

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 813 367**

51 Int. Cl.:

|                   |           |
|-------------------|-----------|
| <b>C07K 14/47</b> | (2006.01) |
| <b>C12N 9/22</b>  | (2006.01) |
| <b>C12N 15/62</b> | (2006.01) |
| <b>C12N 15/09</b> | (2006.01) |
| <b>C12N 9/64</b>  | (2006.01) |
| <b>C12N 15/86</b> | (2006.01) |
| <b>A61K 48/00</b> | (2006.01) |

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **09.12.2014 PCT/US2014/069352**
- 87 Fecha y número de publicación internacional: **18.06.2015 WO15089077**
- 96 Fecha de presentación y número de la solicitud europea: **09.12.2014 E 14869719 (6)**
- 97 Fecha y número de publicación de la concesión europea: **03.06.2020 EP 3080274**

54 Título: **Métodos y composiciones para ingeniería genómica**

30 Prioridad:

**09.12.2013 US 201361913838 P**  
**24.02.2014 US 201461943884 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**23.03.2021**

73 Titular/es:

**SANGAMO THERAPEUTICS, INC. (100.0%)**  
**501 Canal Blvd., Suite A100**  
**Richmond, CA 94804, US**

72 Inventor/es:

**MILLER, JEFFREY C.;**  
**PASCHON, DAVID;**  
**REBAR, EDWARD J.;**  
**WECHSLER, THOMAS y**  
**ZHANG, LEI**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 813 367 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Métodos y composiciones para ingeniería genómica

5 **Campo técnico**

La presente divulgación se refiere al campo de la modificación génica y el tratamiento de la hemofilia.

10 **Antecedentes**

15 Una aplicación especialmente atractiva de la terapia génica implica el tratamiento de trastornos que son causados por una insuficiencia de un producto génico secretado o que se pueden tratar mediante la secreción de una proteína terapéutica. Dichos trastornos son potencialmente abordables mediante la administración de un transgén terapéutico a un número modesto de células, siempre que cada célula receptora exprese un alto nivel del producto genético terapéutico. En dicho contexto, el alivio de la necesidad de la administración de genes a un gran número de células puede permitir el desarrollo con éxito de terapias génicas para indicaciones de otra manera intratables. Dichas aplicaciones requerirían niveles permanentes, seguros y muy altos de expresión transgénica. Por lo tanto, el desarrollo de un sitio seguro que exhiba estas propiedades proporcionaría una utilidad sustancial en el campo de la terapia génica.

20 Las hemofilias como la hemofilia A y la hemofilia B, son trastornos genéticos del sistema de coagulación de la sangre, que se caracterizan por hemorragia en las articulaciones y los tejidos blandos, y por hemorragia excesiva en cualquier sitio que experimente un traumatismo o se someta a cirugía. La hemofilia A es clínicamente indistinguible de la hemofilia B, pero el factor VIII (FVIII o F8) es deficiente o está ausente en la hemofilia A, mientras que el factor IX (FIX o F.IX) es deficiente o está ausente en pacientes con hemofilia B. El gen F8 codifica una glicoproteína plasmática que circula en asociación con el factor de von Willebrand en su forma inactiva. Tras una lesión superficial, se inicia la cascada de coagulación intrínseca y el factor VIII se libera del complejo y se activa. La forma activada actúa con el Factor IX para activar el Factor X para convertirse en el Xa activado, lo que finalmente conduce al cambio de fibrinógeno a fibrina y a la inducción de coágulos. Véase, Levinson *et al.* (1990) *Genomics*7(1):1-11. El 40-50 % de los pacientes con hemofilia A tienen una inversión cromosómica que afecta al intrón 22 de F8 (también conocida como IVS22). La inversión está causada por un evento de recombinación intracromosómica entre una secuencia de 9,6 kb dentro del intrón 22 del gen F8 y una de las dos secuencias orientadas inversamente estrechamente relacionadas ubicadas aproximadamente a 300 kb distal al gen F8, lo que tiene como resultado una inversión de los exones 1 a 22 con respecto a los exones 23 a 26. Véase, *Textbook of Hemophilia*, Lee *et al.* (eds) 2005, Blackwell Publishing. Otros pacientes con hemofilia A tienen defectos en F8 incluyendo mutaciones de sitio activo y mutaciones sin sentido y de sentido erróneo. Por su parte, el Factor IX (F.IX) codifica una de las serina proteasas implicadas en el sistema de coagulación, y se ha demostrado que la restauración de incluso el 3 % de los niveles circulantes normales de la proteína del Factor IX de tipo silvestre puede prevenir la hemorragia espontánea. Otras hemofilias están asociadas con la expresión aberrante de otros factores de la coagulación. Por ejemplo, la deficiencia del factor VII es un rasgo autosómico recesivo que ocurre en aproximadamente 1 de cada 300.000 a 500.000 personas y se asocia con niveles inadecuados de factor VII en el paciente. De manera similar, una deficiencia de Factor X también es un rasgo autosómico recesivo que ocurre en 1 de cada 500.000 a 1 millón de personas, y es causada por variantes genéticas del gen *FX*. La deficiencia de factor X puede tener diversos grados de gravedad en la población de pacientes.

45 Los tratamientos actuales para la hemofilia B se basan en infusiones intravenosas crónicas, repetidas de Factor IX recombinante purificado y tienen una serie de inconvenientes. Esto incluye la necesidad de infusiones intravenosas repetidas, está asociado a la formación de inhibidores y es profiláctico en lugar de curativo.

50 Se ha descrito la terapia génica para pacientes con hemofilia A o B, que implica la introducción de plásmidos y otros vectores (p. ej., AAV) que codifican proteínas F8 o F.IX funcionales. Véase, *p.ej.*, las patentes de Estados Unidos números 6.936.243; 7.238.346 y 6.200.560; Shi *et al.* (2007) *J Thromb Haemost.*(2):352-61; Lee *et al.* (2004) *Pharm. Res.* 7:1229-1232; Graham *et al.* (2008) *Genet Vaccines Ther.* 3:6-9; Manno *et al.* (2003) *Blood* 101(8): 2963-72; Manno *et al.* (2006) *Nature Medicine* 12(3): 342-7; Nathwani *et al.* (2011) *Molecular Therapy* 19(5): 876-85; Nathwani *et al.* (2011); *N Engl J Med.* 365(25): 2357-65. Sin embargo, en estos protocolos, la formación de anticuerpos inhibidores anti-factor VIII o IX (anti-F8 o anti-F.IX) y anticuerpos contra el vehículo de administración sigue siendo una complicación importante del tratamiento de reemplazo de F8 y F.IX para la hemofilia. Véase, p. ej., Scott & Lozier (2012) *Br J Haematol.* 156(3):295-302.

60 Las enfermedades de almacenamiento lisosomal (LSD) son un grupo de enfermedades monogénicas metabólicas raras que se caracterizan por la falta de proteínas lisosomales funcionales individuales que normalmente intervienen en la descomposición de los lípidos, glucoproteínas y mucopolisacáridos de desecho. Estas enfermedades se caracterizan por una acumulación de estos compuestos en la célula, ya que no puede procesarlos para su reciclaje debido al mal funcionamiento de una enzima específica. Los ejemplos más frecuentes son las enfermedades de Gaucher (deficiencia de glucocerebrosidasa - nombre del gen: GBA), Fabry (deficiencia de galactosidasa  $\alpha$  - GLA), Hunter (deficiencia de iduronato-2-sulfatasa-IDS), Hurler (deficiencia de alfa-L iduronidasa-IDUA) y Niemann-Pick (deficiencia de esfingomielina fosfodiesterasa 1 - SMPD1). Si se agrupan todas, las LSD tienen una incidencia en la

población de aproximadamente 1 de cada 7000 nacimientos. Estas enfermedades tienen efectos devastadores en las personas afectadas por ellas. Habitualmente se diagnostican por primera vez en bebés que pueden tener patrones de crecimiento facial y corporal característicos y pueden tener retraso mental de moderado a severo. Las opciones de tratamiento incluyen la terapia de reemplazo enzimático (ERT) donde se administra al paciente la enzima que falta, generalmente mediante inyección intravenosa en grandes dosis. Tal tratamiento es solo para tratar los síntomas y no es curativo, por lo tanto, el paciente debe recibir una dosis repetida de estas proteínas durante el resto de sus vidas, y potencialmente puede desarrollar anticuerpos neutralizantes para la proteína inyectada. A menudo, estas proteínas tienen una semivida sérica corta, por lo que el paciente también debe soportar infusiones frecuentes de la proteína. Por ejemplo, los pacientes con enfermedad de Gaucher que reciben el producto Cerezyme® (imiglucerasa) deben recibir infusiones tres veces por semana. La producción y purificación de las enzimas también es problemática, por lo que los tratamientos son muy costosos (>100.000 dólares al año por paciente).

Se han descrito varios métodos y composiciones para la escisión dirigida del ADN genómico. Tales eventos de escisión dirigida pueden usarse, por ejemplo, para inducir mutagénesis dirigida, inducir deleciones dirigidas de secuencias de ADN celular y facilitar la recombinación dirigida en un locus cromosómico predeterminado. Véase, *p. ej.*, las patentes de Estados Unidos números 8.623.618; 8.034.598; 8.586.526; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.067.317; 7.262.054; 7.888.121; 7.972.854; 7.914.796; 7.951.925; 8.110.379; 8.409.861; las publicaciones de patente de Estados Unidos 20030232410; 20050208489; 20050026157; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y la patente de Estados Unidos Nº 9.873.894. Estos métodos a menudo implican el uso de sistemas de escisión modificados para inducir una rotura de la doble cadena (DSB) o una muesca en una secuencia de ADN diana de tal manera que la reparación de la rotura por un proceso propenso a errores, como la unión de extremos no homólogos (NHEJ) o la reparación usando un molde de reparación (reparación dirigida por homología o HDR) puede provocar la inactivación de un gen o la inserción de una secuencia de interés (integración dirigida). Esta técnica también se puede utilizar para introducir cambios específicos del sitio en la secuencia del genoma mediante el uso de un oligonucleótido donante, incluyendo la introducción de deleciones específicas de regiones genómicas, o de mutaciones puntuales específicas o alteraciones localizadas (también conocida como corrección génica). La escisión puede producirse mediante el uso de nucleasas específicas tales como las nucleasas de dedo de zinc (ZFN) modificadas, las nucleasas efectoras del tipo activador de la transcripción (TALEN) o el uso del sistema CRISPR/Cas con un ARNcr/ARNtracr modificado ('ARN de guía única') para guiar la escisión específica. Además, se están desarrollando nucleasas dirigidas basadas en el sistema Argonauta (*p.ej.*, de *T. thermophilus*, conocidas como 'TtAgo', véase Swarts *et al* (2014) Nature 507(7491): 258-261), que también pueden tener potencial para usos en la edición del genoma y la terapia génica.

Este enfoque de inserción transgénica dirigida por mediación de nucleasas ofrece la posibilidad de una mejor expresión transgénica, mayor seguridad y durabilidad de la expresión, en comparación con los enfoques de integración clásicos, ya que permite un posicionamiento transgénico exacto con un riesgo mínimo de silenciamiento génico o activación de oncogenes cercanos.

La integración dirigida de un transgén puede estar en su locus afín, por ejemplo, la inserción de un transgén de tipo silvestre en el locus endógeno para corregir un gen mutante. Como alternativa, el transgén puede insertarse en un locus no afín, un locus de "sitio seguro". Se han descrito varios loci de sitio seguro, incluyendo *CCR5*, *HPRT*, *AAVS1*, *Rosa* y albúmina. Véase, *p. ej.*, las patentes de Estados Unidos números 7.951.925 y 8.110.379; las publicaciones de patente Estados Unidos números 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y la patente de Estados Unidos Nº 9.873.894. Por ejemplo, la publicación de patente de Estados Unidos Nº 20110027235 se refiere a la integración de proteínas funcionales en células madre aisladas y la publicación de patente de Estados Unidos Nº 20120128635 describe métodos de tratamiento de la hemofilia B. Además, la publicación de patente de Estados Unidos Nº 2014-0017212 y 2014-0112896 describe métodos de tratamiento de enfermedades de almacenamiento lisosomal. Véase también Li *et al* (2011) Nature 475 (7355):217-221 y Anguela *et al* (2013) Blood 122:3283-3287.

Sin embargo, sigue existiendo la necesidad de composiciones y métodos adicionales para proporcionar proteínas terapéuticas a un sujeto con una enfermedad o trastorno en el que faltan una o más proteínas, son deficientes y/o se expresan de manera aberrante.

El documento WO 2013/044008 divulga nucleasas y métodos de uso de estas nucleasas para expresar un transgén de un locus de sitio seguro en un tejido secretor, y clones y animales derivados del mismo.

Mandell *et al.* Nucl. Acids Res., 34, W516-W523 (2006) divulga una utilidad basada en la web llamada ZF Tools que supuestamente se puede utilizar para buscar sitios diana en una secuencia de ADN suministrada por el usuario que sea apropiada para la regulación génica o dirigida a nucleasas utilizando proteínas que contienen dominios de dedo de zinc. También se divulga una utilidad de modificación inversa para predecir el sitio de unión para una proteína de dedo de zinc de secuencia conocida.

Bhakta & Segal, Methods Mol. Biol., 649, 3-30 (2010) describe métodos y orientación para la creación de proteínas de dedo de zinc mediante el ensamblaje modular y también describe un ensayo de recombinación con hibridación de cadena sencilla para la prueba inicial de nucleasas de dedo de zinc.

## Sumario

5 La presente invención proporciona una pareja de nucleasas de dedo de zinc (ZFN) que comprende una ZFN izquierda y una ZFN derecha, comprendiendo cada ZFN un dominio de escisión o semidominio de escisión de tipo silvestre o modificado y cinco o seis dominios de dedo de zinc designados y ordenados como F1 a F5 o F1 a F6, comprendiendo F1 a F5 o F1 a F6 las regiones de la hélice de reconocimiento de la siguiente manera:

- 10 (i) comprendiendo F1 a F5 las SEQ ID NO: 109, 104, 110, 106 y 107, respectivamente (SBS N° 47171);  
 (ii) comprendiendo F1 a F6 las SEQ ID NO: 14, 114, 111, 15, 16 y 12, respectivamente (SBS N° 47898);  
 (iii) comprendiendo F1 a F6 las SEQ ID NO: 14, 9, 10, 15, 16 y 12, respectivamente (SBS N° 47079 o SBS N° 47169);  
 (iv) comprendiendo F1 a F6 las SEQ ID NO: 14, 9, 111, 15, 16 y 12, respectivamente (SBS N° 47931);  
 (v) comprendiendo F1 a F5 las SEQ ID NO: 2, 112, 110, 106 y 113, respectivamente (SBS N° 47863);  
 15 (vi) comprendiendo F1 a F5 las SEQ ID NO: 2, 112, 110, 106 y 107, respectivamente (SBS N° 47192);

en donde la ZFN izquierda comprende las regiones de la hélice de reconocimiento de los dedos de zinc seleccionadas de (i), (v) y (vi), y la ZFN derecha comprende las regiones de la hélice de reconocimiento de los dedos de zinc seleccionadas de (ii), (iii), y (iv), como se define en las reivindicaciones.

20 La presente invención también proporciona uno o más polinucleótidos, una composición farmacéutica, un método *in vitro* o *ex vivo* para escindir un gen de la albúmina en una célula, uno o más vectores de expresión para usar en un método de tratamiento de un paciente, un vector de expresión o composición farmacéutica para usar en un método de tratamiento de un paciente con una enfermedad de almacenamiento lisosomal y un kit, todo tal como se define en las reivindicaciones.

En el presente documento se divulgan métodos y composiciones que se pueden usar para expresar un transgén bajo el control de un promotor de albúmina *in vivo* (p. ej., promotor de albúmina endógeno o exógeno). En algunos aspectos, el transgén puede codificar una proteína terapéutica de interés. El transgén puede codificar una proteína de tal manera que los métodos de la invención se puedan usar para la producción de la proteína que es deficiente o inexistente (p. ej., "reemplazo de la proteína"). En algunos casos, la proteína puede estar implicada en el tratamiento de una enfermedad de almacenamiento lisosomal. Se pueden expresar otras proteínas terapéuticas, incluyendo proteínas terapéuticas para afecciones tan diversas como la epidermolisis ampollosa, diabetes, cáncer, trastornos de la coagulación o enfisema deficiente en AAT. En otros aspectos, el transgén puede comprender secuencias (p. ej., secuencias modificadas) de modo que la proteína expresada tiene características que le confieren características novedosas y deseables (semivida aumentada, características de aclaramiento plasmático modificadas, etc.). Las secuencias modificadas también pueden incluir aminoácidos derivados de la secuencia de albúmina. En algunos aspectos, los transgenes codifican proteínas terapéuticas, hormonas terapéuticas, proteínas plasmáticas, anticuerpos y similares. En algunos aspectos, los transgenes pueden codificar proteínas implicadas en trastornos sanguíneos, tales como trastornos de la coagulación. En algunos aspectos, los transgenes codifican ácidos nucleicos estructurales (ARNhc, ARNi, miARN y similares).

45 En un aspecto, se divulgan aquí métodos y composiciones para la integración dirigida de una secuencia que codifica una proteína de factor de coagulación funcional (p. ej., Factor VII, Factor VIII, Factor IX y/o Factor X). La expresión de una proteína funcional de Factor VIII ("F8") funcional y/o Factor IX ("F.IX" o "FIX") puede dar como resultado, por ejemplo, el tratamiento y/o prevención de la hemofilia A (F8) y/o la hemofilia B (F.IX), mientras que la expresión de un Factor VII o Factor X funcional puede tratar o prevenir las hemofilias asociadas con la deficiencia de Factor VII y/o Factor X.

50 En otro aspecto, se divulgan en el presente documento métodos y composiciones para la integración dirigida de una secuencia que codifica una proteína funcional inexistente en un sujeto con una enfermedad de almacenamiento lisosomal. Las nucleasas, las meganucleasas modificadas, las nucleasas de dedo de zinc (ZFN), las nucleasas TALE (TALEN que incluyen fusiones de dominios efectores TALE con dominios de nucleasas de endonucleasas de restricción y/o de meganucleasas (como mega TALE y TALEN compactas)), el sistema Ttago y/o los sistemas CRISPR/nucleasa Cas se usan para escindir el ADN en un locus de un gen 'sitio seguro' (p. ej., *CCR5*, *AAVS1*, *HPRT*, *Rosa* o albúmina) en la célula en la cual se inserta el gen. La inserción dirigida de un transgén donante puede realizarse mediante mecanismos de reparación dirigida por homología (HDR) o de no homología (p. ej., captura de donante NHEJ). La nucleasa puede inducir una rotura de la cadena doble (DSB) o de la cadena sencilla (muesca) en el ADN diana. En algunas realizaciones, se usan dos nickasas para crear una DSB mediante la introducción de dos muescas. En algunos casos, la nickasa es una ZFN, mientras que en otros, la nickasa es una TALEN o una CRISPR/nickasa Cas.

65 En un aspecto, se describe en el presente documento una proteína de dedo de zinc (ZFP) no natural que se une a un sitio diana en una región de interés (p. ej., un gen de la albúmina) en un genoma, en donde la ZFP comprende uno o más dominios de unión de dedo de zinc modificados. En una realización, la ZFP es una nucleasa de dedo de zinc (ZFN) que escinde una región genómica diana de interés, en donde la ZFN comprende uno o más dominios de unión

de dedos de zinc modificados y un dominio o semidominio de escisión de nucleasa. Los dominios y semidominios de escisión se pueden obtener, por ejemplo, a partir de diversas endonucleasas de restricción y/o endonucleasas orientadoras (homíng). En una realización, los semidominios de escisión derivan de una endonucleasa de restricción de Tipo IIS (*p.ej.*, Fok I). En ciertas partes de la divulgación, el dominio de dedo de zinc reconoce un sitio diana en un gen de la albúmina, por ejemplo una proteína de dedo de zinc con los dominios de la hélice de reconocimiento ordenados como se muestra en una sola fila de la Tabla 5.

En otro aspecto, se describe en el presente documento una proteína efector de tipo activador de la transcripción (TALE) que se une a un sitio diana en una región de interés (*p.ej.*, un gen de la albúmina) en un genoma, en donde la TALE comprende uno o más dominios de unión de TALE modificados. En una parte de la divulgación, la TALE es una nucleasa (TALEN) que escinde una región genómica diana de interés, en donde la TALEN comprende uno o más dominios de unión al ADN de TALE modificados y un dominio de escisión o semidominio de escisión de nucleasa. Los dominios y semidominios de escisión se pueden obtener, por ejemplo, a partir de diversas endonucleasas de restricción y/o endonucleasas orientadoras (meganucleasa). Los semidominios de escisión pueden derivarse de una endonucleasa de restricción de Tipo IIS (*p.ej.*, Fok I). También se divulga en el presente documento, el dominio escisión derivado de una meganucleasa, meganucleasa cuyo dominio también puede exhibir una funcionalidad de unión al ADN.

En otro aspecto, se describen en el presente documento un sistema CRISPR/Cas que se une al sitio diana en una región de interés (*p.ej.*, un gen de la albúmina) en un genoma, en donde el sistema CRISPR/Cas comprende uno o más ARN guía único modificado o un equivalente funcional, así como una nucleasa Cas9.

Las nucleasas (*p.ej.*, ZFN, sistema CRISPR/Cas, Ttago y/o TALEN) como se describe en el presente documento pueden unirse y/o escindir la región de interés en una región codificante o no codificante dentro o adyacente al gen, tal como, por ejemplo, una secuencia líder, una secuencia de remolque o intrón, o dentro de una secuencia no transcrita, en dirección 5' o 3' de la región codificante. En determinadas realizaciones, la nucleasa (*p.ej.*, ZFN) se une a y/o escinde un gen de la albúmina.

En otro aspecto, en el presente documento se describe un polinucleótido que codifica una o más nucleasas (*p.ej.*, ZFN, sistemas CRISPR/Cas, Ttago y/o las TALEN descritos en el presente documento). El polinucleótido puede ser, por ejemplo, ARNm. En algunos aspectos, el ARNm puede estar modificado químicamente (Véase *p. ej.*, Kormann *et al.*, (2011) Nature Biotechnology 29(2):154-157).

En otro aspecto, se describe en el presente documento un vector de expresión de ZFN, sistema CRISPR/Cas, Ttago y/o TALEN que comprende un polinucleótido, que codifica una o más nucleasas (*p.ej.*, ZFN, sistemas CRISPR/Cas, Ttago y/o TALEN) como se describe en el presente documento, unido de forma operativa a un promotor. En una realización, el vector de expresión es un vector viral. En un aspecto, el vector viral exhibe tropismo específico de tejido.

En otro aspecto, se describe en el presente documento una célula hospedadora que comprende uno o más vectores de expresión de nucleasa (*p.ej.*, ZFN, sistemas CRISPR/Cas, Ttago y/o TALEN).

En otro aspecto, se proporcionan composiciones farmacéuticas que comprenden un vector de expresión como se describe en el presente documento. En algunas realizaciones, la composición farmacéutica puede comprender más de un vector de expresión. En algunas realizaciones, la composición farmacéutica comprende un primer vector de expresión que comprende un primer polinucleótido y un segundo vector de expresión que comprende un segundo polinucleótido. En algunas realizaciones, el primer polinucleótido y el segundo polinucleótido son diferentes. En algunas realizaciones, el primer polinucleótido y el segundo polinucleótido son sustancialmente iguales. La composición farmacéutica puede comprender además una secuencia donante (por ejemplo, un transgén que codifica una proteína inexistente o deficiente en una enfermedad o trastorno tal como un LSD o una hemofilia). En algunas realizaciones, la secuencia donante está asociada con un vector de expresión.

En algunas realizaciones, se proporciona una proteína de fusión que comprende una proteína de dedo de zinc y un dominio de escisión o semidominio de escisión de tipo silvestre o modificado.

Como se divulga en el presente documento, se proporciona una composición farmacéutica que comprende: (i) un primer polinucleótido (*p.ej.*, plásmido, ARNm, vector de Ad, vector AAV, etc.) que codifica una nucleasa de dedo de zinc, comprendiendo la nucleasa de dedo de zinc un dominio de escisión FokI y una proteína de dedo de zinc que comprende 5 o 6 dominios de dedo de zinc ordenados de F1 a F5 o de F1 a F6, en donde cada dominio de dedo de zinc comprende una región de la hélice de reconocimiento y en donde las regiones de la hélice de reconocimiento de la proteína de dedo de zinc se muestran en una única fila de la Tabla 1, 2 o 5; (ii) un segundo polinucleótido (*p.ej.*, plásmido, ARNm, vector de Ad, vector AAV, etc.) que codifica una nucleasa de dedo de zinc, comprendiendo la nucleasa de dedo de zinc un dominio de escisión FokI y una proteína de dedo de zinc que comprende 5 o 6 dominios de dedo de zinc ordenados de F1 a F5 o de F1 a F6, en donde cada dominio de dedo de zinc comprende una región de la hélice de reconocimiento y en donde las regiones de la hélice de reconocimiento de la proteína de dedo de zinc se muestran en una única fila de la Tabla 1, 2 o 5; y (iii) un tercer polinucleótido (*p.ej.*, plásmido, ARNm, vector de Ad, vector AAV, etc.) comprendiendo el vector un donante que codifica una proteína inexistente o deficiente en una

enfermedad o trastorno (*p.ej.*, LDS o hemofilia). Las ZFN de las dos ZFN pueden ser iguales o diferentes. De manera similar, los dominios de escisión de las dos ZFN pueden ser iguales o diferentes (*p.ej.*, pueden ser mutantes que forman heterodímeros obligados). En algunas realizaciones, (i), (ii), y (iii) se proporcionan en una relación de aproximadamente 1:1:1, aproximadamente 1:1:2, aproximadamente 1:1:3, aproximadamente 1:1:4, aproximadamente 1:1:5, aproximadamente 1:1:6, aproximadamente 1:1:7, aproximadamente 1:1:8, aproximadamente 1:1:9, aproximadamente 1:1:10, aproximadamente 1:1:11, aproximadamente 1:1:12, aproximadamente 1:1:13, aproximadamente 1:1:14, aproximadamente 1:1:15, aproximadamente 1:1:16, aproximadamente 1:1:17, aproximadamente 1:1:18, aproximadamente 1:1:19, o aproximadamente 1:1:20.

En un aspecto, los métodos y composiciones comprenden células genéticamente modificadas que comprenden un transgén que expresa una versión funcional de una proteína que se expresa de manera aberrante en una hemofilia (proteína Factor VII, F8, F.IX y/o Factor X), en donde el transgén está integrado en un gen de sitio seguro endógeno (*p.ej.*, gen de la albúmina) del genoma de la célula. En otro aspecto, los métodos y composiciones comprenden células genéticamente modificadas que comprenden un transgén que expresa una versión funcional de una proteína inexistente o que se expresa de manera anormal en un sujeto con una enfermedad de almacenamiento lisosomal. En determinadas realizaciones, el transgén se integra de una manera específica del sitio (dirigida) usando al menos una nucleasa. En ciertas partes de la divulgación, la nucleasa (*p.ej.*, ZFN, TALEN, Ttago y/o sistemas CRISPR/Cas) es específica de un gen de sitio seguro (*p.ej.*, *CCR5*, *HPRT*, *AAVS1*, *Rosa* o albúmina. Véase, *p. ej.*, las patentes de Estados Unidos números 7.951.925 y 8.110.379; las publicaciones de patente Estados Unidos números 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y la solicitud de patente de Estados Unidos N° 9.873.894). En algunas realizaciones, el sitio seguro es un gen de la albúmina.

