas realizaciones, un vector vírico comprende un brazo de homología en 5' que comprende la secuencia de SEQ ID NO: 62 y un brazo de homología en 3' que comprende la secuencia de SEQ ID NO: 67.

En algunas realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender una secuencia del brazo de homología en 5' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 63 y una secuencia de brazo de homología en 3' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos

85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 66. En algunas realizaciones, un vector vírico comprende un brazo de homología en 5' que comprende la secuencia de SEQ ID NO: 63 y un brazo de homología en 3' que comprende la secuencia de SEQ ID NO: 66.

5 En algunas realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender una secuencia del brazo de homología en 5' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 64 y una secuencia de brazo de homología en 3' que tenga una
10 identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 65. En algunas realizaciones, un vector vírico comprende un brazo de homología en 5' que comprende la secuencia de SEQ ID NO: 64 y un brazo de homología en 3' que comprende la secuencia de SEQ ID NO: 65.

15 En algunas realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender una secuencia del brazo de homología en 5' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 68 y una secuencia de brazo de homología en 3' que tenga una
20 identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 69. En algunas realizaciones, un vector vírico comprende un brazo de homología en 5' que comprende la secuencia de SEQ ID NO: 68 y un brazo de homología en 3' que comprende la secuencia de SEQ ID NO: 69. En algunas
25 realizaciones, los vectores víricos proporcionados en el presente documento pueden comprender una secuencia del brazo de homología en 5' que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 70 y una secuencia de brazo de homología en 3' que tenga una identidad de

secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con la SEQ ID NO: 69.

Medición de la integración del sitio diana

Como se describe en otra parte del presente documento, uno de los problemas que surgen con el uso tradicional de nucleasas para introducir material de ácido nucleico en las células, es la posibilidad significativa de integración inespecífica. En consecuencia, es importante verificar la correcta integración mediante uno o más ensayos específicamente dirigidos, como se describe a continuación.

De acuerdo con diversas realizaciones, la tasa de integración puede medirse en cualquiera de una variedad de puntos temporales. En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más días. En algunas realizaciones, las tasas de integración del sitio diana se miden después de una o más semanas. En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más meses. En algunas realizaciones, las tasas de integración del sitio diana se miden después de uno o más años. En algunas realizaciones, las tasas de integración del sitio diana se miden a través de la evaluación de uno o más biomarcadores (p. ej., biomarcadores que comprenden péptido 2A). En algunas realizaciones, las tasas de integración del sitio diana se miden a través de la evaluación de uno o más ácidos nucleicos aislados (p. ej., ARNm ARNg). En algunas realizaciones, las tasas de integración del sitio diana se miden a través de la evaluación de la expresión génica (p. ej., mediante tinción inmunohistoquímica).

Tabla 8: Métodos ilustrativos para evaluar la integración del sitio diana

Tipo de ensayo	Muestra analizada	Método ilustrativo	Protocolo ilustrativo
Tasa de integración	Hígado (congelado)	qPCR	Biopsia de hígado sometida a extracción de ADN genómico. Para detectar el porcentaje de alelo (p. ej.,

5	de ADN genómico (% de Int. de ADNg)			albúmina) que contiene inserción específica, se ejecuta el método de qPCR (<i>quantitative polymerase chain reaction</i> , reacción en cadena de la polimerasa cuantitativa).
10	ARNm fusionado	Hígado (congelado)	ddPCR	Biopsia de hígado sometida a extracción de ARN. Para cuantificar el número de copias de ARNm fusionado (ARNm quimérico único transcrito a partir del alelo editado), se ejecuta el método de ddPCR (<i>droplet digital PCR</i> , PCR digital en nanogotas). Este ensayo mide la actividad transcripcional después de la inserción de la diana.
15				
20	ALB-2A	Plasma	ELISA	Se extrae sangre y se procesa para obtener plasma. ELISA patentado utilizado para medir la albúmina marcada con 2A (biomarcador circulante universal para la integración dirigida) Este ensayo mide la expresión total de la proteína después de la inserción de la diana.
25	% de edición de hepatocitos	Sección fijada de hígado	IHQ	Hígado fijado, seccionado y teñido contra el transgén. Se contaron las células transgénicas positivas y se utilizaron para calcular el porcentaje de edición de hepatocitos. Para la integración dirigida en un sitio de
30				

5				integración diana en el locus de albúmina, la expresión del transgén debe ser específica de hepatocitos. Este ensayo se centra en la integración de dianas por célula y es ortogonal al % de Int. de ADNg, que se centra en la integración de dianas por alelo.
10	Expresión de GFP	Células (p. ej., HepG2) fijadas y/o sección de tejido (p. ej., hígado) fijada	ICQ/IHQ	Células fijadas teñidas con colorante nuclear de contraste. Obtención de imágenes de células GFP+ (positivas a GFP) directamente o teñidas utilizando un anticuerpo anti-etiqueta de HA (hemaglutinina). Este ensayo mide el porcentaje de células que expresan el transgén GFP y es un indicador de la eficacia de la edición del vector vírico.
15				
20	Expresión de UGT1A1	Células (p. ej., HepG2) fijadas y/o sección de tejido (p. ej., hígado) fijada	ICQ/IHQ	Células fijadas teñidas con colorante nuclear de contraste. Células teñidas utilizando anticuerpo anti-UGT1A1. Este ensayo mide el porcentaje de células que expresan UGT1A1 y es un indicador de la eficacia de la edición del vector vírico.
25				

Métodos de tratamiento

Las composiciones y construcciones desveladas en el presente documento pueden usarse en cualquier aplicación *in vitro* o *in vivo* para provocar o mejorar la expresión de una carga útil (p. ej., un transgén) de un sitio de integración diana particular en una célula mientras se mantiene la expresión de genes endógenos en

el sitio de integración diana y alrededor de éste. Por ejemplo, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar un trastorno, una enfermedad o una dolencia en un sujeto (p. ej., mediante genoterapia).

5 En algunas realizaciones, el tratamiento comprende obtener o mantener un efecto farmacológico y/o fisiológico deseado. En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender la prevención total o parcial de una enfermedad (p. ej., prevención de los síntomas de la enfermedad). En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede
10 comprender la curación total o parcial de una enfermedad (p. ej., curación de efectos adversos asociados a una enfermedad). En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender la prevención de la recidiva de una enfermedad. En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender la reducción del avance de una enfermedad.
15 En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender el alivio de los síntomas de una enfermedad. En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender la prevención de la remisión de una enfermedad. En algunas realizaciones, un efecto farmacológico y/o fisiológico deseado puede comprender la estabilización y/o reducción de los
20 síntomas asociados a una enfermedad.

 En algunas realizaciones, el tratamiento comprende administrar una composición antes, durante o después de la aparición de una enfermedad (p. ej., antes, durante o después de la aparición de síntomas asociados a una enfermedad). En algunas realizaciones, el tratamiento comprende combiterapia (p. ej., con una o
25 más terapias, incluyendo diferentes tipos de terapias).

Enfermedades de interés

 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar cualquier enfermedad de interés que incluya una carencia o anomalía genética como componente de la enfermedad.
30

Como ejemplo específico, en algunas realizaciones, las composiciones y construcciones, tales como las desveladas en el presente documento, pueden usarse para tratar acidurias orgánicas de cadena ramificada (p. ej., cetoaciduria de cadena ramificada (MSUD, *Maple Syrup Urine Disease*, enfermedad de la orina con olor a jarabe de arce), acidemia metilmalónica (MMA), acidemia propiónica (PA), acidemia isovalérica (IVA), aciduria argininosuccínica). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., complejo BCKDH (subunidades E1a, E1b y E2), metilmalonil-CoA mutasa, propionil-CoA carboxilasa (subunidades alfa y beta), isovaleril CoA deshidrogenasa, argininosuccinato liasa (ASL), y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a acidurias orgánicas de cadena ramificada. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados con acidurias orgánicas de cadena ramificada (p. ej., hipotonía, retraso del desarrollo, convulsiones, atrofia óptica, encefalopatía aguda, hiperventilación, disnea, inestabilidad de la temperatura, vómitos recurrentes, cetoacidosis, pancreatitis, estreñimiento, neutropenia, pancitopenia, hemofagocitosis secundaria, arritmia cardíaca, miocardiopatía, insuficiencia renal crónica, dermatitis, pérdida de audición).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos de oxidación de ácidos grasos (p. ej., carencia de proteína trifuncional, carencia de L-3 hidroxiacl-CoA deshidrogenasa de cadena larga (LCAD), carencia de acil-CoA deshidrogenasa de cadena media (MCHAD), carencia de acil-CoA deshidrogenasa de cadena muy larga (VLCHAD)). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., HADHA, HADHB, LCHAD, ACADM, ACADVL, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a trastornos de oxidación de ácidos grasos. En algunas realizaciones, el tratamiento comprende la

reducción de signos y/o síntomas asociados a trastornos de oxidación de ácidos grasos (p. ej., hígado agrandado, retraso en el desarrollo mental y físico, debilidad muscular cardíaca, arritmia cardíaca, daño nervioso, función hepática anómala, rabdomiólisis, mioglobulinuria, hipoglucemia, acidosis metabólica, disnea, hepatomegalia, hipotonía, miocardiopatía).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la glucogenosis (enfermedades por almacenamiento de glucógeno, *glycogen storage disease*) (p. ej., glucogenosis de tipo 1 (GSD1), glucogenosis de tipo 2 (enfermedad de Pompe, GSD2). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., G6PC (GSD1a), G6PT1 (GSD1b), SLC17A3, SLC37A4 (GSD1c), alfa-glucosidasa ácida, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a glucogenosis. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a glucogenosis (p. ej., hepatomegalia, hipoglucemia, debilidad muscular, calambres musculares, cansancio, retraso en el desarrollo, obesidad, trastornos hemorrágicos, función hepática anómala, función renal anómala, función respiratoria anómala, función cardíaca anómala, úlceras bucales, gota, cirrosis, fibrosis, tumores hepáticos).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos del ciclo de la carnitina. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., OCTN2, CPT1, CACT, CPT2, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a trastornos del ciclo de la carnitina. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a trastornos del ciclo de la carnitina (p. ej., hipoglucemia hipocetósica,

miocardiopatía, debilidad muscular, cansancio, retraso del desarrollo motor, edema).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos del ciclo de la urea. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., CPS1, ARG1, ASL, OTC, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a trastornos del ciclo de la urea. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a trastornos del ciclo de la urea (p. ej., vómitos, náuseas, anomalías del comportamiento, cansancio, coma, psicosis, letargo, vómitos cíclicos, miopía, hiperamonemia, niveles elevados de ornitina).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la homocistinuria (HCU). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., cistationina beta sintasa (CBS), y/o variantes de la misma). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a HCU. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a HCU (p. ej., ectopia (o luxación) del cristalino, miopía, iridodonesis, cataratas, atrofia óptica, glaucoma, desprendimiento de retina, daño retiniano, retraso de los hitos del desarrollo, discapacidad intelectual, depresión, ansiedad, trastorno obsesivo-compulsivo, dolicoftenomelia, rodilla valga, pie cavo, escoliosis, tórax en quilla, tórax en embudo, osteoporosis, aumento del desarrollo de coágulos, tromboembolia, embolia pulmonar, piel frágil, hipopigmentación, rubor facial, hernia inguinal, pancreatitis, cifosis, neumotórax espontáneo).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar el síndrome de Crigler-Najjar. En

algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., UGT1A1, y/o variantes del mismo). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas al síndrome de Crigler-Najjar. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados al síndrome de Crigler-Najjar (*p. ej.*, ictericia, kernícterus (ictericia nuclear), letargo, vómitos, fiebre, reflejos anómalos, espasmos musculares, opistótonos, espasticidad, hipotonía, atetosis, niveles elevados de bilirrubina, diarrea, habla farfullante, confusión, disfagia, convulsiones).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la tirosinemia hereditaria. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., FAH, y/o variantes del mismo). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la tirosinemia hereditaria. En algunas realizaciones, el tratamiento comprende la reducción de los signos y/o síntomas asociados a la tirosinemia hereditaria (p. ej., hepatomegalia, ictericia, hepatopatía, cirrosis, hepatocarcinoma, fiebre, diarrea, melena, vómitos, esplenomegalia, edema, coagulopatía, función renal anómala, raquitismo, debilidad, hipertensión, íleo, taquicardia, hipertensión, crisis neurológicas, insuficiencia respiratoria, miocardiopatía).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la epidermólisis ampullosa. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (*p. ej.*, COL7A1, COL17A1, MMP1, KRT5, LAMA3, LAMB3, LAMC2, ITGB4, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la epidermólisis ampullosa. En algunas realizaciones, el tratamiento comprende la reducción de los signos y/o síntomas asociados a la epidermólisis ampullosa (p. ej., piel frágil,

crecimiento anómalo de las uñas, ampollas, piel engrosada, alopecia cicatricial, cicatrices atróficas, milia, problemas dentales, disfagia, picor y dolor en la piel).

5 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la carencia de alfa-1 antitripsina (A1ATD). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., alfa-1 antitripsina (A1AT), y/o variantes de la misma). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la carencia de alfa-1 antitripsina. En algunas realizaciones, 10 el tratamiento comprende la reducción de signos y/o síntomas asociados con A1ATD (p. ej., enfisema, tos crónica, producción de flema, sibilancias, infecciones respiratorias crónicas, ictericia, hepatomegalia, hemorragias, acumulación anómala de líquidos, enzimas hepáticas elevadas, disfunción hepática, hipertensión portal, cansancio, edema, hepatitis crónica activa, cirrosis, hepatocarcinoma, paniculitis).

15 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la enfermedad de Wilson. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., ATP7B, y/o variantes del mismo). En algunas realizaciones, el tratamiento comprende la 20 reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la enfermedad de Wilson. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados con la enfermedad de Wilson (p. ej., cansancio, falta de apetito, dolor abdominal, ictericia, anillos de Kayser-Fleischer, edema, problemas del habla, disfagia, pérdida de coordinación física, movimientos 25 incontrolados, rigidez muscular, hepatopatía, anemia, depresión, esquizofrenia, amenorrea, esterilidad, cálculos renales, daño tubular renal, artritis, osteoporosis, osteofitos)

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar enfermedades hemáticas (p. ej., 30 hemofilia A, hemofilia B). En algunas realizaciones, el tratamiento comprende la

introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., Factor IX (FIX), Factor VIII (FVIII), y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a enfermedades hemáticas.

5 En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a enfermedades hemáticas (p. ej., sangrado excesivo, hematomas anómalos, dolor e hinchazón articular, orina con sangre, heces con sangre, hemorragias nasales anómalas, cefalea, letargo, vómitos, visión doble, debilidad, convulsiones, crisis epilépticas).

10 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar el angioedema hereditario. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., inhibidor de C1 esterasa (C1-inh)). En algunas realizaciones, el tratamiento comprende la
15 reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a angioedema hereditario. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a angioedema hereditario (p. ej., edema, prurito, urticaria, náuseas, vómitos, dolor abdominal agudo, disfagia, disfonía, estridor).

20 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la enfermedad de Parkinson. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., dopamina descarboxilasa (DDC)). En algunas realizaciones, el tratamiento comprende la
25 reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a enfermedad de Parkinson. En algunas realizaciones, el tratamiento comprende la reducción de los signos y/o síntomas asociados a la enfermedad de Parkinson (p. ej., temblores, bradicinesia, rigidez muscular, deterioro de la postura y el equilibrio, pérdida de movimientos automáticos, cambios en el habla, cambios en la escritura).

30

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar enfermedades musculares. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., distrofias musculares, distrofia muscular de Duchenne (DMD), distrofia muscular de cintura y extremidades). miopatía miotubular ligada al cromosoma X). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a enfermedad musculares. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados con enfermedades musculares (p. ej., dificultad de movimiento, músculos de la pantorrilla agrandados, dolor y rigidez muscular, retraso en el desarrollo, problemas de aprendizaje, marcha inusual, escoliosis, problemas respiratorios, disfagia, arritmia, miocardiopatía, función articular anómala, hipotonía, disnea, ausencia de reflejos).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la mucopolisacaridosis (MPS) (p. ej., MPS IH, MPS IH/S, MPS IS, MPS II, MPS IIIA, MPS IIIB, MPS IIIC, MPS IIID, MPS IVA, MPS IVB, MPS V, MPS VI, MPS VII, MPS IX). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., IDUA, IDS, SGSH, NAGLU, HGSNAT, GNS, GALNS, GLB1, ARSB, GUSB, HYAL1). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la mucopolisacaridosis. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas asociados a MPS (p. ej., anomalías cardíacas, irregularidades respiratorias, hepatomegalia, esplenomegalia, anomalías neurológicas, retrasos en el desarrollo, infecciones recurrentes, secreción nasal persistente, respiración ruidosa, opacidad de la córnea, macroglosia, deformidades de la columna vertebral, rigidez articular, túnel carpiano, insuficiencia valvular aórtica, pérdida progresiva de la audición, convulsiones,

marcha inestable, acumulación de heparán sulfato, carencias enzimáticas, anomalías musculoesqueléticas, cardiopatía, quistes, masas de tejidos blandos).

