

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 613 691**

51 Int. Cl.:

C12N 9/22 (2006.01)

A61K 48/00 (2006.01)

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.03.2013 PCT/US2013/032381**

87 Fecha y número de publicación internacional: **16.01.2014 WO2014011237**

96 Fecha de presentación y número de la solicitud europea: **15.03.2013 E 13715059 (5)**

97 Fecha y número de publicación de la concesión europea: **16.11.2016 EP 2872625**

54 Título: **Métodos y composiciones para el tratamiento de enfermedades por almacenamiento lisosomal**

30 Prioridad:

11.07.2012 US 201261670463 P
21.09.2012 US 201261704072 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.05.2017

73 Titular/es:

SANGAMO BIOSCIENCES, INC. (100.0%)
Point Richmond Tech. Center 501 Canal
Boulevard Suite A100
Richmond, CA 94804, US

72 Inventor/es:

REBAR, EDWARD, J.

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 613 691 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).

Métodos y composiciones para el tratamiento de enfermedades por almacenamiento lisosomal

Descripción

5 CAMPO TÉCNICO

La presente descripción se encuentra en el campo del tratamiento de las enfermedades por almacenamiento lisosomal (EAL) y la genoterapia.

10 ANTECEDENTES

La genoterapia constituye una fuente enorme de agentes terapéuticos para una nueva era de tratamientos para los humanos. Estas metodologías permitirán el tratamiento de afecciones que hasta entonces no se habían podido abordar con la práctica médica estándar. Un área que es especialmente prometedora es la capacidad de añadir un transgén a una célula para hacer que la célula exprese un producto que no se había producido antes en dicha célula. Los ejemplos de los usos de esta tecnología incluyen la inserción de un gen que codifica una proteína terapéutica, la inserción de una secuencia codificante que codifica una proteína que por algún motivo está ausente en la célula o en el individuo, y la inserción de una secuencia que codifica un ácido nucleico estructural, tal como un microARN.

Los transgenes se pueden introducir en una célula de muchas maneras, de tal manera que el transgén se integra en el propio genoma de la célula y se mantiene ahí. En los últimos años se ha desarrollado una estrategia para la integración de transgenes que utiliza la escisión con nucleasas específicas de sitio para dirigir la inserción en un locus genómico elegido (véase, p. ej., patente de los EE. UU. en copropiedad 7.888.121). Las nucleasas, tales como las nucleasas con dedos de cinc (NDZn), las nucleasas efectoras de tipo activador de la transcripción (TALEN) o los sistemas de nucleasas tales como el sistema CRISPR/Cas (que utiliza un ARN guía manipulado genéticamente) son específicas de los genes deseados y se pueden utilizar de tal manera que la construcción con el transgén se inserta mediante reparación dirigida por homología (HDR) o bien por los procesos impulsados por la captura de los extremos durante la unión de extremos no homólogos (NHEJ).

Los locus destinatarios incluyen los locus de «refugio», tales como los genes de AAVS1, HPRT y CCR5 en las células humanas y Rosa26 en las células murinas (véanse, p. ej., las publicaciones de patente de los Estados Unidos en copropiedad n.ºs 20080299580; 20080159996 y 201000218264, y la solicitud de patente de los Estados Unidos 13/660.821). La integración mediada por nucleasa ofrece la perspectiva de una mejor expresión del transgén, incremento de la seguridad y de la duración de la expresión, en comparación con las estrategias clásicas de integración que se basan en la integración aleatoria del transgén, lo que permite colocar el transgén con exactitud para que sea mínimo el riesgo de silenciamiento génico o de activación de oncogenes cercanos.

Mientras que la inserción del transgén en la célula destinataria es un obstáculo que se debe superar para implantar completamente esta tecnología, otro problema que se debe solventar es la garantía de que después de insertar el transgén en la célula y de que se exprese, el producto génico así codificado debe alcanzar la localización necesaria en el organismo y se debe producir a una concentración local que sea suficiente para resultar eficaz. Para las enfermedades caracterizadas por la ausencia de una proteína o por la presencia de una proteína afuncional aberrante, puede ser muy útil la introducción de una proteína de tipo silvestre codificada por un transgén.