En otro aspecto, se describe en el presente documento un método para modificar genéticamente una célula, *in vitro* y/o *in vivo*, para producir una proteína terapéutica (*p.ej.*, una proteína inexistente en una enfermedad o trastorno, tal como una hemofilia (Factor VII, F8, F.IX y/o Factor X) o una enfermedad de almacenamiento lisosomal (IDS, IDUA, etc.), comprendiendo el método la escisión de un gen de sitio seguro en la célula usando una o más nucleasas (*p.ej.*, ZFN, TALEN, CRISPR/Cas) de tal manera que un transgén que codifica la proteína terapéutica se integra en el locus de sitio seguro y se expresa en la célula. En el presente documento el gen de sitio seguro divulgado es *aCCR5*, *HPRT*, *AAVS1*, *Rosa* o un gen de la albúmina. En un aspecto adicional, se describe en el presente documento un método para modificar genéticamente una célula, *in vitro*, para producir una proteína inexistente en una enfermedad de almacenamiento lisosomal. Los ejemplos más frecuentes de estas son la deficiencia de glucocerebrosidasa (nombre del gen: GBA), asociada a la enfermedad de Gaucher, deficiencia de  $\alpha$  galactosidasa (nombre del gen: GLA), asociada a la enfermedad de Fabry, deficiencia de iduronato-2-sulfatasa (nombre del gen: IDS), asociada a la enfermedad de Hunter, deficiencia de alfa-L iduronidasa (nombre del gen: IDUA), asociada a la enfermedad de Hurler, y deficiencia de esfingomielina fosfodiesterasa 1 (nombre del gen: SMPD1), asociada a la enfermedad de Niemann-Pick. En determinadas realizaciones, la célula es una célula de mamífero. En determinadas realizaciones, la célula es una célula de primate. En determinadas realizaciones, la célula es una célula humana. En un conjunto de realizaciones, se proporcionan métodos para la escisión de un gen de la albúmina en una célula (*p. ej.*, una célula hepática) que comprende introducir, en la célula, uno o más vectores de expresión divulgados en el presente documento en condiciones tales que se expresen la una o más proteínas y se escinda el gen de albúmina. El gen de la albúmina se puede modificar, por ejemplo, mediante la integración de una secuencia donante en el gen de la albúmina escindido. En ciertas partes de la divulgación el método comprende modificar genéticamente una célula para producir un factor de coagulación o una proteína inexistente en una enfermedad de almacenamiento lisosomal, comprendiendo el método administrar a la célula las nucleasas de dedo de zinc (ZFN) mostradas en la Tabla 5 (o polinucleótidos que codifican estas ZFN) y un donante. Las ZFN y el donante pueden estar en el mismo o diferentes vectores en cualquier combinación, por ejemplo en 3 vectores separados (*p. ej.*, vectores de AAV) llevando cada uno de ellos uno de los componentes; llevando un vector dos de los componentes y llevando un vector separado el 3<sup>er</sup> componente; o llevando un vector los 3 componentes.

En otros aspectos, la divulgación comprende la administración de un ácido nucleico donante a una célula diana. El donante se puede administrar antes, después, o junto con el ácido nucleico que codifica la nucleasa(s). El ácido nucleico donante puede comprender una secuencia exógena (transgén) para integrarse en el genoma de la célula, por ejemplo, un locus endógeno. En algunas realizaciones, el donante puede comprender un gen de longitud completa o un fragmento del mismo flanqueado por regiones de homología con el sitio de escisión dirigido. En algunas realizaciones, el donante carece de regiones homólogas y está integrado en un locus diana mediante un mecanismo independiente de homología (es decir, NHEJ). El donante puede comprender cualquier secuencia de ácido nucleico, por ejemplo un ácido nucleico que, cuando se usa como sustrato para la reparación dirigida por homología de la rotura de doble cadena inducida por la nucleasa, conduce a la generación de una delección especificada por el donante en el locus cromosómico endógeno o, como alternativa (o además de), nuevas formas alélicas de (*p. ej.*, mutaciones puntuales que eliminan un sitio de unión al factor de transcripción) el locus endógeno a crear. En algunos aspectos, el ácido nucleico donante es un oligonucleótido en donde la integración conduce a un evento de corrección génica o una delección dirigida. En algunos aspectos, el donante comprende una proteína terapéutica, por ejemplo, un factor de coagulación.

En algunas realizaciones, el polinucleótido que codifica la proteína de unión a ADN es un ARNm. En algunos aspectos,

el ARNm puede estar modificado químicamente (Véase *p.ej.*, Kormann *et al.*, (2011) Nature Biotechnology 29(2):154-157). En otros aspectos, el ARNm puede comprender una caperuza ARCA (véanse las patentes de Estados Unidos 7.074.596 y 8.153.773). En realizaciones adicionales, el ARNm puede comprender una mezcla de nucleótidos no modificados y modificados (véase la publicación de patente de Estados Unidos 2012-0195936).

5 En otro aspecto, se divulgan en el presente documento métodos para proporcionar una o más proteínas funcionales inexistentes o deficientes en un mamífero, o en un primate, tal como un primate humano, tal como un paciente humano, con una LSD y/o una hemofilia, por ejemplo para el tratamiento de la enfermedad suministrado la proteína(s) inexistente o deficiente en el sujeto. En otro aspecto, en el presente documento se divulgan métodos para proporcionar una proteína funcional (*p.ej.*, F.IX) inexistente o deficiente en un mamífero, o en un primate, tal como un primate humano, tal como un paciente humano con hemofilia B, por ejemplo para el tratamiento de la hemofilia B. En otro aspecto, en el presente documento se divulgan métodos para proporcionar una proteína funcional (*p.ej.*, Factor VII,) a un mamífero, o a un primate, tal como un primate humano, tal como un paciente humano, para el tratamiento de la hemofilia asociada a la deficiencia de Factor VII. En otro aspecto, en el presente documento se divulgan métodos para proporcionar una proteína funcional (*p.ej.*, Factor X), para el tratamiento de la hemofilia asociada a la deficiencia de Factor X. En determinadas realizaciones, los métodos comprenden el uso de nucleasas para integrar una secuencia que codifica una proteína de Factor VII, F8, F.IX y/o Factor X funcional en una célula en un sujeto que lo necesita. En otras realizaciones, los métodos comprenden el uso de nucleasas para integrar una secuencia que codifica una proteína funcional inexistente o deficiente en una enfermedad de almacenamiento lisosomal. En otras partes de la divulgación el método comprende administrar una célula genéticamente modificada (que expresa una versión funcional de una proteína que se expresa aberrantemente en un sujeto con hemofilia) al sujeto. Por lo tanto, una célula aislada puede introducirse en el sujeto (terapia celular *ex vivo*). También se divulga el uso de los donantes y/o nucleasas descritos en el presente documento para el tratamiento de una hemofilia (*p.ej.*, hemofilia A con donante del Factor VIII, hemofilia B con donante del Factor IX, deficiencia del Factor VII con Factor VII, deficiencia del Factor X con Factor X, enfermedad de Gaucher con donante de GBA, enfermedad de Fabry con donante de GLA, enfermedad de Hunter con donante de IDS, enfermedad de Hurler con donante de IDUA, y/o enfermedad de Niemann-Pick con donante de SMPD1), por ejemplo, para la preparación de un medicamento para el tratamiento de una enfermedad. En determinadas realizaciones, la proteína F8 comprende una delección del dominio B. En determinadas realizaciones, la secuencia que codifica F8 y/o F.IX se administra usando un vector viral, un vector no viral (*p.ej.*, plásmido) y/o combinaciones de los mismos.

En cualquiera de las composiciones y métodos descritos, la nucleasa(s) y/o transgén(es) puede llevarse a cabo en un vector AAV, incluyendo, pero sin limitación, AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV9 y AAVrh10 o AAV pseudotipados tales como AAV2/8, AAV8.2, AAV2/5 y AAV2/6 y similares. En determinadas realizaciones, las nucleasas y los donantes de transgenes se administran usando los mismos tipos de vectores de AAV. En otras realizaciones, las nucleasas y los donantes de transgenes se administran usando diferentes tipos de vectores de AAV. Las nucleasas y los transgenes se pueden administrar usando uno o más vectores, por ejemplo, un vector lleva ambos, el transgén y la nucleasa(s); dos vectores donde uno lleva la nucleasa(s) (*p. ej.*, ZFN izquierda y derecha o una pareja de ZFN, por ejemplo con un péptido 2A) y uno lleva el transgén; o tres vectores donde un vector lleva una nucleasa de una pareja de nucleasas (*p. ej.*, ZFN izquierda), un vector separado lleva la otra nucleasa de una pareja de nucleasas (*p. ej.*, ZFN derecha) y un tercer vector separado lleva el transgén. Véase, Figura 2. En realizaciones en las que se usan dos o más vectores, los vectores se pueden usar en las mismas concentraciones o en diferentes proporciones, por ejemplo, el vector(es) de transgenes se puede administrar en concentraciones 2 veces, 3 veces, 4 veces, 5 veces o más superiores que las del vector(es) de nucleasa. En determinadas realizaciones, las nucleasas y/o donantes de transgenes se administran por vía intravenosa (*p.ej.*, vena intraportal) en el hígado de un animal intacto.

En cualquiera de las composiciones y métodos descritos en el presente documento, la proteína codificada por el transgén puede comprender una proteína F8, por ejemplo un Factor VIII con el dominio B eliminado (BDD-F8). En otras realizaciones, la proteína codificada por el transgén comprende una proteína F.IX. En otras realizaciones, la proteína codificada por el transgén comprende una proteína Factor VII. En otras realizaciones, la proteína codificada por el transgén comprende una glucocerebrosidasa. En otras partes de la divulgación, la proteína codificada por el transgén comprende una  $\alpha$  galactosidasa. En partes adicionales de la divulgación, la proteína codificada por el transgén comprende una iduronato-2-sulfatasa. En algunas partes de la divulgación, la proteína codificada por el transgén comprende una alfa-L-iduronidasa. En partes adicionales de la divulgación, la proteína codificada por el transgén comprende esfingomielina fosfodiesterasa. En cualquiera de las composiciones o métodos descritos en el presente documento, el transgén también comprende un regulador de la transcripción mientras que en otros, no y la transcripción está regulada por un regulador endógeno. En otro aspecto, los métodos de la invención comprenden una composición para el tratamiento terapéutico de un sujeto que lo necesita. En algunas partes de la divulgación, la composición comprende células madre modificadas que comprenden una nucleasa específica de sitio seguro, y un donante de transgén que codifica la proteína Factor VII, F8, F.IX, Factor X, GBA, GLA, IDS, IDUA y/o SMPD1 o un fragmento funcional y/o truncamiento de la misma. En otras partes de la divulgación, la composición comprende células madre modificadas que se han modificado y expresan un donante de transgén que codifica el Factor VII, F8, F.IX, Factor X, GBA, GLA, IDS, IDUA y/o SMPD1 o un fragmento funcional y/o truncamiento de la misma.

En cualquiera de las composiciones o métodos descritos en el presente documento, la célula puede ser una célula

eucariota. Los ejemplos no limitativos de células adecuadas incluyen células eucariotas o líneas celulares tales como células secretoras (*p. ej.*, células hepáticas, células de la mucosa, células de la glándula salival, células hipofisarias, etc.), células sanguíneas (glóbulos rojos), células precursoras de glóbulos rojos, células hepáticas, miocitos, células madre (*p. ej.*, células madre embrionarias, células madre pluripotentes inducidas, células madre hepáticas, células madre hematopoyéticas (*p. ej.*, CD34+)) o células endoteliales (*p. ej.*, células vasculares, glomerulares, y tubulares). Por lo tanto, las células diana pueden ser células de primate, por ejemplo células humanas, o las células diana pueden ser células de mamífero, (incluyendo animales domésticos), por ejemplo especialmente primates no humanos y mamíferos de los órdenes *Rodenta* (ratones, rata, hámster), *Lagomorpha* (conejos), *Carnivora* (gatos, perros), y *Arteriodactyla* (vacas, cerdos, ovejas, cabras, caballos). En algunos aspectos, las células diana comprenden un tejido (*p. ej.*, hígado). En algunos aspectos, la célula diana es una célula madre (*p. ej.*, una célula madre embrionaria, una célula madre pluripotente inducida, una célula madre hepática, etc.) o un embrión animal no humano por cualquiera de los métodos descritos en el presente documento, y a continuación el embrión se implanta de manera que nazca un animal vivo con la condición de que el método no comprenda procesos para modificar la identidad genética de la línea germinal de los seres humanos. El animal no humano se cría hasta la madurez sexual y se le permite producir descendencia en donde al menos parte de la descendencia comprende la modificación genómica. La célula también puede comprender una célula embrionaria no humana, por ejemplo, de un ratón, rata, conejo u otra célula de mamífero, con la condición de que no sea una célula germinal humana. La célula puede ser de cualquier organismo, por ejemplo de un ser humano, primate no humano, ratón, rata, conejo, gato, perro u otras células de mamífero, con la condición de que no sea una célula germinal humana. La célula puede estar aislada o formar parte de un organismo (*p. ej.*, sujeto), con la condición de que no sea una célula germinal humana.

En cualquiera de los métodos o composiciones descritos en el presente documento, la célula que contiene el locus modificado (*p. ej.*, locus de la albúmina) puede ser una célula madre que puede ser útil para fines terapéuticos. Los tipos de células madre específicos que se pueden usar con los métodos y composiciones de la invención incluyen células madre embrionarias (ESC), incluidas células madre pluripotentes inducidas (iPSC) y células madre hepáticas o de hígado. Las iPSC se pueden obtener de muestras de pacientes y de controles normales en donde las iPSC derivadas del paciente pueden mutar a una secuencia génica normal en el gen de interés, o las células normales se pueden alterar al alelo de la enfermedad conocido en el gen de interés. De manera similar, las células madre hepáticas pueden aislarse de un paciente. Estas células se modifican a continuación para expresar el transgén de interés, expandirse y después reintroducirse en el paciente.

En cualquiera de los métodos y composiciones descritos en el presente documento, el transgén se puede integrar en el gen de sitio seguro endógeno de tal modo que algunos, todos o ningunos de los genes endógenos exprese, por ejemplo una proteína de fusión con el transgén integrado. En algunas realizaciones, el gen endógeno de sitio seguro es un gen de la albúmina y las secuencias endógenas son las secuencias de la albúmina. El gen endógeno puede estar presente en la porción amino (N)-terminal de la proteína exógena y/o en la porción carboxilo (C)-terminal de la proteína exógena. Las secuencias de la albúmina pueden incluir secuencias de albúmina de tipo silvestre o mutantes de longitud completa o, como alternativa, pueden incluir secuencias de aminoácidos de la albúmina parciales. En determinadas realizaciones, las secuencias de la albúmina (de longitud completa o parcial) sirven para aumentar la semivida en suero del polipéptido expresado por el transgén al que se fusiona y/o como un vehículo. En otras realizaciones, el transgén comprende las secuencias de la albúmina y está dirigido para la inserción en otro sitio seguro dentro de un genoma. Asimismo, el transgén puede incluir un promotor exógeno (*p. ej.*, promotor constitutivo o inducible) que dirige su expresión o su expresión puede ser dirigida por secuencias de control endógenas (*p. j.*, promotor de la albúmina endógeno). En algunas realizaciones, el donante incluye modificaciones adicionales, incluyendo, aunque no de forma limitativa, optimización de codones, adición de sitios de glicosilación y similares.

En cualquiera de las composiciones o métodos descritos en el presente documento, la escisión puede producirse mediante el uso de nucleasas específicas tales como las nucleasas de dedo de zinc (ZFN) modificadas, nucleasas efectoras del tipo activador de la transcripción (TALEN) o el uso del sistema Ttago o CRISPR/Cas con un ARNcr/ARNtracr modificado ('ARN de guía única') para guiar la escisión específica. En algunas realizaciones, se usan dos nickasas para crear una DSB mediante la introducción de dos muescas. En algunos casos, la nickasa es una ZFN, mientras que en otros, la nickasa es una TALEN o un sistema CRISPR/Cas. La integración dirigida puede producirse mediante mecanismos de reparación dirigida por homología (HDR) y/o mediante mecanismos de reparación sin homología (*p. ej.*, captura de donante NHEJ). Las nucleasas como se describe en el presente documento pueden unirse y/o escindir la región de interés en una región codificante o no codificante dentro o adyacente al gen, tal como, por ejemplo, una secuencia líder, una secuencia de remolque o intrón, o dentro de una secuencia no transcrita, en dirección 5' o 3' de la región codificante. En determinadas realizaciones, la nucleasa activa la secuencia diana en o cerca del sitio de unión (*p. ej.*, el sitio de unión mostrado en la Tabla 5). La escisión puede tener como resultado la modificación del gen, por ejemplo, mediante inserciones, deleciones o combinaciones de las mismas. En determinadas realizaciones, la modificación está en o cerca del sitio de unión de la nucleasa(s) y/o sitio(s) de unión, por ejemplo, en 1-300 (o cualquier valor entre los mismos) pares de bases en dirección 5' o 3' del sitio(s) de escisión, más preferentemente entre 1-100 pares de bases (o cualquier valor entre los mismos) de cada lado del sitio(s) de unión y/o escisión, aún más preferentemente entre 1 y 50 pares de bases (o cualquier valor entre los mismos) en cada lado del sitio(s) de unión y/o escisión.

El vector de expresión o composición farmacéutica puede usarse para tratar o prevenir una hemofilia en un sujeto que

lo necesita. En algunas realizaciones, las composiciones comprenden vectores y se usan para dirigirse a las células hepáticas. También se divulgan en el presente documento las composiciones que comprenden células madre modificadas y se administran a un paciente como un trasplante de médula ósea. En algunos casos, los pacientes se someten a inmunoblación parcial o completa antes del trasplante. En otros casos, los pacientes se tratan con uno o más agentes inmunosupresores antes, durante y/o después de la modificación mediada por nucleasas de un gen endógeno (*p.ej.*, integración dirigida de un transgén en un locus de la albúmina). Además, cualquiera de los métodos descritos en el presente documento puede comprender etapas adicionales, que incluyen hepatectomía parcial o tratamiento con agentes secundarios que mejoran la transducción y/o inducen a las células hepáticas a entrar en el ciclo celular. Los ejemplos de agentes secundarios incluyen irradiación gamma, irradiación UV, nucleótidos tritidos tales como timidina, cisplatino, etopósido, hidroxurea, afidicolina, prednisolona, tetracloruro de carbono y/o adenovirus.

Los métodos descritos en el presente documento se pueden poner en práctica *in vitro*, o *ex vivo*. En determinadas realizaciones, las composiciones para usar en un método de tratamiento se introducen en un mamífero vivo, intacto. El mamífero puede estar en cualquier etapa del desarrollo en el momento de la administración, *p.ej.*, embrionario, fetal, neonatal, infantil, juvenil o adulto. Adicionalmente, las células diana pueden ser sanas o estar afectadas. En determinadas realizaciones, una o más de las composiciones se administran por vía intravenosa (*p.ej.*, en el hígado a través de la vena intraportal, por ejemplo inyección en la vena de la cola), por vía intraarterial, por vía intraperitoneal, por vía intramuscular, en el parénquima hepático (*p.ej.*, mediante inyección), en la arteria hepática (*p.ej.*, mediante inyección), y/o a través del árbol biliar (*p.ej.*, mediante inyección).

En un aspecto particular, los métodos y composiciones descritos en el presente documento incluyen una composición terapéutica que comprende (i) un transgén donante que codifica FVIII (ii) una nucleasa (*p.ej.*, ZFN, TALEN, Ttgo o sistema CRISPR/Cas dirigido a un locus de un gen endógeno diferente a FVIII, respectivamente, por ejemplo, dirigido al gen de la albúmina endógeno de un mamífero, o primate o ser humano, tal como un paciente con hemofilia. En determinadas realizaciones, la composición terapéutica comprende el transgén donante de FVIII y la nucleasa específica del gen de la albúmina en vectores separados, independientes, tales como vectores de AAV separados, en cantidades diferentes, que se pueden administrar conjuntamente (*p.ej.*, mezclados en una solución única o administrarse simultáneamente) o, como alternativa, que se pueden administrar por separado (por ejemplo, se pueden administrar en soluciones separadas con un retraso sustancial, *p. ej.*, de 10 minutos o más, de 30 minutos o más, 1 hora o más, 2 horas o más, 3 horas o más, o más tiempo entre las respectivas administraciones). En determinadas realizaciones, la composición terapéutica se administra para proporcionar la integración del transgén donante de FVIII en el locus no FVIII y la posterior expresión del FVIII integrado para lograr un nivel terapéutico de FVIII en el plasma del mamífero, o primate o ser humano, o paciente con hemofilia. En determinadas realizaciones, un nivel terapéutico de FVIII puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, o más de una concentración plasmática de FVIII normal clínicamente aceptable. Como alternativa o adicionalmente, un nivel terapéutico de FVIII puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, más de 30 %, más de 35 % o más de la concentración plasmática de FVIII funcional medida en el mamífero individual, o primate o ser humano, o paciente con hemofilia antes de la administración del transgén donante de FVIII y la nucleasa del gen de la albúmina a ese individuo.

En un aspecto particular, los métodos y composiciones descritos en el presente documento incluyen una composición terapéutica que comprende (i) un transgén donante que codifica una proteína deficiente en una enfermedad de almacenamiento lisosomal (ii) una nucleasa (*p.ej.*, ZFN, TALEN, Ttgo o sistema CRISPR/Cas) dirigido a un locus de un gen endógeno diferente al gen de la proteína deficiente en una enfermedad de almacenamiento lisosomal, respectivamente, por ejemplo, dirigido al gen de la albúmina endógeno de un mamífero, o primate o ser humano, tal como un sujeto con una enfermedad de almacenamiento lisosomal. En ciertas partes de la divulgación, la composición terapéutica comprende el transgén donante seleccionado de GBA, GLA, IDS, IDUA y/o SMPD1 y la nucleasa específica del gen de la albúmina en vectores separados, independientes, tales como vectores de AAV separados, en cantidades diferentes, que se pueden administrar conjuntamente (*p.ej.*, mezclados en una solución única o administrarse simultáneamente) o, como alternativa, que se pueden administrar por separado (por ejemplo, se pueden administrar en soluciones separadas con un retraso sustancial, *p.ej.*, de 10 minutos o más, de 30 minutos o más, 1 hora o más, 2 horas o más, 3 horas o más, o más tiempo entre las respectivas administraciones). En ciertas partes de la divulgación, la composición terapéutica se administra para proporcionar la integración del transgén donante de GBA, GLA, IDS, IDUA y/o SMPD1 en un locus que no es el locus que codifica GBA, GLA, IDS, IDUA y/o SMPD1, respectivamente, y la expresión posterior de GBA, GLA, IDS, IDUA y/o SMPD1 integrado para lograr un nivel terapéutico de GBA, GLA, IDS, IDUA y/o SMPD1 en el plasma del mamífero, o primate o ser humano, o paciente con hemofilia. En ciertas partes de la divulgación, un nivel terapéutico de GBA, GLA, IDS, IDUA y/o SMPD1 puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, o más de una concentración plasmática de GBA normal clínicamente aceptable, GLA, IDS, IDUA y/o SMPD1. Como alternativa o adicionalmente, un nivel terapéutico de GBA, GLA, IDS, IDUA y/o SMPD1 puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, más de 30 %, más de 35 %, o más de la concentración plasmática de GBA, GLA, IDS, IDUA y/o SMPD1 funcionales medido en el mamífero individual, o primate o ser humano, o paciente antes de la administración del transgén donante de GBA, GLA, IDS, IDUA y/o SMPD1 y la

nucleasa del gen de la albúmina a ese individuo.

En otro aspecto particular, los métodos y composiciones descritos en el presente documento incluyen una composición terapéutica que comprende (i) un transgén donante que codifica FIX (ii) una nucleasa (p.ej., ZFN, TALEN, Ttgo o sistema CRISPR/Cas dirigida a un locus de un gen endógeno diferente a FIX, respectivamente, por ejemplo, dirigido al gen de la albúmina endógeno de un mamífero, o primate o ser humano, tal como un paciente con hemofilia. En ciertas partes de la divulgación, la composición terapéutica comprende el transgén donante de FIX y la nucleasa específica del gen de la albúmina en vectores separados, independientes, tales como vectores de AAV separados, en cantidades diferentes, que se pueden administrar conjuntamente (p. ej., mezclados en una solución única o administrarse simultáneamente) o, como alternativa, que se pueden administrar por separado ((por ejemplo, se pueden administrar en soluciones separadas con un retraso sustancial, p. ej., de 10 minutos o más, de 30 minutos o más, 1 hora o más, 2 horas o más, 3 horas o más, o más tiempo entre las respectivas administraciones). En ciertas partes de la divulgación, la composición terapéutica se administra para proporcionar la integración del transgén donante de FIX en el locus no FIX y la posterior expresión del FIX integrado para lograr un nivel terapéutico de FIX en el plasma del mamífero, o primate o ser humano, o paciente con hemofilia. En ciertas partes de la divulgación, un nivel terapéutico de FIX puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, o más de una concentración plasmática de FIX normal clínicamente aceptable. Como alternativa o adicionalmente, un nivel terapéutico de FIX puede incluir, por ejemplo, más de 2 %, más de 4 %, más de 5 %, más de 6 %, más de 8 %, más de 10 %, más de 12 %, más de 15 %, más de 20 %, más de 25 %, más de 30 %, más de 35 % o más de la concentración plasmática de FIX funcional medida en el mamífero individual, o primate o ser humano, o paciente con hemofilia antes de la administración del transgén donante de FIX y la nucleasa del gen de la albúmina a ese individuo.

Para el direccionamiento de las composiciones a un tipo de célula particular, p. ej., plaquetas, fibroblastos, hepatocitos, etc., una o más de las composiciones administradas puede estar asociada con un agente de orientación (homing) que se une específicamente a un receptor de la superficie de la célula. Por ejemplo, el vector puede conjugarse con un ligando (p.ej., galactosa) para el cual tienen receptores ciertas células del sistema hepático. La conjugación puede ser covalente, p.ej., un agente de reticulación, tal como glutaraldehído, o no covalente, p.ej., la unión de un ligando avidinado a un vector biotinilado. Se proporciona otra forma de conjugación covalente modificando el plásmido auxiliar de AAV utilizado para preparar el stock de vectores de modo que una o más de las proteínas de la cubierta codificadas sean un híbrido de una proteína de la cubierta del AAV nativa y un péptido o ligando de proteína, de modo que el ligando sea expuesto en la superficie de la partícula viral.

También se divulga un kit, que comprende las composiciones (p.ej., células modificadas genéticamente, ZFP, el sistema CRISPR/Cas y/o TALEN y opcionalmente donantes de transgenes). El kit puede comprender ácidos nucleicos que codifican las nucleasas, (p.ej., moléculas de ARN o genes que codifican nucleasas contenidas en un vector de expresión adecuado), moléculas donantes, líneas celulares hospedadoras adecuadas, instrucciones para llevar a cabo los métodos de la invención y similares.

Estos y otros aspectos serán fácilmente evidentes para el experto en la técnica a la luz de la divulgación en su totalidad.

### Breve descripción de los dibujos

La **Figura 1** es un esquema que representa una inserción mediada por nucleasa de dedo de zinc de un transgén que codifica una proteína (p.ej., proteína terapéutica) en el locus endógeno de la albúmina. "SA" se refiere a un sitio aceptor de corte y empalme; "pA" se refiere a una señal de poliadenilación; y "Alb Ex1" se refiere al exón 1 del locus endógeno de la albúmina.

Las **Figuras 2A a 2C** son esquemas que representan ejemplos de nucleasas y diseños de donante para un donante de transgén de F8. La Figura 2A muestra un ejemplo de diseño de donante y la Figura 2B muestra otro ejemplo de diseño ("donante optimizado") que incluye la optimización de codones, el uso de diferentes señales de poliadenilación y/o la adición de un motivo de glucosilación putativo ("péptido V3"). Los donantes representados en las Figuras 2A y 2B carecen de regiones promotoras/potenciadoras y tienen un tamaño de aproximadamente 4,4 a 4,7 kb, lo que es ideal para empaquetar en AAV. La Figura 2C es un diseño esquemático que representa vectores separados para cada ZFN de la pareja utilizada para la escisión de sitio seguro para la integración dirigida del transgén de F8 y/o F9.

Las **Figuras 3A a 3D** representan los resultados de la integración dirigida mediada por nucleasas de dedo de zinc de una secuencia que codifica una proteína F8 humana (hFVIII) en un locus de la albúmina endógeno que muestra actividad de hFVIII en plasma de ratones HA/CD4<sup>-/-</sup> tras la administración de ZFN y la administración del donante usando vectores AAV2/8. La Figura 3A es un esquema que representa el donante de transgén AAV2/8. La Figura 3B y 3C son gráficos que representan niveles plasmáticos de hFVIII como un porcentaje de los niveles normales en ratones HA/CD4<sup>-/-</sup> inyectados con los vectores ZFN y donantes de transgén ("Donante ZFN+") o vectores vacíos (sin secuencia de ZFN) y los donantes de transgén ("Simulado+Donante"). Los resultados se muestran 2 semanas después de la administración (Figura 3B) o 2 y 8 semanas después de la administración (Figura 3C). Se muestran los vectores y las cantidades administradas. La Figura 3D es un gráfico que muestra los niveles de modificación del gen de la albúmina en ratones a los que se les administró un solo vector que codifica tanto la ZFN izquierda como la derecha ("fusión 2A") o vectores separados que codifican cada uno una ZFN de la pareja como se muestra

en la Figura 2C ("ZFN individuales"). Los vectores usados fueron AAV2/8 y la dosis se representa a lo largo del eje horizontal (Genomas virales ("VGs") por ratón).