5 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la carencia de lipasa ácida lisosomal. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., LIPA y/o variantes del mismo). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la carencia de lipasa ácida lisosomal. En algunas realizaciones, el
10 tratamiento comprende reducciones de signos y/o síntomas asociados a la carencia de lipasa ácida lisosomal (p. ej., vómitos, diarrea, hinchazón del abdomen y falta de aumento de peso, pérdida de peso, ictericia, fiebre, calcificación, anemia, disfunción o insuficiencia hepática, caquexia, mala absorción, problemas de las vías biliares, cardiopatía, accidente cerebrovascular).

15 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos asociados al metabolismo, transporte y/o colestasis de ácidos biliares. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., PFIC1, PFIC2, PFIC3, ABCB4, y/o
20 variante de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas al metabolismo, transporte y/o colestasis de ácidos biliares. En algunas realizaciones, el tratamiento comprende reducciones de signos y/o síntomas asociados con el metabolismo, transporte y/o colestasis de ácidos biliares (p. ej., prurito, ictericia,
25 retraso del crecimiento, hipertensión portal, hepatoesplenomegalia, diarrea, pancreatitis, carcinoma hepatocelular).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar la fenilcetonuria. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de
30 polinucleótidos que codifica uno o más transgenes de interés (p. ej., fenilalanina

hidroxilasa (PAH) y/o variante de la misma). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la fenilcetonuria. En algunas realizaciones, el tratamiento comprende reducciones de signos y/o síntomas asociados a la fenilcetonuria (p. ej., olor a

5 humedad en el aliento, piel y/u orina, convulsiones, erupciones cutáneas, microcefalia, hiperactividad, discapacidad intelectual, asma, eccema, anemia, aumento de peso, insuficiencia renal, osteoporosis, gastritis, esofagitis y carencias renales, cálculos renales, hipertensión, problemas psiquiátricos, mareos).

En algunas realizaciones, las composiciones y construcciones desveladas en

10 el presente documento pueden usarse para tratar la hiperoxaluria primaria. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., AGT, AGXT, GRHPR, HOGA1, y/o variante de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no

15 funcionales) asociadas a hiperoxaluria primaria. En algunas realizaciones, el tratamiento comprende reducciones de signos y/o síntomas asociados a la fenilcetonuria (p. ej., dolor en fosa lumbar, oxalosis, cálculos renales y/o cálculos en otras partes de las vías urinarias como la vejiga o la uretra, nefrocalcinosis, hematuria, disuria, necesidad de orinar con frecuencia, cólico renal, obstrucción de

20 las vías urinarias, infecciones repetidas de las vías urinarias, daño renal, insuficiencia renal, retraso del desarrollo).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar porfirias. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de

25 polinucleótidos que codifica uno o más transgenes de interés (p. ej., ALAD, HMBS, UROS, UROD, CPOX, PPOCX, FECH, ALAS2, y/o variante de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a porfirias. En algunas realizaciones, el tratamiento comprende reducciones de signos y/o síntomas

30 asociados a porfirias (p. ej., dolor abdominal, dolor en brazos y piernas, debilidad

generalizada, vómitos, confusión, estreñimiento, taquicardia, fluctuación de la presión arterial, retención urinaria, psicosis, alucinaciones, convulsiones, abrasiones, ampollas, erosiones de la piel, lesiones cutáneas, náuseas, aumento de la presión arterial, confusión).

5 En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos asociados a la producción de anticuerpos (p. ej., trastornos autoinmunitarios). En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés (p. ej., POLB, HLA-
10 DRB1, IL7R, CYP27B1, TNFRSF1A, HLA-B, HLA-DPB1, HLA-DRB1, IRF5, PTPN22, RBPJ, RUNX1, STAT4, y/o variantes de los mismos). En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la producción de anticuerpos. En algunas realizaciones, el tratamiento comprende la reducción de signos y/o síntomas
15 asociados a la producción de anticuerpos (p. ej., inflamación articular, rigidez articular, cansancio, fiebre, pérdida de apetito, problemas de visión, temblor, marcha inestable, mareos, erupción cutánea, lesiones, hiperalgesia).

En algunas realizaciones, las composiciones y construcciones desveladas en el presente documento pueden usarse para tratar trastornos asociados a la
20 producción de proteínas secretadas. En algunas realizaciones, el tratamiento comprende la introducción de una secuencia de polinucleótidos que codifica uno o más transgenes de interés. En algunas realizaciones, el tratamiento comprende la reducción de proteínas aberrantes (p. ej., proteínas no funcionales) asociadas a la producción de proteínas secretadas. En algunas realizaciones, el tratamiento
25 comprende la reducción de los signos y/o síntomas asociados a la producción de proteínas secretadas.

Integración dirigida

En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil
30 (p. ej., un transgén y/o ácido nucleico funcional) en un sitio de integración diana (p.

ej., un gen endógeno). En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil (p. ej., un transgén y/o ácido nucleico funcional) en un sitio de integración diana (p. ej., un gen endógeno) que se encuentra únicamente en un tejido específico. En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil (p. ej., un transgén y/o ácido nucleico funcional) en un sitio de integración diana (p. ej., un gen endógeno) que se encuentra en un sujeto en células que están presentes en más de un tejido (p. ej., en 2, 3, 4, 5 o en todos los tejidos). En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil (p. ej., un transgén y/o ácido nucleico funcional) en un sitio de integración diana (p. ej., un gen endógeno) que puede permitir la expresión inducible de una carga útil. En algunas realizaciones, la expresión inducible puede controlarse a través de medios artificiales (p. ej., administración de un fármaco u otra señal exógena) o a través de medios naturales (p. ej., expresión de IgH a partir de linfocitos B). En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil en un sitio de integración diana en un tipo celular específico (p. ej., locus específicos de tejido). En algunas realizaciones, la integración de una carga útil se produce en un tejido específico (p. ej., hígado, sistema nervioso central (SNC), músculo, riñón, vascular, pulmón). En algunas realizaciones, la integración de una carga útil se produce en múltiples tejidos específicos (p. ej., hígado, sistema nervioso central (SNC), músculo, riñón, vascular, pulmón).

En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil en un sitio de integración diana que se considera que es un sitio de puerto seguro (p. ej., albúmina, apolipoproteína A2 (ApoA2), haptoglobina, IgH (p. ej., linfocitos B), β 2 microglobulina, β -actina (beta-actina), GAPDH). En algunas realizaciones, un sitio de integración diana puede seleccionarse de entre cualquier sitio genómico

apropiado para su uso con los métodos y las composiciones que se proporcionan en el presente documento. En algunas realizaciones, un sitio de integración diana codifica un polipéptido. En algunas realizaciones, un sitio de integración diana codifica un polipéptido que está altamente expresado en un sujeto (p. ej., un sujeto que no padece una enfermedad, trastorno o afección, o un sujeto que padece una enfermedad, trastorno o afección). En algunas realizaciones, la integración de una carga útil se produce en un extremo 5' o 3' de uno o más genes endógenos (p. ej., genes que codifican polipéptidos). En algunas realizaciones, la integración de una carga útil se produce entre un extremo 5' o 3' de uno o más genes endógenos (p. ej., genes que codifican polipéptidos).

En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil en un sitio de integración diana con mínima o ninguna integración inespecífica (p. ej., integración en un locus no diana). En algunas realizaciones, las composiciones y construcciones proporcionadas en el presente documento dirigen la integración de una carga útil en un sitio de integración diana con integración inespecífica reducida en comparación con una composición o construcción de referencia (p. ej., en relación con una composición o construcción sin secuencias de homología flanqueantes).

En algunas realizaciones, la integración de un transgén en un sitio de integración diana, permite la expresión de una carga útil sin que se altere la expresión del gen endógeno. En algunas realizaciones, la integración de un transgén en un sitio de integración diana, permite la expresión de una carga útil a partir de un promotor endógeno. En algunas realizaciones, la integración de un transgén en un sitio de integración diana, altera la expresión del gen endógeno. En algunas realizaciones, la integración de un transgén en un sitio de integración diana, altera la expresión del gen endógeno sin que afecte negativamente a una célula y/o a un sujeto diana (p. ej., dirigiéndose a un sitio de puerto seguro). En algunas realizaciones, la integración de un transgén en un sitio de integración diana no requiere el uso de una nucleasa (p. ej., nucleasas Cas, TALEN, ZFN). En algunas

realizaciones, la integración de un transgén en un sitio de integración diana se facilita mediante el uso de una nucleasa (p. ej., nucleasas Cas, TALEN, ZFN).

En algunas realizaciones, la integración de un transgén en un sitio de integración diana confiere una ventaja selectiva (p. ej., aumento de la tasa de supervivencia en una pluralidad de células en relación con otras células en un tejido). En algunas realizaciones, una ventaja selectiva puede producir un mayor porcentaje de células en uno o más tejidos que expresan un transgén.

Sitio de corte

En algunas realizaciones, la integración de un transgén en un sitio de integración diana se mejora mediante el uso de una nucleasa (p. ej., nucleasas Cas, TALEN, ZFN) dirigida a un sitio de corte específico para producir una rotura monocatenaria o bicatenaria de ADN. En algunas realizaciones, la integración de un transgén en un sitio de integración diana se mejora mediante el uso de una nucleasa dirigida a un sitio de corte en una secuencia de ADN que es distal con respecto al sitio de integración diana. En algunas realizaciones, la integración de un transgén en un sitio de integración diana se facilita mediante el uso de una nucleasa dirigida a un sitio de corte en una secuencia de ADN que no se solapa con el sitio de integración diana. En algunas realizaciones, la integración de un transgén en un sitio de integración diana se mejora mediante el uso de una nucleasa dirigida a un sitio de corte en una secuencia de ADN que tiene una separación de al menos aproximadamente 100 pb (p. ej., aproximadamente 100 pb, 150 pb, 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 450 pb, 500 pb, 550 pb, 600 pb, 650 pb, 700 pb, 750 pb, 800 pb, 850 pb, 900 pb, 950 pb, 1000 pb, 1050 pb, 1100 pb, 1150 pb, 1200 pb, 1250 pb, 1300 pb, 1350 pb, 1400 pb, 1450 pb, 1500 pb, 1550 pb, 1600 pb, 1650 pb, 1700 pb, 1800 pb, 1850 pb, 1900 pb, 1950 pb, 2000 pb) de un sitio de integración diana. En algunas realizaciones, la integración de un transgén en un sitio de integración diana se mejora mediante el uso de una nucleasa dirigida a un sitio de corte en una secuencia de ADN que está a una distancia de aproximadamente 100 pb (p. ej., aproximadamente 1 pb, 3 pb, 5 pb, 10 pb, 20 pb, 30 pb, 40 pb, 50 pb, 60 pb, 70 pb, 80 pb, 90 pb, 100 pb) de un sitio de integración diana.

En algunas realizaciones, un sitio de corte se encuentra dentro de una secuencia de ADN no codificante (p. ej., un intrón, una región no traducida, un potenciador, un promotor, un silenciador o un aislante). En algunas realizaciones, un sitio de corte se encuentra dentro de una secuencia no codificante de un gen humano. En algunas realizaciones, un sitio de corte se encuentra en un intrón de un gen humano (p. ej., colágeno, actina, albúmina, $\beta 2$ microglobulina, IgH, GAPDH, G6PC). En algunas realizaciones, un sitio de corte se encuentra en el intrón 13 o el intrón 14 de un gen de albúmina humano. En algunas realizaciones, un sitio de corte se encuentra en el intrón 5 o en un sitio después del exón 6 de un gen de actina humano. En algunas realizaciones, un sitio de corte se encuentra en el intrón 2 o el intrón 3 de un gen de $\beta 2$ microglobulina humano. En algunas realizaciones, un sitio de corte se encuentra dentro de IGHJ6 o entre IGHJ6 e IGHM de un gen de IgH humano.

En algunas realizaciones, un sitio de corte dentro de una o más secuencias no codificantes de un gen humano (p. ej., intrón 13 y 14 de un gen de albúmina) puede proporcionar una integración mejorada del transgén cuando se combina con componentes GENERIDE™ y una o más nucleasas apropiadas. En algunas realizaciones, la selección de un sitio de corte puede basarse en efectos inespecíficos (*off-target*) esperados (p. ej., efectos inespecíficos reducidos) utilizando una base de datos externa o una herramienta informática de predicción/cálculo conocida en la técnica (p. ej., Benchling, CHOPCHOP, IDT, CRISPOR, E-CRISP, TureDesign, CRISPick). En algunas realizaciones, un sitio de corte puede reducir efectos inespecíficos no deseados (p. ej., citotoxicidad, carcinogenicidad, inmunogenicidad y similares). En algunas realizaciones, un sitio de corte puede estar dentro de un locus de puerto seguro para reducir los posibles efectos de la integración alterada.

Composiciones

En algunas realizaciones, las composiciones pueden producirse usando métodos y construcciones proporcionados en el presente documento (p. ej., vectores víricos). En algunas realizaciones, las composiciones incluyen

composiciones líquidas, sólidas y gaseosas. En algunas realizaciones, las composiciones comprenden ingredientes adicionales (p. ej., diluyentes, estabilizante, excipientes, adyuvantes). En algunas realizaciones, los ingredientes adicionales pueden comprender tampones (p. ej., fosfato, citrato, tampones de ácido orgánico), antioxidantes (p. ej., ácido ascórbico), polipéptidos de bajo peso molecular (p. ej., de menos de 10 restos), diversas proteínas (p. ej., seroalbúmina, gelatina, inmunoglobulinas), polímeros hidrófilos (p. ej., polivinilpirrolidona), aminoácidos (p. ej., glicina, glutamina, asparagina, arginina, lisina), hidratos de carbono (p. ej., monosacáridos, disacáridos, glucosa, manosa, dextrinas), agentes quelantes (p. ej., EDTA), alcoholes de azúcar (p. ej., manitol, sorbitol), contraiones formadores de sal (p. ej., sodio, potasio) y/o tensioactivos no iónicos (p. ej., Tween™, Pluronic™, polietilenglicol (PEG)), entre otros. En algunas realizaciones, el transportador acuoso es una solución acuosa de pH tamponado.

En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en un intervalo de dosificaciones. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en una sola dosis. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en múltiples dosis. En algunas realizaciones, las composiciones se proporcionan durante un periodo de tiempo. En algunas realizaciones, las composiciones se proporcionan a intervalos específicos (p. ej., intervalos variables, intervalos establecidos). En algunas realizaciones, las dosis pueden variar en función de la forma farmacéutica y de la vía de administración. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en dosis entre $1e11$ y $1e14$ vg/kg. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en dosis entre $1e12$ y $1e13$ vg/kg. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en dosis entre $1e12$ y $1e14$ vg/kg. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en dosis entre $1e14$ y $1e15$ vg/kg. En algunas realizaciones, las

composiciones proporcionadas en el presente documento pueden proporcionarse en dosis no superiores a 1×10^{14} vg/kg. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden proporcionarse en dosis no superiores a 1×10^{15} vg/kg.

5 En algunas realizaciones, un experto familiarizado con la técnica será capaz de idear un nivel de dosificación y una pauta posológica apropiados utilizando las composiciones farmacéuticas descritas en el presente documento para el tratamiento de diversas afecciones en diversos pacientes. Por ejemplo, en algunas realizaciones, una dosis seleccionada depende de un efecto terapéutico deseado, de una vía de administración y de una duración de tratamiento deseada. En algunas
10 realizaciones, se administran niveles de dosificación de aproximadamente 0.001 mg a aproximadamente 6 mg de ácido nucleico por kg de peso corporal durante cada dosificación a un sujeto (p. ej., animal, ser humano). En algunas realizaciones, los niveles de dosificación de ácidos nucleicos dentro de las nanopartículas lipídicas desveladas son de aproximadamente 0.1 mg/kg a aproximadamente 1.0 mg/kg. En
15 algunas realizaciones, los niveles de dosificación de ácidos nucleicos dentro de las nanopartículas lipídicas desveladas son de aproximadamente 0.1 mg/kg a aproximadamente 3.0 mg/kg. En algunas realizaciones, a un sujeto (p. ej., animal, ser humano), se le administran niveles de dosificación de las nanopartículas
20 lipídicas desveladas de aproximadamente 0.2 mg a aproximadamente 100 mg del total de uno o más componentes (p. ej., lípidos ionizables, esteroides, lípidos conjugados con enlazador, fosfolípidos)/kg de peso corporal. En algunas realizaciones, los niveles de dosificación de las nanopartículas lipídicas desveladas son de aproximadamente 0.5 mg/kg a aproximadamente 6 mg/kg del total de uno o
25 más componentes (p. ej., lípidos ionizables, esteroides, lípidos conjugados con enlazador, fosfolípidos)/kg de peso corporal.

 En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto en un punto temporal particular (p. ej., edad de un sujeto). En algunas realizaciones, las composiciones proporcionadas en
30 el presente documento pueden administrarse a un sujeto recién nacido. En algunas

- realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto neonatal. En algunas realizaciones, un sujeto ratón neonatal tiene entre 0 y 7 días de vida. En algunas realizaciones, el sujeto humano neonatal tiene una edad de entre 0 días y 1 mes. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto con una edad de entre 7 y 30 días. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto con una edad de entre 3 meses y 1 año. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto con una edad de entre 1 y 5 años. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto con una edad de entre 4 y 7 años. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto a partir de los 5 años.
- En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto en un punto temporal en función del estadio de crecimiento (p. ej., porcentaje del tamaño o peso estimado/promedio de un adulto) de un tejido u órgano particular. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto en donde un tejido u órgano (p. ej., hígado, músculo, SNC, pulmón, y similares) es al menos el 10 %, al menos el 20 %, al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, o al menos el 99 % del tamaño o peso estimado/promedio de un adulto. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto en donde un tejido u órgano tiene aproximadamente un 20 % (+/- 5 %) del tamaño o peso estimado/promedio de un adulto. En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto en donde un tejido u órgano tiene aproximadamente un 50 % (+/- 5 %) del tamaño o peso estimado/promedio de un adulto. En algunas realizaciones, las composiciones proporcionadas en el presente

documento pueden administrarse a un sujeto en donde un tejido u órgano tiene aproximadamente un 60 % (+/- 5 %) del tamaño o peso estimado/promedio de un adulto. En algunas realizaciones, el tamaño o peso estimado/promedio de un adulto, de un tejido u órgano particular, puede determinarse como se describe en la técnica (Véase, Noda *et al.* Pediatric radiology, 1997; Johnson *et al.* Liver transplantation, 2005; y Szpinda *et al.* Biomed research international, 2015, cada uno de los cuales se incorpora en el presente documento por referencia en su totalidad).

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 49), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65,

SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 50), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un

casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 51), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 52), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad

de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 53), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al

menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 54), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65,

SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 55), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un

casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 56), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 57), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad

de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 58), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al

menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 59), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65,

SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

En algunas realizaciones, una composición como la descrita en el presente documento puede comprender (i) una nucleasa o secuencia de polinucleótidos que codifique una nucleasa capaz de inducir una rotura bicatenaria y/o una rotura monocatenaria en un sitio de corte distal de un sitio de integración diana; y (ii) un casete polinucleotídico que comprenda un casete de expresión que comprenda una primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 6 (p. ej., SEQ ID NO: 60), una segunda secuencia de ácido nucleico situada en la secuencia en posición 5' o 3' en la primera secuencia de ácido nucleico que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 5 (p. ej., SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48) una tercera secuencia de ácido nucleico situada en posición 5' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 68 o SEQ ID NO: 70) y una cuarta secuencia de ácido nucleico situada en posición 3' en el casete de expresión y que comprenda una secuencia que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % o del 100 % con una secuencia de la Tabla 7 (p. ej., SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 o SEQ ID NO: 69). En algunas realizaciones, (i) y (ii) pueden administrarse como composiciones distintas.

Vías de administración

En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden administrarse a un sujeto a través de cualquiera (o más) de una variedad de vías conocidas en la técnica (p. ej., parenteral, subcutánea, intravenosa,

intracraneal, intraespinal, intraocular, intramuscular, intravaginal, intraperitoneal, epicutánea, intradérmica, rectal, pulmonar, intraósea, oral, bucal, intraportal, intraarterial, intratraqueal o nasal). En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden introducirse en células, que después se introducen en un sujeto (p. ej., células de hígado, músculo, sistema nervioso central (SNC), pulmón, hemáticas). En algunas realizaciones, las composiciones proporcionadas en el presente documento pueden introducirse mediante métodos de suministro conocidos en la técnica (p. ej., inyección, catéter).

En algunas realizaciones, la reescritura genómica con la plataforma GENERIDE™ difiere de la genoterapia convencional porque utiliza HR para suministrar un gen corrector en una ubicación específica en el genoma. En algunas realizaciones, GENERIDE™ inserta un gen corrector de una manera exacta, lo que lleva a la integración específica de sitio en el genoma.

En algunas realizaciones, las composiciones proporcionadas comprenden uno o más brazos de homología, un transgén, y un ácido nucleico que promueve la producción de dos productos génicos independientes. En algunas realizaciones, las composiciones y métodos de la presente divulgación comprenden una primera secuencia de ácido nucleico que codifica un transgén. En algunas realizaciones, las composiciones y métodos de la presente divulgación comprenden un segundo ácido nucleico que promueve la producción de dos productos génicos independientes (p. ej., un péptido 2A). En algunas realizaciones, la presente divulgación proporciona un casete de expresión que comprende una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico como se describe en el presente documento.

En algunas realizaciones, un segundo ácido nucleico comprende una secuencia de ácido nucleico que codifica un péptido 2A; una secuencia de ácido nucleico que codifica un sitio interno de entrada al ribosoma (IRES); una secuencia de ácido nucleico que codifica una región de corte y empalme aminoterminal de inteína y una región de corte y empalme carboxiterminal de inteína; y/o una secuencia de ácido nucleico que codifica un donante de corte y empalme y un

aceptor de corte y empalme. En algunas realizaciones, las composiciones y métodos de la presente divulgación comprenden un casete polinucleotídico que comprende un casete de expresión que comprende dicho primer ácido nucleico y dicho segundo ácido nucleico. En algunas realizaciones, las composiciones y métodos de la presente divulgación comprenden una tercera secuencia de ácido nucleico que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica. En algunas realizaciones, las composiciones y métodos de la presente divulgación comprenden una cuarta secuencia de ácido nucleico que comprende una secuencia que es sustancialmente homóloga a una secuencia genómica. En algunas realizaciones, dicha tercera secuencia de ácido nucleico se sitúa en posición 5' en el casete de expresión y comprende una secuencia que es sustancialmente homóloga a una secuencia genómica en posición 5' de un sitio de integración diana en un genoma de una célula. En algunas realizaciones, dicha cuarta secuencia de ácido nucleico se sitúa en posición 3' en el casete de expresión y comprende una secuencia que es sustancialmente homóloga a una secuencia genómica en posición 3' de un sitio de integración diana en el genoma de la célula.

En algunas realizaciones, una o más composiciones descritas en el presente documento se administran en combinación. En algunas realizaciones, una primera composición puede administrarse simultáneamente con una segunda composición. En algunas realizaciones, una primera composición y una segunda composición pueden administrarse secuencialmente (p. ej., con minutos, horas, días, semanas o meses de diferencia). En algunas realizaciones, una o más composiciones pueden administrarse por la misma vía (p. ej., parenteral, subcutánea, intravenosa, intracraneal, intraespinal, intraocular, intramuscular, intravaginal, intraperitoneal, epicutánea, intradérmica, rectal, pulmonar, intraósea, oral, bucal, intraportal, intraarterial, intratraqueal o nasal). En algunas realizaciones, una o más composiciones pueden administrarse por diferentes vías (p. ej., parenteral, subcutánea, intravenosa, intracraneal, intraespinal, intraocular, intramuscular, intravaginal, intraperitoneal, epicutánea, intradérmica, rectal, pulmonar, intraósea, oral, bucal, intraportal, intraarterial, intratraqueal o nasal).

En algunas realizaciones, una o más composiciones administradas en combinación pueden comprender: (i) una primera composición que comprende una carga útil (p. ej., uno o más ARNm y/o proteínas de nucleasas (p. ej., proteínas Cas (p. ej., Cas9 de *Staphylococcus aureus* (saCas9), *Streptococcus pyogenes* (spCas9), Nucleasa de AZ, HF1-Cas9, HF2-Cas9 o HiFi-Cas9), endonucleasas, meganucleasas, TALEN, ZFN)); y (ii) una segunda composición que comprende una segunda carga útil distinta (p. ej., una secuencia de polinucleótidos que comprende una región que codifica un transgén). En algunas realizaciones, la primera y segunda composición se suministra con el mismo sistema de suministro (p. ej., vector vírico, nanopartícula lipídica, y similares). En algunas realizaciones, la primera y segunda composición se suministra con diferentes sistemas de suministro (p. ej., vector vírico, nanopartícula lipídica, y similares). En algunas realizaciones, la primera composición se administra antes que la segunda composición (p. ej., con una diferencia de minutos, horas, días, semanas o meses). En algunas realizaciones, la segunda composición se administra antes que la primera composición (p. ej., con una diferencia de minutos, horas, días, semanas o meses). En algunas realizaciones, la primera y segunda composición se administran simultáneamente. En algunas realizaciones, la primera y segunda composición se combinan antes de la administración.

En algunas realizaciones, la primera y/o segunda composición se administra una sola vez a una dosis particular (p. ej., una dosis fija o una dosis basada en el peso). En algunas realizaciones, la primera y/o segunda composición se administra más de una vez a una dosis particular (p. ej., una dosis fija o una dosis basada en el peso). En algunas realizaciones, cuando se administra más de una dosis (p. ej., una dosis fija o una dosis basada en el peso) la primera y/o segunda composición puede administrarse de manera simultánea, sustancialmente de manera simultánea o consecutiva. En algunas realizaciones, se administran múltiples dosis (p. ej., una dosis fija o una dosis basada en el peso) en un periodo de tiempo especificado (p. ej., en minutos, horas, días, semanas o meses).

En algunas realizaciones, la primera y/o segunda composición se administra en respuesta a un biomarcador (p. ej., un biomarcador circulante como se describe en el documento WO2020214582A1). Por ejemplo, la primera y/o segunda composición se administra a una dosis particular (p. ej., una dosis fija o una dosis basada en el peso) y en un periodo de tiempo especificado (p. ej., en minutos, horas, días, semanas o meses) se controlan los niveles de un biomarcador (p. ej., como se describe en el documento WO2020214582A1). Si los niveles de un biomarcador (p. ej., como se describe en el documento WO2020214582A1) son bajos (p. ej., en comparación con los de una referencia apropiada (p. ej., niveles de un biomarcador antes de la administración)), entonces la primera y/o segunda composición se administra a una dosis particular (p. ej., una dosis fija o una dosis basada en el peso). Si los niveles de un biomarcador (p. ej., como se describe en el documento WO2020214582A1) son altos (p. ej., en comparación con los de una referencia apropiada (p. ej., niveles de un biomarcador después de una administración inicial)), entonces puede reevaluarse la dosificación posterior (p. ej., una dosis fija o una dosis basada en el peso) de la primera y/o segunda composición (p. ej., suspensión del tratamiento, reducción de la dosis fija o dosis basada en el peso).

Métodos de producción de vectores víricos

Producción de vectores víricos

En algunas realizaciones, la producción de vectores víricos (p. ej., vectores víricos de AAV) puede incluir tanto etapas previas para generar vectores víricos (p. ej., cultivo basado en células) como etapas posteriores para procesar los vectores víricos (p. ej., purificación, formulación, y similares). En algunas realizaciones, las etapas previas pueden comprender uno o más de ampliación celular, cultivo celular, transfección celular, lisis celular, producción de vectores víricos y/o recogida de vectores víricos.

En algunas realizaciones, las etapas posteriores pueden comprender una o más de separación, filtración, concentración, clarificación, purificación, cromatografía (p. ej., afinidad, intercambio iónico, hidrófobo, modo mixto), centrifugación (p. ej., ultracentrifugación), y/o formulación.

En algunas realizaciones, las construcciones y métodos descritos en el presente documento están diseñados para aumentar los rendimientos de vectores víricos (p. ej., rendimientos de vectores de AAV), reducir los niveles de vectores víricos competentes en cuanto a su replicación (p. ej., AAV competentes en cuanto a su replicación (rcAAV, *replication competent AAV*)), mejorar la eficacia de empaquetamiento de vectores víricos (p. ej., empaquetamiento de la cápside de vectores de AAV), y/o cualquier combinación de los mismos, en relación con una construcción o método de referencia, por ejemplo los de Xiao et al. 1998 y Grieger et al. 2015, cada uno de los cuales se incorpora en el presente documento por referencia en su totalidad.

Líneas celulares y reactivos de transfección

En algunas realizaciones, la producción de vectores víricos comprende el uso de células (p. ej., cultivo celular). En algunas realizaciones, la producción de vectores víricos comprende el uso del cultivo celular de una o más líneas celulares (p. ej., líneas celulares de mamífero). En algunas realizaciones, la producción de vectores víricos comprende el uso de líneas celulares HEK293 o variantes de las mismas (p. ej., líneas celulares HEK293T, HEK293F). En algunas realizaciones, las células pueden crecer en suspensión. En algunas realizaciones, las células están compuestas por células adherentes. En algunas realizaciones, las células pueden crecer en medios que no contienen componentes animales (p. ej., suero animal). En algunas realizaciones, las células pueden crecer en medios aséricos (p. ej., medio F17, medio Expi293). En algunas realizaciones, la producción de vectores víricos comprende la transfección de células con construcciones de expresión (p. ej., plásmidos). En algunas realizaciones, las células se seleccionan con respecto a una alta expresión de vectores víricos (p. ej., vectores de VAA). En algunas realizaciones, las células se seleccionan con respecto a una alta eficacia de empaquetamiento de vectores víricos (p. ej., empaquetamiento de la cápside de vectores de AAV). En algunas realizaciones, las células se seleccionan para mejorar la eficacia de la transfección (p. ej., con reactivos químicos de transfección, incluyendo moléculas catiónicas). En algunas realizaciones, las células se diseñan

con respecto a una alta expresión de vectores víricos (p. ej., vectores de VAA). En algunas realizaciones, las células se diseñan con respecto a una alta eficacia de empaquetamiento de vectores víricos (p. ej., empaquetamiento de la cápside de vectores de AAV). En algunas realizaciones, las células se diseñan para mejorar la eficacia de transfección (p. ej., con reactivos químicos de transfección, incluyendo moléculas catiónicas). En algunas realizaciones, las células pueden ser diseñarse o seleccionarse con respecto a dos o más de los atributos anteriores. En algunas realizaciones, las células se ponen en contacto con una o más construcciones de expresión (p. ej., plásmidos). En algunas realizaciones, las células se ponen en contacto con uno o más reactivos de transfección (p. ej., reactivos químicos de transfección, incluyendo lípidos, polímeros y moléculas catiónicas) y una o más construcciones de expresión. En algunas realizaciones, las células se ponen en contacto con una o más moléculas catiónicas (p. ej., un lípido catiónico, un reactivo PEI) y con una o más construcciones de expresión. En algunas realizaciones, las células se ponen en contacto con un reactivo PEIMAX y con una o más construcciones de expresión. En algunas realizaciones, las células se ponen en contacto con un reactivo FectoVir-AAV y con una o más construcciones de expresión. En algunas realizaciones, las células se ponen en contacto con uno o más reactivos de transfección y con una o más construcciones de expresión a proporciones particulares. En algunas realizaciones, las proporciones de los reactivos de transfección con respecto a las construcciones de expresión mejoran la producción de los vectores víricos (p. ej., mejora del rendimiento del vector, mejora de la eficacia del empaquetamiento y/o mejora de la eficacia de la transfección).

Construcciones de expresión

En algunas realizaciones, las construcciones de expresión son o comprenden una o más secuencias de polinucleótidos (p. ej., plásmidos). En algunas realizaciones, las construcciones de expresión comprenden elementos de secuencia de polinucleótidos particulares (p. ej., cargas útiles, promotores, genes víricos y similares). En algunas realizaciones, las construcciones de expresión

comprenden secuencias de polinucleótidos que codifican genes víricos (p. ej., un gen o variante de gen *rep* o *cap*, uno o más genes o variantes génicas de virus auxiliares). En algunas realizaciones, las construcciones de expresión de un tipo particular comprenden combinaciones específicas de elementos de secuencia de polinucleótidos. En algunas realizaciones, las construcciones de expresión de un tipo particular no comprenden combinaciones específicas de elementos de secuencia de polinucleótidos. En algunas realizaciones, una construcción de expresión particular no comprende elementos de secuencia de polinucleótidos que codifican genes *rep* y *cap* y/o variantes génicas.

En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican genes víricos de tipo silvestre (p. ej., genes *rep*, genes *cap*, genes auxiliares víricos o combinaciones de los mismos de tipo silvestre). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican genes o variantes génicas de virus auxiliar (p. ej., genes o variantes génicas de herpesvirus, genes o variantes génicas de adenovirus). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican una o más copias de genes que expresan una o más proteínas Rep de tipo silvestre (p. ej., 1 copia, 2 copias, 3 copias, 4 copias, 5 copias, y similares). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican una única copia génica que expresa una o más proteínas Rep de tipo silvestre (p. ej., Rep68, Rep40, Rep52, Rep78, o combinaciones de las mismas). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican una o más proteínas Rep de tipo silvestre (p. ej., Rep68, Rep40, Rep52, Rep78, o combinaciones de las mismas). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican al menos cuatro proteínas Rep de tipo silvestre (p. ej., Rep68, Rep40, Rep52, Rep78). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos que codifican cada una de Rep68, Rep40, Rep52 y Rep78. En algunas realizaciones, las construcciones de

expresión comprenden secuencias de polinucleótidos que codifican una o más proteínas auxiliares adenovíricas de tipo silvestre (p. ej., E2 y E4).