La **Figura 4** es un gráfico que representa los niveles plasmáticos de hFVIII como un porcentaje de los niveles normales en ratones HA/CD4<sup>-/-</sup> a los que se les inyectó ZFN dirigidas a la albúmina y la construcción del donante optimizada (V3) en la Figura 2B. Se usó AAV2/8-ZFN ( $5 \times 10^{10}$  vg de cada ZFN) + Donante de AAV2/8 ( $1 \times 10^{11}$  vg/ratón).

Las **Figuras 5A y 5B** muestran la integración mediada por nucleasa de un transgén de F.IX en el locus de la albúmina. La Figura 5A es un esquema de la construcción F.IX del transgén donante. El donante comprende brazos de donantes que son homólogos al locus F9 humano ("Brazo humano izquierdo" y "Brazo humano derecho"), por lo que no se espera que promuevan HDR en este experimento. Por lo tanto, la inserción del donante depende de NHEJ a través de la captura final. La Figura 5B muestra los niveles circulantes de hFIX después de la administración de ZFN dirigidas a un locus de la albúmina de ratón endógeno ("mAlb ZFN") o al locus del Factor IX humano ("hF9 ZFN") y un transgén donante de hF9 ("Donante") a ratones de tipo silvestre. Los vectores de AAV y las cantidades administradas fueron las siguientes: AAV2/8-ZFN a  $1 \times 10^{11}$  vg/ratón y AAV2/8-donante a  $5 \times 10^{11}$  vg/ratón para mAlb ZFN y donante de F.IX y AAV2/8-ZFN a  $1 \times 10^{11}$  vg/ratón y AAV2/8-Donante a  $5 \times 10^{11}$  vg/ratón para hF9 ZFN y el donante de F.IX. Obsérvese que la hF9 ZFN no escinde el locus de F9 endógeno de ratón.

La **Figura 6** es un gráfico que representa niveles de hF.IX tras la administración de ZFN dirigidas a la albúmina y el donante de hF.IX a ratones. Se administraron AAV2/8-ZFN a la dosis indicada y el donante AAV2/8 a 5x la dosis de las ZFN. La edición del genoma es proporcional a la dosis de AAV en más de tres órdenes de magnitud.

Las **Figuras 7A y 7B** son gráficos que representan tiempos de coagulación en ratones HB tratados con ZFN y donantes de hF.IX. La Figura 7A muestra niveles plasmáticos de hF.IX en los animales indicados y la Figura 7B muestra el tiempo de tromboplastina parcial activado (aPTT(s)). Las dosis de AAV se muestran en la parte inferior.

Las **Figuras 8A a 8D** son gráficos que muestran los datos de ELISPOT en el día 65 post-administración de las ZFN en primates no humanos que reciben solo ZFN dirigidas a la albúmina. "LN" se refiere a ganglio linfático. Tal como se muestra, no existe ninguna respuesta inmunitaria contra la cápside de AAV8 o las ZFN.

La **Figura 9** es un gráfico que muestra el direccionamiento al locus de la albúmina mediado por nucleasa en hepatocitos humanos primarios. Los hepatocitos humanos primarios se transdujeron *in vitro* con el donante AAV2/6 hF9 (MOI  $9 \times 10^5$  vg/célula) y 24 horas más tarde con 500 ng del ARNm de hALB ZFN. Panel inferior: % Índels medido mediante el análisis MiSeq. Los sobrenadantes recogidos el día 5 (barra del extremo más a la izquierda de cada grupo), 7 (barra del medio de cada grupo) y 9 (barra del extremo más a la derecha de cada grupo) se analizaron para determinar los niveles de proteína hFIX mediante ELISA. Barras de error= e.e.m. Los datos son representativos de al menos 2 experimentos independientes.

Las **Figuras 10A a 10D** muestran la caracterización de mAlb ZFN *in vitro* e *in vivo*. La Figura 10A muestra la actividad de ZFN medida por índels en células Hepa 1-6 transfectadas con la cantidad indicada de ARNm de ZFN o GFP. Se aisló el ADN genómico y la secuencia diana se amplificó por PCR para la secuenciación con el sistema Illumina MiSeq. Los porcentajes indican lecturas que contienen inserciones y/o deleciones concordantes con la escisión y la reparación de NHEJ. La Figura 10B muestra que los niveles de hF.IX en ratones tratados permanecieron estables durante más de un año después de la inyección IV con  $5 \times 10^{11}$  vg de donante AAV8-hF9 y  $1 \times 10^{10}$  vg de AAV8-mALB-ZFN. La Figura 10C muestra que los valores de ALT en plasma después del tratamiento no se desviaron del intervalo normal (área sombreada). La Figura 10D muestra el uso de PCR cuantitativa para determinar la abundancia relativa del ARNm de albúmina "F9-mAlb híbrido" frente al de tipo silvestre. Se inyectó a los ratones una proporción 1:1 de ZFN: Donante a las dosis indicadas. El ARN total se aisló de hígados de ratones 2 semanas después de la inyección. Como control negativo, se administró luciferasa (simulada) + donante a la dosis más alta de  $5 \times 10^{11}$  vg de cada uno. Se usó una prueba de Mann-Whitney de 2 colas para comparar 2 grupos.  $n = 6-8$  ratones / grupo. Barras de error= e.e.m. \*\* $P < 0,01$  frente a simulado.

La **Figura 11** es un esquema que muestra el diseño del estudio para un estudio de primates no humanos (macaco Rhesus) de inserción mediada por nucleasas de transgenes F.IX. "ALB ZFN" se refiere a ZFN dirigidas a la albúmina como se describe en la publicación de Estados Unidos N° 20130177983. "FokI-eHi / fi" se refiere a dominios de escisión de FokI modificados que forman heterodímeros obligados tal como se usan en las ZFN. Véase, p. ej., patentes de Estados Unidos números 7.914.796, 8.034.598; publicaciones de Estados Unidos números 20110201055 y 20120142062. Los vectores de AAV y serotipos se usaron como se muestra en la dosis alta (ZFN único:  $1,5 \times 10^{11}$  vg/kg; Donante:  $1,5 \times 10^{14}$  vg/kg) o dosis baja (ZFN único:  $5 \times 10^{12}$  vg/kg; Donante:  $5 \times 10^{13}$  vg/kg). Las **Figuras 12A y 12B** muestran secuencias de aminoácidos de ejemplos de ZFN SBS N° 30724 y SBS N° 30725, respectivamente, utilizadas para dirigirse al intrón 1 del locus de la albúmina de ratón. La Figura 12A ZFN<sup>Izquierda</sup> (SEQ ID NO:115) y la Figura 12B ZFN<sup>Derecha</sup> (SEQ ID NO:116) muestran las secuencias de aminoácidos usadas en el estudio. La etiqueta 3xFLAG está indicada en cursiva. La secuencia de localización nuclear del antígeno T grande del SV40 se indica subrayada. El dominio FokI se indica en negrita. Las regiones de la hélice de reconocimiento se muestran en doble subrayado y en cursiva.

Las **Figuras 13A a 13D** muestran la expresión de proteínas deficientes en enfermedades de almacenamiento lisosomal y la presencia de actividad enzimática en el sobrenadante de las células después de la modificación con ZFN específica de la albúmina y un donante. La Figura 13A muestra la expresión enzimática de la proteína alfa-L iduronidasa (IDUA) a través del análisis de transferencia Western (Figura 13B) y la presencia de actividad enzimática IDUA en el sobrenadante de las células. Se tomaron muestras de los cultivos en los días 3 y 6 después de la transfección y en dosis bajas o altas de ZFN y donante de IDUA. Las Figuras 13C y 13D muestran un conjunto similar de datos que miden la presencia de proteína iduronato-2-sulfatasa (IDS) deficiente y actividad enzimática después de la transfección de ZFN y un donante de IDS.

## Descripción detallada

En el presente documento se divulgan composiciones y métodos para modificar una célula para producir una o más proteínas cuya expresión o secuencia génica, antes de la modificación, es aberrante y está asociada a una enfermedad o trastorno, por ejemplo, una hemofilia o una enfermedad de almacenamiento lisosomal (LSD). La célula se modifica mediante la inserción dirigida de un transgén que codifica una o más proteínas funcionales en un gen de sitio seguro (p. ej., albúmina) de la célula. En algunas realizaciones, el transgén se inserta en un gen de la albúmina endógeno. El transgén puede codificar cualquier proteína o péptido implicado en la hemofilia, por ejemplo Factor VII, F8, F.IX, Factor X, GBA, GLA, IDS, IDUA, SMPD1 y/o fragmentos funcionales de los mismos. También se divulgan métodos de tratamiento de un trastorno en el cual existe una o más proteínas inexistentes o deficientes (p. ej., una hemofilia o una enfermedad de almacenamiento lisosomal) usando una célula como se ha descrito en el presente documento y/o modificando una célula (*ex vivo* o *in vivo*) como se describe en el presente documento. Se describen además composiciones que comprenden ácidos nucleicos que codifican nucleasas y moléculas donantes para modificar una célula y métodos para modificar la célula *in vivo* o *ex vivo*. Adicionalmente, se describen composiciones que comprenden células que han sido modificadas por los métodos y composiciones de la invención.

Las células genómicamente modificadas descritas en el presente documento se modifican normalmente mediante la integración dirigida mediada por nucleasas (ZFN, TALEN y/o CRISPR/Cas) para insertar una secuencia que codifica una proteína terapéutica (p.ej., Factor VII, Factor VIII (F8), Factor IX, Factor X, glucocerebrosidasa,  $\alpha$  galactosidasa, iduronato-2-sulfatasa (IDS), alfa-L iduronidasa (IDUA) y/o esfingomielina 1), en donde la proteína, cuyo gen en un estado alterado o aberrante, está asociado a una enfermedad, en el genoma de una o más células del sujeto (*in vivo* o *ex vivo*), de modo que las células producen la proteína *in vivo*. En ciertas partes de la divulgación, los métodos pueden comprender además inducir las células del sujeto, particularmente células hepáticas, para proliferar (entrar en el ciclo celular), por ejemplo, mediante hepatectomía parcial y/o mediante administración de uno o más compuestos que inducen a las células hepáticas a entrar en el ciclo celular. Los sujetos incluyen, pero sin limitación, seres humanos, primates no humanos, animales domésticos tales como gatos, perros, conejos, ratas, ratones, cobayas, vacas, cerdos, caballos, cabras y similares.

## General

La puesta en práctica de los métodos, así como la preparación y el uso de las composiciones divulgadas en el presente documento emplean, a menos que se indique otra cosa, técnicas convencionales en biología molecular, bioquímica, estructura y análisis de la cromatina, química computacional, cultivo de células, ADN recombinante y campos relacionados como se conocen dentro del campo de la técnica. Estas técnicas se explican detalladamente en la bibliografía. Véase, por ejemplo, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edición, Cold Spring Harbor Laboratory Press, 1989 y Tercera edición, 2001; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, 1987 y actualizaciones periódicas; la serie METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Tercera edición, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; y METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

## Definiciones

La expresión "ácido nucleico", y los términos "polinucleótido", y "oligonucleótido" se usan indistintamente y se refieren a un polímero desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma de cadena sencilla o doble. Para los fines de la presente divulgación, estos términos no deben interpretarse como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que se modifican en los restos de la base, el azúcar y/o el fosfato (p.ej., cadenas principales de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo de A se emparejará con T.

Los términos "polipéptido", "péptido" y "proteína" se utilizan indistintamente en el presente documento para referirse a un polímero de restos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más restos de aminoácido son análogos químicos o derivados modificados de los correspondientes aminoácidos de origen natural.

"Unión" se refiere a la interacción específica de la secuencia, no covalente entre macromoléculas (p.ej., entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión tienen que ser específicos de la secuencia (p.ej., los contactos con los restos de fosfato en una cadena principal de ADN), siempre que la interacción en su totalidad sea específica de la secuencia. Dichas interacciones se caracterizan generalmente por una constante de disociación ( $K_d$ ) de  $10^{-6}$  M<sup>-1</sup> o menos. La "afinidad" se refiere a la fuerza de la unión: correlacionándose una afinidad de unión aumentada con una  $K_D$  menor.

Una "proteína de unión" es una proteína que es capaz de unirse no covalentemente a otra molécula. Una proteína de

unión puede unirse a, por ejemplo, una molécula de ADN (una proteína de unión al ADN), una molécula de ARN (una proteína de unión a ARN) y/o una molécula de proteína (una proteína de unión a proteína). En el caso de una proteína de unión a proteína, esta se puede unir consigo mismo (para formar homodímeros, homotrímeros, etc.) y/o se puede unir a una o más moléculas de una proteína o proteínas diferentes. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas de dedo de zinc tienen actividad de unión al ADN, unión al ARN y unión a proteínas.

Una "proteína de unión al ADN de dedo de zinc" (o dominio de unión) es una proteína, o un dominio con una proteína más grande, que se une al ADN de forma específica de la secuencia a través de uno o más dedos de zinc, que son regiones de la secuencia de aminoácidos dentro del dominio de unión cuya estructura se estabiliza a través de la coordinación de un ion de zinc. La expresión proteína de unión al ADN de dedo de zinc frecuentemente se abrevia como proteína de dedo de zinc o ZFP.

Un "dominio de unión al ADN de TALE" o "TALE" es un polipéptido que comprende uno o más de dominios/unidades repetidas TALE. Los dominios repetidos están implicados en la unión de TALE a su secuencia de ADN diana afin. Una "unidad de repetición" (también denominada como una "repetición") tiene normalmente de 33-35 aminoácidos de longitud y exhibe al menos homología de secuencia con otras secuencias de repetición TALE dentro de una proteína TALE de origen natural. Véase, *p. ej.*, la patente de Estados Unidos N° 8.586.526.

Los dominios de unión de dedos de zinc y TALE se pueden "modificar" para unirse a una secuencia de nucleótidos predeterminada, por ejemplo mediante ingeniería genética (alterando uno o más aminoácidos) de la región de la hélice de reconocimiento de un dedo de zinc o proteína TALE de origen natural. Por lo tanto, las proteínas de unión al ADN modificadas (dedos de zinc o TALE) son proteínas que no son de origen natural. Ejemplos no limitativos de métodos para modificar proteínas de unión al ADN son el diseño y la selección. Una proteína de unión al ADN modificada es una proteína que no existe en la naturaleza y cuyo diseño/composición es el resultado principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y algoritmos computarizados para procesar información en una base de datos que almacena información de diseños ZFP y/o TALE existentes y datos de unión. Véase, por ejemplo, las patentes de Estados Unidos 6.140.081; 6.453.242; 6.534.261; y 8.586.526, véase también los documentos WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

Una proteína de dedo de zinc o TALE "seleccionada" es una proteína que no existe en la naturaleza cuya producción resulta principalmente de un proceso empírico, tal como la presentación en fagos, trampa de interacción o selección de híbridos. Véanse *p. ej.*, los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970, WO 01/88197 y WO 02/099084.

"TtAgo" es una proteína Argonauta procariota que se cree que está implicada en el silenciamiento génico. TtAgo deriva de la bacteria *Thermus thermophilus*. Véase, *p. ej.*, Swarts *et al*, *ibid*, G. Sheng *et al.*, (2013) Proc. Natl. Acad. Sci. U.S.A. 111, 652). Un "sistema TtAgo" tiene todos los componentes requeridos incluyendo, por ejemplo, ADN guías para la escisión por una enzima TtAgo.

"Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos, que incluyen, pero sin limitación, captura de donantes por unión de extremos no homólogos (NHEJ) y recombinación homóloga. Para los fines de esta divulgación, "recombinación homóloga (HR)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble cadena en las células a través de mecanismos de reparación dirigidos por homología. Este proceso requiere una homología de secuencia de nucleótidos, utiliza una molécula "donante" para reparar el molde de una molécula "diana" (es decir, la que experimentó la rotura de doble cadena), y es conocido de diversas formas como "conversión génica sin entrecruzamiento" o "conversión génica de tramo corto", porque conduce a la transferencia de información genética del donante a la diana. Sin desear quedar ligado a ninguna teoría en particular, dicha transferencia puede implicar una corrección de emparejamiento incorrecto del ADN heterodúplex que se forma entre la diana rota y el donante, y/o la "hibridación dependiente de la síntesis", en la cual se usa el donante para volver a sintetizar información genética, que se convertirá en parte de la diana y/o procesos relacionados. Dicha HR especializada a menudo da como resultado una alteración de la secuencia de la molécula diana de tal manera que parte o toda la secuencia del polinucleótido donante se incorpora al polinucleótido diana.

En los métodos de la divulgación, se pueden introducir en la célula una o más nucleasas dirigidas como se describe en el presente documento creando una rotura de la doble cadena en la secuencia diana (*p. ej.*, cromatina celular en un sitio predeterminado, y un polinucleótido "donante", que tiene homología con la secuencia de nucleótidos en la región de la rotura. Se ha demostrado que la presencia de la rotura de la doble cadena facilita la integración de la secuencia del donante. La secuencia del donante puede estar físicamente integrada o, como alternativa, el polinucleótido del donante se usa como molde para reparar la rotura mediante recombinación homóloga, lo que da como resultado la introducción de toda o parte de la secuencia de nucleótidos como en el donante en la cromatina celular. Por lo tanto, una primera secuencia en la cromatina celular puede alterarse y, en determinadas realizaciones, puede convertirse en una secuencia presente en un polinucleótido donante. Por lo tanto, el uso de los términos "reemplazar" o

"reemplazo" representa el reemplazo de una secuencia de nucleótidos por otra, (es decir., el reemplazo de una secuencia en el sentido de información), y no necesariamente requiere reemplazo físico o químico de un polinucleótido por otro.

- 5 En cualquiera de los métodos descritos en el presente documento, se pueden usar parejas adicionales de proteínas con dedos de zinc o TALEN para la escisión adicional de doble cadena de sitios diana adicionales dentro de la célula.

10 En ciertas realizaciones de métodos para la recombinación dirigida y/o reemplazo y/o alteración de una secuencia en una región de interés en la cromatina celular, una secuencia cromosómica se altera por recombinación homóloga con una secuencia de nucleótidos "donante" exógena. Dicha recombinación homóloga es estimulada por la presencia de una rotura de la doble cadena en la cromatina celular, si están presentes secuencias homólogas a la región de la rotura.

15 En cualquiera de los métodos descritos en el presente documento, la primera secuencia de nucleótidos (la "secuencia donante") puede contener secuencias que son homólogas, pero no idénticas, a las secuencias genómicas en la región de interés, estimulando así la recombinación homóloga para insertar una secuencia no idéntica en la región de interés. Por lo tanto, las porciones de la secuencia donante que son homólogas a las secuencias de la región de interés exhiben entre aproximadamente el 80 y el 99 % (o cualquier número entero entre estos) de identidad de secuencia con la secuencia genómica que se reemplaza. En otras situaciones, la homología entre el donante y la secuencia genómica es superior al 99 %, por ejemplo si solo 1 nucleótido difiere entre el donante y las secuencias genómicas de más de 100 pares de bases contiguas. En determinados casos, una porción no homóloga de la secuencia donante puede contener secuencias que no están presentes en la región de interés, de modo que se introducen nuevas secuencias en la región de interés. En estos casos, la secuencia no homóloga generalmente está flanqueada por secuencias de 50-1.000 pares de bases (o cualquier valor entero entre ellas) o cualquier número de pares de bases mayor que 1.000, que son homólogas o idénticas a las secuencias en la región de interés. En otros casos, la secuencia donante no es homóloga a la primera secuencia, y se inserta en el genoma mediante mecanismos de recombinación no homóloga.

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

35 Asimismo, los métodos de integración dirigida como se describe en el presente documento también se pueden usar para integrar una o más secuencias exógenas. La secuencia de ácido nucleico exógeno puede comprender, por ejemplo, uno o más genes o moléculas de ADNc, o cualquier tipo de secuencia codificante o no codificante, así como uno o más elementos de control (p. ej., promotores). Además, la secuencia de ácido nucleico exógeno puede producir una o más moléculas de ARN (p.ej., ARN de horquilla corta (ARNhc), ARN inhibidores (ARNis), microARN (miARN), etc.).

40 "Escisión" se refiere a la rotura de la cadena principal de una molécula de ADN. La escisión puede iniciarse mediante varios métodos que incluyen, pero sin limitación, hidrólisis enzimática o química de un enlace fosfodiéster. Es posible la escisión de cadena sencilla y la escisión de cadena doble y la escisión de cadena doble puede producirse como resultado de dos eventos de escisión de cadena sencilla. La escisión del ADN puede dar como resultado la producción de extremos romos o extremos escalonados. En determinadas realizaciones, los polipéptidos de fusión se usan para la escisión dirigida del ADN de doble cadena.

50 Un "semidominio de escisión" es una secuencia de polipéptidos que, junto con un segundo polipéptido (ya sea idéntico o diferente) forma un complejo que tiene actividad de escisión (preferentemente actividad de escisión de doble cadena). Las expresiones los términos "primer y segundo semidominios de escisión"; "semidominios de escisión + y -" y "semidominios de escisión derecho e izquierdo" se usan indistintamente para referirse a parejas de semidominios de escisión que se dimerizan.

55 Un "semidominio de escisión modificado" es un semidominio de escisión que se ha modificado para formar heterodímeros obligados con otro semidominio de escisión (p.ej., otro semidominio de escisión modificado). Véase, también, las patentes de Estados Unidos números 7.914.796; 8.034.598; y 8.623.618.

60 El término "secuencia" se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser ADN o ARN; puede ser lineal, circular o ramificado y puede ser de cadena sencilla o de doble cadena. El término "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede tener cualquier longitud, por ejemplo entre 2 y 10.000 nucleótidos de longitud (o cualquier valor entero entre los mismos o superior a estos), preferentemente entre aproximadamente 100 y 1.000 nucleótidos de longitud (o cualquier número entero entre los mismos), más preferentemente entre aproximadamente 200 y 500 nucleótidos de longitud.

65 "Cromatina" es la estructura de nucleoproteína que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN y proteínas, incluidas las histonas y las proteínas cromosómicas no histonas. La

mayoría de la cromatina celular eucariota existe en forma de nucleosomas, en donde un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociados a un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y el ADN conector (de longitud variable dependiendo del organismo) se extiende entre núcleos de nucleosomas. Una molécula de histona HI generalmente se asocia con el ADN conector. Para los fines de la presente divulgación, el término "cromatina" pretende abarcar todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye cromatina cromosómica y episomal.

Un "cromosoma", es un complejo de cromatina que comprende toda o parte del genoma de una célula. El genoma de una célula a menudo se caracteriza por su cariotipo, que es la colección de todos los cromosomas que comprenden el genoma de la célula. El genoma de una célula puede comprender uno o más cromosomas.

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

Un "sitio diana" o "secuencia diana" es una secuencia de ácido nucleico que define una porción de un ácido nucleico a la que se unirá una molécula de unión, siempre que existan condiciones suficientes para la unión.

Una molécula "exógena" es una molécula que normalmente no está presente en una célula, pero puede introducirse en una célula mediante uno o más métodos genéticos, bioquímicos u otros. La "presencia normal en la célula" se determina con respecto al estadio de desarrollo particular y condiciones ambientales de la célula. Por lo tanto, por ejemplo, una molécula que está presente solo durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular adulta. De manera similar, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula no sometida a choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión funcional de una molécula endógena que funciona mal o una versión que funciona mal de una molécula endógena que funciona normalmente.

Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, como la que se genera mediante un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, hidrato de carbono, lípidos, glucoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprenda una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser de cadena sencilla o doble; pueden ser lineales, ramificados o circulares; y pueden ser de cualquier longitud. Los ácidos nucleicos incluyen aquellos capaces de formar dúplex, así como los ácidos nucleicos formadores de tríplex. Véase, por ejemplo, las patentes de Estados Unidos números 5.176.996 y 5.422.251. Las proteínas incluyen, pero sin limitación, proteínas de unión al ADN, factores de transcripción, factores de remodelación de la cromatina, proteínas de unión al ADN metiladas, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, cinasas, fosfatasa, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.

Una molécula exógena puede ser del mismo tipo de molécula que una molécula endógena, p. ej., una proteína o ácido nucleico exógenos. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma viral infectante, un plásmido o episoma introducido en una célula, o un cromosoma que está normalmente presente en la célula. Los expertos en la materia conocen los métodos para la introducción de moléculas exógenas en las células e incluyen, pero sin limitación, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato cálcico, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores virales. Una molécula exógena también puede ser del mismo tipo de molécula que una molécula endógena pero derivada de una especie diferente de la que se deriva la célula. Por ejemplo, una secuencia de ácido nucleico humano puede introducirse en una línea celular derivada originalmente de un ratón o hámster.

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

Una molécula de "fusión" es una molécula en donde dos o más moléculas de subunidad están unidas, preferentemente de forma covalente. Las moléculas de subunidad pueden ser del mismo tipo químico de molécula, o pueden ser diferentes tipos químicos de moléculas. Los ejemplos del primer tipo de molécula de fusión incluyen, pero sin limitación, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión al ADN de ZFP o TALE y uno o más dominios de activación) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita *supra*). Los ejemplos del segundo tipo de molécula de fusión incluyen, pero sin limitación, una fusión entre un ácido nucleico formador de tríplex y un polipéptido, y una fusión entre un ligando de unión al surco menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede ser el resultado la administración de la proteína de fusión a la célula o de la administración de un polinucleótido que codifica la proteína de fusión a una célula, en donde el polinucleótido se transcribe y el transcrito se traduce, para generar la proteína de fusión. El corte y empalme en trans,

la escisión de polipéptidos y la ligadura de polipéptidos también pueden estar implicados en la expresión de una proteína en una célula. Los métodos para la administración de polinucleótidos y polipéptidos a las células se presentan en otra parte de esta divulgación.

- 5 Un "gen", para los fines de la presente divulgación, incluye una región de ADN que codifica un producto génico (véase *infra*), así como todas las regiones de ADN que regulan la producción del producto génico, independientemente de que tales secuencias reguladoras sean o no adyacentes a las secuencias de codificación y/o transcritas. En consecuencia, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias reguladoras de la traducción, como los sitios de unión al ribosoma y los sitios internos de entrada al ribosoma, potenciadores, silenciadores, aisladores, aisladores de la cromatina, orígenes de replicación, sitios de unión de matriz y regiones de control de locus.

15 "Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto transcripcional directo de un gen (*p.ej.*, ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por la traducción de un ARNm. Los productos génicos también incluyen ARN que se modifican, mediante procesos como la adición de una caperuza, poliadenilación, metilación y edición, y proteínas modificadas mediante, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristilación y glicosilación.

20 La "modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitación, activación génica y represión génica. Se puede usar la edición del genoma (*p.ej.*, escisión, alteración, inactivación, mutación aleatoria) para modular la expresión. La inactivación génica se refiere a cualquier reducción en la expresión génica en comparación con una célula que no incluye una ZFP como se describe en el presente documento. Por lo tanto, la inactivación génica puede ser parcial o completa.

25 Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante dentro de o adyacente a un gen, en la cual es deseable unir una molécula exógena. La unión puede ser para fines de escisión dirigida del ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma de orgánulos (*p.ej.*, mitocondrial, de cloroplastos), o un genoma viral infectante, por ejemplo. Una región de interés puede estar presente en la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias líderes, secuencia de remolque o intrones, o dentro de regiones no transcritas, en dirección 5' o 3' de la región codificante. Una región de interés puede ser tan pequeña como un solo par de nucleótidos o tener hasta 2.000 pares de nucleótidos de longitud, o cualquier valor integral de pares de nucleótidos.

35 Las células "eucariotas" incluyen, pero sin limitación, células fúngicas, (tales como células de levadura), células vegetales, células animales, células de mamífero y células humanas (*p. ej.*, linfocitos T).