En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos de tipo silvestre que codifican genes víricos de tipo silvestre (p. ej., genes *rep*, genes *cap*, genes auxiliares). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos modificadas (p. ej., con optimización de codones) que codifican genes víricos de tipo silvestre (p. ej., genes *rep*, genes *cap*, genes auxiliares). En algunas realizaciones, las construcciones de expresión comprenden secuencias de polinucleótidos modificadas que codifican genes víricos modificados (p. ej., genes *rep*, genes *cap*, genes auxiliares). En algunas realizaciones, los genes víricos modificados se diseñan y/o genomodifican para obtener determinadas mejoras (p. ej., mejora de la transducción, especificidad tisular, tamaño reducido, respuesta inmunitaria reducida, mejora del empaquetamiento, niveles de rcAAV reducidos, y similares).

De acuerdo con diversas realizaciones, las construcciones de expresión desveladas en el presente documento pueden ofrecer una mayor flexibilidad y modularidad en comparación con tecnologías anteriores. En algunas realizaciones, las construcciones de expresión desveladas en el presente documento pueden permitir el intercambio de diversas secuencias de polinucleótidos (p. ej., diferentes genes *rep*, genes *cap*, cargas útiles, genes auxiliares, promotores, y similares) al tiempo que proporcionan determinadas mejoras (p. ej., mayor rendimiento del vector vírico, mayor empaquetamiento, niveles reducidos de rcAAV, y similares). En algunas realizaciones, las construcciones de expresión desveladas en el presente documento son compatibles con diversos procesos de producción anteriores (p. ej., diferentes condiciones de cultivo celular, diferentes reactivos de transfección, y similares) al tiempo que proporcionan determinadas mejoras (p. ej., mayor rendimiento del vector vírico, mayor empaquetamiento, niveles reducidos de rcAAV, y similares)

En algunas realizaciones, las construcciones de expresión de diferentes tipos comprenden diferentes combinaciones de secuencias de polinucleótidos. En algunas realizaciones, una construcción de expresión de un tipo comprende uno o más elementos de secuencia de polinucleótidos (p. ej., cargas útiles, promotores, genes víricos y similares) que no está(n) presente(s) en una construcción de expresión de un tipo diferente. En algunas realizaciones, una construcción de expresión de un tipo comprende elementos de secuencia de polinucleótidos que codifican un gen vírico (p. ej., un gen o variante de gen *rep* o *cap*) y elementos de secuencia de polinucleótidos que codifican una carga útil (p. ej., un transgén y/o ácido nucleico funcional). En algunas realizaciones, una construcción de expresión de un tipo comprende elementos de secuencia de polinucleótidos que codifican uno o más genes víricos (p. ej., un gen o variante de gen *rep* o *cap* y/o uno o más genes de virus auxiliares). En algunas realizaciones, una construcción de expresión de un tipo comprende elementos de secuencia de polinucleótidos que codifican uno o más genes víricos, en donde los genes víricos son de uno o más tipos de virus (p. ej., genes o variantes génicas de AAV y adenovirus). En algunas realizaciones, los genes víricos de adenovirus son genes y/o variantes génicas. En algunas realizaciones, los genes víricos de adenovirus son uno o más de E2A (p. ej., proteína de unión a ADN (DBP, *DNA Binding Protein*) E2A), E4 (p. ej., marco abierto de lectura 2 (ORF, *Open Reading Frame*) de E4, ORF3, ORF4, ORF6/7), VA, y/o variantes de los mismos.

En algunas realizaciones, las construcciones de expresión se usan para la producción de vectores víricos (p. ej., mediante cultivo celular). En algunas realizaciones, las construcciones de expresión se ponen en contacto con las células junto con uno o más reactivos de transfección (p. ej., reactivos químicos de transfección). En algunas realizaciones, las construcciones de expresión se ponen en contacto con las células a proporciones particulares junto con uno o más reactivos de transfección. En algunas realizaciones, las construcciones de expresión de diferentes tipos se ponen en contacto con las células a proporciones (p. ej., proporciones en peso) particulares junto con uno o más reactivos de

transfección. En algunas realizaciones, las construcciones de expresión de diferentes tipos se ponen en contacto con las células a una proporción (p. ej., proporción en peso) de aproximadamente 10:1, 9:1, 8:1, 7:1, 6:1, 5:1, 4:1, 3:1, 2:1, 1.5:1, 1:1, 1:1.5, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, o 1:10. En algunas realizaciones,

5 una primera construcción de expresión que comprende uno o más genes auxiliares víricos y una segunda construcción de expresión que comprende una o más cargas útiles se ponen en contacto con las células a una proporción (p. ej., proporción en peso) de aproximadamente 10:1, 9:1, 8:1, 7:1, 6:1, 5:1, 4:1, 3:1, 2:1, 1.5:1, 1:1, 1:1.5, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, o 1:10 de la primera construcción de expresión

10 con respecto a la segunda construcción de expresión. En algunas realizaciones, una primera construcción de expresión que comprende una o más cargas útiles y una segunda construcción de expresión que comprende uno o más genes auxiliares víricos se ponen en contacto con las células a una proporción (p. ej., proporción en peso) de aproximadamente 10:1, 9:1, 8:1, 7:1, 6:1, 5:1, 4:1, 3:1, 2:1, 1.5:1, 1:1, 1:1.5, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, o 1:10 de la primera construcción de expresión

15 con respecto a la segunda construcción de expresión. En algunas realizaciones, las proporciones particulares de las construcciones de expresión mejoran la producción de los AAV (p. ej., aumento de los rendimientos del vector vírico, aumento de la eficacia del empaquetamiento y/o aumento de la eficacia de la transfección. En algunas realizaciones, las células se ponen en contacto con una o más construcciones de expresión (p. ej., de manera secuencial o sustancialmente de manera simultánea). En algunas realizaciones, tres o más construcciones de expresión se ponen en contacto con las células. En algunas realizaciones, las construcciones de expresión comprenden uno o más promotores (p. ej., uno o más

20 promotores exógenos). En algunas realizaciones, los promotores son o comprenden CMV, RSV, CAG, EF1alfa, PGK, A1AT, C5-12, MCK, desmina, p5, p40, o combinaciones de los mismos. En algunas realizaciones, las construcciones de expresión comprenden uno o más promotores por delante de un elemento de secuencia de polinucleótidos particular (p. ej., un gen o variante de gen *rep* o *cap*).

25 En algunas realizaciones, las construcciones de expresión comprenden uno o más

30

promotores por detrás de un elemento de secuencia de polinucleótidos particular (p. ej., un gen o variante de gen *rep* o *cap*).

5 En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias de polinucleótidos que codifican elementos (p. ej., marcadores de selección, orígenes de replicación) necesarios para el cultivo celular (p. ej., cultivo de células bacterianas, cultivo de células de mamífero). En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias de polinucleótidos que codifican genes de resistencia a antibióticos (p. ej., genes de resistencia a kanamicina, genes de resistencia a ampicilina). En algunas
10 realizaciones, las construcciones de expresión comprenden una o más secuencias de polinucleótidos que codifican un origen de replicación bacteriano (p. ej., origen de replicación *colE1*).

En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias de terminación de la transcripción (p. ej., una secuencia de *poliA*).
15 En algunas realizaciones, las construcciones de expresión comprenden uno o más de BGH *poliA*, FIX *poliA*, SV40 *poliA*, *poliA* sintética, o combinaciones de los mismos. En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias de terminación de la transcripción por detrás de un elemento de secuencia particular (p. ej., un gen o variante de gen *rep* o *cap*). En algunas
20 realizaciones, las construcciones de expresión comprenden una o más secuencias de terminación de la transcripción por delante de un elemento de secuencia particular (p. ej., un gen o variante de gen *rep* o *cap*).

En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas. En algunas realizaciones, las construcciones de
25 expresión comprenden uno o más intrones de diferentes orígenes (p. ej., genes conocidos), incluyendo pero sin limitación, el intrón FIX, el intrón de Albúmina, o combinaciones de los mismos. En algunas realizaciones, las construcciones de expresión comprenden uno o más intrones de diferentes longitudes (p. ej., de 133 pb a 4 kb). En algunas realizaciones, las construcciones de expresión comprenden
30 una o más secuencias intrónicas por delante de un elemento de secuencia particular

(p. ej., un gen o variante de gen *rep* o *cap*). En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas dentro de un elemento de secuencia particular (p. ej., un gen o variante de gen *rep* o *cap*). En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas por detrás de un elemento de secuencia particular (p. ej., un gen o variante de gen *rep* o *cap*). En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas después de un promotor (p. ej., un promotor p5). En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas antes de un gen o variante de gen *rep*. En algunas realizaciones, las construcciones de expresión comprenden una o más secuencias intrónicas entre un promotor y gen o variante de gen *rep*.

Ejemplos ilustrativos

Ejemplo 1: La rotura bicatenaria inducida por Cas9 puede mejorar la eficacia de GENERIDE™

Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante de un locus humano endógeno (p. ej., albúmina (ALB) humana) junto con una o más enzimas Cas9 dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana), puede mejorar la eficacia de integración de transgenes.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., UGT1A1), un flaqueo equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de MOI (*Multiplicity of infection*, multiplicidad de infección)=1E⁵. Después de la transducción, las células se transfectaron con spCas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del

fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 3). Cuarenta y ocho (48) horas después de la transfección, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado (p. ej., véase la Tabla 8).

5 Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, que comprenden un transgén (p. ej., UGT1A1) junto con spCas9, pueden mejorar la actividad de edición. En algunas realizaciones, spCas9 puede inducir una rotura bicatenaria (de doble cadena, DSB) en un sitio de corte dentro del intrón 13 y/o 14 del gen ALB (Fig. 1A y 1B). En algunas
10 realizaciones, como se demuestra en la Fig. 2A, la DSB inducida en el intrón 13 del gen de ALB humano, puede mejorar la eficacia de integración del transgén en al menos 400 veces en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB). En algunas realizaciones, como se
15 demuestra en la Fig. 2B, la DSB inducida en el intrón 14 del gen de ALB humano, puede mejorar la eficacia de integración del transgén en al menos 200 veces en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB). En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una DSB en cualquier secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes en
20 relación con una construcción GENERIDE™ sola. En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una DSB en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración del transgén o puede ser comparable en comparación con una DSB en una región alternativa de una secuencia no
25 codificante (Fig. 2C). En algunas realizaciones, una DSB se localiza en una secuencia no codificante distal al sitio de integración del transgén.

 Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., GFP), un flanco equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología
30 se diseñaron para que fueran complementarios a un sitio de integración diana de

albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de $MOI=1E^5$. Después de la transducción, las células se transfectaron con spCas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 3). Cuarenta y ocho (48) horas después de la transfección, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado. Además, se realizó inmunohistoquímica y/o inmunocitoquímica (por ejemplo, véase la Tabla 8) Las imágenes se evaluaron utilizando un programa informático de procesamiento de imágenes (p. Ej., ImageJ).

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, que comprenden un transgén (p. ej., GFP) junto con spCas9, pueden mejorar la actividad de edición y la expresión de proteínas. Como se demuestra en las Figs. 2D y 2E, los DSB inducidos en el intrón 13 y/o 14 pueden mejorar al menos 5 veces la expresión de proteínas (p. ej., % de células positivas a GFP). Además, como se muestra en la Figura 2F, el aumento de los niveles de ARNm fusionado puede estar asociado a un aumento de la expresión de proteínas.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., GFP), un flanqueo equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de $MOI=1E^5$. Después de la transducción, las células se transfectaron con ARNm de spCas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en

el intrón 13 o 14 (Tabla 8). Setenta y dos (72) horas después de la transfección, las células se fijaron y se obtuvieron imágenes. Las células positivas a GFP se cuantificaron utilizando programas informáticos (p. ej., Image J o CellProfiler) (Fig. 3A).

Entre otras cosas, este ejemplo demuestra que el ARNm de spCas9 también puede mejorar la actividad de edición de los vectores víricos, como se describe en el presente documento (Fig. 3B). Como se demuestra en la Figura 4, una proporción de aproximadamente 5 de ARNm de spCas9 por ARN guía, puede mejorar la actividad de edición de los vectores víricos. En algunas realizaciones, la proporción de ARNm spCas9 y ARN guía puede optimizarse para mejorar la eficacia de integración del transgén.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., GFP, GFP conjugada con etiqueta de HA o UGT1A1), un flanqueo equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de $MOI=1E^5$. Después de la transducción, las células se transfectaron con Cas9 (p. ej., spCas9-HF1 y/o spCas9-HF2) y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 3). Cuarenta y ocho (48) horas después de la transfección, se realizó inmunohistoquímica y/o inmunocitoquímica utilizando un anticuerpo. Las imágenes se evaluaron utilizando un programa informático de procesamiento de imágenes (p. ej., ImageJ).

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con diferentes spCas9 (p. ej., spCas9-HF1 y/o spCas9-HF2) que utilizan ARNg diseñado para dirigirse al intrón 13 y/o 14 de la albúmina, pueden mejorar la actividad de edición de GENERIDE™. Como se

demuestra en la Figura 5, la DSB inducida por spCas9-HF1 y/o spCas9-HF2 en el intrón 14, puede mejorar la expresión de proteínas (p. ej., % de células positivas a GFP) en comparación con la DSB inducida en el intrón 13.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico conjugado con HA (p. ej., GFP, GFP conjugado con HA y/o UGT1A1), un flanqueo equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de $MOI=1E^5$. Después de la transducción, las células se transfectaron con saCas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 3). Se realizó inmunohistoquímica y/o inmunocitoquímica (p. ej., véase la Tabla 8) Las imágenes se evaluaron utilizando un programa informático de procesamiento de imágenes (p. ej., ImageJ).

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con saCas9 que utilizan ARNg diseñado para dirigirse al intrón 13 y/o 14 de la albúmina, pueden mejorar la actividad de edición de GENERIDE™. Como se demuestra en la Figura 6, la DSB inducida por saCas9 puede aumentar la expresión de proteínas (p. ej., % de células positivas HA+).

Ejemplo 2: La rotura de ADN inducida por nucleasas Cas puede aumentar la eficacia de GENERIDE™

Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante de un locus humano endógeno (albúmina humana) junto con una o más enzimas Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) dirigidas a un sitio de

corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana), puede mejorar la eficacia de integración de transgenes.

Se construyó una primera composición que comprendía vectores víricos que comprendían una cápside vírica, una secuencia P2A, un gen transgénico y longitudes de brazos de homología en 5' y 3' flanqueantes equilibrados o no equilibrados. Los brazos de homología se diseñan para que sean complementarios a un sitio de integración diana (p. ej., albúmina genómica humana).

Se construye una segunda composición que comprende una enzima Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) y un ARNg diseñado para dirigirse a una secuencia no codificante de un locus humano endógeno (p. ej., intrón ALB humana). En la técnica se conocen diversos métodos para suministrar dicha composición (p. ej., transfección con plásmidos o ARNm que codifican una enzima Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) y un ARNg, transducción con vectores víricos que codifican una enzima Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) y un ARNg, y/o suministro intracelular de una enzima Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares)-ARNg ribonucleoproteína (RNP))

Después de la construcción de cada composición, se prueban diversas condiciones experimentales. En un experimento, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición. En otro experimento, las células (p. ej., HepG2) se transducen secuencialmente con una primera composición y después se transducen y/o transfectan con una segunda composición. En otro experimento, las células (p. ej., HepG2) se transducen y/o transfectan secuencialmente con una segunda composición y después se transducen con una primera composición. En otro experimento, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición, y después se transducen y/o transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. En otro experimento, las células se transducen secuencialmente con una primera composición y después se transducen y/o

transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. Después de la transfección y/o transducción, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado.

5 Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con enzimas Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) pueden mejorar la actividad de edición. En algunas realizaciones, las enzimas Cas (p. ej., Cas9, Cas13, Cas12a, Cas9 nicasa, y similares) pueden inducir una rotura de ADN (p. ej., una rotura bicatenaria (DSB) y/o una rotura monocatenaria (SSB)) dentro de una secuencia no codificante de un gen de ALB humano. En algunas realizaciones, la rotura de ADN inducida (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano puede mejorar la eficacia de integración de transgenes en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB y/o SSB).

10 En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en cualquier secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes en relación con una construcción GENERIDE™ sola. En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una

15 rotura de ADN inducida (p. ej., DSB y/o SSB) en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes o puede ser comparable en comparación con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región alternativa de una secuencia no codificante. En algunas realizaciones, una rotura de ADN inducida (p. ej., DSB y/o SSB) se localiza en una secuencia no codificante que es distal con respecto al

20 sitio de integración del transgén.

25

Ejemplo 3: La rotura de ADN inducida por ZFN puede aumentar la eficacia de GENERIDE™

30 Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante

de un locus humano endógeno (p. ej., albúmina (ALB) humana) junto con una o más nucleasas de dedos de zinc (ZFN) (p. ej., pares de ZFN y/o Nicasa de ZFN (ZFNicasa) dirigida a un sitio de corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de albúmina humana), puede mejorar la

5 eficacia de integración de transgenes.

Se construyó una primera composición que comprendía vectores víricos que comprendían una cápside vírica, una secuencia P2A, un gen transgénico y longitudes de brazos de homología en 5' y 3' flanqueantes equilibrados o no equilibrados. Los brazos de homología se diseñan para que sean complementarios a un sitio de integración diana (p. ej., albúmina genómica humana).

10

Se construye una segunda composición que comprende una o más ZFN (p. ej., pares de ZFN y/o Nicasa de ZF) diseñadas para dirigirse a una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana). Una ZFN/ZFNicasa puede prepararse de forma que pueda unirse a una región específica de un gen (por ejemplo, el intrón 13 y/o 14 de un gen de albúmina humana) y comprender un dominio nucleasa (por ejemplo, FokI) capaz de cortar (p. ej., ZFNicasa) o escindir (p. ej., pares de ZFN) la secuencia diana. En la técnica se conocen diversos métodos para suministrar dicha composición (p. ej., transfección con plásmidos o ARNm que codifican una o más ZFN (p. ej., pares de ZFN y/o

15 Nicasa de ZF), transducción con vectores víricos que codifican una o más ZFN (p. ej., pares de ZFN y/o Nicasa de ZF) y/o electroporación).

20

Después de la construcción de cada composición, se prueban diversas condiciones experimentales. En un experimento, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen secuencialmente con una primera composición y después se transducen y/o transfectan con una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen y/o transfectan secuencialmente con una segunda composición y después se transducen con una primera

25

30 composición. En un experimento posterior, las células (p. ej., HepG2) se transducen

y/o transfectan simultáneamente con una primera composición y una segunda composición, y después se transducen y/o transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. En un experimento posterior, las células se transducen secuencialmente con una primera composición y después se transducen y/o transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. Después de la transfección y/o transducción, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado.

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con una o más ZFN (p. ej., pares de ZFN y/o Nicasa de ZF) pueden proporcionar una actividad de edición mejorada. En algunas realizaciones, una o más ZFN (p. ej., pares de ZFN y/o Nicasa de ZF (ZFNicasa)) pueden inducir una rotura de ADN (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano. En algunas realizaciones, la rotura de ADN inducida (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano puede mejorar la eficacia de integración de transgenes en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB y/o SSB). En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en cualquier secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes en relación con una construcción GENERIDE™ sola. En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes o puede ser comparable en comparación con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región alternativa de una secuencia no codificante. En algunas realizaciones, una rotura de ADN inducida (p. ej., DSB y/o SSB) se localiza en una secuencia no codificante que es distal con respecto al sitio de integración del transgén.

Ejemplo 4: La rotura de ADN inducida por meganucleasas puede aumentar la eficacia de GENERIDE™

Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante de un locus humano endógeno (p. ej., albúmina (ALB) humana) junto con una o más meganucleasas dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de albúmina humana), puede mejorar la eficacia de integración de transgenes.

Se construyó una primera composición que comprendía vectores víricos que comprendían una cápside vírica, una secuencia P2A, un gen transgénico y longitudes de brazos de homología en 5' y 3' flanqueantes equilibrados o no equilibrados. Los brazos de homología se diseñan para que sean complementarios a un sitio de integración diana (p. ej., albúmina genómica humana).

Se construye una segunda composición que comprende una o más meganucleasas diseñadas para dirigirse a una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana). En la técnica se conocen diversos métodos para suministrar dicha composición (p. ej., transfección con plásmidos o ARNm que codifican una o más meganucleasas y/o transducción con vectores víricos que codifican una o más meganucleasas).

Después de la construcción de cada composición, se prueban diversas condiciones experimentales. En un experimento, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen secuencialmente con una primera composición y después se transducen y/o transfectan con una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen y/o transfectan secuencialmente con una segunda composición y después se transducen con una primera composición. En un experimento posterior, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición, y después se transducen y/o transfectan varias veces (p. ej., al menos

más de una vez) con una segunda composición. En un experimento posterior, las células se transducen secuencialmente con una primera composición y después se transducen y/o transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. Después de la transfección y/o transducción, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado.

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con una o más meganucleasas pueden mejorar la actividad de edición. En algunas realizaciones, una o más meganucleasas pueden inducir una rotura de ADN (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano. En algunas realizaciones, la rotura de ADN inducida (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano puede mejorar la eficacia de integración de transgenes en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB y/o SSB). En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en cualquier secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes en relación con una construcción GENERIDE™ sola. En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes o puede ser comparable en comparación con una rotura de ADN inducida (p. ej., DSB y/o SSB). En algunas realizaciones, una rotura de ADN inducida (p. ej., DSB y/o SSB) se localiza en una secuencia no codificante que es distal con respecto al sitio de integración del transgén.

Ejemplo 5: La rotura de ADN inducida por una nucleasa ilustrativa puede mejorar la eficacia GENERIDE™

Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante

de un locus humano endógeno (p. ej., albúmina (ALB) humana) junto con una o más nucleasas ilustrativas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de albúmina humana), puede mejorar la eficacia de integración de transgenes.

Se construyó una primera composición que comprendía vectores víricos que comprendían una cápside vírica, una secuencia P2A, un gen transgénico y longitudes de brazos de homología en 5' y 3' flanqueantes equilibrados o no equilibrados. Los brazos de homología se diseñan para que sean complementarios a un sitio de integración diana (p. ej., albúmina genómica humana).

Se construye una segunda composición que comprende una o más nucleasas ilustrativas, como se describe en el presente documento, diseñadas para dirigirse a una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana). En la técnica se conocen diversos métodos para suministrar dicha composición (p. ej., transfección con plásmidos o ARNm que codifican una o más nucleasas ilustrativas, como se describe en el presente documento, transducción con vectores víricos que codifican una o más nucleasas ilustrativas, como se describe en el presente documento, y/o suministro intracelular de una o más nucleasas RNP ilustrativas, como se describe en el presente documento).

Después de la construcción de cada composición, se prueban diversas condiciones experimentales. En un experimento, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen secuencialmente con una primera composición y después se transducen y/o transfectan con una segunda composición. En un experimento posterior, las células (p. ej., HepG2) se transducen y/o transfectan secuencialmente con una segunda composición y después se transducen con una primera composición. En un experimento posterior, las células (p. ej., HepG2) se transducen y/o transfectan simultáneamente con una primera composición y una segunda composición, y después se transducen y/o transfectan varias veces (p. ej., al menos

más de una vez) con una segunda composición. En un experimento posterior, las células se transducen secuencialmente con una primera composición y después se transducen y/o transfectan varias veces (p. ej., al menos más de una vez) con una segunda composición. Después de la transfección y/o transducción, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado.

Entre otras cosas, este ejemplo demuestra que los vectores víricos, como se describe en el presente documento, junto con una o más nucleasas ilustrativas, como se describe en el presente documento, pueden mejorar la actividad de edición. En algunas realizaciones, una o más nucleasas ilustrativas, como se describe en el presente documento, pueden inducir una rotura de ADN (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano. En algunas realizaciones, la rotura de ADN inducida (p. ej., DSB y/o SSB) dentro de una secuencia no codificante de un gen de ALB humano puede mejorar la eficacia de integración de transgenes en comparación con una referencia apropiada (p. ej., administración de vector sin inducción de DSB y/o SSB). En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en cualquier secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes en relación con una construcción GENERIDE™ sola. En algunas realizaciones, la administración de una construcción GENERIDE™ junto con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes o puede ser comparable en comparación con una rotura de ADN inducida (p. ej., DSB y/o SSB) en una región alternativa de una secuencia no codificante. En algunas realizaciones, una rotura de ADN inducida (p. ej., DSB y/o SSB) se localiza en una secuencia no codificante que es distal con respecto al sitio de integración del transgén.

Ejemplo 6: La rotura de ADN inducida puede aumentar la eficacia de GENERIDE™ *in vivo* o *ex vivo*

Este ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante de un locus endógeno (p. ej., albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno (p. ej., intrón de albúmina humana), puede mejorar la eficacia de integración de transgenes en un sujeto (p. ej., un sujeto que padece una enfermedad).

En algunas realizaciones, este ejemplo incluye el uso de una o más construcciones GENERIDE™, p. ej., vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (p. ej., P2A), un transgén, como se describe en el presente documento, y brazos de homología humanos flanqueantes (p. ej., equilibrados o no equilibrados).

En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece una enfermedad). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (p. ej., una célula hepática) *ex vivo*, opcionalmente seguido de la administración de la célula a un sujeto (p. ej., un sujeto que padece una enfermedad) después de la integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se administran antes de la administración con una o más endonucleasas. En algunas realizaciones, las construcciones GENERIDE™ se administran después de la administración con una o más endonucleasas. En algunas realizaciones, las construcciones GENERIDE se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece una enfermedad), opcionalmente seguido de la administración con una o más endonucleasas.

Entre otras cosas, este ejemplo proporciona una idea de que la rotura inducida de ADN (p. ej., DSB y/o SSB) puede mejorar la eficacia de GENERIDE™

in-vivo. En alguna realización, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura de ADN inducida (*p. ej.*, DSB y/o SSB). En algunas realizaciones, los niveles de ALB-2A pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura inducida de ADN (*p. ej.*, DSB y/o SSB). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB).

Entre otras cosas, este ejemplo proporciona una visión de que las construcciones GENERIDE™ administradas de manera secuencial o simultánea con una o más endonucleasas puede proporcionar una ventaja selectiva mejorada para las células que han integrado con éxito un transgén de interés.

Ejemplo 7: La rotura de ADN inducida con GENERIDE™ puede mejorar el fenotipo de la enfermedad de Wilson

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™, que comprende un transgén terapéutico que codifica ATP7B (*p. ej.*, un ATP7B truncado), a un sitio de integración diana en una región codificante de un locus endógeno (*p. ej.*, albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno (*p. ej.*, intrón de albúmina humana), puede usarse para tratar o prevenir la enfermedad de Wilson (*p. ej.*, mediante la reducción de efectos fenotípicos y/o síntomas *in vivo*) en un sujeto (*p. ej.*, un sujeto que padece una enfermedad de Wilson).

Se construyen construcciones GENERIDE™, *p. ej.*, vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (*p. ej.*, P2A), un transgén ATP7B humano (*p. ej.*, un ATP7B truncado), y brazos de homología humanos flanqueantes (*p. ej.*, equilibrados o no equilibrados). Las construcciones GENERIDE™ se administran (*p. ej.*, por vía

intravenosa) a una dosis (*p. ej.*, a una dosis fija o una dosis basada en el peso) a un sujeto (*p. ej.*, un sujeto que padece la enfermedad de Wilson). En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (*por ejemplo*, un sujeto que padece la enfermedad de Wilson). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (*p. ej.*, una célula hepática) *ex vivo*, opcionalmente seguido de la administración de la célula a un sujeto (*p. ej.*, un sujeto que padece la enfermedad de Wilson) después de la integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (*p. ej.*, un sujeto que padece la enfermedad de Wilson), opcionalmente seguido de la administración con una o más endonucleasas.

Entre otras cosas, este ejemplo proporciona una idea de que la rotura inducida de ADN (*p. ej.*, DSB y/o SSB) puede mejorar la eficacia de GENERIDE™ *in-vivo*. En algunas realizaciones, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura de ADN inducida (*p. ej.*, DSB y/o SSB)). En algunas realizaciones, los niveles de ALBA-2A pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura inducida de ADN (*p. ej.*, DSB y/o SSB)). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB)).

Entre otras cosas, el presente ejemplo demuestra que el tratamiento de un sujeto (*por ejemplo*, un sujeto que padece la enfermedad de Wilson) con los vectores víricos de la presente divulgación, puede aumentar una tasa y/o un nivel de integración (*por ejemplo*, medida a través del porcentaje de células positivas) en

relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece la enfermedad de Wilson) con los vectores víricos de la presente divulgación, puede permitir o restaurar el crecimiento normal (*p. ej.*, medido a través de cambios en el porcentaje de peso corporal a lo largo del tiempo) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad de Wilson) con los vectores víricos de la presente divulgación, puede proporcionar niveles reducidos de un biomarcador (*p. ej.*, cobre, hAAT, alanina aminotransferasa (ALT)) asociado a la enfermedad de Wilson. En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad de Wilson) con los vectores víricos de la presente divulgación, puede proporcionar una o más de una mejora de la función tisular (*p. ej.*, hepática) (*p. ej.*, medida mediante evaluación de marcadores de función tisular (*p. ej.*, función hepática)), crecimiento normal (*p. ej.*, medido mediante cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia (*p. ej.*, sin tratar o con vehículo), niveles reducidos de un biomarcador (*p. ej.*, cobre hepático o urinario, hAAT, alanina aminotransferasa (ALT)) asociado a la enfermedad de Wilson.

Ejemplo 8: La rotura de ADN inducida con GENERIDE™ puede mejorar el fenotipo de la Tirosinemia Hereditaria

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™, que comprende un transgén terapéutico que codifica FAH y/o una variante del mismo, a un sitio de integración diana en una región codificante de un locus endógeno (*p. ej.*, albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno (*p. ej.*, intrón de albúmina humana), puede usarse para tratar o prevenir la tirosinemia hereditaria (*p. ej.*, mediante la reducción de efectos fenotípicos y/o síntomas *in vivo*) en un sujeto (*p. ej.*, un sujeto que padece una tirosinemia hereditaria).

Se construyen construcciones GENERIDE™, p. ej., vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (p. ej., P2A), un transgén FAH humano y/o una variante del mismo, y brazos de homología humanos flanqueantes (p. ej., equilibrados o no equilibrados).

5 Las construcciones GENERIDE se administran (p. ej., por vía intravenosa) a una dosis (p. ej., a una dosis fija o una dosis basada en el peso) a un sujeto (p. ej., un sujeto que padece tirosinemia hereditaria). En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En
10 algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (por ejemplo, un sujeto que padece tirosinemia hereditaria). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (p. ej., una célula hepática) *ex vivo*, opcionalmente seguido de la administración de la célula a un
15 sujeto (p. ej., un sujeto que padece una tirosinemia hereditaria) después de la integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece tirosinemia hereditaria), opcionalmente seguido de la administración con una o más endonucleasas.

20 Entre otras cosas, este ejemplo proporciona una idea de que la rotura inducida de ADN (p. ej., DSB y/o SSB) puede mejorar la eficacia de GENERIDE™ *in-vivo*. En algunas realizaciones, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (p. ej., administración de vector sin rotura de ADN inducida (p. ej., DSB
25 y/o SSB). En algunas realizaciones, los niveles de ALBA-2A pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (p. ej., administración de vector sin rotura inducida de ADN (p. ej., DSB y/o SSB). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia

apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB).

Entre otras cosas, este ejemplo demuestra que el tratamiento de un sujeto (*p. ej.*, un sujeto que padece tirosinemia hereditaria) con los vectores víricos de la presente divulgación, puede aumentar una tasa y/o un nivel de integración (*p. ej.*, medida a través del porcentaje de células positivas) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece tirosinemia hereditaria) con los vectores víricos de la presente divulgación, puede permitir o restaurar el crecimiento normal (*p. ej.*, medido a través de cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una tirosinemia hereditaria) con los vectores víricos de la presente divulgación, puede proporcionar niveles reducidos de un biomarcador (*p. ej.*, ALT, bilirrubina, succinilacetona (SUAC), alfa-fetoproteína (AFP)) asociado a tirosinemia hereditaria. En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una tirosinemia hereditaria) con los vectores víricos de la presente divulgación, puede proporcionar una o más de una mejora de la función tisular (*p. ej.*, hepática) (*p. ej.*, medida mediante evaluación de marcadores de función tisular (*p. ej.*, función hepática)), crecimiento normal (*p. ej.*, medido a través de cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia (*p. ej.*, sin tratar o con vehículo), niveles reducidos de un biomarcador (*p. ej.*, ALT, bilirrubina, succinilacetona (SUAC), alfa-fetoproteína (AFP)) a tirosinemia hereditaria.

Ejemplo 9: La rotura de ADN inducida con GENERIDE™ puede mejorar el síndrome de Crigler-Najjar

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™, que comprende un gen UGT1A1 y/o una variante del mismo, a un sitio de integración diana en una región codificante de un locus endógeno (*p. ej.*, albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, que se dirige a un sitio de corte dentro de una

secuencia no codificante de un locus endógeno (p. ej., intrón de albúmina humana), puede usarse para tratar o prevenir el síndrome de Crigler-Najjar (p. ej., mediante la reducción de efectos fenotípicos y/o síntomas *in vivo*) en un sujeto (p. ej., un sujeto que padece el síndrome de Crigler-Najjar).

5 Se construyen construcciones GENERIDE™, p. ej., vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (p. ej., P2A), un transgén UGT1A1 y/o una variante del mismo, y brazos de homología humanos flanqueantes (p. ej., equilibrados o no equilibrados). Las construcciones GENERIDE™ se administran (p. ej., por vía intravenosa) a una
10 dosis (p. ej., a una dosis fija o una dosis basada en el peso) a un sujeto (p. ej., un sujeto que padece el síndrome de Crigler-Najjar). En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una
15 o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece el síndrome de Crigler-Najjar). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (p. ej., una célula hepática) *ex vivo*, opcionalmente seguido de la administración de la célula a un sujeto (p. ej., un sujeto que padece el síndrome de Crigler-Najjar) después de la
20 integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece el síndrome de Crigler-Najjar), opcionalmente seguido de la administración con una o más endonucleasas.