40 Las expresiones "enlace operativo" y "unido operativamente" (o "unido de forma operativa") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en donde los componentes están dispuestos de manera que ambos componentes funcionen normalmente y permite la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está operativamente unida a una secuencia codificante si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción generalmente está operativamente unida en *cis* a una secuencia codificante, pero no necesita estar directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está operativamente unida a una secuencia codificante, aunque no sean contiguas.

50 Con respecto a los polipéptidos de fusión, la expresión "unido operativamente" puede referirse al hecho de que cada uno de los componentes realiza la misma función estando unido con el otro componente como lo haría si no estuviera tan unido. Por ejemplo, con respecto a un polipéptido de fusión en donde un dominio de unión al ADN de ZFP se fusiona con un dominio de activación, el dominio de unión al ADN de ZFP y el dominio de activación están en unión operativa si, en el péptido de fusión, la porción del dominio de unión al ADN de ZFP puede unirse al sitio diana y/o su sitio de unión, mientras que el dominio de activación puede regular positivamente la expresión génica. Cuando un polipéptido de fusión en donde un dominio de unión al ADN de ZFP se fusiona con un dominio de escisión, el dominio de unión al ADN de ZFP y el dominio de escisión están en unión operativa si, en el péptido de fusión, la porción del dominio de unión al ADN de ZFP puede unirse al sitio diana y/o su sitio de unión, mientras que el dominio de escisión es capaz de escindir el ADN en las proximidades del sitio diana.

65 Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la de la proteína, polipéptido o ácido nucleico de longitud completa, aunque todavía mantiene la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede poseer más, menos, o el mismo número de restos que la molécula nativa correspondiente, y/o puede mantener uno o más aminoácidos o sustituciones de nucleótidos. Los métodos para determinar la función de un ácido nucleico

- (p. ej., función de la codificación, capacidad para hibridar con otro ácido nucleico) son bien conocidos en la técnica. De manera similar, los métodos para determinar la función de las proteínas son bien conocidos. Por ejemplo, la función de unión al ADN de un polipéptido se puede determinar, por ejemplo, por unión al filtro, cambio de movilidad electroforética o ensayos de inmunoprecipitación. La escisión del ADN se puede analizar mediante electroforesis en gel. Véase Ausubel *et al.*, *supra*. La capacidad de una proteína para interactuar con otra proteína puede determinarse, por ejemplo, mediante inmunoprecipitación conjunta, ensayos de dos híbridos o complementación, tanto genéticos como bioquímicos. Véase, por ejemplo, Fields *et al.* (1989) *Nature*340:245-246; patente de Estados Unidos N° 5.585.245 y PCT WO 98/44350.
- 10 Un "vector" es capaz de transferir secuencias génicas a células diana. Normalmente, "construcción de vectores", "vector de expresión", y "vector de transferencia génica", significan cualquier construcción de ácido nucleico capaz de dirigir la expresión de un gen de interés y que puede transferir secuencias de genes a células diana. Por lo tanto, la expresión incluye vehículos de clonación y expresión, así como vectores de integración.
- 15 Un locus de "sitio seguro" es un locus dentro del genoma en donde se puede insertar un gen sin ningún efecto nocivo sobre la célula hospedadora. Lo más beneficioso es un locus seguro en donde la expresión de la secuencia del gen insertado no se vea perturbada por ninguna expresión de ultralectura de genes vecinos. Los ejemplos no limitativos de loci de sitio seguro que son diana de nucleasa(s) incluyen CCR5, HPRT, AAVS1, Rosa y albúmina. Véase, p.ej., las patentes de Estados Unidos números 7.951.925 y 8.110.379; las publicaciones de patente Estados Unidos números 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y la solicitud provisional de Estados Unidos N° 61/823.689)

### Nucleasas

- 25 En el presente documento se describen composiciones, particularmente nucleasas que son útiles en la integración de una secuencia que codifica una proteína funcional que es inexistente, deficiente o que se expresa aberrantemente en un sujeto con una enfermedad o trastorno (p. ej., una proteína que es inexistente o deficiente en un sujeto con un LSD y/o proteína de factor de coagulación (p.ej., F8 y/o F.IX) en el genoma de una célula o en un sujeto con hemofilia A o B). La nucleasa es de origen no natural, es decir, está modificada en el dominio de unión al ADN y/o el dominio de escisión. Por ejemplo, el dominio de unión al ADN de una nucleasa natural puede alterarse para unirse a un sitio diana seleccionado (p.ej., una meganucleasa que ha sido modificada para unirse a un sitio diferente al sitio de unión afín). En otras realizaciones, la nucleasa comprende dominios de unión y escisión del ADN heterólogo (p.ej., nucleasas de dedo de zinc; proteínas de unión al ADN del dominio efector TAL; dominios de unión al ADN de meganucleasa con dominios de escisión heterólogos) y/o un sistema CRISPR/Cas que utiliza un ARN de guía único modificado).

#### A. Dominios de unión al ADN

- Se puede usar cualquier dominio de unión al ADN en las composiciones y métodos divulgados en el presente documento, incluyendo aunque no de forma limitativa un dominio de unión al ADN de dedo de zinc, un dominio de unión al ADN de TALE, la porción de unión al ADN de una nucleasa CRISPR/Cas o un dominio de unión al ADN de una meganucleasa.

- Como se divulga en el presente documento, la nucleasa es una meganucleasa (endonucleasa orientadora) de origen natural o modificada (no natural). Los ejemplos de endonucleasas orientadoras incluyen *I-SceI*, *I-CeuI*, *PI-PspI*, *PI-Sce*, *I-SceIV*, *I-Csml*, *I-PanI*, *I-Scell*, *I-Ppol*, *I-ScellIII*, *I-Crel*, *I-TevI*, *I-TevII* y *I-TevIII*. Se conocen sus secuencias de reconocimiento. Véase también la patente de Estados Unidos N° 5.420.032; la patente de Estados Unidos N° 6.833.252; Belfort *et al.*(1997) *Nucleic Acids Res.*25:3379-3388; Dujon *et al.* (1989) *Gene* 82:115-118; Perler *et al.*(1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble *et al.* (1996) *J. Mol. Biol.*263:163-180; Argast *et al.* (1998) *J. Mol. Biol.* 280:345-353 y el catálogo de New England Biolabs. Las meganucleasas modificadas se describen por ejemplo en la publicación de patente de Estados Unidos N° 20070117128. Los dominios de unión al ADN de las endonucleasas y meganucleasas orientadoras pueden alterarse en el contexto de la nucleasa en su conjunto (es decir, de modo que la nucleasa incluya el dominio de escisión relacionado) o puede fusionarse a un dominio de escisión heterólogo. Los dominios de unión al ADN de meganucleasas también pueden exhibir actividad nucleasa (p. ej., cTALEN).

- También se divulga en el presente documento, que el dominio de unión al ADN comprende un dominio de unión al ADN del efector TAL natural o modificado (no natural). Véase, p. ej., la patente de Estados Unidos N° 8.586.526. Las bacterias patógenas de plantas del género *Xanthomonas* son conocidas por provocar muchas enfermedades en plantas de cultivo importantes. La patogenicidad de *Xanthomonas* depende de un sistema conservado de secreción de tipo III (T3S) que inyecta más de 25 proteínas efectoras diferentes en la célula vegetal. Entre estas proteínas inyectadas se encuentran los efectores similares a los activadores de la transcripción (TAL) que imitan a los activadores transcripcionales de la planta y manipulan el transcriptoma de la planta (véase Kay *et al.* (2007) *Science* 318:648-651). Estas proteínas contienen un dominio de unión al ADN y un dominio de activación de la transcripción. Uno de los efectores TAL mejor caracterizados es AvrBs3 de *Xanthomonas campestris* pv. *Vesicatoria* (véase Bonaset *al.* (1989) *Mol Gen Genet* 218: 127-136 y WO2010079430). Los efectores TAL contienen un dominio centralizado de repeticiones en tándem, conteniendo cada repetición aproximadamente 34 aminoácidos, los cuales

son clave para la especificidad de unión al ADN de estas proteínas. Además, contienen una secuencia de localización nuclear y un dominio de activación de la transcripción ácido (para una revisión véase Schornack S, *et al* (2006) J Plant Physiol 163(3): 256-272). Además, en la bacteria fitopatogena *Ralstonia solanacearum* se han descubierto dos genes, designados brg11 y hpx17 que son homólogos con la familia AvrBs3 de *Xanthomonas* en la cepa GMI1000 de *R. solanacearum* biovar 1 y en la cepa RS1000 biovar 4 (Véase Heuer *et al* (2007) Appl and Envir Micro 73(13): 4379-4384). Estos genes son 98,9 % idénticos en la secuencia de nucleótidos entre sí, pero difieren en una delección de 1.575 pares de bases en el dominio repetido de hpx17. Sin embargo, ambos productos génicos tienen menos del 40 % de identidad de secuencia con las proteínas de la familia AvrBs3 de *Xanthomonas*. Véase, p. ej., la patente de Estados Unidos N° 8.586.526.

La especificidad de estos efectores TAL depende de las secuencias encontradas en las repeticiones en tándem. La secuencia repetida comprende aproximadamente 102 pares de bases y las repeticiones son normalmente 91-100% homólogas entre sí (Bonas *et al, ibid*). El polimorfismo de las repeticiones generalmente se encuentra en las posiciones 12 y 13 y parece haber una correspondencia biunívoca entre la identidad de dos restos hipervariables (dos restos variables repetidos o región RVD) en las posiciones 12 y 13 con la identidad de los nucleótidos contiguos en la secuencia diana del efector TAL (véase Moscou y Bogdanove, (2009) Science 326:1501 y Boch *et al* (2009) Science 326:1509-1512). Experimentalmente, el código natural para el reconocimiento del ADN de estos efectores TAL se ha determinado de manera que una secuencia de HD en las posiciones 12 y 13 (repetición de dos restos variables o RVD) conduce a una unión a la citosina (C), NG se une a T, NI a A, C, G o T, NN se une a A o G, e ING se une a T. Estas repeticiones de unión al ADN se han ensamblado en proteínas con nuevas combinaciones y números de repeticiones, para crear factores de transcripción artificiales que pueden interactuar con nuevas secuencias y activar la expresión de un gen indicador no endógeno en células vegetales (Boch *et al, ibid*). Las proteínas TAL modificadas se han relacionado con un semidominio de escisión *FokI* para producir una fusión de nucleasa con el dominio efector TAL (TALEN), incluyendo TALEN con RVD atípicas. Véase, p. ej., la patente de Estados Unidos N° 8.586.526.

En algunas partes de la divulgación, la TALEN comprende un dominio de escisión o un semidominio de escisión de endonucleasa (por ejemplo, *FokI*). En otras realizaciones, la nucleasa TALE es una mega TAL. Estas meganucleasas TAL son proteínas de fusión que comprenden un dominio de unión al ADN de TALE y un dominio de escisión de meganucleasa. El dominio de escisión de meganucleasa está activo como un monómero y no requiere dimerización para la actividad. (Véase Boissel *et al.*, (2013) Nucl Acid Res: 1-13, doi: 10.1093/nar/gkt1224).

También se divulga en el presente documento, la nucleasa que comprende una TALEN compacta. Estas son proteínas de fusión de cadena sencilla que unen un dominio de unión al ADN de TALE a un dominio de nucleasa *TevI*. La proteína de fusión puede actuar como una nickasa localizada por la región TALE, o puede crear una rotura de doble cadena, dependiendo de dónde se encuentre el dominio de unión al ADN de TALE con respecto al dominio de nucleasa *TevI* (véase Beurdeley *et al* (2013) Nat Comm: 1-8 DOI: 10.1038/ncomms2782). Además, el dominio de nucleasa también puede exhibir una funcionalidad de unión al ADN. Se puede usar cualquier TALEN en combinación con TALEN adicionales (p.ej., una o más TALEN (cTALEN o *FokI*-TALEN) con uno o más mega-TALE).

En ciertas partes de la divulgación, el dominio de unión al ADN comprende una proteína de dedo de zinc. Preferentemente, la proteína de dedo de zinc es una proteína de origen no natural en el sentido de que está modificada para unirse a un sitio diana de elección. Véase, por ejemplo, Beerli *et al.* (2002) Nature Biotechnol. 20:135-141; Pabo *et al.* (2001) Ann. Rev. Biochem.70:313-340; Isalan *et al.* (2001) Nature Biotechnol. 19:656-660; Segal *et al.* (2001) Curr. Opin. Biotechnol. 12:632-637; Choo *et al.* (2000) Curr. Opin. Struct. Biol. 10:411-416; patentes de Estados Unidos números 6.453.242; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.030.215; 6.794.136; 7.067.317; 7.262.054; 7.070.934; 7.361.635; 7.253.273; y las publicaciones de patente de Estados Unidos números 2005/0064474; 2007/0218528; 2005/0267061.

Un dominio de unión de dedo de zinc modificado puede tener una especificidad de unión nueva, en comparación con una proteína de dedo de zinc de origen natural. Los métodos de modificación incluyen, pero sin limitación, un diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, la utilización de bases de datos que comprenden secuencias de nucleótidos de triplete (o cuadruplete) y secuencias de aminoácidos de dedo de zinc individuales, en las cuales cada secuencia de nucleótidos de triplete o cuadruplete se asocia con una o más secuencias de aminoácidos de dedos de zinc que se unen a la secuencia triplete o cuadruplete particular. Véase, por ejemplo, en las patentes estadounidenses de cotitularidad números 6.453.242 y 6.534.261.

Los ejemplos de métodos de selección, incluyendo la presentación en fagos y los sistemas de dos híbridos, se divulgan en las patentes de Estados Unidos 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como en el documento WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, la mayor especificidad de unión para dominios de unión de dedos de zinc se ha descrito, por ejemplo, en el documento de cotitularidad WO 02/077227.

Además, como se divulga en estas y otras referencias, los dominios de dedos de zinc y/o las proteínas de dedos de zinc de múltiples dedos pueden unirse entre sí usando cualquier secuencia de conector adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véase, también, las patentes de Estados Unidos números 6.479.626; 6.903.185; y 7.153.949 para ejemplos de secuencias de conector de 6 o más aminoácidos de longitud. Las

proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína.

5 Selección de sitios diana; Las ZFP y los métodos para el diseño y la construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos por un experto en la materia y se describen en detalle en las patentes de Estados Unidos números 6.140.081; 5.789.538; 6.453.242; 6.534.261; 5.925.523; 6.007.988; 6.013.453; 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

10 Además, como se divulga en estas y otras referencias, los dominios de dedos de zinc y/o las proteínas de dedos de zinc de múltiples dedos pueden unirse entre sí usando cualquier secuencia de conector adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véase, también, las patentes de Estados Unidos números 6.479.626; 6.903.185; y 7.153.949 para ejemplos de secuencias de conector de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de zinc individuales de la proteína.

### B. Dominios de escisión

20 Cualquier dominio de escisión adecuado puede unirse operativamente a un dominio de unión al ADN para formar una nucleasa, tal como una nucleasa de dedo de zinc, una TALEN o un sistema CRISPR/nucleasa Cas. Véase, p. ej., las patentes de Estados Unidos números 7.951.925; 8.110.379 y 8.586.526; las publicaciones de patente Estados Unidos números 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 y 20130177960 y la solicitud provisional de Estados Unidos N° 61/823.689).

25 Tal como se ha indicado anteriormente, el dominio de escisión puede ser heterólogo respecto al dominio de unión al ADN, por ejemplo, un dominio de unión al ADN de dedo de zinc y un dominio de escisión de una nucleasa o un dominio de unión al ADN de TALEN y un dominio de escisión, o un dominio de unión al ADN de meganucleasa y un dominio de escisión de una nucleasa diferente. Los dominios de escisión heterólogos se pueden obtener de cualquier endonucleasa o exonucleasa. Los ejemplos de endonucleasas a partir de las cuales se puede derivar un dominio de escisión incluyen, pero sin limitación, endonucleasas de restricción y endonucleasas orientadoras. Véase, por ejemplo, el catálogo 2002-2003, New England Biolabs, Beverly, MA; y Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388. Se conocen otras enzimas que escinden el ADN (p. ej., nucleasa S1; nucleasa de judía mungo; ADNasa pancreática I; nucleasa microcócica; endonucleasa HO de levadura; véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o fragmentos funcionales de las mismas) pueden usarse como

30 fuente de dominios de escisión y semidominios de escisión.

De manera similar, un semidominio de escisión puede derivarse de cualquier nucleasa o porción de la misma, tal como se ha expuesto anteriormente, que requiera la dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Como alternativa,

40 se puede usar una única proteína que comprende dos semidominios de escisión. Los dos semidominios de escisión pueden derivarse de la misma endonucleasa (o fragmentos funcionales de la misma), o cada semidominio de escisión puede derivarse de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios diana para las dos proteínas de fusión están dispuestos preferentemente, uno con respecto al otro, de modo que la unión de las dos proteínas de fusión a sus sitios diana respectivos coloca los semidominios de escisión en una orientación espacial

45 entre sí que permite que los semidominios de escisión formen un dominio de escisión funcional, *p.ej.*, mediante dimerización. Por lo tanto, en ciertas partes de la divulgación, los bordes cercanos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número de nucleótidos o pares de nucleótidos puede intervenir entre dos sitios diana (por ejemplo, de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios diana.

50 Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unirse al ADN de secuencia específica (en un sitio de reconocimiento), y escindir el ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (p. ej., Tipo IIS) escinden el ADN en sitios eliminados del sitio de reconocimiento y tienen dominios de unión y escisión separables. Por ejemplo, la enzima Tipo IIS *Fok I* cataliza la escisión de doble cadena del ADN, a 9 nucleótidos de su sitio de reconocimiento en una cadena y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véase, por ejemplo, las patentes de Estados Unidos 5.356.802; 5.436.150 y 5.487.994; así como Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* 91:883-887; Kim *et al.* (1994b) *J. Biol. Chem.* 269:31.978-31.982. Por lo tanto, en una parte de la divulgación, las proteínas de fusión comprenden el dominio de escisión (o semidominio de escisión) de al menos una enzima de restricción de Tipo IIS y uno o más dominios de unión de dedos de zinc, que pueden o no estar modificados.

Un ejemplo de enzima de restricción de Tipo IIS, cuyo dominio de unión es separable del dominio de unión, es *Fok I*. Esta enzima particular es activa como un dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:10.570-10.575.

65 Por consiguiente, para los fines de la presente divulgación, la parte de la enzima *Fok I* utilizada en las proteínas de fusión divulgadas se considera un semidominio de escisión. Por lo tanto, para la escisión dirigida de la doble cadena

y/o el reemplazo dirigido de secuencias celulares usando fusiones de *FokI*-dedo de zinc, se pueden usar dos proteínas de fusión, comprendiendo cada una un semidominio de escisión de *FokI*, para reconstituir catalíticamente un dominio de escisión catalíticamente activo. Como alternativa, también se puede usar una única molécula de polipéptido que contiene un dominio de unión de dedo de zinc y dos semidominios de escisión de *FokI*. Los parámetros para la escisión dirigida y la alteración de la secuencia dirigida usando las fusiones dedo de zinc-*FokI* se proporcionan en otro lugar de esta divulgación.

Un dominio de escisión o semidominio de escisión puede ser cualquier parte de una proteína que retiene la actividad de escisión, o que retiene la capacidad de multimerizarse (por ejemplo, dimerizarse) para formar un dominio de escisión funcional.

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

En ciertas partes de la divulgación, el dominio de escisión comprende uno o más semidominios de escisión modificados (también denominados mutantes del dominio de dimerización) que minimizan o evitan la homodimerización, como se ha descrito, por ejemplo, en las publicaciones de patente de Estados Unidos números 20050064474; 20060188987; 20070305346 y 20080131962. Los restos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, y 538 de *FokI* son todos dianas para la influencia en la dimerización de los semidominios de escisión de *FokI*.

Los ejemplos de semidominios de escisión modificados de *FokI* que forman heterodímeros obligados incluyen un par en el cual un primer semidominio de escisión incluye mutaciones en los restos de aminoácidos en las posiciones 490 y 538 de *FokI* y un segundo semidominio de escisión incluye mutaciones en los restos de aminoácidos 486 y 499.

Por lo tanto, en una parte de la divulgación, una mutación en 490 reemplaza Glu (E) con Lys (K); la mutación en 538 reemplaza a Iso (I) con Lys (K); la mutación en 486 reemplaza Gln (Q) con Glu (E); y la mutación en la posición 499 reemplaza Iso (I) con Lys (K). Específicamente, los semidominios de escisión modificados descritos en el presente documento se prepararon mutando las posiciones 490 (E→K) y 538 (I→K) en un semidominio de escisión para producir un semidominio de escisión modificado designado "E490K:I538K" y mutando las posiciones 486 (Q→E) y 499 (I→L) en otro semidominio de escisión para producir un semidominio de escisión modificado designado "Q486E:I499L". Los semidominios de escisión modificados descritos en el presente documento son mutantes heterodímeros obligados en los que la escisión aberrante se minimiza o se elimina. Véase, p. ej., las patentes de Estados Unidos números 7.914.796 y 8.034.598. En ciertas partes de la divulgación, el semidominio de escisión modificado comprende mutaciones en las posiciones 486, 499 y 496 (numeradas en relación con *FokI* de tipo silvestre), por ejemplo las mutaciones que reemplazan el resto de Gln (Q) de tipo silvestre en la posición 486 con un resto Glu (E), el resto de Iso (I) de tipo silvestre en la posición 499 con un resto Leu (L) y el resto de Asn (N) de tipo silvestre en la posición 496 con un resto Asp (D) o Glu (E) (también denominados dominios "ELD" y "ELE", respectivamente). En otras realizaciones, el semidominio de escisión modificado comprende mutaciones en las posiciones 490, 538 y 537 (numeradas en relación con *FokI* de tipo silvestre), por ejemplo las mutaciones que reemplazan el resto de Glu (E) de tipo silvestre en la posición 490 con un resto Lys (K), el resto de Iso (I) de tipo silvestre en la posición 538 con un resto Lys (K), y el resto de His (H) de tipo silvestre en la posición 537 con un resto de Lys (K) o un resto de Arg (R) (también denominados dominios "KKK" y "KKR", respectivamente). En otras partes de la divulgación, el semidominio de escisión modificado comprende mutaciones en las posiciones 490 y 537 (numeradas en relación con *FokI* de tipo silvestre), por ejemplo, mutaciones que reemplazan el resto de Glu (E) de tipo silvestre en la posición 490 con un resto de Lys (K) y el resto His (H) de tipo silvestre en la posición 537 con un resto de Lys (K) o un resto de Arg (R) (también denominados dominios "KIK" y "KIR", respectivamente). (Véase la patente de Estados Unidos N° 8.623.618). En otras partes de la divulgación, el semidominio de escisión modificado comprende las mutaciones "Sharkey" y/o "Sharkey" (véase Guo *et al.*, (2010) *J. Mol. Biol.* 400(1):96-107).

Los semidominios de escisión modificados descritos en el presente documento pueden prepararse usando cualquier método adecuado, por ejemplo, mediante mutagénesis dirigida al sitio de semidominios de escisión de tipo silvestre (*FokI*) como se describe en las patentes de Estados Unidos números 7.888.121; 7.914.796; 8.034.598 y 8.623.618.

Como alternativa, las nucleasas se pueden ensamblar *in vivo* en el sitio diana ácido utilizando la denominada tecnología "enzima dividida" (véase, p.ej., la publicación de patente de Estados Unidos N° 20090068164). Los componentes de tales enzimas divididas se pueden expresar en construcciones de expresión separadas, o se pueden unir en un marco de lectura abierto donde los componentes individuales están separados, por ejemplo, por un péptido 2A auto-escindible o una secuencia IRES. Los componentes pueden ser dominios de unión de dedos de zinc individuales o dominios de un dominio de unión de ácido nucleico de meganucleasa.

Las nucleasas pueden seleccionarse para detectar actividad antes de su uso, por ejemplo en un sistema cromosómico basado en levadura como se describe en los documentos WO 2009/042163 y 20090068164. Las construcciones de expresión de nucleasas se pueden diseñar fácilmente usando métodos conocidos en la técnica. Véase, p. ej., las publicaciones de patente de Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474;

20060188987; 20060063231; y la publicación internacional WO 07/014275. La expresión de la nucleasa puede estar bajo el control de un promotor constitutivo o de un promotor inducible, por ejemplo el promotor de galactocinasa que se activa (desreprimido) en presencia de rafinosa y/o galactosa y reprimido en presencia de glucosa.

5 Como se divulga en el presente documento, la nucleasa comprende un sistema CRISPR/Cas. El locus CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas), que codifica los componentes del ARN del sistema, y el locus cas (asociado a CRISPR), que codifica las proteínas (Jansen *et al.*, 2002. Mol. Microbiol. 43: 1565-1575; Makarova *et al.*, 2002. Nucleic Acids Res. 30: 482-496; Makarova *et al.*, 2006. Biol. Direct 1: 7; Haft *et al.*, 2005. PLoS Comput. Biol. 1: e60) constituyen las secuencias génicas del sistema CRISPR/nucleasa Cas. Los loci  
10 CRISPR en los hospedadores microbianos contienen una combinación de genes (Cas) asociados a CRISPR, así como elementos de ARN no codificantes capaces de programar la especificidad de la escisión del ácido nucleico mediada por CRISPR.

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

30 La proteína Cas puede ser un "derivado funcional" de una proteína Cas natural. Un "derivado funcional" de un polipéptido de secuencia nativa es un compuesto que tiene una propiedad biológica cualitativa en común con un polipéptido de secuencia nativa. Los "derivados funcionales" incluyen, pero sin limitación, fragmentos de una secuencia nativa y derivados de un polipéptido de secuencia nativa y sus fragmentos, siempre que tengan una actividad biológica en común con un polipéptido de secuencia nativa correspondiente. Una actividad biológica contemplada en el presente documento es la capacidad del derivado funcional para hidrolizar un sustrato de ADN en fragmentos. El término "derivado" abarca variantes de secuencia de aminoácidos del polipéptido, modificaciones  
35 covalentes y fusiones de los mismos. Los derivados adecuados de un polipéptido Cas o un fragmento del mismo incluyen, pero sin limitación, mutantes, fusiones, modificaciones covalentes de la proteína Cas o un fragmento de la misma. La proteína Cas, que incluye la proteína Cas o un fragmento de la misma, así como derivados de la proteína Cas o un fragmento de los mismos, pueden obtenerse de una célula o sintetizarse químicamente o mediante una combinación de estos dos procedimientos. La célula puede ser una célula que produce naturalmente proteína Cas, o  
40 una célula que produce naturalmente proteína Cas y está genéticamente modificada para producir la proteína Cas endógena a un nivel de expresión más alto o para producir una proteína Cas a partir de un ácido nucleico introducido de forma exógena, ácido nucleico que codifica una Cas que es igual o diferente de la Cas endógena. En algún caso, la célula no produce naturalmente la proteína Cas y está genéticamente modificada para producir una proteína Cas.

45 Los ejemplos de sistemas de nucleasa CRISPR/Cas dirigidos a un sitio seguro y otros genes se divulgan, por ejemplo, en la solicitud provisional de Estados Unidos N° 61/823.689.

50 Por lo tanto, la nucleasa comprende un dominio de unión al ADN que se une específicamente a un sitio diana en cualquier gen en donde se desea insertar un donante (transgén).

### Sitios diana

55 Tal como se describe en detalle anteriormente, los dominios de unión al ADN pueden modificarse para unirse a cualquier secuencia de elección, por ejemplo, en un locus seguro, tal como el de la albúmina. Un dominio de unión al ADN modificado puede tener una especificidad de unión nueva, en comparación con un dominio de unión al ADN natural. Los métodos de modificación incluyen, pero sin limitación, un diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, la utilización de bases de datos que comprenden secuencias de nucleótidos de triplete (o cuadruplete) y secuencias de aminoácidos de dedo de zinc individuales, en las cuales cada secuencia de nucleótidos de triplete o cuadruplete se asocia con una o más secuencias de aminoácidos de dedos de zinc que se  
60 unen a la secuencia triplete o cuadruplete particular. Véase, por ejemplo, en las patentes estadounidenses de cotitularidad números 6.453.242 y 6.534.261. También se puede realizar el diseño racional de dominios efectores TAL. Véase, p. ej., la patente de Estados Unidos N° 8.586.526.