 Entre otras cosas, este ejemplo proporciona una idea de que la rotura
25 inducida de ADN (p. ej., DSB y/o SSB) puede mejorar la eficacia de GENERIDE™ *in-vivo*. En algunas realizaciones, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (p. ej., administración de vector sin rotura de ADN inducida (p. ej., DSB y/o SSB). En algunas realizaciones, los niveles de ALBA-2A pueden aumentar en
30 un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia

apropiada (*p. ej.*, administración de vector sin rotura inducida de ADN (*p. ej.*, DSB y/o SSB). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB).

Entre otras cosas, este ejemplo demuestra que el tratamiento de un sujeto (*p. ej.*, un sujeto que padece el síndrome de Crigler-Najjar) con los vectores víricos de la presente divulgación, puede aumentar una tasa y/o un nivel de integración (*p. ej.*, medido a través del porcentaje de células positivas) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece el síndrome de Crigler-Najjar) con los vectores víricos de la presente divulgación, puede permitir o restaurar el crecimiento normal (*p. ej.*, medido a través de cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece el síndrome de Crigler-Najjar) con los vectores víricos de la presente divulgación, puede proporcionar niveles reducidos de un biomarcador asociado al síndrome de Crigler-Najjar. En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece el síndrome de Crigler-Najjar) con los vectores víricos de la presente divulgación, puede proporcionar una o más de una mejora de la función tisular (*p. ej.*, hepática) (*p. ej.*, medida mediante evaluación de marcadores de función tisular (*p. ej.*, función hepática)), crecimiento normal (*p. ej.*, medido a través de cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia (*p. ej.*, sin tratar o con vehículo), niveles reducidos de un biomarcador asociado al síndrome de Crigler-Najjar.

Ejemplo 10: La rotura de ADN inducida con GENERIDE™ puede mejorar la carencia de alfa-1 antitripsina

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™, que comprende un transgén SERPINA1 y/o variantes del mismo, a un sitio de integración diana en una región codificante de un

locus endógeno (p. ej., albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno (p. ej., intrón de albúmina (ALB) humana), puede usarse para tratar o prevenir la carencia de alfa-1 antitripsina (p. ej., mediante la reducción de efectos fenotípicos y/o síntomas *in vivo*) en un sujeto (p. ej., un sujeto que padece carencia de alfa-1 antitripsina).

Se construyen construcciones GENERIDE™, p. ej., vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (p. ej., P2A), un transgén SERPINA1 y/o una variante del mismo, y brazos de homología humanos flanqueantes (p. ej., equilibrados o no equilibrados). Las construcciones GENERIDE™ se administran (p. ej., por vía intravenosa) a una dosis (p. ej., una dosis fija o una dosis basada en el peso) a un sujeto (p. ej., un sujeto que padece carencia de alfa-1 antitripsina). En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece carencia de alfa-1 antitripsina). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (p. ej., célula hepática y/o pulmonar) *ex vivo*, opcionalmente seguido de la administración de la célula a un sujeto (p. ej., un sujeto que padece carencia de alfa-1 antitripsina) después de la integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece carencia de alfa-1 antitripsina), opcionalmente seguido de la administración con una o más endonucleasas.

Entre otras cosas, este ejemplo proporciona una idea de que la rotura inducida de ADN (p. ej., DSB y/o SSB) puede mejorar la eficacia de GENERIDE™ *in-vivo*. En algunas realizaciones, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (p. ej., administración de vector sin rotura de ADN inducida (p. ej., DSB

y/o SSB). En algunas realizaciones, los niveles de ALBA-2A pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura inducida de ADN (*p. ej.*, DSB y/o SSB). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB)).

Entre otras cosas, este ejemplo demuestra que el tratamiento de un sujeto (*p. ej.*, un sujeto que padece carencia de alfa-1 antitripsina) con los vectores víricos de la presente divulgación, puede aumentar una tasa y/o un nivel de integración (*p. ej.*, medido a través del porcentaje de células positivas) en relación con una referencia adecuada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece carencia de alfa-1 antitripsina) con los vectores víricos de la presente divulgación, puede permitir o restaurar la función pulmonar en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece carencia de alfa-1 antitripsina) con los vectores víricos de la presente divulgación, puede proporcionar niveles reducidos de un biomarcador asociado a la carencia de alfa-1 Antitripsina. En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece carencia de alfa-1 antitripsina) con los vectores víricos de la presente divulgación, puede proporcionar una o más de una mejora de la función tisular (*p. ej.*, pulmonar) (*p. ej.*, medida mediante evaluación de marcadores de función tisular (*p. ej.*, pulmonar)), función respiratoria normal o reducción del empeoramiento de la función respiratoria, niveles reducidos de un biomarcador asociado a la carencia de alfa-1 antitripsina.

Ejemplo 11: La rotura bicatenaria inducida con GENERIDE™ puede mejorar la función tisular

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™, que comprende un transgén terapéutico y/o una variante del mismo, a un sitio de integración diana en una región codificante de un

locus endógeno (p. ej., albúmina humana) junto con una o más endonucleasas, como se describe en el presente documento, dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno (p. ej., intrón de albúmina humana), puede usarse para tratar o prevenir enfermedades (p. ej., mediante la reducción de efectos fenotípicos y/o síntomas *in vivo*) en un sujeto (p. ej., un sujeto que padece una enfermedad).

Se construyen construcciones GENERIDE™, p. ej., vectores víricos que comprenden una cápside vírica, como se describe en el presente documento, una secuencia 2A (p. ej., P2A), un transgén terapéutico y/o una variante del mismo, y brazos de homología humanos flanqueantes (p. ej., equilibrados o no equilibrados). Las construcciones GENERIDE se administran (p. ej., por vía intravenosa) a una dosis (p. ej., a una dosis fija o una dosis basada en el peso) a un sujeto (p. ej., un sujeto que padece una enfermedad). En algunas realizaciones, las construcciones GENERIDE™ se administran de manera secuencial o simultánea con una o más endonucleasas, como se describe en el presente documento. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece una enfermedad). En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a una célula (p. ej., una célula hepática) *ex vivo*, opcionalmente seguido de la administración de la célula a un sujeto (p. ej., un sujeto que padece una enfermedad) después de la integración del transgén. En algunas realizaciones, las construcciones GENERIDE™ se coadministran con una o más endonucleasas a un sujeto *in vivo* (p. ej., un sujeto que padece una enfermedad), opcionalmente seguido de la administración con una o más endonucleasas.

Entre otras cosas, este ejemplo proporciona una idea de que la rotura inducida de ADN (p. ej., DSB y/o SSB) puede mejorar la eficacia de GENERIDE™ *in-vivo*. En alguna realización, los niveles de ARNm fusionado pueden aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (p. ej., administración de vector sin rotura de ADN inducida (p. ej., DSB y/o SSB)). En algunas realizaciones, los niveles de ALBA-2A pueden aumentar en

un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*p. ej.*, administración de vector sin rotura inducida de ADN (*p. ej.*, DSB y/o SSB). En algunas realizaciones, una expresión transgénica puede aumentar en un sujeto, y/o en una célula, y/o en un tejido en comparación con una referencia apropiada (*por ejemplo*, administración de vector sin rotura de ADN inducida (*por ejemplo*, DSB y/o SSB).

Entre otras cosas, este ejemplo demuestra que el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad) con los vectores víricos de la presente divulgación, puede aumentar una tasa y/o un nivel de integración (*p. ej.*, medido a través del porcentaje de células positivas) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad) con los vectores víricos de la presente divulgación, puede permitir o restaurar el crecimiento normal (*p. ej.*, medido a través de cambios en el porcentaje de peso corporal a lo largo del tiempo) en relación con una referencia apropiada (*p. ej.*, sin tratar o con vehículo). En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad) con los vectores víricos de la presente divulgación, puede proporcionar niveles reducidos de un biomarcador asociado a la enfermedad. En algunas realizaciones, el tratamiento de un sujeto (*p. ej.*, un sujeto que padece una enfermedad) con los vectores víricos de la presente divulgación, puede proporcionar una o más de una mejora de la función tisular (*p. ej.*, hepática) (*p. ej.*, medida mediante evaluación de marcadores de función tisular (*p. ej.*, función hepática)), crecimiento normal (*p. ej.*, medido a través de cambios porcentuales del peso corporal a lo largo del tiempo) en relación con una referencia (*p. ej.*, sin tratar o con vehículo), niveles reducidos de un biomarcador asociado a la enfermedad.

Ejemplo 12: Las alineaciones de homología más largas predicen mayor eficacia la integración

El presente ejemplo demuestra que, entre otras cosas, la administración de una construcción GENERIDE™ con alineaciones de homología más largas junto con una o más enzimas Cas9 dirigidas a un sitio de corte dentro de una secuencia

no codificante de un locus humano endógeno (p. ej., intrón de albúmina (ALB) humana), puede mejorar la eficacia de integración de transgenes.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., GFP), diferentes longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de $MOI=1E^5$. Después de la transducción, las células se transfectaron con spCas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 8). Cuarenta y ocho (48) horas después de la transfección, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado. Además, se realizó inmunohistoquímica y/o inmunocitoquímica utilizando anticuerpos anti-GFP. Las imágenes se evaluaron utilizando un programa informático de procesamiento de imágenes (p. ej., ImageJ).

Entre otras cosas, este ejemplo demuestra que las alineaciones de homología más largas pueden proporcionar una mayor eficacia de integración. En algunas realizaciones, como se demuestra en la Figura 4B, una alineación de homología de al menos 800 pb puede mejorar la eficacia de integración de transgenes (medida como niveles de ARNm fusionado) en comparación con una referencia apropiada. En algunas realizaciones, como se demuestra en la Figura 4C, la alineación de homología de al menos 800 pb puede mejorar la eficacia de integración del transgén (medida como niveles de células positivas a GFP). En algunas realizaciones, las alineaciones de homología pueden optimizarse para mejorar la eficacia de integración del transgén.

Ejemplo 13: La rotura bicatenaria inducida por Cas9 con ARNg apropiado también demuestra mejora de la eficacia de GENERIDE™

Este ejemplo demuestra además que la combinación de una construcción GENERIDE™ en un sitio de integración diana en una región codificante de un locus humano endógeno (p. ej., albúmina (ALB) humana) con una o más enzimas Cas9 dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus humano endógeno (p. ej., intrón de ALB humana) puede mejorar la eficacia de integración de transgenes.

Se construyeron vectores víricos que comprendían una cápside vírica (p. ej., LK03), la secuencia P2A, un gen transgénico (p. ej., GFP), un flanqueo equilibrado (1kb/1kb) de longitudes de brazos de homología en 5' y 3'. Los brazos de homología se diseñaron para que fueran complementarios a un sitio de integración diana de albúmina genómica humana. Células HepG2, una línea celular de carcinoma hepatocelular de hígado humano, se transdujeron con vectores víricos a una dosis de MOI=1E⁵. Después de la transducción, las células se transfectaron con Cas9 y ARNg (diseñado usando el algoritmo de IDT para ARNg) utilizando Lipofectamine™ CRISPRMAX (ThermoFisher Scientific, Inc) de acuerdo con las instrucciones del fabricante. Los ARNg se diseñaron para dirigirse al gen de ALB en el intrón 13 o 14 (Tabla 9). Cuarenta y ocho (48) horas después de la transfección, se aisló el ARN total de las células utilizando el kit RNeasy (Qiagen) siguiendo las instrucciones del proveedor y se evaluaron los niveles de ARNm fusionado (p. ej., véase la Tabla 8). La Tabla 9 muestra cuales son las secuencias espaciadoras del ARNg que son idénticas a las de la secuencia análoga en la ALB.

Tabla 9: Secuencias ilustrativas de ARNg para spCas9.

Locus diana	Cadena	Secuencia	SEQ ID NO:
Albúmina humana Intrón 13_11	-	ACAGTATGGCACAATAGAGC	71
Albúmina humana Intrón 13_12	-	AGTTACTGTCAGGCCATAAA	72

5	Albúmina humana Intrón 13_13	-	AAGTTACTGTCAGGCCATAA	73
	Albúmina humana Intrón 13_14	+	AGGGATTTATATATCAAAGG	74
	Albúmina humana Intrón 13_15	+	AAAGGAGGCTTTGTACATGT	75
	Albúmina humana Intrón 13_16	-	TTACTACTATAATCTCTGGA	76
10	Albúmina humana Intrón 13_17	+	GGCTTTGTACATGTGGGACA	77
	Albúmina humana Intrón 14_11	-	AAAGGATGAGTTGAACTATT	78
15	Albúmina humana Intrón 14_12	-	ATAAGTCAGAACAGGACAAT	79
	Albúmina humana Intrón 14_13	+	ACATGAAAGAGTGCTTTATA	80
	Albúmina humana Intrón 14_14	-	AAATAGTTTAGAATAGTGGT	81
20	Albúmina humana Intrón 14_15	+	GCTCAGCAGTGGAATACTCT	82
	Albúmina humana Intrón 14_16	+	CTGTTCTGACTTATATGATG	83
25	Albúmina humana Intrón 14_17	+	AGTAAGTGATGGCTCAGCAG	84
	Albúmina humana Intrón 14_18	-	TATAAGTCAGAACAGGACAA	85
	Albúmina humana Intrón 14_19	-	AATGGGCAACACTGAACTT	86

Los vectores víricos, como se describe en el presente documento, que comprenden un transgén (p. ej., GFP) junto con Cas9, mejoran la actividad de

edición. Como se demuestra en la Fig. 9A, la DSB inducida en el intrón 13 y/o 14 mejoró la expresión de proteínas (p. ej., aumento del % de células positivas a GFP) al menos 5 veces en relación con una construcción GENERIDE™ sola. Además, como se demuestra en la Fig. 9B, la DSB inducida en el intrón 13 y/o 14 mejoró la expresión de proteínas en relación con una construcción GENERIDE™ sola. El aumento de los niveles de ARNm fusionado puede estar asociado a un aumento de la expresión de proteínas. Hubo al menos un aumento de 600 veces en cuanto al ARNm fusionado de los ARNg de Cas9 probados junto con GENERIDE™ en relación con una construcción GENERIDE™ sola.

Por tanto, este ejemplo demuestra que una construcción GENERIDE™ junto con una DSB en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes.

Ejemplo 14: La administración de Cas9 puede aumentar la eficacia de GENERIDE™

El presente ejemplo demuestra que la administración de una construcción GENERIDE™ a un sitio de integración diana en una región codificante de un locus endógeno de ratón (p. ej., albúmina (ALB) de ratón) junto con una o más enzimas Cas9 dirigidas a un sitio de corte dentro de una secuencia no codificante de un locus endógeno de ratón (p. ej., intrón de ALB de ratón), puede mejorar la eficacia de integración de transgenes *in-vivo*.

Estudio con animales

Los animales se adquirieron en Jackson Laboratories. Los procedimientos generales para el cuidado y alojamiento de los animales fueron aprobados por el Comité Institucional para el Cuidado y Uso de Animales de LogicBio Therapeutics. Animales FvB macho de cuatro semanas de vida se trataron con rAAV.DJ-GR-eGFP a la dosis de 3e13 vg/kg vía seno retro-orbital con anestesia, seguido de la dosificación de ARNm de spCas9 formulado con LNP y ARNg. SpCas9 14_4 LNP se trató 1, 3 o 7 días después de la dosificación del AAV. Durante el estudio, se tomaron muestras periódicas de los animales mediante sangrado submentoniano y se recogió plasma, que se conservó a -80°C hasta su posterior análisis. En los

sacrificios, se recogió sangre para plasma mediante punción cardíaca. Para la disección hepática, se fijó un lóbulo del hígado con formol al 10 % y el resto se congeló rápidamente y se conservó a -80°C. Al día siguiente, el hígado se transfirió a etanol al 70 % para su inclusión en parafina.

5 *Cuantificación de la proteína de fusión albúmina-2A en plasma*

La Albúmina-2A de ratón en plasma se midió mediante ELISA de quimioluminiscencia, utilizando para la captura, un anticuerpo policlonal de conejo anti-2A patentado y para la detección, un anticuerpo policlonal de cabra anti-Albúmina de ratón marcado con HRP (Abcam ab19195). La albúmina-2A
10 recombinante de ratón expresada en células de mamífero y purificada por afinidad se utilizó para construir la curva patrón en plasma de ratón de control al 1 % para tener en cuenta los efectos de la matriz. Se utilizó leche al 1 % (Cell Signaling 9999S) en PBS para el bloqueo y BSA al 1 % para la dilución de muestras en PBST.

Inmunohistoquímica

15 La inmunohistoquímica se realizó en una plataforma robótica (Ventana™ DISCOVERY ULTRA Staining Module, Ventana Co., Tucson, AZ). Las secciones de tejido (4 µm) se desparafinaron y se sometieron a recuperación de antígenos termoinducida durante 64 minutos. Las peroxidasas endógenas se bloquearon con inhibidor de peroxidasa (CM1) durante 8 min antes de incubar la sección con
20 anticuerpo anti-GFP (Novus Biologicals, Centennial, CO) a una dilución de 1:800 durante 60 min a temperatura ambiente. A continuación, se detectó el complejo antígeno-anticuerpo mediante DISC. OmniMap anti-rabbit multimer RUO detection system y DISCOVERY ChromoMap DAB Kit Ventana Co., Tucson, AZ). Posteriormente, todos los portaobjetos se tiñeron con hematoxilina de contraste
25 (Fisher Sci, Waltham, MA), se deshidrataron, se lavaron y se montaron para escanearlos con un escáner digital de portaobjetos (Hamamatsu, Bridgewater, NJ). Las imágenes escaneadas se evaluaron con ocultación utilizando el programa informático ImageJ para cuantificar la zona de tinción positiva.