65 Los ejemplos de métodos de selección aplicables a los dominios de unión al ADN, incluyendo la presentación en fagos y los sistemas de dos híbridos, se divulgan en las patentes de Estados Unidos 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como en el documento WO 98/37186; WO 98/53057;

WO 00/27878; WO 01/88197 y GB 2.338.237. Además, la mayor especificidad de unión para dominios de unión de dedos de zinc se ha descrito, por ejemplo, en el documento de cotitularidad WO 02/077227.

5 Selección de sitios diana; las nucleasas y los métodos para el diseño y la construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos por un experto en la materia y se describen en detalle en las publicaciones de solicitud de patente de Estados Unidos números 20050064474 y 20060188987.

10 Además, como se divulga en estas y otras referencias, los dominios de unión al ADN (p. ej., proteínas de dedo de zinc de múltiples dedos) pueden unirse entre sí usando cualquier secuencia de conector adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos. Véase, p. ej., las patentes de Estados Unidos números 6.479.626; 6.903.185; y 7.153.949 para ejemplos de secuencias de conector de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dominios de unión al ADN individuales de la proteína. Véase, *también*, la patente de Estados Unidos N° 8.586.526.

15 Para el tratamiento de la hemofilia mediante la inserción dirigida de una secuencia que codifica una proteína F8 y/o F.IX funcional, cualquier sitio de inserción deseado en el genoma del sujeto se escinde con una nucleasa, que estimula la inserción dirigida del polinucleótido donante que porta la secuencia codificante de F8 y/o F.IX. Los dominios de unión al ADN de las nucleasas pueden dirigirse a cualquier sitio deseado en el genoma. En determinadas realizaciones, el dominio de unión al ADN de la nucleasa está dirigido a un locus de sitio seguro endógeno, por ejemplo  
20 un locus de la albúmina endógeno.

Se puede insertar cualquier donante mediante la integración dirigida mediada por nucleasas como se describe en el presente documento. En determinadas realizaciones, el donante comprende un polinucleótido (transgén) que codifica una proteína terapéutica, por ejemplo una proteína inexistente, deficiente y/o expresada de manera aberrante en un  
25 sujeto con una enfermedad o trastorno. Los ejemplos no limitativos de dichos trastornos incluyen, epidermolisis ampollosa, diabetes, cáncer, trastornos de la coagulación o enfisema deficiente en AAT, trastornos de la coagulación y/o enfermedades de almacenamiento lisosomal.

30 Para el tratamiento de la hemofilia, la secuencia del donante (también llamada "secuencia exógena" o "donante" o "transgén") comprende una secuencia que codifica una proteína del factor de coagulación funcional, o una parte de la misma, para dar como resultado una secuencia que codifica y expresa un factor de coagulación funcional tras la integración del donante. Los ejemplos no limitativos de transgenes de proteínas del factor de coagulación incluyen el Factor VIII y/o el Factor IX, incluyendo fragmentos funcionales de estas proteínas. En determinadas realizaciones, se elimina el dominio B de la proteína F8. Véase, *p.ej.*, Chuah *et al.* (2003) 101(5):1734-1743. En partes de la divulgación,  
35 otro, el transgén comprende una secuencia que codifica una proteína F.IX funcional, o parte de la misma, para dar como resultado una secuencia que codifica y expresa una proteína F.IX funcional después de la integración del donante. De manera similar, para tratar un LSD, la secuencia del donante codifica una o más proteínas inexistentes en un sujeto con un LSD. Los ejemplos no limitativos de tales proteínas incluyen la glucocerebrosidasa (GBA), que es deficiente en la enfermedad de Gaucher; la  $\alpha$  galactosidasa (GLA), que es deficiente en la enfermedad de Fabry; la iduronato-2-sulfatasa (IDS), que es deficiente en la enfermedad de Hunter; la alfa-L iduronidasa (IDUA), que es deficiente en la enfermedad de Hurler; la esfingomielina fosfodiesterasa 1 (SMPD1), que es deficiente en la enfermedad de Niemann-Pick.

45 La molécula donante puede insertarse en un gen endógeno de modo que se exprese todo, parte o nada del gen endógeno. Por ejemplo, un transgén que comprende secuencias de la proteína de factor de coagulación funcional (p. ej., F8 y/o F.IX) tal como se describe en el presente documento puede insertarse en un locus de la albúmina endógeno de modo que parte o nada de la albúmina endógena se exprese con el transgén.

50 La secuencia del donante (transgén) puede introducirse en la célula antes de, simultáneamente con o posteriormente a, la expresión de la proteína(s) de fusión (*p.ej.*, nucleasas). El polinucleótido donante puede contener suficiente homología (regiones continuas o discontinuas) con una secuencia genómica para soportar la recombinación homóloga (o reparación dirigida por la homología) entre él y la secuencia genómica a la que lleva homología o, como alternativa, las secuencias del donante pueden integrarse a través de mecanismos HDR (*p.ej.*, captura de donante NHEJ), en cuyo caso el polinucleótido donante (*p.ej.*, vector) no necesita contener secuencias que sean homólogas a la región  
55 de interés en la cromatina celular. Véase, p. ej., las patentes de Estados Unidos N° 7.888.121 y 7.972.843 y la publicación de patente de Estados Unidos N° 20110281361; 20100047805 y 20110207221.

60 El polinucleótido donante puede ser ADN o ARN, de cadena sencilla, de doble cadena o parcialmente de cadena sencilla y parcialmente de doble cadena y se puede introducir en una célula en forma lineal o circular (p. ej., minicírculo). Véase, p. ej., las publicaciones de patente de Estados Unidos números 20100047805, 20110281361, 20110207221. Si se introduce en forma lineal, los extremos de la secuencia del donante pueden protegerse (por ejemplo, de la degradación exonucleolítica) por métodos conocidos por los expertos en la materia. Por ejemplo, se añaden uno o más restos de didesoxinucleótidos al extremo 3' de una molécula lineal y/o los oligonucleótidos autocomplementarios se ligan a uno o ambos extremos. Véase, por ejemplo, Chang *et al.* (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls *et al.* (1996) Science 272:886-889. Los métodos adicionales para proteger los  
65 polinucleótidos exógenos de la degradación incluyen, aunque no de forma limitativa, la adición de grupo(s) amino

terminal y el uso de enlaces internucleotídicos modificados tales como, por ejemplo, fosforotioatos, fosforamidatos y restos de O-metil ribosa o desoxirribosa. Se puede introducir un polinucleótido en una célula como parte de una molécula de vector que tiene secuencias adicionales tales como, por ejemplo, orígenes de replicación, promotores y genes que codifican resistencia a antibióticos. Además, los polinucleótidos donantes se pueden introducir como ácido nucleico desnudo, como ácido nucleico formando un complejo con un agente tal como un liposoma o poloxámero, o pueden ser administrados por virus (*p.ej.*, adenovirus, AAV, herpesvirus, retrovirus, lentivirus).

El donante generalmente se inserta de modo que su expresión sea dirigida por el promotor endógeno en el sitio de integración (*p. ej.*, el promotor de la albúmina endógeno cuando el donante se integra en el locus de la albúmina del paciente). Por lo tanto, el transgén generalmente carece de elementos de control (*p. ej.*, promotor y/o potenciador) que dirijan su expresión (*p. ej.*, también denominada "construcción sin promotor"). No obstante, será evidente que el donante puede comprender un promotor y/o potenciador, por ejemplo, un promotor constitutivo o un promotor inducible o específico de tejido (por ejemplo, específico de hígado o plaquetas) que dirige la expresión de la proteína funcional tras la integración.

La secuencia del donante puede integrarse específicamente en cualquier sitio diana de elección, eliminando así los problemas asociados a la integración aleatoria en la terapia génica tradicional.

Cuando las secuencias de la albúmina (endógenas o parte del transgén) se expresan con el transgén, las secuencias de la albúmina pueden ser secuencias de longitud completa (tipo silvestre o mutante) o secuencias parciales. Preferentemente las secuencias de la albúmina son funcionales. Los ejemplos no limitativos de la función de estas secuencias de la albúmina de longitud completa o parciales incluyen aumentar la semivida en suero del polipéptido expresado por el transgén (por ejemplo, gen terapéutico) y/o actuar como un vehículo.

Asimismo, aunque no se requiere para la expresión, las secuencias exógenas también pueden incluir secuencias reguladoras de la transcripción o de la traducción, por ejemplo, promotores, potenciadores, aisladores, sitios de entrada del ribosoma internos, secuencias que codifican péptidos 2A y/o señales de poliadenilación.

Cualquiera de las secuencias de donantes puede incluir una o más de las siguientes modificaciones: optimización de codones (*p.ej.*, en codones humanos) y/o adición de uno o más sitios de glicosilación. Véase, *p. ej.*, McIntosh *et al.* (2013) *Blood* (17):3335-44. Las secuencias exógenas también pueden comprender secuencias peptídicas que permiten la administración dirigida de una proteína terapéutica. Por ejemplo, las secuencias de ácido nucleico que codifican el polipéptido p97 humano y/o fragmentos del mismo pueden unirse a una secuencia exógena del donante de manera que la proteína de fusión tendrá el potencial de atravesar la barrera hematoencefálica (véase *p. ej.*, la solicitud provisional de patente de Estados Unidos Nº 20130183368 y Karkan *et al.* (2008) *PLOS One*. DOI: 10.1371/journal.pone.0002469) o pueden usarse otros péptidos para dirigir una proteína codificada por un donante de transgén a orgánulos intracelulares tales como la mitocondria (*p.ej.*, Jacotot *et al.* (2006) *Biochim Biophys Acta Bioenerg* 1757: 1312-1323).

#### 40 Administración

Las nucleasas, los polinucleótidos que codifican estas nucleasas, los polinucleótidos donantes y las composiciones que comprenden las proteínas y/o polinucleótidos descritos en el presente documento pueden administrarse *in vivo* o *ex vivo* mediante cualquier medio adecuado.

Los métodos para administrar nucleasas como se describe en el presente documento se describen, por ejemplo, en las patentes de Estados Unidos números 8.586.526; 6.453.242; 6.503.717; 6.534.261; 6.599.692; 6.607.882; 6.689.558; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824.

Las construcciones de nucleasas y/o donantes como se describe en el presente documento también pueden administrarse usando vectores que contienen secuencias que codifican una o más de las proteínas de dedo de zinc, proteína(s) TALEN y/o sistema CRISPR/Cas. Se puede usar cualquier sistema de vectores incluyendo, pero sin limitación, vectores plasmídicos, vectores retrovirales, vectores lentivirales, vectores de adenovirus, vectores de poxvirus; vectores de herpesvirus y vectores de virus adenoasociados, etc. Véase, *también*, las patentes de Estados Unidos números 6.534.261; 6.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824. Asimismo, será evidente que cualquiera de estos vectores puede comprender una o más de las secuencias necesarias. Por lo tanto, cuando una o más nucleasas y una construcción donante se introducen en la célula, las nucleasas y/o el polinucleótido donante pueden transportarse en el mismo vector o en diferentes vectores. Cuando se usan múltiples vectores, cada vector puede comprender una secuencia que codifica una o múltiples nucleasas y/o construcciones de donantes. En determinadas realizaciones, un vector se usa para transportar tanto el transgén como la nucleasa(s). En otras realizaciones, se usan dos vectores (el mismo tipo o diferentes tipos de vectores), donde un vector lleva la nucleasa(s) (*p. ej.*, las ZFN izquierda y derecha o una pareja de ZFN, por ejemplo con un péptido 2A) y una lleva el transgén. En otras realizaciones adicionales, se usan tres vectores donde el primer vector lleva una nucleasa de una pareja de nucleasas (*p. ej.*, la ZFN izquierda), el segundo vector lleva la otra nucleasa de una pareja de nucleasas (*p. ej.*, la ZFN derecha) y el tercer vector lleva el transgén. Véase, Figura 2.

Los donantes y/o la nucleasa pueden usarse a cualquier concentración adecuada. En ciertas partes de la divulgación, el donante y el vector(es) de nucleasa separados se usan con la misma concentración. En otras partes de la divulgación, el donante y el vector(es) de nucleasa separados se usan con diferentes concentraciones, por ejemplo, 2-, 3-, 4-, 5-, 10- o más veces de un vector que del otro (*p.ej.*, más vector(es) donantes que vector(es) de nucleasa. Cuando se usan vectores de AAV para la administración, por ejemplo, el vector o vectores virales que comprenden el donante y/o la nucleasa están entre  $1 \times 10^8$  y  $1 \times 10^{13}$  partículas por dosis (*p. ej.*, célula o animal).

Se pueden usar métodos convencionales de transferencia génica virales y no virales para introducir ácidos nucleicos que codifican nucleasas y construcciones de donantes en células (*p.ej.*, células de mamífero) y tejidos diana. Los sistemas de administración de vectores no virales incluyen plásmidos de ADN, ácido nucleico desnudo y ácido nucleico en forma de complejo con un vehículo de administración tal como un liposoma o un poloxámero. Los sistemas de administración de vectores virales incluyen virus de ADN y ARN, que tienen genomas episómicos o integrados después de la administración a la célula. Para una revisión de la administración *in vivo* de proteínas de unión al ADN modificadas y de las proteínas de fusión que comprenden estas proteínas de unión, véase, *p. ej.*, Rebar (2004) Expert Opinion Invest. Drugs 13(7):829-839; Rossi *et al.* (2007) Nature Biotech. 25(12):1444-1454 así como las referencias generales de administración de genes tales como Anderson, Science 256:808-813 (1992); Nabel & Felgner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10):1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada *et al.*, en Current Topics in Microbiology and Immunology Doerfler and Böhm (eds.) (1995); y Yu *et al.*, Gene Therapy 1:13-26 (1994).

Los métodos de administración no viral de ácidos nucleicos incluyen electroporación, lipofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados de policaciones o ácidos nucleicos, ADN desnudo, viriones artificiales, y captación de ADN reforzada con un agente. La sonoporación que usa, *p. ej.*, el sistema Sonitron 2000 (Rich-Mar) también se puede usar para la administración de ácidos nucleicos.

Los ejemplos de sistemas de administración de ácido nucleico incluyen aquellos proporcionados por AmaxaBiosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) y Copernicus Therapeutics Inc, (véase por ejemplo el documento US-6.008.336). La lipofección se describe en *p. ej.*, las patentes de Estados Unidos números 5.049.386; 4.946.787; y 4.897.355) y los reactivos de lipofección se venden comercialmente (*p.ej.*, Transfectam™ y Lipofectin™). Los lípidos catiónicos y neutros que son adecuados para la lipofección eficiente de polinucleótidos con reconocimiento de receptor incluyen los de Felgner, los documentos WO 91/17424, WO 91/16024.

La preparación de los complejos lípido:ácido nucleico, incluidos los liposomas dirigidos tales como los complejos de inmunolípidos, es bien conocida por un experto en la materia (véase, *p.ej.*, Crystal, Science 270:404-410 (1995); Blaese *et al.*, Cancer Gene Ther. 2:291-297 (1995); Behr *et al.*, Bioconjugate Chem. 5:382-389 (1994); Remy *et al.*, Bioconjugate Chem. 5:647-654 (1994); Gao *et al.*, Gene Therapy 2:710-722 (1995); Ahmad *et al.*, Cancer Res. 52:4817-4820 (1992); patentes de Estados Unidos números 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028, y 4.946.787).

Los métodos adicionales de administración incluyen el uso del empaquetamiento de los ácidos nucleicos para ser administrados en los vehículos de administración EnGeneIC (EDV). Estos EDV se administran específicamente a los tejidos diana utilizando anticuerpos biespecíficos donde un brazo del anticuerpo tiene especificidad por el tejido diana y el otro tiene especificidad por el EDV. El anticuerpo lleva los EDV a la superficie de la célula diana y luego el EDV es introducido en la célula por endocitosis. Una vez en la célula, el contenido se libera (véase MacDiarmid *et al.* (2009) Nature Biotechnology 27(7):643).

El uso de sistemas basados en virus de ARN o ADN para la administración de ácidos nucleicos que codifican nucleasas y/o donantes modificados aprovecha procesos altamente evolucionados para dirigir un virus a células específicas en el cuerpo y transportar la carga viral hacia el núcleo. Los vectores virales se pueden administrar directamente a los pacientes (*in vivo*) o pueden usarse para tratar células *in vitro* y las células modificadas se administran a pacientes (*ex vivo*). Los sistemas convencionales basados en virus para la administración de nucleasas y/o donantes incluyen, pero sin limitación, vectores retrovirales, de lentivirus, adenovirales, adeno asociados, del virus vaccinia y del herpes simple para la transferencia génica. La integración en el genoma del hospedador es posible con los métodos de transferencia génica de retrovirus, lentivirus, y virus adenoasociados, lo que a menudo tiene como resultado la expresión a largo plazo del transgén insertado. Adicionalmente, se han observado altas eficiencias de transducción en muchos tipos diferentes de células y tejidos diana.

El tropismo de un retrovirus puede alterarse incorporando proteínas de la envoltura extrañas, expandiendo la población diana potencial de células diana. Los vectores lentivirales son vectores retrovirales que son capaces de transducir o infectar células que no se dividen y normalmente producen altos títulos virales. La selección de un sistema de transferencia génica retroviral depende del tejido diana. Los vectores retrovirales están compuestos por repeticiones terminales largas que actúan en cis con capacidad de empaquetamiento de hasta 6-10 kb de secuencia extraña. Las LTR que actúan en cis mínimas son suficientes para la replicación y el empaquetamiento de los vectores, que se

utilizan a continuación para integrar el gen terapéutico en la célula diana para proporcionar una expresión transgénica permanente. Los vectores retrovirales ampliamente utilizados incluyen aquellos basados en el virus de la leucemia murina (MuLV), el virus de leucemia del gibón (GaLV), el virus de la inmunodeficiencia de simio (SIV), el virus de la inmunodeficiencia humana (VIH) y combinaciones de los mismos (véase, *p.ej.*, Buchscher *et al.*, J. Virol. 66:2731-2739 (1992); Johann *et al.*, J. Virol. 66:1635-1640 (1992); Sommerfelt *et al.*, Virol. 176:58-59 (1990); Wilson *et al.*, J. Virol. 63:2374-2378 (1989); Miller *et al.*, J. Virol. 65:2220-2224 (1991); PCT/US94/05700).

En aplicaciones en las que se prefiere la expresión transitoria, se pueden usar sistemas basados en adenovirus. Los vectores basados en adenovirus son capaces de una eficiencia de transducción muy alta en muchos tipos de células y no requieren división celular. Con tales vectores, se ha obtenido un alto título y altos niveles de expresión. Este vector se puede producir en grandes cantidades en un sistema relativamente simple. Los vectores de virus adenoasociados ("AAV") también se usan para transducir células con ácidos nucleicos diana, *p. ej.*, en la producción *in vitro* de ácidos nucleicos y péptidos, y para procedimientos de terapia génica *in vivo* y *ex vivo* (véase, *p.ej.*, West *et al.*, Virology 160:38-47 (1987); la patente de Estados Unidos N° 4.797.368; el documento WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). La construcción de vectores de AAV recombinantes se describe en varias publicaciones, incluyendo la patente de Estados Unidos. N° 5.173.414; Tratschin *et al.*, Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, *et al.*, Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); y Samulski *et al.*, J. Virol. 63:03822-3828 (1989).

Actualmente existen al menos seis enfoques de vectores virales para la transferencia génica en ensayos clínicos, que utilizan enfoques que implican la complementación de vectores defectuosos mediante genes insertados en líneas celulares auxiliares para generar el agente transductor.

pLASN y MFG-S son ejemplos de vectores retrovirales que se han usado en ensayos clínicos (Dunbar *et al.*, Blood 85:3048-305 (1995); Kohn *et al.*, Nat. Med. 1:1017-102 (1995); Malech *et al.*, PNAS 94:22 12133-12138 (1997)). PA317/pLASN fue el primer vector terapéutico usado en un ensayo de terapia génica. (Blaese *et al.*, Science 270:475-480 (1995)). Se han observado eficiencias de transducción del 50 % o más en vectores empaquetados en MFG-S. (Ellem *et al.*, Immunol Immunother. 44(1):10-20 (1997); Dranoff *et al.*, Hum. Gene Ther. 1:111-2 (1997).

Los vectores de virus adenoasociados recombinantes (rAAV) son sistemas prometedores de administración de genes alternativos basados en el virus adenoasociado parvovirus de tipo 2 defectuoso y no patógeno. Todos los vectores se derivan de un plásmido que conserva solo las repeticiones terminales invertidas de 145 pares de bases del AAV que flanquean el casete de expresión transgénica. La transferencia eficiente de genes y la administración estable de transgenes debido a la integración en los genomas de la célula transducida son características clave de este sistema de vectores. (Wagner *et al.*, Lancet 351:9117 1702-3 (1998), Kearns *et al.*, Gene Ther. 9:748-55 (1996)). Otros serotipos de AAV, incluyendo AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV9 y AAVrh10 o AAV pseudotipados tales como AAV2/8, AAV8.2, AAV2/5 y AAV2/6 y cualquier serotipo de AAV novedoso también se pueden usar de acuerdo con la presente invención.

Los vectores adenovirales (Ad) recombinantes deficientes en replicación pueden producirse a un título alto e infectar fácilmente varios tipos de células diferentes. La mayoría de los vectores de adenovirus están modificados de tal manera que un transgén reemplaza los genes Ad E1a, E1b, y/o E3; posteriormente el vector defectuoso de replicación se propaga en células 293 humanas que suministran la función del gen eliminado en *trans*. Los vectores Ad pueden transducir múltiples tipos de tejidos *in vivo*, incluyendo células que no se dividen, células diferenciadas, tales como las que existen en el hígado, riñón y músculo. Los vectores Ad convencionales tienen una gran capacidad de carga. Un ejemplo del uso de un vector Ad en un ensayo clínico incluyó la terapia con polinucleótidos para la inmunización antitumoral con inyección intramuscular (Sternan *et al.*, Hum. Gene Ther. 7:1083-9 (1998)). Los ejemplos adicionales del uso de vectores para la transferencia génica en ensayos clínicos incluyen Rosenecker *et al.*, Infection 24:1 5-10 (1996); Sternan *et al.*, Hum. Gene Ther. 9:7 1083-1089 (1998); Welsh *et al.*, Hum. Gene Ther. 2:205-18 (1995); Alvarez *et al.*, Hum. Gene Ther. 5:597-613 (1997); Topf *et al.*, Gene Ther. 5:507-513 (1998); Sternan *et al.*, Hum. Gene Ther. 7:1083-1089 (1998).

Las células de empaquetamiento se usan para formar partículas de virus que son capaces de infectar una célula hospedadora. Dichas células incluyen células 293, que empaquetan adenovirus, y células  $\psi$ 2 o células PA317, que empaquetan retrovirus. Los vectores virales utilizados en la terapia génica generalmente son generados por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula viral. Los vectores normalmente contienen las secuencias virales mínimas requeridas para el empaquetamiento y la posterior integración en un hospedador (si corresponde), siendo reemplazadas otras secuencias virales por un casete de expresión que codifica la proteína a expresar. Las funciones virales que faltan se suministran en *trans* por la línea celular de empaquetamiento. Por ejemplo, los vectores de AAV utilizados en terapia génica generalmente poseen secuencias de repetición terminal invertida (ITR) del genoma de AAV que se requieren para el empaquetamiento y la integración en el genoma del hospedador. El ADN viral se empaqueta en una línea celular, que contiene un plásmido auxiliar que codifica los otros genes de AAV, concretamente *rep* y *cap*, pero que carece de las secuencias ITR. La línea celular también se infecta con adenovirus como un auxiliar. El virus auxiliar promueve la replicación del vector AAV y la expresión de genes de AAV a partir del plásmido auxiliar. El plásmido auxiliar no está empaquetado en cantidades significativas debido a la falta de secuencias ITR. La contaminación con adenovirus puede reducirse mediante, *p. ej.*,

tratamiento térmico al que el adenovirus es más sensible que el AAV.

En muchas aplicaciones de terapia génica, es deseable que el vector de terapia génica se administre con un alto grado de especificidad a un tipo de tejido particular. En consecuencia, un vector viral puede modificarse para tener especificidad para un tipo de célula dado al expresar un ligando como una proteína de fusión con una proteína de cubierta viral en la superficie externa del virus. El ligando se elige de modo que tenga afinidad por un receptor que se sabe que está presente en el tipo de célula de interés. Por ejemplo, en Han *et al.*, Proc. Natl. Acad. Sci. USA 92:9747-9751 (1995) se describe que el virus de la leucemia murina de Moloney se puede modificar para expresar heregulina humana fusionada a gp70, y el virus recombinante infecta ciertas células de cáncer de mama humano que expresan el receptor del factor de crecimiento epidérmico humano. Este principio puede extenderse a otros pares de virus-célula diana, en los cuales la célula diana expresa un receptor y el virus expresa una proteína de fusión que comprende un ligando para el receptor de la superficie celular. Por ejemplo, los fagos filamentosos se pueden modificar para presentar fragmentos de anticuerpos (*p.ej.*, FAB o Fv) que tienen afinidad de unión específica para prácticamente cualquier receptor celular elegido. Aunque la descripción anterior se aplica principalmente a los vectores virales, los mismos principios se pueden aplicar a los vectores no virales. Dichos vectores pueden modificarse para contener secuencias de captación específicas que favorecen la captación por células diana específicas.

Los vectores de terapia génica pueden administrarse *in vivo* mediante la administración a un paciente individual, normalmente mediante administración sistémica (*p.ej.*, intravenosa, intraperitoneal, intramuscular, subdérmica, o infusión intracraneal) o administración tópica, como se describe a continuación. Como alternativa, los vectores se pueden administrar a células *ex vivo*, tal como las células explantadas de un paciente individual (*p.ej.*, linfocitos, aspirados de médula ósea, biopsia de tejido) o células madre hematopoyéticas de donantes universales, seguido de reimplantación de las células en un paciente, generalmente después de la selección de células que han incorporado el vector.

Los vectores (*p.ej.*, retrovirus, adenovirus, liposomas, etc.) que contienen nucleasas y/o construcciones de donante también se pueden administrar directamente a un organismo para la transducción de las células *in vivo*. Como alternativa, se puede administrar ADN desnudo. La administración se realiza por cualquiera de las vías que normalmente se utilizan para introducir una molécula en contacto final con las células sanguíneas o tisulares incluyendo, pero sin limitación, inyección, infusión, aplicación tópica y electroporación. Los métodos adecuados para administrar tales ácidos nucleicos están disponibles y son bien conocidos por los expertos en la materia y, aunque se puede usar más de una vía para administrar una composición particular, una vía particular a menudo puede proporcionar un camino más inmediato y más eficaz que otra vía.

Los vectores adecuados para la introducción de polinucleótidos (por ejemplo, codificadores y/o donantes de nucleasas) descritos en el presente documento incluyen vectores de lentivirus no integrantes (IDLV). Véase, por ejemplo, Ory *et al.* (1996) Proc. Natl. Acad. Sci. USA 93:11382-11388; Dull *et al.* (1998) J. Virol.72:8463-8471; Zuffery *et al.* (1998) J. Virol.72:9873-9880; Follenzi *et al.* (2000) Nature Genetics 25:217-222; publicación de patente de Estados Unidos N° 2009/054985.

Los vehículos farmacéuticamente aceptables pueden determinarse en parte por la composición particular que se administre, así como por el método particular usado para administrar la composición. En consecuencia, existe una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas disponibles, como se describe a continuación (véase, *p. ej.*, Remington's Pharmaceutical Sciences, 17<sup>o</sup> ed., 1989).

La administración *in vitro* e *in vivo* puede realizarse mediante el uso de nanopartículas. Muchas nanopartículas que se investigan actualmente están compuestas de moléculas terapéuticas que se autoensamblan con lípidos o polímeros en nanoestructuras. Estas partículas tienen el potencial de administrar dosis terapéuticas de ácidos nucleicos a los tejidos diana (*p. ej.*, células tumorales, órganos específicos, etc.). Véase *p. ej.*, Rink *et al.* (2013), Curr Opin Oncol: 25(6): p. 646-651.