30 Como se demuestra en la Fig. 10, la administración de spCas9 (ARNm y ARNg) formulada con LNP ya sea 1, 3 o 7 días después de la dosificación de

GENERIDE™, mejoró la eficacia de integración del transgén *in-vivo*. Los niveles de Albúmina-2A (ALB-2A) aumentaron al menos 54 veces en los ratones a los que se había administrado spCas9 formulada con LNP, 1 día después de la dosificación de GENERIDE™, en relación con la administración de una construcción GENERIDE™ sola. Este efecto fue similar en los ratones a los que se había administrado spCas9 formulada con LNP, 3 días después de la dosificación de GENERIDE™. La administración de spCas9 formulada con LNP, 7 días después de la dosificación de GENERIDE™, aumentó los niveles de ALB-2A al menos 20 veces en relación con la administración de una construcción GENERIDE™ sola. La FIG. 11 muestra imágenes que representan que la administración *in vivo* de Cas9 y de ARN guía apropiados puede aumentar la eficacia de integración.

Las FIGS. 22A-22E muestran resultados de experimentos realizados para optimizar una estrategia de dosificación de GENERIDE™ y de la nanopartícula lipídica (LNP) GenVoy-ILM en células hepáticas, para confirmar el suministro de Cas9 por la LNP GenVoy-ILM, y determinar la cronología de los componentes para la integración máxima. A diferentes grupos de ratones se les administraron diferentes combinaciones de GR (3e13 vg/kg) solo, GR (3e13 vg/kg) junto con Cas9 (1 mg/kg) después de 1, 3 o 7 días, o solo LNP que contenía Cas9 + ARNg (1 mg/kg). El vector de la herramienta GENERIDE™ que se utilizó fue DJ-GFP (LB-Vt-0298-001). Los INDELES se analizaron mediante secuenciación de Sanger seguido de análisis TIDE. La FIG. 22A muestra imágenes de IHQ (IHQ dirigida a Cas9-Flag) de células hepáticas que confirman la expresión de Cas9 después del suministro de la LNP que contenía Cas9, con un control de WT sin tratamiento previo. La FIG. 22B muestra una serie de gráficos que miden los niveles de ALB-2A (µg/ml) hasta 6 semanas después de la dosificación de LNP (panel izquierdo) y que miden los niveles de ALB-2A (µg/ml) a las 3 semanas después de la dosificación de LNP (panel derecho) para GR dosificado con LNP en diferentes puntos temporales (D1, D3 y D7), con un control de GR únicamente. Se realizó un ensayo para medir los niveles del biomarcador PD de GENERIDE™ (ALB-2A). La FIG. 22C muestra una imagen IHQ de hepatocitos que confirma una mejor integración en GR2.0

(Cas9) después del suministro de GR (3e13) + LNP (1 mg/kg) que contenía Cas9. Para este experimento, ratones FvB/NJ recibieron dosis de 3e13 vk de GR-GFP (1.0/1.0) y después, el Día 1, de LNP-Cas9 GenVoy-ILM a 1 mg/kg. La FIG. 22D muestra un gráfico de barras que mide los niveles de ALB-2A ($\mu\text{g/ml}$) 1 semana después de la administración para confirmar la potencia de la LNP almacenada cuando la LNP se ha almacenado durante 1 o 2 meses. La FIG. 22E muestra una serie de gráficos que cuantifican la frecuencia de indel (inserción/delección) mediante un ensayo TIDE de D4-1 (panel superior) y D4-2 (panel inferior).

Por tanto, este ejemplo demuestra que una construcción GENERIDE™ junto con una DSB en una región particular de una secuencia no codificante dentro de un locus endógeno, puede mejorar la eficacia de integración de transgenes *in-vivo*. La administración de spCas9 formulada con LNP, 1 o 3 días después de la dosificación de GENERIDE™, puede conferir una mayor mejora de la eficacia de integración del transgén que la administración de spCas9 formulada con LNP 7 días después de la dosificación de GENERIDE™.

Este Ejemplo también demuestra que Cas9 se expresó en la mayoría de las células hepáticas usando la LNP GenVoy-ILM y que la dosificación de GR 1 día después es un punto temporal optimizado para realizar la dosificación de LNP ya que la dosificación de LNP el D1 y D3 mostró niveles de ALB-2A comparables y niveles de ALB-2A más altos que la dosificación el D7. Este ejemplo también muestra que la ALB-2A mejoró significativamente (~50X) con la aplicación de Cas9 con LNP y a las 5 semanas, hay una disminución de ALB-2A, posiblemente debido a una respuesta inmunitaria a la GOI, sin embargo, el grupo de sólo GR no demuestra la misma disminución de expresión. Esto último puede deberse a los menores niveles generales de integración. Este ejemplo confirmó la mejora de la integración en GR2.0 (Cas9) y que la potencia de la LNP seguía siendo la misma cuando la LNP se almacenaba durante 2 meses. La actual plataforma GR2.0 (Cas9) (dosificación de GR-GFP (1.0/1.0) a 3E13 vg/kg) dio lugar a tasas de edición de hepatocitos del 1.9 %. Históricamente, la misma dosis de GR-MMUT (1.0/1.0) editaba un 0.05 % de los hepatocitos, lo que sugiere un aumento de 40 veces con

la plataforma actual. Por último, este ejemplo ilustra una frecuencia de indel del 9.2 % en la muestra D4-1 y una frecuencia de indel del 9.0 % en la muestra D4-2.

Ejemplo 15: El desequilibrio de la longitud del brazo de homología (0.4-0.8 kb) no afectó significativamente a la mejora de la integración GENERIDE™ mediada por Cas9

El experimento descrito en este Ejemplo demuestra que un desequilibrio de la longitud del brazo de homología (0.4-0.8 kb) no afecta significativamente a la mejora de integración GENERIDE™ mediada por Cas9.

La FIG. 12 muestra los resultados de un experimento en el que se administraron dos construcciones GENERIDE™ (GR-hATP7B o GR-GFP) solas o junto con Cas9, con un control sin tratamiento previo. La construcción GR-hATP7b incluía brazos de homología de distintos tamaños (0.4/0.8 kb; 0.8/0.4 kb o 0.6/0.6 kb), mientras que la GR-GFP tenía brazos de homología de 1.0/1.0 kb. GOI, gen de interés (*gene of interest*). Células HEPG2 se sembraron a 100 k/pocillo en formato de 24 pocillos. Al día siguiente, el medio se sustituyó por DMEM que contenía suero bovino fetal (FBS) al 2.5 % con o sin las construcciones (GR-hATP7B o GR-GFP) a una MOI de 1E5. El ARNm de Cas9 (1 µg/pocillo) y el ARNg (250 ng/pocillo) se transfectaron utilizando el reactivo Lipofectamine™ MessengerMAX™ de acuerdo con el protocolo del proveedor. Las células se recogieron 3 días después de la transfección y se extrajo el ARNm para el análisis. La FIG. 12 muestra una serie de gráficos que miden los niveles de ARNm fusionado (copias/20 mg) (panel izquierdo) y ARNm fusionado normalizado (norm. a 0.6/0.6 GR-hATP7B) (panel derecho). GR-GFP tenía un valor inicial más alto, lo que, sin querer ligarse a ninguna teoría en particular, puede deberse a brazos de homología más largos o a un GOI más pequeño. La FIG. 13 muestra los resultados de un experimento en el cual el factor de aumento de ARNm fusionado entre los GOI (genes de interés) se midió como ARNm fusionado normalizado (norm. a GR solo) para las construcciones indicadas. El nivel de ARNm fusionado es una medida de la eficacia de integración.

Por tanto, este ejemplo demuestra que un desequilibrio de la longitud del brazo de homología no tiene ningún impacto significativo en la mejora mediada por Cas9 cuando se administra una construcción GENERIDE™ junto con Cas9.

Ejemplo 16: La eficacia de integración, medida por el nivel de ARNm fusionado, es mayor cuando GENERIDE™ se administra con ARNm de spCas9

El experimento descrito en este Ejemplo demuestra que la eficacia de integración del gen hAAT es mayor cuando GENERIDE™ se administra con ARNm (de spCas9) de *Streptococcus pyogenes*.

La FIG. 14 muestra los resultados de un experimento en el cual se administró una construcción GENERIDE™ (GR-hAAT) sola o junto con ARNm de spCas9, con un control sin tratamiento previo y H2O. Células HEPG2 se sembraron a 100 k/pocillo en formato de 24 pocillos. Al día siguiente, el medio se sustituyó por DMEM que contenía suero bovino fetal (FBS) al 2.5% con o sin la construcción (GR-hAAT) a una MOI de 1E5. El ARNm de Cas9 (1 µg/pocillo) y el ARNgu (250 ng/pocillo) se transfectaron utilizando el reactivo Lipofectamine™ MessengerMAX™ de acuerdo con el protocolo del proveedor. Las células se recogieron 3 días después de la transfección y se extrajo el ARNm para el análisis. La FIG. 14 muestra un gráfico que mide los niveles de ARNm fusionado (copias/100 ng). El ARNm de spCas9 aumentó la eficacia de integración de GR-hAAT.

Por tanto, este ejemplo demuestra que la eficacia de integración, medida por los niveles de ARNm fusionado, son mayores cuando GENERIDE™ se administra con ARNm de spCas9.

Ejemplo 17: La integración de plásmidos circulares y lineales puede mejorarse con la adición de Cas9

El experimento descrito en este Ejemplo demuestra que la integración de plásmidos tanto circulares como lineales puede mejorarse con la adición de Cas9.

Las FIGS. 15A y 15B muestran los resultados de un experimento en el que se midió la eficacia de integración de plásmidos de tipo silvestre (WT, *wild type*) frente a plásmidos mutantes, en formato circular o lineal con o sin Cas9. Los plásmidos tenían sitios de corte de enzimas de restricción para linealizar el

plásmido, y las mutaciones en el sitio de corte pueden mejorar aún más la integración. La FIG. 15A muestra un gráfico que mide los niveles de ARNm fusionado (copias/100 ng) en presencia o ausencia de Cas9 de plásmidos de WT y mutantes, en formato circular y lineal, con un control sin tratamiento previo (panel izquierdo). La FIG. 15B muestra un gráfico que mide los niveles de ARNm fusionado en presencia de Cas9 (norm. a C-WT+Cas9) de plásmidos de WT y mutantes, tanto en formato circular como lineal (panel derecho).

Por tanto, este ejemplo demuestra que Cas9 mejora la integración tanto de plásmidos circulares como lineales, siendo los plásmidos lineales los que muestran una mejor eficacia.

Ejemplo 18: Medición de la eficacia de edición, la reparación dirigida por homología (HDR) y la unión de extremos no homólogos (NHEJ) después de la transfección con ARNm de Nucleasa de AZ y de SpCas9 en células HEPG2

El experimento descrito en este Ejemplo midió la eficacia de edición, la HDR y la NHEJ después de la transfección con ARNm de Nucleasa de AZ y de SpCas9 en células HEPG2 humanas.

Las FIGS. 16A y 16B muestran los resultados de un experimento en el que se midió la eficacia de edición, KI_HDR y NHEJ_Fw+Rv después de la transfección con ARNm de Nucleasa de AZ (FIG. 16A) o de SpCas9 (FIG. 16B) de diferentes proporciones entre donante y ARNg respectivamente. Las células HepG2 se trataron con SpCas9 o Nucleasa de AZ junto con el ARNgu designado. Para los datos de HDR, se utilizó un donante de ADNmc corto con brazos de homología flanqueando el sitio de corte. Para los datos de NHEJ, se utilizó un donante de ODNbc (oligodesoxinucleótido bicatenario) corto. En todos los casos, se utilizó NGS para definir la tasa de INDELES, la integración de ADNmc y la integración ODNbc en relación con las lecturas totales. Las flechas indican solapamiento de los ARNg para SpCas9 y Nucleasa de AZ; los asteriscos indican los ARNg con el mayor KI_HDR para la Nucleasa de AZ. La FIG. 16A muestra una serie de gráficos en los que se midió la eficacia de edición (%) (panel superior), la KI_HDR (%) (panel central) y la NHEJ_Fw+Rv (%) (panel inferior) después de la transfección con ARNm de

Nucleasa de AZ para diferentes proporciones entre donante y ARNg en células HEPG2. La media de KI_HDR (%) fue mayor para el ADNmc (monocatenario) que para el ADNbc (bicatenario) y viceversa en el caso de NHEJ_Fw+Rv (%). La FIG. 16B muestra una serie de gráficos en los que se midió la eficacia de edición (%) (panel superior), la KI_HDR (%) (panel central) y la NHEJ_Fw+Rv (%) (panel inferior) después de la transfección con ARNm de SpCas9 para diferentes proporciones entre donante y ARNg en células HEPG2. La media de KI_HDR (%) fue mayor para el ADNmc que para el ADNbc viceversa en el caso de NHEJ_Fw+Rv (%). Por tanto, este ejemplo demuestra que tanto el ARNm de Nucleasa de AZ como de SpCas9 pueden dar lugar a la integración de GENERIDE™ en células humanas HEPG2.

Ejemplo 19: Comparación de los ARNg en células HEPG2 humanas

El experimento descrito en este Ejemplo comparó ARNg humanos de spCas9 y Nucleasa de AZ Cas9 en células HEPG2 y compara la eficacia de la Nucleasa de AZ Cas9 cuando realiza el corte dentro del brazo de homología (HA) frente a cuando lo hace fuera del mismo, en células HEPG2 humanas.

Las FIGS. 17A y 17B muestran los resultados de un experimento en donde se midió la eficacia de integración del transgén de spCas9 de diferentes ARNg humanos en células HEPG2. Células HEPG2 se sembraron a 100 k/pocillo en formato de 24 pocillos. Al día siguiente, el medio se sustituyó por DMEM que contenía suero bovino fetal (FBS) al 2.5 % con o sin las construcciones (GR-hATP7B o GR-GFP) a una MOI de 1E5. El ARNm de Cas9 (1 µg/pocillo) y el ARNgu (250 ng/pocillo) se transfectaron utilizando el reactivo Lipofectamine™ MessengerMAX™ de acuerdo con el protocolo del proveedor. Las células se recogieron 3 días después de la transfección y se extrajo el ARNm para el análisis. La FIG. 17A muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de spCas9 en células HEPG2, con un control PD128 (panel izquierdo). La FIG. 17B muestra un gráfico que mide los niveles de ARNm fusionado normalizados (norm. a vt solamente) en presencia de spCas9 en células HEPG2, con un control de PD128 (panel izquierdo). Las FIGS. 18A y 18B muestran

los resultados de un experimento que evalúa el efecto de la ubicación del corte (es decir, corte dentro del brazo de homología (HA) GR, en donde tanto el HA en 5' como el HA en 3' pueden experimentar HDR, frente al corte fuera del brazo de homología GR, en donde sólo un extremo puede experimentar HDR) para spCas9

5 en células humanas. La FIG. 18A muestra un gráfico en donde se comparan los niveles de ARNm fusionado (norm.) cuando el corte se realiza fuera del HA (1 brazo) frente a cuando el corte se realiza dentro del HA (2 brazos); el corte dentro del HA dio lugar a una eficacia de integración significativamente más alta para spCas9 (panel izquierdo). La FIG. 18B muestra un gráfico que mide los niveles de ARNm

10 fusionado (norm.) en función de la longitud alineada (intrón 12 y 13) cuando el corte se realiza fuera del HA y dentro del HA (panel derecho). Las FIGS. 19A-19C muestran los resultados de un experimento en donde se midió la eficacia de integración de transgenes de GENERIDE™ con Nucleasa de AZ Cas9 para diferentes ARNg humanos en células HEPG2. Obsérvese que la 'referencia 14_2'

15 es un ARNgu de SpCas9, que se utiliza con SpCas9 como control positivo. La FIG. 19A muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-mut y Nucleasa de AZ Cas9 en células HEPG2, con un control (PD128) de GR únicamente (panel izquierdo). La FIG. 19B muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-GFP y

20 Nucleasa de AZ Cas9 en células HEPG2, con un control sin tratamiento previo y un control de GR únicamente (vt-0290) (panel central). La FIG. 19C muestra un gráfico que mide los niveles de ARNm fusionado (copias/2.5 ng) en presencia de GR-ATP7B y Nucleasa de AZ Cas9 en células HEPG2, con un control sin tratamiento previo y un control de GR únicamente (vt-235) (panel derecho). Las FIGS. 20A y

25 20B muestran los resultados de un experimento que mide la eficacia de integración GENERIDE™ con la Nucleasa de AZ Cas9 cuando corta dentro del HA frente a cuando corta fuera del HA en células HEPG2. La FIG. 20A muestra un gráfico en donde se comparan los niveles de ARNm fusionado (norm. a vt únicamente) cuando el corte se realiza fuera del HA (1 brazo) frente a cuando el corte se realiza dentro

30 del HA (2 brazos) (panel izquierdo). La FIG. 20B muestra un gráfico que mide los

niveles de ARNm fusionado (norm.) en función de la longitud alineada (intrón 12 y 13) cuando el corte se realiza fuera del HA y dentro del HA (panel derecho).