Será evidente que las secuencias que codifican la nucleasa y las construcciones del donante pueden administrarse utilizando el mismo sistema o sistemas diferentes. Por ejemplo, las nucleasas y los donantes se pueden transportar en el mismo vector (*p.ej.*, AAV). Como alternativa, un polinucleótido donante puede ser transportado por un plásmido, mientras que una o más nucleasas pueden ser transportadas por un vector diferente (*p. ej.*, vector AAV). Asimismo, los diferentes vectores pueden administrarse por la misma o diferentes vías (inyección intramuscular, inyección en la vena de la cola, otra inyección intravenosa, administración intraperitoneal y/o inyección intramuscular. Los vectores pueden administrarse simultáneamente o en cualquier orden secuencial.

Por lo tanto, la presente divulgación incluye el tratamiento y/o la prevención *in vivo* o *ex vivo* de una enfermedad o trastorno en el cual una proteína es inexistente o deficiente. Por ejemplo, la hemofilia A se puede tratar, mediante la integración mediada por nucleasas de la secuencia que codifica F8. La divulgación también incluye el tratamiento *in vivo* o *ex vivo* de la hemofilia B, mediante la integración mediada por nucleasas de la secuencia que codifica F.IX. De manera similar, la divulgación incluye el tratamiento de las hemofilias relacionadas con la deficiencia de Factor VII y la deficiencia de Factor X mediante la integración mediada por nucleasas de una secuencia codificante del Factor VII o del Factor X, respectivamente. Además, la divulgación incluye el tratamiento de uno o más LSD mediante la

integración mediada por nucleasas de una o más proteínas inexistentes o deficientes en el LSD. Las composiciones se administran a un paciente humano en una cantidad efectiva para obtener la concentración deseada del polipéptido terapéutico en el suero, el hígado o las células diana. La administración puede ser por cualquier medio en el cual los polinucleótidos se administren a las células diana deseadas. Por ejemplo, se contemplan los métodos *in vivo* y *ex vivo*.

5 La inyección intravenosa en la vena porta es un método preferido de administración. Otros modos de administración *in vivo* incluyen, por ejemplo, inyección directa en los lóbulos del hígado o el conducto biliar e inyección intravenosa distal al hígado, incluyendo a través de la arteria hepática, inyección directa en el parénquima hepático, inyección a través de la arteria hepática y/o inyección retrógrada a través del árbol biliar. Los modos de administración *ex vivo* incluyen transducción *in vitro* de hepatocitos resecados u otras células del hígado, seguido de infusión de los

10 hepatocitos resecados, transducidos de nuevo a la vasculatura portal, el parénquima hepático o el árbol biliar del paciente humano, véase p. ej., Grossman *et al.* (1994) *Nature Genetics*, 6:335-341. Otros modos de administración incluyen la inserción mediada por nucleasas *ex vivo* de un transgén que codifica el Factor VII, F8, F.IX, Factor X, glucocerebrosidasa,  $\alpha$  galactosidasa, iduronato-2-sulfatasa y/o alfa-L iduronidasa en una ubicación de sitio seguro en las células madre del paciente o alogénicas. Tras la modificación, las células tratadas se vuelven a infundir en el

15 paciente para el tratamiento de la enfermedad o trastorno (p. ej., LSD y/o hemofilia).

La cantidad eficaz de nucleasa(s) y donante (p. ej., Factor VII, F8, F.IX, Factor X, GBA, GLA, IDS, IDUA, o SMPD1) a administrar variará de un paciente a otro y de acuerdo con el polipéptido terapéutico de interés. En consecuencia, las cantidades efectivas se determinan mejor por el médico que administra las composiciones y las dosis apropiadas se

20 pueden determinar fácilmente por un experto en la materia. Después de permitir suficiente tiempo para la integración y la expresión (por lo general, 4-15 días, por ejemplo), el análisis del suero u otros niveles de tejido del polipéptido terapéutico y la comparación con el nivel inicial antes de la administración determinarán si la cantidad administrada es demasiado baja, está dentro del intervalo correcto o si es demasiado alta. Los regímenes adecuados para las administraciones iniciales y posteriores también son variables, pero están tipificados por una administración inicial seguida de administraciones posteriores si es necesario. Las administraciones posteriores pueden administrarse a

25 intervalos variables, que van desde diariamente a anualmente a cada varios años. El experto en la materia apreciará que se pueden recomendar técnicas inmunosupresoras apropiadas para evitar la inhibición o el bloqueo de la transducción por inmunosupresión de los vectores de administración, véase p.ej., Vilquin *et al.* (1995) *Human Gene Ther.* 6:1391-1401.

30

Las formulaciones para las administraciones *ex vivo* e *in vivo* incluyen suspensiones en líquidos o líquidos emulsionados. Los principios activos se mezclan a menudo con excipientes que son farmacéuticamente aceptables y compatibles con el principio activo. Excipientes adecuados incluyen, por ejemplo, agua, solución salina, dextrosa, glicerol, etanol o similares y combinaciones de los mismos. Además, la composición puede contener cantidades

35 menores de sustancias auxiliares, tales como, agentes humectantes o emulsionantes, agentes tamponadores del pH, agentes estabilizantes u otros reactivos que mejoran la eficacia de la composición farmacéutica.

### Aplicaciones

40 Los métodos y composiciones de la invención se pueden usar en cualquier circunstancia en donde se desee administrar un transgén que codifique una o más proteínas de tal manera que la proteína o las proteínas sean secretadas de la diana. Por lo tanto, esta tecnología es útil en una afección en donde un paciente tiene deficiencia de algunas proteínas debido a problemas (p. ej., problemas en el nivel de expresión o problemas con la proteína expresada como subfuncional o no funcional). Adicionalmente, los trastornos por deficiencia de A1AT, tales como la EPOC o el daño hepático, u otros trastornos, afecciones o enfermedades que pueden mitigarse mediante la

45 administración de proteínas exógenas por un órgano secretor pueden tratarse con éxito mediante los métodos y composiciones de esta invención. Las enfermedades de almacenamiento lisosomal pueden tratarse mediante los métodos y composiciones de la invención, así como las enfermedades metabólicas tales como la diabetes.

50 Las proteínas que son útiles terapéuticamente y que se administran normalmente por inyección o infusión también son útiles con los métodos y composiciones de la invención. A modo de ejemplos no limitativos, la producción de un péptido C (p. ej., Ersatta™ por Cebix) o insulina para su uso en terapia diabética. Otra aplicación incluye el tratamiento de la epidermólisis ampollosa mediante la producción de colágeno VII. La expresión de IGF-1 en el tejido secretor como se describe en el presente documento puede usarse para aumentar los niveles de esta proteína en pacientes con cirrosis

55 hepática y deficiencia de lipoproteína lipasa mediante la expresión de la lipoproteína lipasa. Los anticuerpos también pueden secretarse para beneficio terapéutico, por ejemplo, para el tratamiento de cánceres, enfermedades autoinmunitarias y otras enfermedades. Los ejemplos de anticuerpos terapéuticos incluyen anticuerpos contra TNF- $\alpha$ , EpCAM, CD20, CD19, VEGFR, CD52 y similares. Otras proteínas relacionadas con la coagulación podrían producirse en el tejido secretor, incluyen fibrinógeno, protrombina, factor tisular, Factor V, Factor XI, Factor XII (factor de Hageman), Factor XIII (factor estabilizante de la fibrina), Factor de von Willebrand, precalicreína, quinínogeno de alto

60 peso molecular (factor de Fitzgerald), fibronectina, antitrombina III, cofactor II de la heparina, proteína C, proteína S, proteína Z, inhibidor de la proteasa dependiente de la proteína Z, plasminógeno, alfa 2-antiplasmina, activador del plasminógeno tisular, urocinasa, inhibidor del activador de plasminógeno-1, e inhibidor del activador del plasminógeno-2.

65

Los métodos y composiciones de la invención también se pueden usar en cualquier circunstancia en donde se desee

5 administrar y expresar un transgén que codifique uno o más ácidos nucleicos no codificantes o estructurales (p. ej., ARNhc o ARNi). Tales ARN pueden formar estructuras inhibitorias y ser útiles en el tratamiento de enfermedades tales como trastornos de los lípidos (donde la diana es p. ej., ApoB-100, ApoC-III, ANGPTL3, PCSK9); cardiopatía coronaria (donde la diana es p. ej., CRP, Apo(a)); trastornos de la coagulación y sanguíneos (donde la diana es p. ej., el F.XI, FVII, antitrombina, TMPRSS6); enfermedades autoinmunitarias (donde la diana esp. ej., ICAM-1, GCCR, GCGR, PTP-1B, VLA-4); amiloidosis por TTR; enfermedades musculares (donde la diana esp. ej., SMN2, GHR, DMPK); enfermedad inflamatoria (donde la diana esp. ej., PKK); obesidad (donde la diana es p. ej., FGFR4); enfermedad hepática (donde la diana esp. ej., DGAT2, ALAS-1, C5, AAT); Cáncer (donde la diana es p. ej., clusterina, eIF-4E, Hsp27, AR); enfermedad fibrótica (donde la diana esp. ej., CTGF); enfermedad ocular (donde la diana es p. ej. cinasa C-raf); o una enfermedad infecciosa (donde la diana es p. ej., aminoglucósido, hepcidina, RG-101).

15 Los siguientes ejemplos se refieren a realizaciones ilustrativas de la presente divulgación en donde la nucleasa comprende una nucleasa de dedo de zinc (ZFN). Se apreciará que esto es solo para fines ilustrativos y que se pueden usar otras nucleasas, por ejemplo endonucleasas orientadoras (meganucleasas) con dominios de unión al ADN modificados y/o fusiones de dominios de unión al ADN de endonucleasas orientadoras (meganucleasas) modificadas de origen natural y dominios de escisión heterólogos, TALEN y/o un sistema CRISPR/Cas que comprende un ARN guía único modificado.

## 20 Ejemplos

### Ejemplo 1: Integración dirigida de un transgén de F8 *in vivo*

25 A ratones HA/CD4<sup>-/-</sup> se les administró (1) vectores de control AAV2/8 y transgenes de donantes AAV2/8 que codifican F8, ("Simulado + Donante") o 2) vectores AAV que codifican pares de ZFN dirigidas al locus de la albúmina (como se describe en la publicación de patente de Estados Unidos 20130177983 y en la Figura 10A y 10B) y los transgenes del donante AAV que codifican F8 ("ZFN + Donante"), ambos por inyección en la vena de la cola como se describe en la publicación de patente de Estados Unidos N° 20120128635. Los donantes se usan como se muestra esquemáticamente en las Figuras 1 y 2 e incluyen un ADNc del Factor VIII con dominio B eliminado (BDD-F8) sin promotor de aproximadamente 4,4 a 4,7 kb de tamaño. Las dosis se administraron a los ratones en dos niveles de dosificación, un nivel de dosificación "bajo" que comprende AAV8-ZFN (1e<sup>11</sup>vg/ratón) + AAV8-Donante (1e<sup>11</sup>vg/ratón): Figura 3B; y un nivel de dosificación "alto" que comprende AAV8-ZFN (5e<sup>11</sup>vg/ratón) + AAV8-Donante (5e<sup>11</sup>vg/ratón): Figura 3C.

35 Adicionalmente, los donantes de F8 se optimizaron mediante la optimización de codones para la expresión en células de mamíferos según los protocolos estándar y mediante la adición de un conector (V3) con sitios de glicosilación (véase McIntosh *et al.*, (2013) Blood 121:3335). En este experimento, se administró a ratones HA/CD4<sup>-/-</sup> AAV8-ZFN (5e<sup>10</sup> vg de cada ZFN) + AAV8-Donante (1 e<sup>11</sup>vg/ratón): Figura 4.

40 Las concentraciones plasmáticas de F8 se evaluaron usando técnicas estándar.

45 Como se muestra en las Figuras 3 y 4, la integración dirigida de un donante de Factor 8 en el locus de la albúmina de ratón dio como resultado niveles de actividad de hasta el 50 % de lo normal en ratones con hemofilia A cuando se usaba cualquier donante. Sin embargo, cuando se usaba la construcción del donante de F8 optimizada, se observaron concentraciones plasmáticas de F8 comparables utilizando solo el 20 % de la dosis. Además, cuando se usaban dos vectores de ZFN, comprendiendo cada uno de ellos una de las dos ZFN necesarias para la pareja, se observaron niveles más altos de escisión que cuando las ZFN se introducían juntas en un vector de expresión separado por un sitio 2A (Figura 3D).

### 50 Ejemplo 2: Integración dirigida de un transgén de F9 *in vivo*

#### A. Hepatocitos humanos

55 Primero se transdujeron hepatocitos primarios humanos con el vector AAV6 que contiene un donante de F9 junto con la transfección del ARNm que codifica una pareja de ZFN dirigida a un sitio dentro del primer intrón de la albúmina humana.

Como se muestra en la Figura 9, los hepatocitos tratados con donantes y ZFN exhibieron F.IX humano medible en el sobrenadante de cultivo.

#### 60 B. Ratones

65 A continuación intentamos demostrar este enfoque *in vivo* en el ratón. Para conseguir esto, primero diseñamos una pareja de ZFN (que se muestra a continuación en la Tabla 1) dirigida a una ubicación análoga en el intrón 1 de la albúmina de ratón (que se muestra en la Figura 12) como se muestra a continuación y se confirmó la actividad de la pareja *in vitro* en células de hepatoma murino.

Tabla 1: Diseños de nucleasa específica de la albúmina de ratón

| SBS N° , Diana   | Diseño                         |                                |                             |                                |                                 |     |
|--|--------------------------------|--------------------------------|-----------------------------|--------------------------------|---------------------------------|-----|
| ZFN específicas de la albúmina de ratón  |                                |                                |                             |                                |                                 |     |
|  | F1                             | F2                             | F3                          | F4                             | F5                              | F6  |
| <b>SBS N° 30724</b><br>ctGAAGGTgGCAA<br>TGGTTcctctctg<br>ct (SEQ ID<br>NO:117) | TGSLTR<br>(SEQ ID NO:<br>119)  | RSDALST<br>(SEQ ID NO:<br>120) | QSATRTK                     | TSGHLR                         | QSGNLR<br>(SEQ ID NO:<br>2)     | N/A |
| <b>SBS N° 30725</b><br>ttTCCTGTAACGA<br>TCGGgaactggca<br>tc (SEQ ID<br>NO:118) | RSDHLSA<br>(SEQ ID NO:<br>122) | TKSNRTK<br>(SEQ ID NO:<br>123) | DRSNLSR<br>(SEQ ID<br>NO:5) | WRSSLRA<br>(SEQ ID NO:<br>124) | DSSDRKKQ<br>(SEQ ID NO:<br>125) | N/A |

Como se muestra en la Figura 10, la pareja era activa en células de hepatoma murino.

- 5 Además, se administraron a ratones de tipo silvestre (3 animales por grupo) mediante una inyección en la vena de la cola transgenes de donante AAV2/8 que codifican F.IX (véase, Figura 5A) con vectores AAV que codifican parejas de ZFN dirigidas a la albúmina de ratón ("mAlb ZFN") como se describe en la publicación de Estados Unidos N° 20130177983 o parejas de ZFN dirigidas al F.IX humano ("hF9 ZFN") como se describe en la publicación de patente de Estados Unidos N° 20120128635 y se muestra en las Figuras 12A y 12B. Las construcciones y las dosis del vector fueron las siguientes: albúmina AAV2/8-ZFN a  $1 \times 10^{11}$  vg/ratón y AAV8-Donante a  $5 \times 10^{11}$  vg/ratón y F.IX AAV8-ZFN a  $1 \times 10^{11}$  vg/ratón y AAV8-Donante a  $5 \times 10^{11}$  vg/ratón.

15 Las concentraciones plasmáticas de F9 se evaluaron usando kits ELISA estándar disponibles en el mercado usando anticuerpos disponibles en el mercado. Además, se llevaron a cabo ensayos Ce-I (Suveryor™, Transgenomics) como se describe en la publicación de patente de Estados Unidos N° 20120128635 y en Perez *et al*, (2008) Nat. Biotechnol. 26: 808-816 y en Guschin *et al*, (2010) Methods Mol Biol. 649:247-56).

20 Como se muestra en la Figura 5B, se obtuvieron niveles de hFIX circulantes robustos después de la administración de ZFN de albúmina y donante de F9. Las ZFN específicas de F.IX humano no reconocen el locus de F.IX de ratón endógeno, por lo que no hay una integración apreciable del donante de F9 cuando se usa esta pareja de nucleasas. Asimismo, como se muestra en la Figura 6, la edición del genoma es proporcional a la dosis de AAV en más de tres órdenes de magnitud.

25 Además, a ratones hemofílicos (ratones HB) se les administraron los donantes y la ZFN de albúmina como se ha descrito anteriormente (AAV8-mAlb-ZFN a  $1 \times 10^{11}$  vg/ratón y AAV8-donante de F9 a  $5 \times 10^{11}$  vg/ratón, véase también Li *et al* (2011) *ibid* y Anguela *et al* (2013) *ibid*) y también se determinaron las concentraciones plasmáticas de hF.IX y el tiempo(s) de tromboplastina parcial activado (aPTT(s)) mediante kits disponibles en el mercado (p.ej., el kit cromogénico Rox Factor IX de Rossix, y Vitaclot, Vital® Diagnostics).

30 Como se muestra en la Figura 7, la integración mediada por ZFN de un transgén del donante de F9 en el locus de la albúmina de ratones HB dio como resultado altos niveles de F.IX en el plasma y la corrección de los tiempos de coagulación prolongados.

### C. Macacos Rhesus

35 Para probar la modificación del genoma dirigida por ZFN y la inserción de transgenes en animales más grandes, se realizaron dos estudios. Las ZFN utilizadas se muestran a continuación en la Tabla 2. Las mayúsculas en la secuencia diana indican nucleótidos unidos y las minúsculas indican nucleótidos no unidos.

Tabla 2: Diseños de nucleasa específica de la albúmina de macaco Rhesus

| SBS N° , Diana  | Diseño                      |                              |                             |                              |                              |     |
|---|-----------------------------|------------------------------|-----------------------------|------------------------------|------------------------------|-----|
|   | F1                          | F2                           | F3                          | F4                           | F5                           | F6  |
| <b>SBS N° 36806 (rhesus)</b><br>ttAGGGACAGT<br>TATGAAttcaa<br>tttca<br>(SEQ ID<br>NO:1) | QSGNLR<br>(SEQ ID NO:<br>2) | LMQNRNQ<br>(SEQ ID NO:<br>3) | LKHHLTD<br>(SEQ ID<br>NO:4) | DRSNLSR<br>(SEQ ID NO:<br>5) | RSDHLTQ<br>(SEQ ID NO:<br>6) | N/A |

(continuación)

| SBS N° , Diana  | Diseño                      |                              |                              |                               |                              |                              |
|---|-----------------------------|------------------------------|------------------------------|-------------------------------|------------------------------|------------------------------|
|   | F1                          | F2                           | F3                           | F4                            | F5                           | F6                           |
| <b>SBS N° 35396</b><br><b>(humana/rhesus)</b><br>ccTATCCATTG<br>CACTATGCTtt<br>atttaa<br>(SEQ ID<br>NO:7) | QSSDLR<br>(SEQ ID NO:<br>8) | LKWNLRT<br>(SEQ ID NO:<br>9) | DQSNLRA<br>(SEQ ID<br>NO:10) | RPYTLRL<br>(SEQ ID NO:<br>11) | QSSDLR<br>(SEQ ID NO:<br>8)  | HRSNLNK<br>(SEQ ID<br>NO:12) |
| <b>SBS N° 37804</b><br><b>(rhesus)</b><br>ttAGGACAGT<br>TATGAAttcaa<br>tcttca<br>(SEQ ID<br>NO:1)         | QSGNLR<br>(SEQ ID NO:<br>2) | LMQNRNQ<br>(SEQ ID NO:<br>3) | LAHHLVE<br>(SEQ ID<br>NO:13) | DRSNLSR<br>(SEQ ID NO:<br>5)  | RSDHLTQ<br>(SEQ ID NO:<br>6) | N/A                          |

Todos los diseños que se muestran en la Tabla 2 están vinculados a sus sitios diana.

- 5 Se realizaron estudios ilustrativos con la pareja de ZFN 36806 y 35396 (Pareja 2) como sigue. Los monos rhesus (criados a propósito), de 2 a 4 años de edad con pesos de 3 a 4,6 kg se seleccionaron previamente para detectar la presencia de anticuerpos neutralizantes contra rAAV 2/6 y 2/8, el genotipo del locus de la albúmina, la bioquímica sérica normal y la hematología. Los animales fueron alojados socialmente (hasta 3 animales del mismo grupo de dosificación alojados juntos). La administración del vector se realizó mediante infusión IV en una vena periférica a una
- 10 tasa de 1 ml/min, para una duración de la dosificación que variaba de ~10-30 minutos (10 ml de cada para el Estudio 1,29 ml de cada para el Estudio 2). Los monos fueron evaluados durante todo el estudio para determinar la mortalidad/morbilidad, observaciones clínicas de rutina, observaciones secundarias de la jaula y consumo de alimentos (diariamente), pesos corporales (preestudio y semanal), patología clínica que incluye niveles de enzimas hepáticas (ALT y AST), bioquímica clínica y hematología, y coagulación utilizando metodologías de rutina. Se
- 15 realizaron biopsias de hígado y se examinaron los tejidos para determinar su histopatología y la farmacocinética de los vectores rAAV, así como para evaluar la modificación génica mediante el sistema miSEQ (Illumina) y la expresión de ZFN mediante transferencia Western. El análisis de anticuerpos contra fármacos se realizó durante todo el estudio y las PBMC se aislaron de la sangre completa para el análisis EliSpot (ver arriba). La patología macroscópica y microscópica se realiza en la evaluación de tejidos al final del estudio.

- 20 Estudio N° 1: Se administraron a macacos Rhesus ZFN dirigidas a la albúmina en vectores AAV2/8 como se describe en la publicación de patente de Estados Unidos N° 20130177983. En este estudio, también se utilizó una variante del dominio de escisión FokI de tipo silvestre en donde la secuencia se había optimizado para la expresión de mamíferos de acuerdo con técnicas estándar (DNA 2.0). Los grupos de dosificación y las ID de los animales se muestran a
- 25 continuación en la Tabla 3.

**Tabla 3: Estudio en PNH N° 1; Grupos de dosificación**

| Grupo | Descripción  | Dosis            | ID del animal |
|-------|--|------------------|---------------|
| 1     | Control negativo   | Ninguno          | 1001          |
| 2     | Pareja de ZFN 2, FokI WT, ZFN solo-                        | 1,5e+13 cada ZFN | 7001          |
|       |  | 1,5e+13 cada ZFN | 7002          |
| 3     | Pareja de ZFN 2, Optimización de codones Fok1 WT, ZFN solo | 1,5e+13 cada ZFN | 8001          |
|       | Pareja de ZFN 2, Optimización de codones Fok1 WT, ZFN solo | 1,5e+13 cada ZFN | 8002          |

- 30 Los ensayos por inmunoabsorción ligados a enzimas (ELI SPOT, véase Markusic *et al* (2013), EMBO Mol Med 5:1698-1709) se realizaron en el bazo y el tejido de los ganglios linfáticos mesentéricos aislados de los animales en el día 65 y, como se muestra en la Figura 8, no existe ninguna respuesta inmunitaria provocada contra la cápside de AAV8 o contra los transgenes de ZFN en los animales. Los animales 7001 y 7002 (Figura 8, paneles A y C) así como los animales 8001 y 8002 (Figura 8, paneles B y C) fueron todos negativos para la respuesta de anticuerpos.

- 35 Estudio N° 2: En un estudio separado, se evaluaron en tres grupos de dos animales la inserción mediada por ZFN de un transgén F9 en el locus de la albúmina de macacos Rhesus. Véase, Figura 11.

Se obtuvieron resultados ilustrativos usando la Pareja 2 como se describió anteriormente, que comprendía dominios de escisión de nucleasa FokI de tipo silvestre (marcados como "FokI WT") o dominios modificados (marcados como

"FokI eHiFi", véase la patente de Estados Unidos N° 8.623.618) como se indica, en un vector AAV2/8 o AAV2/6. A los animales que recibieron los AAV que contenían el donante se les administró el donante de F9 (con brazos de homología con la albúmina) en un vector AAV2/8. Véase la Tabla 4 a continuación. En la tabla, las dosis "Alta" y "Baja" se refieren a la cantidad total de AAV administrado.

5

**Tabla 4: Estudio en PNH N° 2; Grupos de dosificación**

| Grupo | Descripción  | Serotipo de AAV | Dosis                             | ID del animal |
|-------|--|-----------------|-----------------------------------|---------------|
| 1     | Control negativo   | -               | Ninguno                           | 1001          |
| 6     | Pareja de ZFN 2, FokI-WT, + Donante de F9 (relación ZFN:donante 1:5), Dosis alta | AAV2/6          | 1,5e+13 cada ZFN, 1,5e+14 donante | 6101          |
|       |  |                 | 1,5e+13 cada ZFN, 1,5e+14 donante | 6102          |
| 7     | Solo donante   | -               | 1,5e+14                           | 7001          |

Los animales que solo recibieron ZFN (sin donante) mostraron una escisión robusta (0,4 - 4,1 %) el día 14 después de la administración.

10

Se realizó una transferencia Western en las muestras para evaluar la expresión de ZFN. Además, se detectó la expresión de la proteína F.IX en el plasma en animales que habían recibido tanto las ZFN como los vectores donantes. En presencia de ambos, ZFN y donante, los altos niveles de hFIX en plasma fueron detectables y aumentaron con el tiempo.

15

En conjunto, estos datos muestran que la inserción dirigida de un donante de Factor 8 o Factor 9 en el locus de la albúmina aumenta los niveles de actividad, incluso hasta el 50 % del FVIII normal en ratones con hemofilia A. Asimismo, con la optimización de las construcciones de donante y ZFN (p.ej., optimización de codones, la inclusión de sitios de glicosilación y/o administración de ZFN en vectores separados), la dosis de AAV puede reducirse manteniéndose la actividad de los transgenes. De hecho, una sola coinyección intravenosa de AAV que codifica cada una de las ZFN específicas de la albúmina con un donante de hF.IX dio como resultado la escisión de ADN detectable y la expresión de hF.IX en el plasma de macacos Rhesus.

20

### Ejemplo 3: Diseño, construcción y caracterización general de nucleasas específicas de la albúmina humana

25

Las nucleasas (p. ej., ZFN, TALEN, CRISPR/Cas) dirigidas a la albúmina se describen en las publicaciones de patente de Estados Unidos números 20130177983 y 20130177960 y en la patente de Estados Unidos N° 9.873.894). Para estos experimentos, se usaron ZFN que comprenden las ZFP (operablemente unidas a los dominios de escisión modificados) para escindir el locus de la albúmina endógeno en células humanas. Las parejas específicas de la albúmina humana se muestran a continuación en la Tabla 5. Todas las nucleasas de la Tabla 5 se unen a sus dianas.

30

**Tabla 5: Diseños de nucleasa específica de la albúmina humana**

| SBS N° , Diana   | Diseño                        |                              |                              |                               |                               |                              |
|--|-------------------------------|------------------------------|------------------------------|-------------------------------|-------------------------------|------------------------------|
| ZFN específicas de la albúmina humana  |                               |                              |                              |                               |                               |                              |
|  | F1                            | F2                           | F3                           | F4                            | F5                            | F6                           |
| SBS N° 35396<br>(humana/rhesus)<br>ccTATCCATTGCAC<br>TATGCTTtatttaa<br>(SEQ ID NO:7) | QSSDLSR<br>(SEQ ID<br>NO:8)   | LKWNLR<br>(SEQ ID<br>NO:9)   | DQSNLRA<br>(SEQ ID<br>NO:10) | RPYTLRL<br>(SEQ ID<br>NO:11)  | QSSDLSR<br>(SEQ ID<br>NO:8)   | HRSNLNK<br>(SEQ ID<br>NO:12) |
| SBS N° 39330<br>(humana)<br>ttTGGGATAGTTAT<br>GAAttcaatcttca<br>(SEQ ID<br>NO:103)   | QSGNLR<br>(SEQ ID NO:<br>2)   | LKQNL<br>(SEQ ID<br>NO:104)  | WQSNLQ<br>(SEQ ID<br>NO:105) | TSGNLTR<br>(SEQ ID<br>NO:106) | RQSHLCL<br>(SEQ ID<br>NO:107) | NA                           |
| SBS N° 43116<br>(humana)<br>ccTATCCATTGCAC<br>TATgctttatttaa<br>(SEQ ID NO:7)        | LKWNLR<br>(SEQ ID<br>NO:9)    | DQSNLRA<br>(SEQ ID<br>NO:10) | RNFSLTM<br>(SEQ ID<br>NO:15) | QSSTLDT<br>(SEQ ID<br>NO:108) | HRSNLNK<br>(SEQ ID<br>NO:12)  | NA                           |
| SBS N° 47171<br>(humana)   | QSGNLSR<br>(SEQ ID<br>NO:109) | LKQNL<br>(SEQ ID<br>NO:104)  | WADNLQ<br>(SEQ ID<br>NO:110) | TSGNLTR<br>(SEQ ID<br>NO:106) | RQSHLCL<br>(SEQ ID<br>NO:107) | NA                           |

(continuación)

| SBS N° , Diana  | Diseño                       |                               |                               |                               |                               |                              |
|---|------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|------------------------------|
| ZFN específicas de la albúmina humana   |                              |                               |                               |                               |                               |                              |
|   | F1                           | F2                            | F3                            | F4                            | F5                            | F6                           |
| ttTGGGATAGTTAT<br>GAAttcaatcttca<br>(SEQ ID<br>NO:103)                                      |                              |                               |                               |                               |                               |                              |
| <b>SBS N.º47931<br/>(humana)</b><br>ccTATCCATTGCAC<br>CTATGCTttatttta<br>a (SEQ ID<br>NO:7) | TPQLLDR<br>(SEQ ID<br>NO:14) | LKWNLR<br>(SEQ ID<br>NO:9)    | DQSNLNA<br>(SEQ ID<br>NO:111) | RNFSLTM<br>(SEQ ID<br>NO:15)  | LRHDLDR<br>(SEQ ID<br>NO:16)  | HRSNLNK<br>(SEQ ID<br>NO:12) |
| <b>SBS N° 47863<br/>(humana)</b><br>ttTGGGATAGTTAT<br>GAAttcaatcttca<br>(SEQ ID<br>NO:103)  | QSGNLAR<br>(SEQ ID NO:<br>2) | LIQYLQS<br>(SEQ ID<br>NO:112) | WADNLQN<br>(SEQ ID<br>NO:110) | TSGNLTR<br>(SEQ ID<br>NO:106) | RQSHLSL<br>(SEQ ID<br>NO:113) | NA                           |
| <b>SBS N° 47079<br/>(humana)</b><br>ccTATCCATTGCAC<br>TATGCTttatttta<br>(SEQ ID NO:7)       | TPQLLDR<br>(SEQ ID<br>NO:14) | LKWNLR<br>(SEQ ID<br>NO:9)    | DQSNLRA<br>(SEQ ID<br>NO:10)  | RNFSLTM<br>(SEQ ID<br>NO:15)  | LRHDLDR<br>(SEQ ID<br>NO:16)  | HRSNLNK<br>(SEQ ID<br>NO:12) |
| <b>SBS N.º47192<br/>(humana)</b><br>ttTGGGATAGTTAT<br>GAAttcaatcttca<br>(SEQ ID<br>NO:103)  | QSGNLAR<br>(SEQ ID NO:<br>2) | LIQYLQS<br>(SEQ ID<br>NO:112) | WADNLQN<br>(SEQ ID<br>NO:110) | TSGNLTR<br>(SEQ ID<br>NO:106) | RQSHLCL<br>(SEQ ID<br>NO:107) | NA                           |
| <b>SBS N.º47898<br/>(humana)</b><br>ccTATCCATTGCAC<br>TATGCTttatttta<br>(SEQ ID NO:7)       | TPQLLDR<br>(SEQ ID<br>NO:14) | LKHNL<br>(SEQ ID<br>NO:114)   | DQSNLNA<br>(SEQ ID<br>NO:111) | RNFSLTM<br>(SEQ ID<br>NO:15)  | LRHDLDR<br>(SEQ ID<br>NO:16)  | HRSNLNK<br>(SEQ ID<br>NO:12) |
| <b>SBS N.º47169<br/>(humana)</b><br>ccTATCCATTGCAC<br>TATGCTttatttta<br>(SEQ ID NO:7)       | TPQLLDR<br>(SEQ ID<br>NO:14) | LKWNLR<br>(SEQ ID<br>NO:9)    | DQSNLRA<br>(SEQ ID<br>NO:10)  | RNFSLTM<br>(SEQ ID<br>NO:15)  | LRHDLDR<br>(SEQ ID<br>NO:16)  | HRSNLNK<br>(SEQ ID<br>NO:12) |
| <b>SBS N.º47864<br/>(humana)</b><br>ttTGGGATAGTTAT<br>GAAttcaatcttca<br>(SEQ ID NO:103)     | QSGNLAR<br>(SEQ ID NO:<br>2) | LIQYLQS<br>(SEQ ID<br>NO:112) | WQSNLQN<br>(SEQ ID<br>NO:105) | TSGNLTR<br>(SEQ ID<br>NO:106) | RQSHLCL<br>(SEQ ID<br>NO:107) | N/A                          |
| <b>SBS N.º40477<br/>(humana)</b><br>ccTATCCATTGCAC<br>TATGCTttatttta<br>(SEQ ID NO:7)       | QSSDLR<br>(SEQ ID<br>NO:8)   | LKHNL<br>(SEQ ID<br>NO:114)   | LKHNL<br>(SEQ ID<br>NO:114)   | RPYTLR<br>(SEQ ID<br>NO:11)   | LRPDLR<br>(SEQ ID<br>NO:126)  | HRSNLNK<br>(SEQ ID<br>NO:12) |

En estos experimentos, las ZFN se transfectaron en las células en forma de ARNm y se introdujeron mediante nucleofección BTX mediante métodos estándar. Las concentraciones del ARNm de la ZFN variaron según el experimento. La actividad de NHEJ se midió mediante análisis MiSeq (Illumina), de acuerdo con los métodos conocidos en la técnica.

Para el ensayo en hepatocitos primarios humanos, se usaron 50 ng de ARN que codifica cada ZFN de la pareja. Los resultados se muestran a continuación (Tabla 6) y demuestran que todas las parejas de ZFN tenían actividad. Para el ensayo en HepG2 humanas se usaron 100 ng de ARN que codifica cada ZFN de la pareja. Los resultados se muestran a continuación (Tabla 7) y demuestran que todas las parejas de ZFN tenían actividad. Las células K562 humanas también se ensayaron usando 75 ng (Tabla 8) por duplicado y se demostró que las parejas eran activas.

**Tabla 6: Hepatocitos primarios humanos**

| ZFN derecha | ZFN izquierda |       |       |       |       |  | % Indeles |
|-------------|---------------|-------|-------|-------|-------|--|-----------|
|             | 47162         | 47171 | 47192 | 47863 | 47864 |  |           |
| 40477       | 8,9           | 6,5   | 6,1   | 11,7  | 14,0  |  |           |
| 47079       | 8,2           | 13,0  | 12,3  | 13,7  | 13,7  |  |           |
| 47169       | 11,2          | 9,6   | 11,6  | 20,3  | 14,1  |  |           |
| 47898       | 12,1          | 10,8  | 11,1  | 16,8  | 14,6  |  |           |
| 47931       | 10,1          | 11,8  | 11,5  | 17,2  | 14,7  |  |           |

**Tabla 7: Células HepG2 humanas**

| ZFN derecha | ZFN izquierda |       |       |       |       |  | % Indeles |
|-------------|---------------|-------|-------|-------|-------|--|-----------|
|             | 47162         | 47171 | 47192 | 47863 | 47864 |  |           |
| 40477       | 20,3          | 21,3  | 23,7  | 22,2  | 18,4  |  |           |
| 47079       | 20,8          | 25,8  | 22,9  | 23,5  | 16,0  |  |           |
| 47169       | 21,0          | 22,2  | 22,2  | 21,6  | 17,7  |  |           |
| 47898       | 21,5          | 22,0  | 13,9  | 22,4  | 14,3  |  |           |
| 47931       | 23,1          | 19,4  | 23,1  | 20,9  | 21,3  |  |           |

**Tabla 8: Células K562**

| Pareja de ZFN | 1    | 2    | PROMEDIO%Indeles |
|---------------|------|------|------------------|
| 47171:47931   | 73,9 | 75,7 | 74,8             |
| 47171:47079   | 77,1 | 81,7 | 79,4             |
| 47171:47898   | 75,3 | 68,6 | 72               |
| 47863:47931   | 76,9 | 80,9 | 78,9             |
| 47863:47079   | 86,8 | 86,7 | 86,8             |
| 47863:47898   | 91,1 | 92,7 | 91,9             |
| 47192:47931   | 69,9 | 76,1 | 73               |
| 47192:47079   | 80   | 75,9 | 77,9             |
| 47192:47898   | 70,1 | 72,3 | 71,2             |

5

**Ejemplo 5: Integración de los donantes para LSD usando ZFN específicas de la albúmina humana.**

Para estos experimentos, el donante de ADNc de IDS o IDUA se administró mediante una partícula AAV2/8, donde los transgenes de ADNc comprendían brazos de homología con las regiones de la albúmina que flanquean el sitio de corte. En estas construcciones de donantes, el gen terapéutico estaba flanqueado por secuencias homólogas al gen de la albúmina. En 5' del transgén, las construcciones de donante contienen secuencias homólogas al intrón 1 de la albúmina murina, mientras que en 3' del gen, las construcciones contienen secuencias homólogas con el límite del intrón 1-exón 2 de la albúmina murina (como se describe en la publicación de patente de Estados Unidos 2014-0017212).

10

15

Para integrar los transgenes de ADNc de IDS o IDUA y analizar su expresión, se transfectaron nucleasas de dedos de zinc específicas de la albúmina en forma de ARNm en células HepG2/C3a humanas. Brevemente, se transfectaron 100.000 células mediante administración viral por métodos estándar. La MOI X1000 para zfn:zfn:donante fue 100:100:200 ("baja" o "L") o 300:300:600 ("alta" o "H"). La expresión se analizó ensayando la actividad enzimática de la proteína codificada por el transgén en el sobrenadante celular o realizando transferencias Western en los sedimentos celulares después de 6 días. Los datos, mostrado en la figura 13, demuestran la expresión de la IDS y la IDUA del donante en las células.

20

25

Las descripciones y ejemplos anteriores no deben interpretarse como limitantes.

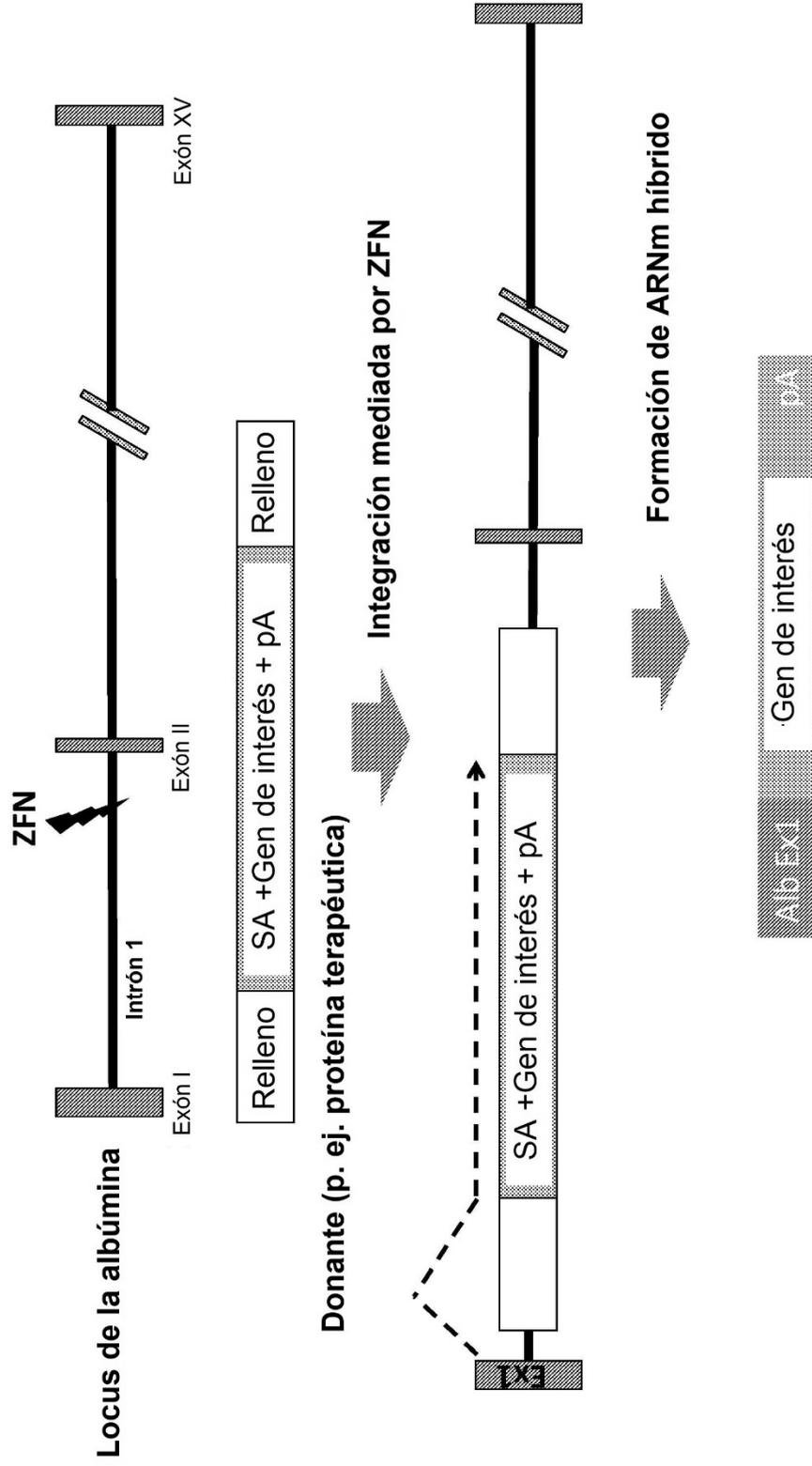
## REIVINDICACIONES

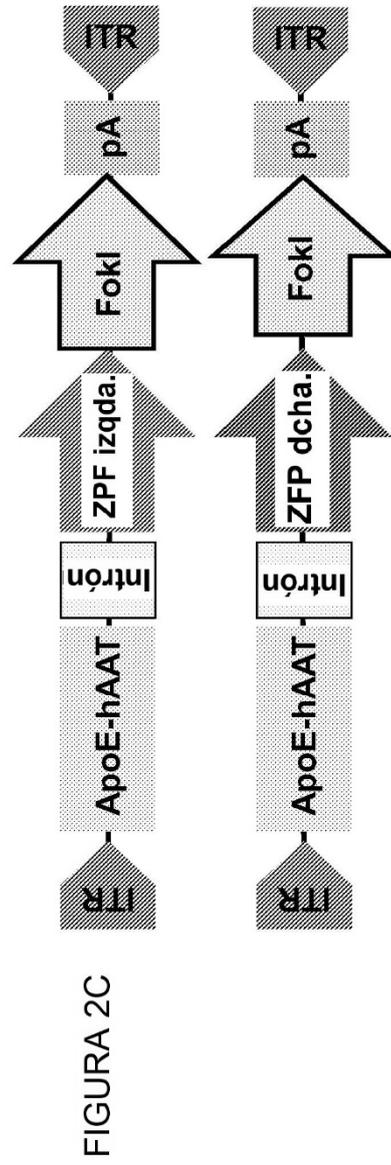
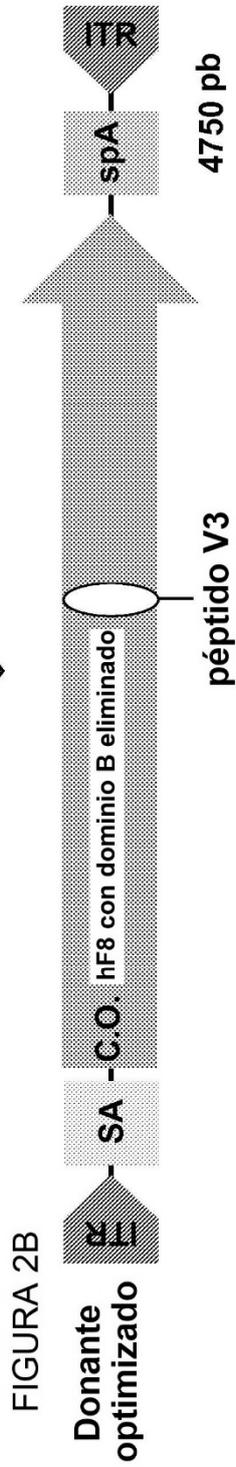
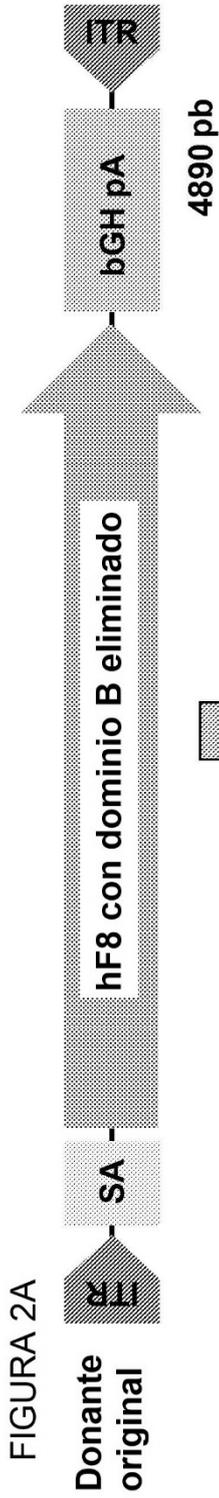
1. Una pareja de nucleasas de dedo de zinc (ZFN) que comprende una ZFN izquierda y una ZFN derecha, comprendiendo cada ZFN un dominio de escisión o semidominio de escisión de tipo silvestre o modificado y cinco o seis dominios de dedo de zinc designados y ordenados como F1 a F5 o F1 a F6, comprendiendo F1 a F5 o F1 a F6 las regiones de la hélice de reconocimiento de la siguiente manera:
- (i) comprendiendo F1 a F5 las SEQ ID NO: 109, 104, 110, 106 y 107, respectivamente (SBS N° 47171);
  - (ii) comprendiendo F1 a F6 las SEQ ID NO: 14, 114, 111, 15, 16 y 12, respectivamente (SBS N° 47898);
  - (iii) comprendiendo F1 a F6 las SEQ ID NO: 14, 9, 10, 15, 16 y 12, respectivamente (SBS N° 47079 o SBS N° 47169);
  - (iv) comprendiendo F1 a F6 las SEQ ID NO: 14, 9, 111, 15, 16 y 12, respectivamente (SBS N° 47931);
  - (v) comprendiendo F1 a F5 las SEQ ID NO: 2, 112, 110, 106 y 113, respectivamente (SBS N° 47863);
  - (vi) comprendiendo F1 a F5 las SEQ ID NO: 2, 112, 110, 106 y 107, respectivamente (SBS N° 47192);
- en donde la ZFN izquierda comprende las regiones de la hélice de reconocimiento de los dedos de zinc seleccionadas de (i), (v) y (vi), y la ZFN derecha comprende las regiones de la hélice de reconocimiento de los dedos de zinc seleccionadas de (ii), (iii) y (iv).
2. La pareja de nucleasas de dedo de zinc de la reivindicación 1, en donde la ZFN izquierda comprende las regiones de la hélice de reconocimiento de los dedos de zinc de (i) y la ZFN derecha comprende las regiones de la hélice de reconocimiento de los dedos de zinc de (ii).
3. Uno o más polinucleótidos que codifican la pareja de nucleasas de dedos de zinc de la reivindicación 1 o la reivindicación 2.
4. Uno o más vectores de expresión que comprenden el uno o más polinucleótidos de acuerdo con la reivindicación 3.
5. El uno o más vectores de expresión de la reivindicación 4, en donde el vector es un vector AAV, el cual es opcionalmente un vector AAV2/8.
6. Una composición farmacéutica que comprende un vector de expresión de las reivindicaciones 4 o 5.
7. Un método *in vitro* o *ex vivo* para escindir un gen de la albúmina en una célula, comprendiendo el método: introducir, en la célula, uno o más vectores de expresión de acuerdo con las reivindicaciones 4 o 5, en condiciones tales que la una o más proteínas se expresen y el gen de la albúmina se escinda, comprendiendo opcionalmente dicho método además integrar una secuencia donante en el gen de la albúmina escindido.
8. El método de la reivindicación 7, en donde la secuencia donante se introduce usando un vector AAV, el cual es opcionalmente un vector AAV2/8.
9. El método de las reivindicaciones 7 u 8, en donde la secuencia donante codifica una proteína inexistente o deficiente en una enfermedad de almacenamiento lisosomal (LSD), y/o en donde la célula es una célula hepática.
10. Uno o más vectores de expresión de acuerdo con las reivindicaciones 4 o 5 para usar en un método de tratamiento de un paciente, en donde dicho método comprende escindir un gen de la albúmina en una célula de dicho paciente e integrar una secuencia que codifica una proteína funcional en el locus de la albúmina.
11. Un vector de expresión de acuerdo con las reivindicaciones 4 o 5, o una composición farmacéutica de acuerdo con la reivindicación 6, para usar en un método para tratar a un paciente con una enfermedad de almacenamiento lisosomal, comprendiendo el método administrar al paciente dicho vector de expresión o composición farmacéutica que media la integración dirigida de un transgén que codifica una proteína funcional inexistente en un sujeto con una enfermedad de almacenamiento lisosomal en un gen de la albúmina endógeno.
12. Un vector de expresión o una composición para usar de acuerdo con la reivindicación 11, en donde el método comprende administrar:
- (i) un primer polinucleótido que codifica una nucleasa de dedo de zinc de acuerdo con la reivindicación 1, comprendiendo la nucleasa de dedo de zinc un dominio de escisión de FokI;
  - (ii) un segundo polinucleótido que codifica una nucleasa de dedo de zinc de acuerdo con la reivindicación 1, comprendiendo la nucleasa de dedo de zinc un dominio de escisión de FokI; y
  - (iii) un vector de expresión que comprende un donante que codifica una proteína inexistente en un sujeto con una enfermedad de almacenamiento lisosomal.
13. Un vector de expresión o una composición para usar de acuerdo con la reivindicación 12, en donde:

el primer polinucleótido comprende una nucleasa de dedo de zinc que comprende las regiones de la hélice de reconocimiento especificadas en (ii) de la reivindicación 1; y  
el segundo polinucleótido comprende una nucleasa de dedo de zinc que comprende las regiones de la hélice de reconocimiento especificadas en (i) de la reivindicación 1.

- 5
14. Un vector de expresión o composición para usar de acuerdo con una cualquiera de las reivindicaciones 11 a 13, en donde el vector o los vectores de expresión o la composición farmacéutica y el transgén se administran por vía intravenosa.
- 10 15. Un kit que comprende un vector de expresión de acuerdo con las reivindicaciones 4 o 5.

Figura 1

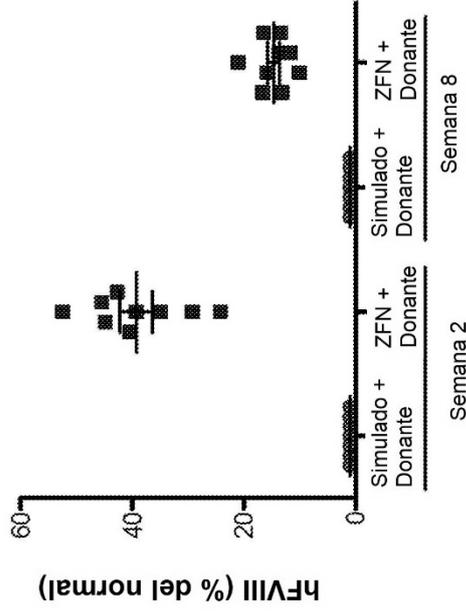




**Figura 3A**

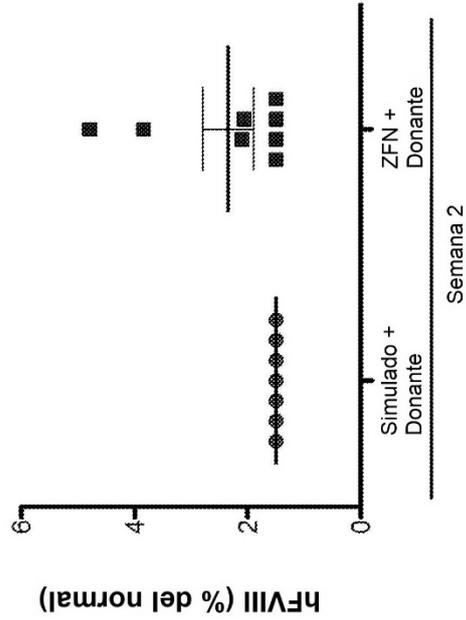


4890 pb



AAV8-ZFN ( $5e^{11}$ vg/ratón) +  
AAV8-Donante ( $5e^{11}$ vg/ratón)

**Figura 3C**



AAV8-ZFN ( $1e^{11}$ vg/ratón) +  
AAV8-Donante ( $1e^{11}$ vg/ratón)