Por tanto, este ejemplo demuestra que el ARNg 14-2 para spCas9 dio lugar a una mayor eficacia de integración en este experimento que otros ARNg, mientras que el ARNgu 163 fue el más activo entre los ARNgu de la Nucleasa de AZ, seguido del 37, 188, 171 y 116. También se descubrió que el ARNgu 188 funciona mejor y de forma más fiable que otros ARNgu de la Nucleasa de AZ para tres vectores diferentes, presumiblemente debido a su mayor proximidad al sitio de integración lo que hace que el sitio DSB esté dentro de los límites del brazo de homología para los tres vectores. Además, este ejemplo ilustra que cuando el corte se realizó dentro del brazo de homología da lugar a una mayor eficacia de integración en el caso de spCas9, pero que esto no se aplica a la Nucleasa de AZ Cas9. Por último, también muestra que el vector ATP7B con brazo de homología corto muestra una menor integración en comparación con el brazo de homología de 1kb tanto para spCas9 como para la Nucleasa de AZ Cas9.

Ejemplo 20: Medición de la eficacia de edición, de la HDR y de la NHEJ, después de la transfección con ARNm de Nucleasa de AZ en células HEPG2

El experimento descrito en este Ejemplo midió la eficacia de edición, la HDR y la NHEJ después de la transfección con ARNm de Nucleasa de AZ en células HEPG2 humanas.

La FIG. 21 muestra los resultados de un experimento en donde se midió la eficacia de edición, KI_HDR y NHEJ_Fw+Rv después de la transfección con ARNm de Nucleasa de AZ de los ARNg y donantes indicados (Sin donante, ADNmc o ADNbc). Las células HepG2 se trataron con SpCas9 o Nucleasa de AZ junto con el ARNgu designado. Para los datos de HDR, se utilizó un donante de ADNmc corto con brazos de homología flanqueando el sitio de corte. Para los datos de NHEJ, se utilizó un donante de ODNbc (oligodesoxinucleótido bicatenario) corto. En todos los casos, se utilizó NGS para definir la tasa de INDELES, la integración de ADNmc y la integración ODNbc en relación con las lecturas totales. La FIG. 21 muestra una serie de gráficos en los que se midió la eficacia de edición (%) (panel superior), la

KI_HDR (%) (panel central) y la NHEJ_Fw+Rv (%) (una medida de integración por NHEJ en orientaciones directa (*forward*, *fw*) o inversa (*reverse*, *rev*); panel inferior) después de la transfección con ARNm de Nucleasa de AZ para los ARNg y donantes indicados (Sin donante, ADNmc o ADNbc) en células HEPG2. La media de KI_HDR (%) fue mayor para el ADNmc (monocatenario) que para el ADNbc (bicatenario) y viceversa en el caso de NHEJ_Fw+Rv (%). Las flechas indican solapamiento de los ARNg para SpCas9 y Nucleasa de AZ; los asteriscos indican los ARNg con el mayor KI-HDR para la Nucleasa de AZ.

Por tanto, este ejemplo demuestra que el ARNm de la Nucleasa de AZ conduce a reescritura genómica, a HDR y a NHEJ en células HEPG2. Por tanto, los datos muestran que el ARNgu promueve la mayor integración por NHEJ, aunque también que la HDR es generalmente más eficaz que la NHEJ, al menos en este escenario. Además, los ARNgu más activos por NHEJ se encuentran a menudo, pero no siempre, entre los ARNgu más activos por HDR. Estas tendencias son más pronunciadas en el caso de la Nucleasa de AZ que en el de SpCas9. Por otro lado, las secuencias de ARNgu ilustrativas de la Tabla 10 fueron las más activas en ambos conjuntos de datos – el conjunto de cribado HDR+NHEJ (datos NGS) y el conjunto de cribado basado en GR (ARN fusionado).

Tabla 10: Secuencias ilustrativas de ARNgu

Espe cie	Nucleas a	Nombre	Región diana	Protoespacia dor	Secuencia completa ordenada
Ser huma no	Nucleas a de AZ	Nucleasa de AZ_65/ Nucleasa de AZ_65	Intrón 12	TAGCAATAG GGAGTTGA GCTTC (SEQ ID NO: 87)	mU*mA*mG*rCrArArUrA rGrGrGrArGrUrUrGrArG rCrUrUrCrGrUrUrUrCrAr GrUrUrGrArArArGrArAr GrUrCrArCrUrCrUrUrAr ArArGrUrGrArGrCrUrGr ArArArUrCrArCrUrArArA rArArUrUrArArGrArUrUr

					GrArArCrCrCrGrGrCrUr ArCrUrGrArCrUrCrUrGr UrCrArUrCrCrG*mG*mG *mU (SEQ ID NO: 93)
5	Ser huma no	Nucleas a de AZ	Nucleasa de AZ_116/ Nucleasa de AZ_116	Intrón 12	AAAATAAGG CTATCCAAA CTCA (SEQ ID NO: 88)
10					mA*mA*mA*rArUrArArG rGrCrUrArUrCrCrArArAr CrUrCrArGrUrUrUrCrAr GrUrUrGrArArArGrArAr GrUrCrArCrUrCrUrUrAr ArArGrUrGrArGrCrUrGr ArArArUrCrArCrUrArArA rArArUrUrArArGrArUrUr GrArArCrCrCrGrGrCrUr ArCrUrGrArCrUrCrUrGr UrCrArUrCrCrG*mG*mG *mU(SEQ ID NO: 94)
15	Ser huma no	Nucleas a de AZ	Nucleasa de AZ_163/ Nucleasa de AZ_163	Intrón 12	TCTCATGTG TGGGGCCA CCATA (SEQ ID NO: 89)
20					mU*mC*mU*rCrArUrGr UrGrUrGrGrGrCrCrAr CrCrArUrArGrUrUrUrCr ArGrUrUrGrArArArGrAr ArGrUrCrArCrUrCrUrUr ArArArGrUrGrArGrCrUr GrArArArUrCrArCrUrArA rArArArUrUrArArGrArUr UrGrArArCrCrCrGrGrCr UrArCrUrGrArCrUrCrUr GrUrCrArUrCrCrG*mG* mG*mU(SEQ ID NO: 95)
25					
30					

5	Ser humano	Nucleasa de AZ	Nucleasa de AZ_188/ Nucleasa de AZ_188	Intrón 13	AAAAGTTAC TGTCAGGC CATAA (SEQ ID NO: 90)	mA*mA*mA*rArGrUrAr rCrUrGrUrCrArGrGrCrCr ArUrArArGrUrUrUrCrAr GrUrUrGrArArArGrArAr GrUrCrArCrUrCrUrUrAr ArArGrUrGrArGrCrUrGr ArArArUrCrArCrUrArArA rArArUrUrArArGrArUrUr GrArArCrCrCrGrGrCrUr ArCrUrGrArCrUrCrUrGr UrCrArUrCrCrG*mG*mG *mU (SEQ ID NO: 96)
15	Ser humano	Nucleasa de AZ	ESOH_Alb_14-2	Intrón 14	AGTTTAGAA TAGTGGTC GGTGC (SEQ ID NO: 91)	mA*mG*mU*rUrUrArGrA rArUrArGrUrGrGrUrCrGr GrUrGrCrGrUrUrUrCrAr GrUrUrGrArArArGrArAr GrUrCrArCrUrCrUrUrAr ArArGrUrGrArGrCrUrGr ArArArUrCrArCrUrArArA rArArUrUrArArGrArUrUr GrArArCrCrCrGrGrCrUr ArCrUrGrArCrUrCrUrGr UrCrArUrCrCrG*mG*mG *mU (SEQ ID NO: 97)
25	Ser humano	SpCas9	spCas9_14_2_Hi Mod	Intrón 14	TTTAGAATA GTGGTCGG TGC (SEQ ID NO: 92)	mU*mU*mU*rArGrArArU rArGrUrGrGrUrCrGrGrUr GrCrGrUrUrUrUrArGrA mGmCmUmAmGmAmA mAmUmAmGmCrArArG

5

					rUrUrArArArArUrArArGr GrCrUrArGrUrCrCrGrUr UrArUrCrAmAmCmUmU mGmAmAmAmAmAmG mUmGmGmCmAmCmC mGmAmGmUmCmGmG mUmGmCmU*mU*mU* mU (SEQ ID NO: 98)
--	--	--	--	--	---

'm' es 2'OMe, * indica enlace fosforotioato y r indica ribo.

10

Ejemplo 22: Evaluación de ARN de IDT para garantizar la viabilidad con la LNP de AZ

15

El experimento descrito en este Ejemplo demuestra un estudio diseñado para evaluar el ARN de IDT para garantizar la viabilidad con la LNP (obtenida de AstraZeneca; "LNP de AZ") donde se evaluó la calidad de los materiales de ARN de IDT para garantizar la viabilidad de trabajo con la LNP y se compararon las proporciones N:P, las preparaciones de formulación y el escalonamiento de la dosificación de LB el Día 1 (D1) frente a 4 horas (H4).

20

25

30

Las FIGS. 23A-23F muestran los resultados de experimentos realizados para evaluar el ARN de IDT para garantizar la viabilidad con la LNP de AZ. El donante GR-HDR utilizado fue DJ-mHA (1.0/1.0) GFP-LB-Vt-0298-001 y la LNP utilizada fue la AZ de AstraZeneca. Se administraron 5 combinaciones diferentes a distintos grupos de ratones de WT: Únicamente GR WT, GR WT + LNP N:P=3 (Co-formulación) escalonado el D1, GR WT + LNP N:P=6 (Co-formulación) escalonado el D1, GR WT + LNP N:P=6 (formulación por sep.) escalonado el D1 y GR WT + LNP N:P=6 (Co-formulación) escalonado en H4. Ratones FvB recibieron DJ-AAV-GFP junto con preparados con una proporción N.P de 3 o 6 (como se indica), con el ARNm y el ARN_{gu} formulados conjuntamente ('Co-formulación') o por separado ('formulación por sep.'). El tiempo de escalonamiento indica si la LNP se dosificó 1 día o 4 horas después de la dosificación del AAV. La FIG. 23A muestra los resultados de un experimento realizado para medir los niveles de ALB-2A (µg/ml)

hasta 3 semanas después de la dosificación de LNP en las 5 combinaciones diferentes enumeradas anteriormente, con el GR únicamente como control. La FIG. 23B es un gráfico de barras que mide las copias de ARNm fusionado/100 ng de ARN de las 5 combinaciones diferentes administradas, con el GR únicamente como control. La FIG. 23C es un conjunto de imágenes de IHQ, realizándose la IHQ para etiquetar y medir células GFP+ después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23D es un gráfico de barras que mide el % de células positivas a GFP después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23E es un gráfico de barras que mide los niveles de ALB-2A ($\mu\text{g/ml}$) después de la administración de las 5 combinaciones enumeradas anteriormente. La FIG. 23F muestra los niveles de ALB-2A ($\mu\text{g/ml}$) en función del % de células positivas a GFP después de la administración de las 5 combinaciones enumeradas.

Por tanto, este ejemplo demuestra que hay una mejor integración, cuantificada por la expresión de ALB-2A y la integración de ARNm fusionado, cuando se utiliza GR2.0 en comparación con GR solo y eso es cierto para todos los grupos de GR2.0, con un rendimiento de LNP de AZ que es mejor que el de la LNP Genvoy-ILM anterior. El N:P6 dio lugar a una mayor expresión de ALB-2A que el N:P3, sin embargo, la combinación de ARN guía y de spCas9 formulados conjuntamente dio lugar a una mayor expresión de ALB-2A en comparación con la combinación de ARN guía y de Cas9 formulada por separado. Hubo una expresión de ALB-2A comparable con el escalonamiento de la dosis en H4 y D1 entre los componentes GR y LNP, con respecto a la administración de los componentes. Este ejemplo mostró que la coformulación LNP N:P6 produjo una integración de ~16 % de células GFP+ en la IHQ que se correlacionó con la expresión de ALB-2A. Además, la tinción de CD4 no mostró diferencias entre los grupos cuando se comparó con GR solo y la integración de ADN_g molecular se correlacionó en general con ARN fusionado y ALB-2A.

Ejemplo 23: Optimización del calendario de tratamiento: El tratamiento con GR un día después de la transfección del ARNm también mejoró la integración

El experimento descrito en este Ejemplo demuestra que el tratamiento con GR un día después de la transfección de ARNm también mejora la integración de GR.

Las FIGS. 24A y 24B muestran resultados de experimentos realizados para optimizar un calendario de tratamiento, en los que las células se trataron de acuerdo con 2 protocolos diferentes para determinar cuál era mejor para mejorar la integración. La FIG. 24A muestra los resultados de experimentos en los que las células se trataron con GR el mismo día que la transfección de ARNm (protocolo original). La siembra se realizó el D0 seguido de la transfección de ARNm de nucleasa junto con el tratamiento GR-GFP el D1 y la recogida se realizó el D4 (panel superior). La FIG. 24A muestra un gráfico de barras que mide el % de células HA+ de las 4 combinaciones diferentes administradas: sin tratamiento previo (control), GR-GFP, GR-GFP + Cas9, y sólo ARNm (panel inferior izquierdo) y otro gráfico de barras que mide el % de células HA+ (norm.) de las mismas 4 combinaciones administradas (panel inferior derecho). La FIG. 24B muestra los resultados de experimentos en los que las células se trataron con GR el día después de la transfección de ARNm. La siembra se realizó el D0, seguido de la transfección del ARNm de la nucleasa el D1, seguido del tratamiento con GR-GFP el D2 y, por último, la recogida se realizó el D5 (panel superior). La FIG. 24B muestra un gráfico de barras que mide el % de células HA+ de las 4 combinaciones diferentes administradas: sin tratamiento previo (control), GR-GFP, GR-GFP + Cas9, y sólo ARNm (panel inferior izquierdo) y otro gráfico de barras que mide el % de células HA+ (norm.) de las mismas 4 combinaciones administradas (panel inferior derecho).

Por tanto, este ejemplo demuestra que el tratamiento con GR un día después de la transfección de ARNm también mejora la integración. Los datos también pueden sugerir que la adición de ODNbc un día después dará lugar a una mejor eficacia de hibridación de ODNbc.

Ejemplo 24: Efecto de los reactivos de transfección sobre la eficacia de transfección y la proporción de edición *in vitro*

El experimento descrito en este Ejemplo evaluó el efecto de diferentes reactivos de transfección sobre la eficacia de transfección y la proporción de edición *in vitro* en células humanas.

La FIG. 25 muestra los resultados de experimentos realizados para comparar la eficacia de transfección de diferentes reactivos mediante inmunofluorescencia. La FIG. 25 muestra 4 imágenes de inmunofluorescencia (panel superior izquierdo) que prueban la eficacia de transfección de Lipofectamine™ 3000 para la transfección de mCherry y eGFP, con un canal DAPI y una imagen fusionada de eGFP-mCherry que muestra la expresión de ambos y 4 imágenes de inmunofluorescencia (panel superior derecho) que prueban la eficacia de transfección de MessengerMax™ para la transfección de mCherry y eGFP, con un canal DAPI y una imagen fusionada de eGFP-mCherry que muestra la expresión de ambos. La FIG. 25 también muestra gráficos de barras que miden el % de eficacia de transfección para eGFP y mCherry tanto para Lipofectamine™ 3000 como para MessengerMax™, paneles inferior izquierdo e inferior derecho, respectivamente. La FIG. 26 muestra un gráfico de barras que mide el % de células HA+ de las 8 combinaciones diferentes administradas: sin tratamiento previo (control), ARNm solo, Vehículo + GR-GFP + ARNm de Cas9, Vehículo + ODNbc + GR-GFP + ARNm de Cas9, ARNg n.º 2 + GR-GFP + ARNm de Cas9, ARNg n.º 2 + ODNbc + GR-GFP + ARNm de Cas9, ARNg n.º 10 + GR-GFP + ARNm de Cas9, y ARNg n.º 10 + ODNbc + GR-GFP + ARNm de Cas9 (panel inferior izquierdo), un gráfico de barras que mide el % de células HA+ (norm.) de las mismas 8 combinaciones administradas (panel inferior central), y otro gráfico de barras que mide el recuento de células/imagen de las mismas 8 combinaciones administradas (panel inferior derecho).

Por tanto, este ejemplo demuestra que MessengerMax™ es superior a Lipofectamine 3000 para la transfección de ARNm, como se evidencia por su

eficacia de transfección mucho mayor que Lipofectamine 3000, pero no mostró mucha separación del valor inicial.

Equivalentes

5 Los expertos en la materia reconocerán o podrán determinar, utilizando tan solo experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención descritas en el presente documento. El alcance de la presente invención no pretende limitarse a la Descripción anterior, sino más bien a lo expuesto en las siguientes reivindicaciones:

10 Se hace constar que con relación a esta fecha, el mejor método conocido por la solicitante para llevar a la práctica la citada invención, es el que resulta claro de la presente descripción de la invención.

15

20

25

30