**Figura 3B**

**Figura 3D**

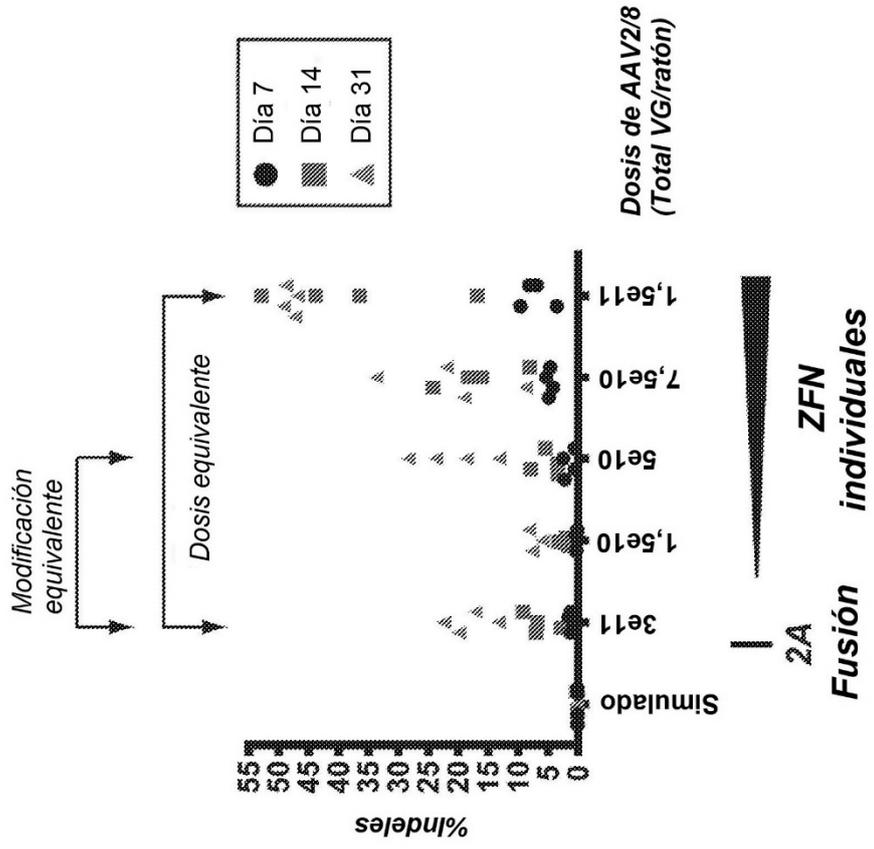


Figura 4

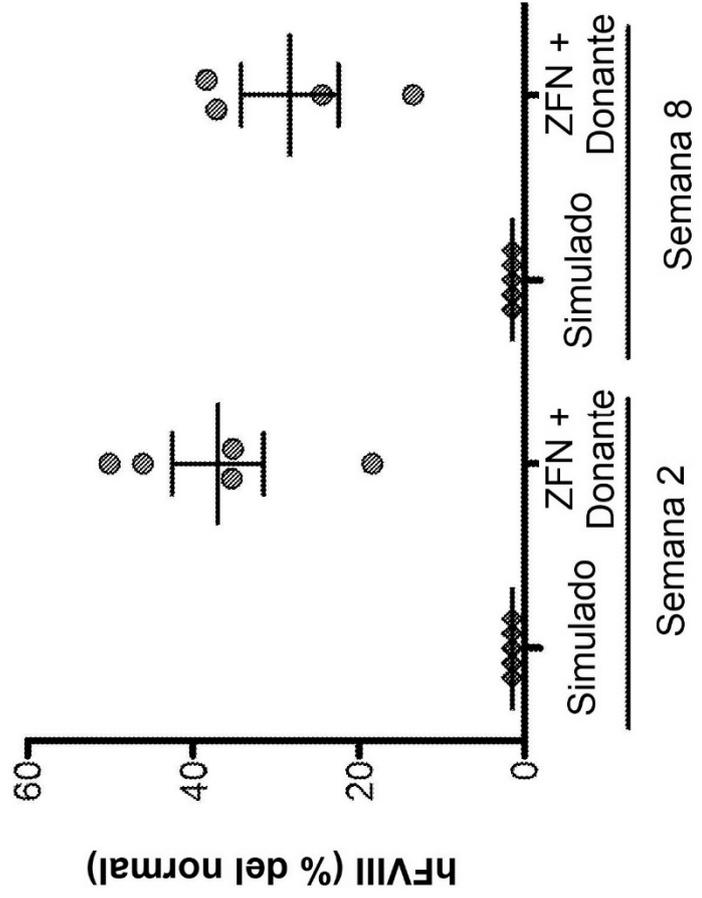


Figura 5A

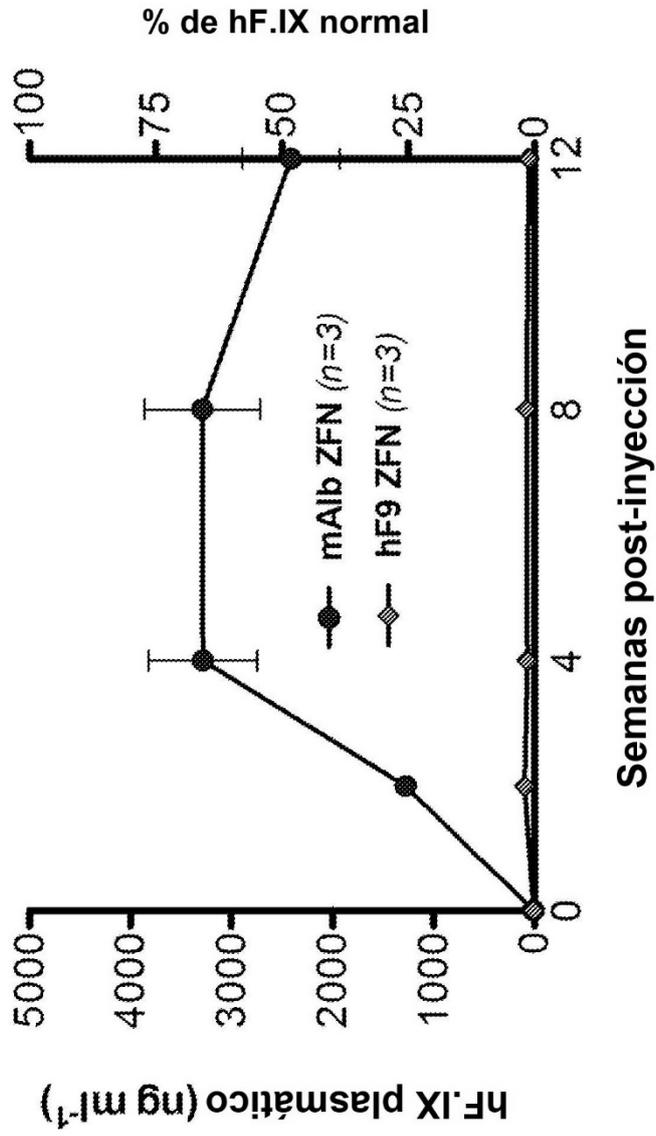
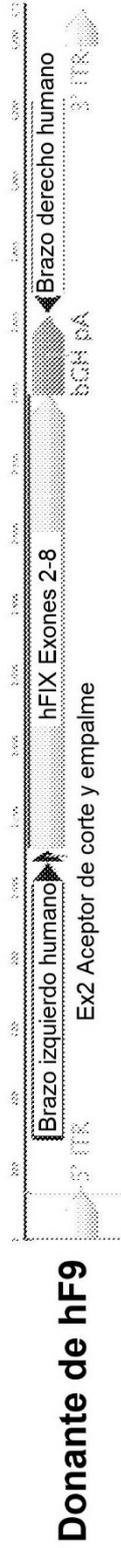
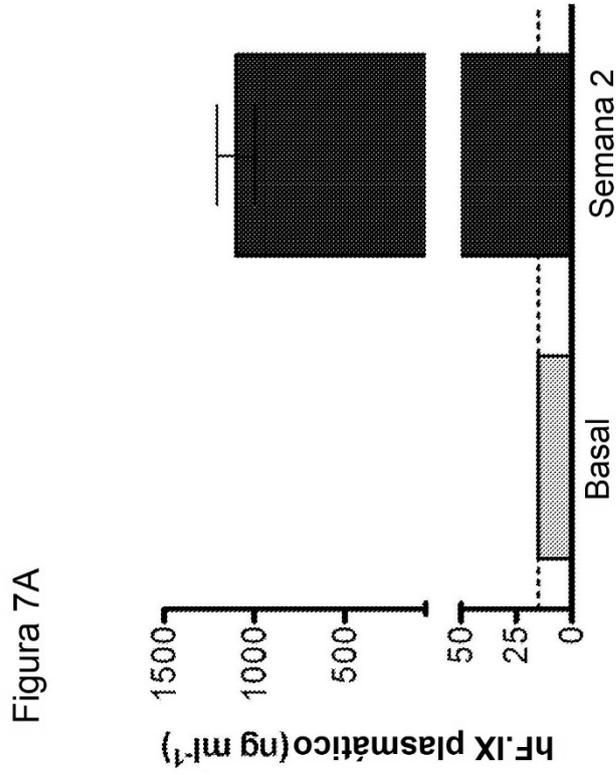
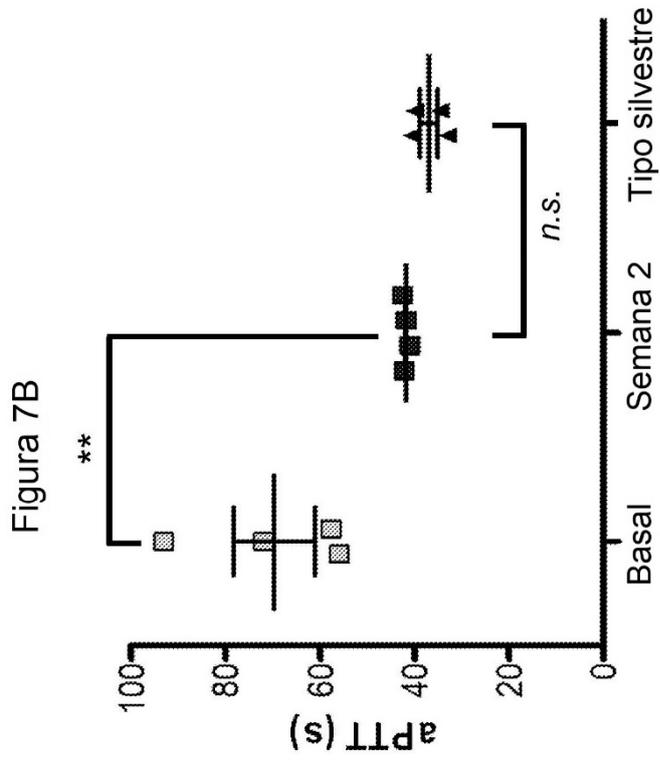


Figura 5B





**mAlb ZFN + Donante: AAV8-ZFN (1x10<sup>11</sup> vg/ratón) + AAV8-Donante (5x10<sup>11</sup> vg/ratón)**

Figura 8A

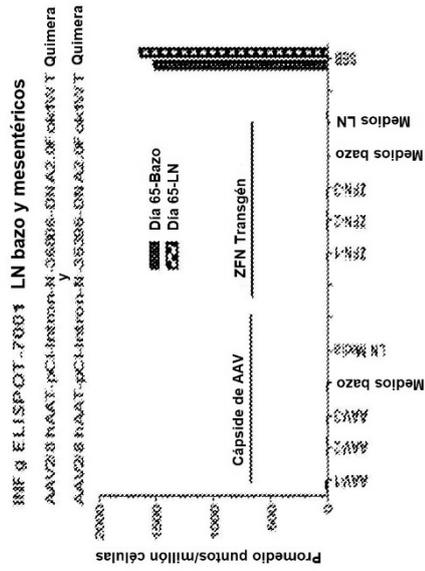


Figura 8B

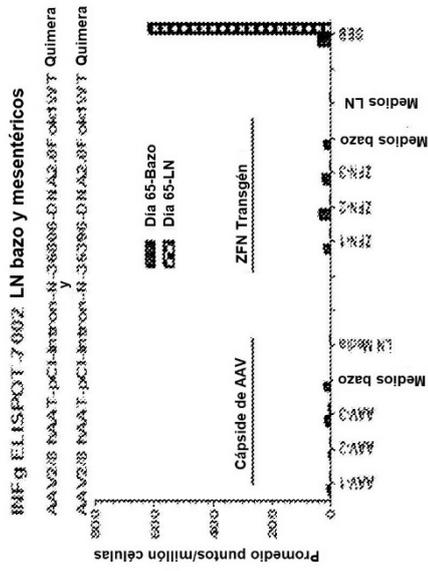


Figura 8C

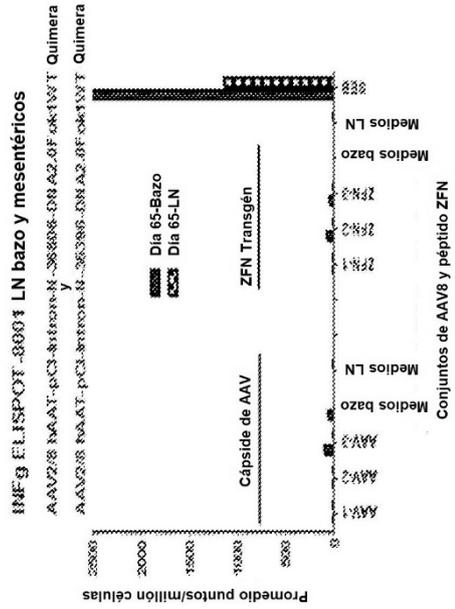
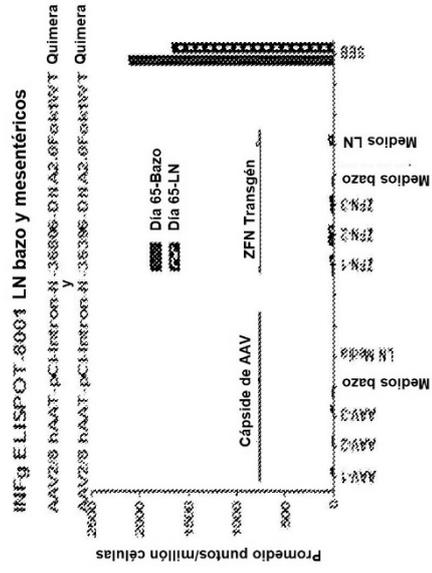
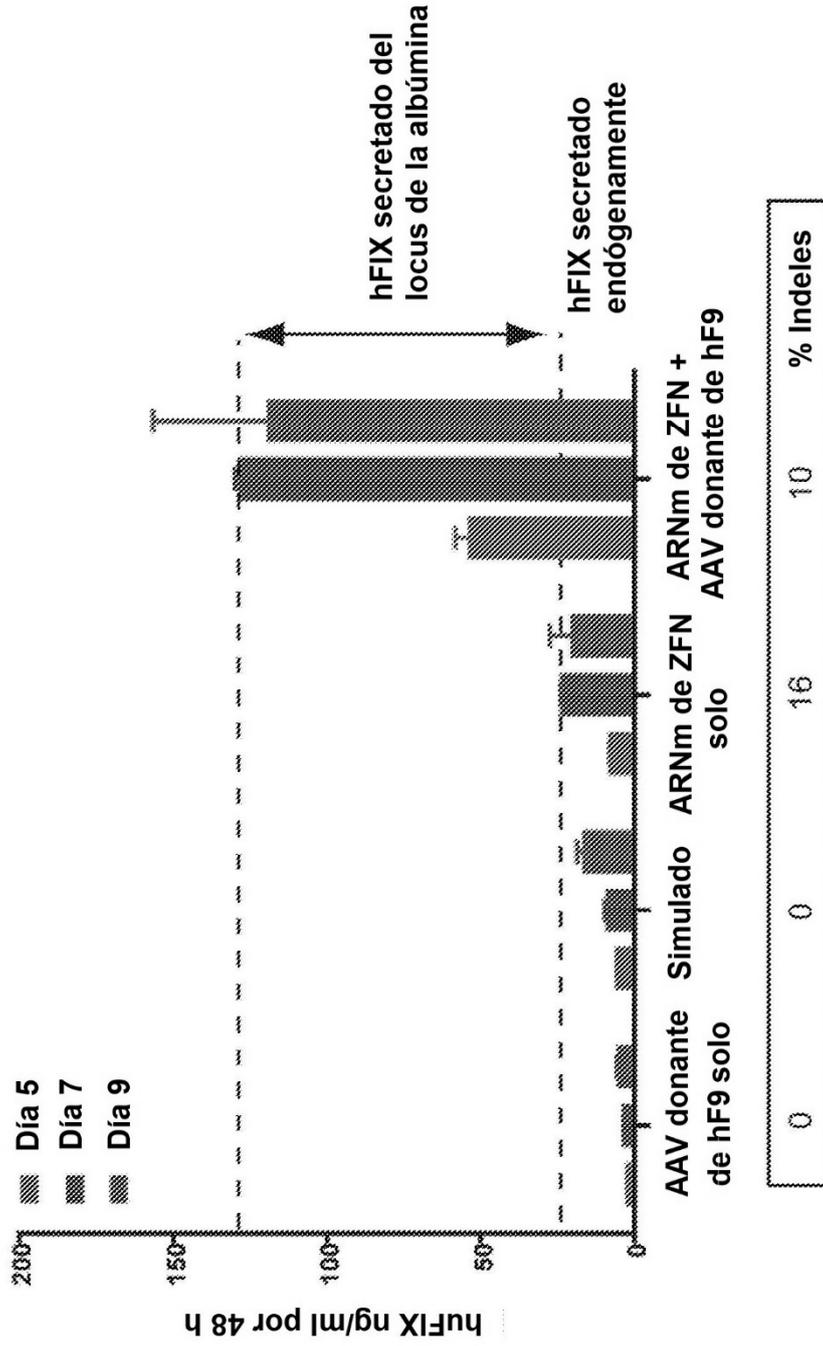


Figura 8D



**FIGURA 9**

**Hepatocitos primarios humanos**



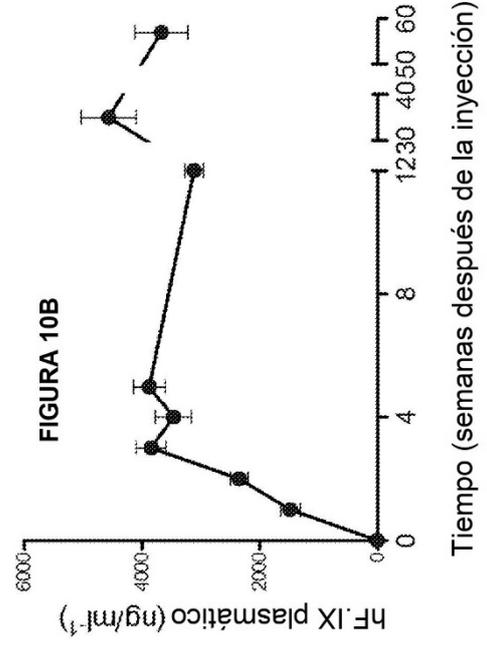
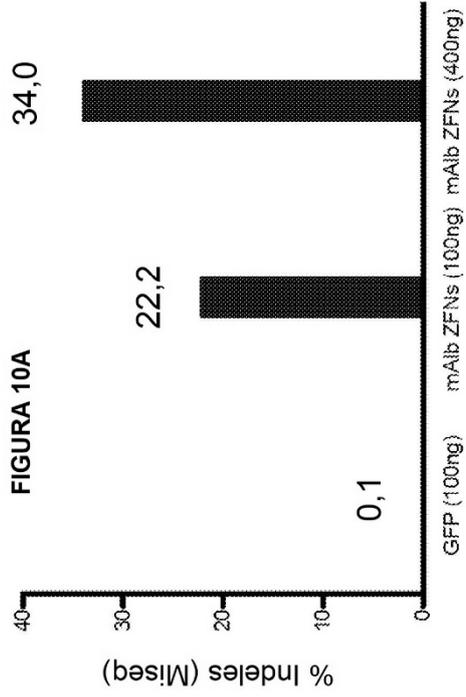
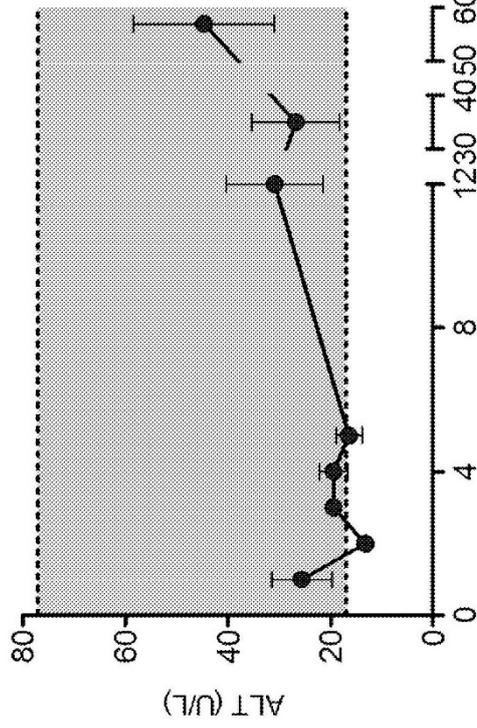


FIGURA 10C



Semanas post-inyección

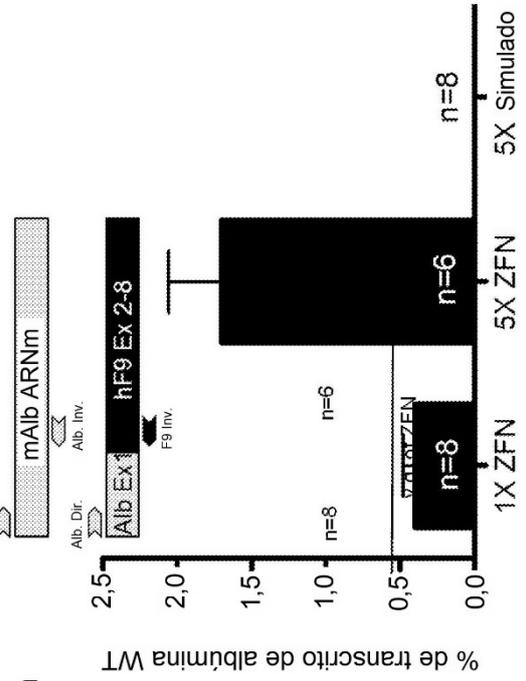
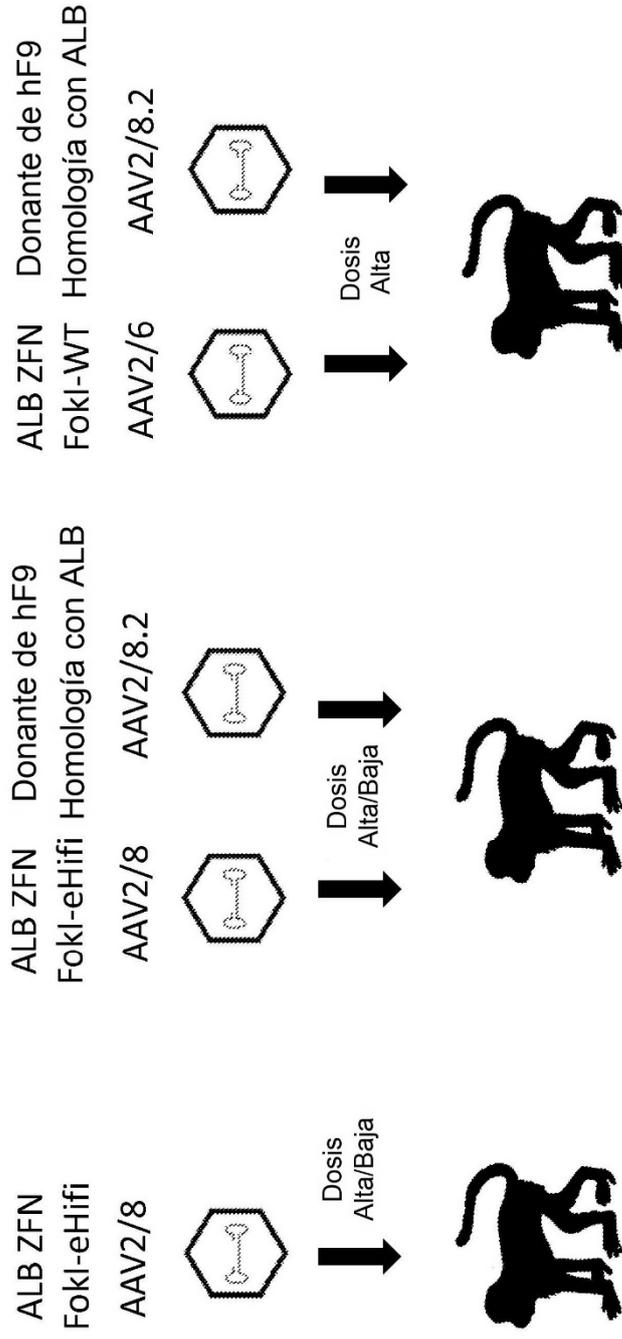


FIGURA 10D

**Figura 11**



**Análisis:**

- Biopsias de hígado para los genomas del vector y corte de ZFN (% indeles por miSEQ)
- Recogida de plasma para ELISA de hFIX
- Niveles de enzimas hepáticas (ALT-AST)
- Histopatología del hígado
- ELISPOT de hígado, bazo y PBMC

**Figura 12A**

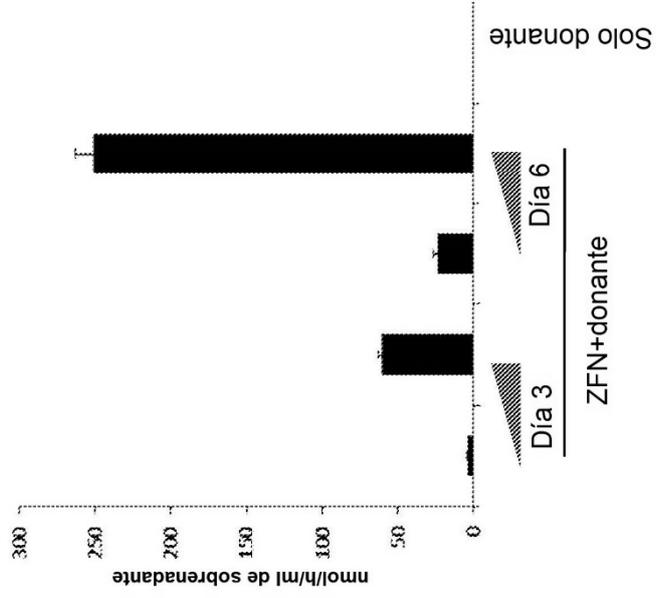
MDYKDHGDYKDHDIDYKDDDDKMAPKKRKYGIHGVPAAEAERPFQCRI  
 CMRKFATSGLTRHTKIHTGEKPFQCRICMRNFSRSDALSTHIRTHTGEK  
 PFACDICGRKFAQSATRTKHTKIHTHPRAPIPKPFQCRICMRNFSTSGHL  
SRHIRTHTGEKPFACDICGRKFAQSGNLARHTKIHLRGSQLVKELEEKK  
 SELRHKLKYPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLG  
 GSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQT  
 RDKHLNPNEWKVPSSVTEFKFLFVSGHFKGNKYKAQLTRLNHITNCNGA  
 VLSVEELIGGEMIKAGTLTLEEVRKFNNGEINF

MDYKDHGDYKDHDIDYKDDDDKMAPKKRKYGIHGVPAAEAERPFQCRI  
 CMRNFSRSDDLSAHIRTHTGEKPFACDICGRKFATKSNRTKHTKIHTGSQ  
 KPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAWRSSLRAHT  
 KIHTGEKPFQCRICMRKFADSSDRKKHTKIHLRGSQLVKELEEKKSELR  
 HKLKYPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGSRK  
 PDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRVYKENVQTRNKH  
 INPNEWKVPSSVTEFKFLFVSGHFKGNKYKAQLTRLNRKTNCGAVLSV  
 EELIGGEMIKAGTLTLEEVRKFNNGEINF

**Figura 12B**

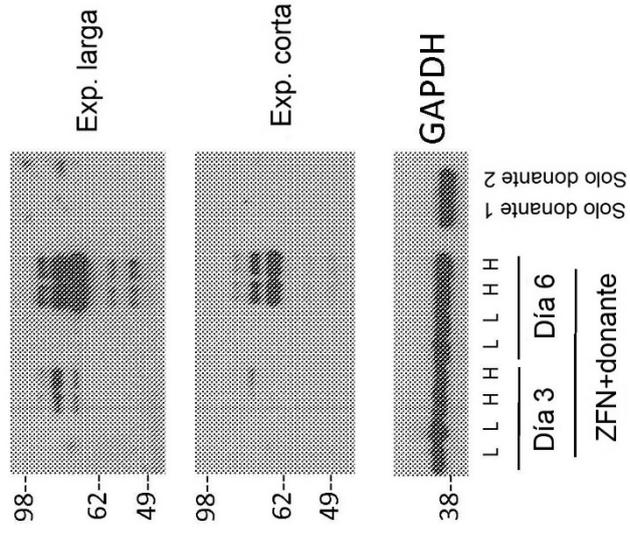
**Figura 13A**

Actividad enzimática  
(sobrenadante)



**Figura 13B**

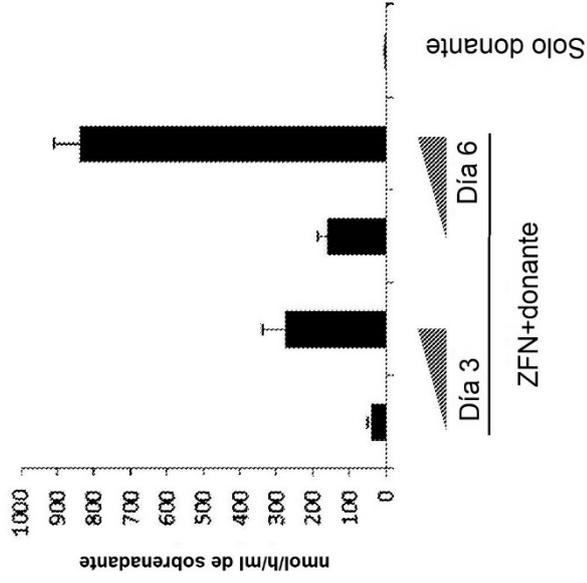
Expresión  
(Transf. Western, sedimento celular)



Donante de IDUA

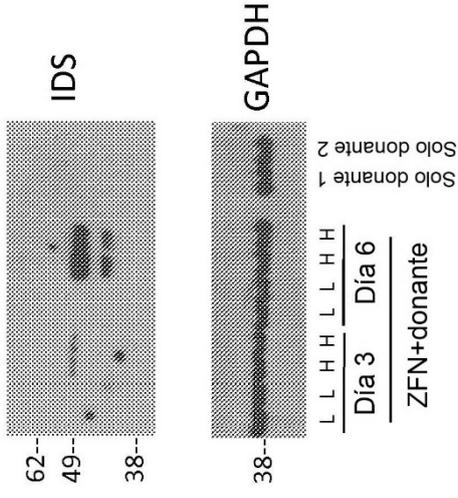
**Figura 13C**

Actividad enzimática  
(sobrenadante)



**Figura 13D**

Expresión  
(Transf. Western, sedimento celular)



Donante de IDS