Las enfermedades por almacenamiento lisosomal (EAL) son un grupo de enfermedades metabólicas monogénicas poco frecuentes que se caracterizan por la ausencia de determinadas proteínas lisosomales funcionales que normalmente están implicadas en la degradación de lípidos, glucoproteínas y mucopolisacáridos de desecho. Estas enfermedades se caracterizan por una acumulación de estos compuestos en la célula al ser incapaz de procesarlos para el reciclaje debido al mal funcionamiento de una enzima específica. Los ejemplos más frecuentes son la enfermedad de Gaucher (deficiencia de la glucocerebrosidasa cuyo gen tiene por nombre GBA), enfermedad de Fabry (deficiencia de la galactosidasa α , GLA), enfermedad de Hunter (deficiencia de la iduronato-2-sulfatasa, IDS), enfermedad de Hurler (deficiencia de la α -L-iduronidasa, IDUA) y enfermedad de Niemann-Pick (deficiencia de la esfingomielina fosfodiesterasa 1, SMPD1). Cuando se agrupan todas juntas, las EAL tienen una incidencia en la población de aproximadamente 1 de cada 7.000 nacimientos. Estas enfermedades tienen efectos devastadores en los que las padecen. Se suelen diagnosticar primero en los niños que podrían tener unos rasgos de crecimiento corporal y faciales característicos, y que podrían tener un retraso mental de moderado a grave. Las opciones de tratamiento incluyen el tratamiento enzimático sustitutivo (ERT), en donde la enzima ausente se da al paciente, normalmente por inyección intravenosa en dosis grandes. Tal tratamiento es sólo para tratar los síntomas, no para curar, con lo que hay que administrar al paciente numerosas dosis de estas proteínas durante el resto de su vida, y es posible que desarrollen anticuerpos neutralizantes contra la proteína inyectada. A menudo, estas proteínas tienen una semivida breve en el suero, por lo que el paciente debe también soportar con frecuencia infusiones de la proteína. Por ejemplo, los pacientes con la enfermedad de Gaucher tratados con el producto Cerezyme® (imiglucerasa) deben recibir infusiones tres veces por semana. También resultan problemáticas la producción y la purificación de las enzimas, y de igual forma los tratamientos son muy costosos (más de 100.000 \$ al año por paciente).

Así pues, existe la necesidad de más métodos y composiciones que se puedan utilizar para tratar una enfermedad

monogénica (p. ej., enfermedades por almacenamiento lisosomal) a través de la edición del genoma, y métodos para introducir un producto génico expresado desde el transgén que lo codifica a un nivel relevante desde el punto de vista terapéutico.

5 COMPENDIO DE LA INVENCION

La invención da a conocer un método *in vitro* para generar una célula que produce una proteína que está ausente en un sujeto con una enfermedad por almacenamiento lisosomal, en donde el método comprende: integrar un transgén que codifica la proteína en un locus de la albúmina dentro de la célula mediante el uso de una nucleasa que no se produce en la naturaleza, de tal manera que la célula produce la proteína que está ausente en la enfermedad por almacenamiento lisosomal.

La invención también da a conocer una célula generada por el método de la invención para ser usada en un método para tratar una enfermedad por almacenamiento lisosomal, en donde el método comprende la administración de la célula a un sujeto que lo necesita.

La invención también da a conocer una o varias nucleasas y uno o varios transgenes para ser usados en un método para tratar una enfermedad por almacenamiento lisosomal, en donde el método comprende la administración de una o varias nucleasas y uno o varios transgenes a un sujeto que lo necesita para generar una célula que produce una proteína que está ausente en el sujeto, en donde el transgén que codifica la proteína está integrado en un locus endógeno de la albúmina mediante el uso de una nucleasa que no se produce en la naturaleza, de tal manera que la célula produce la proteína que está ausente en la enfermedad por almacenamiento lisosomal.

COMPENDIO DE LA DESCRIPCIÓN

Se describen en la presente memoria métodos y composiciones para tratar una enfermedad monogénica. Se describen métodos para insertar una secuencia de transgén en una célula destinataria idónea, en donde el transgén codifica una proteína para el tratamiento de la enfermedad. La proteína terapéutica se puede excretar desde la célula destinataria, de tal manera que es capaz de afectar o ser tomada por otras células que no albergan el transgén. También se describen métodos para producir una célula (p. ej., una célula madura o indiferenciada) que produce gran cantidad de un agente terapéutico, en donde la introducción de una población de estas células alteradas en un paciente suministrará la proteína necesaria para tratar una enfermedad o afección.

Se describe en la presente memoria una proteína con dedos de cinc (PDZn) que se fija al sitio diana en una región de interés (p. ej., un gen asociado a la enfermedad, un gen muy expresado, un gen de la albúmina u otro gen, o un gen de refugio) en un genoma, en donde la PDZn comprende uno o varios dominios de fijación con dedos de cinc manipulados genéticamente. En una realización, la PDZn es una nucleasa con dedos de cinc (NDZn) que escinde una región genómica deseada de interés, en donde la NDZn comprende uno o varios dominios de fijación con dedos de cinc manipulados genéticamente y un dominio de escisión, o semidominio de escisión, de nucleasa. Los dominios de escisión y los semidominios de escisión se pueden obtener, por ejemplo, de diferentes endonucleasas de restricción y/o endonucleasas de asentamiento. En una realización, los semidominios de escisión proceden de una endonucleasa de restricción de tipo IIS (p. ej., *FokI*). En determinadas realizaciones, el dominio con dedos de cinc reconoce un sitio diana en un asociado a la enfermedad o un gen de refugio, tal como la albúmina (p. ej., una proteína con dedos de cinc que tiene 5 o 6 dedos con las regiones de la hélice de reconocimiento se muestran en una sola fila de la tabla 3).

Se describe en la presente descripción una proteína TALE (de tipo activador de la transcripción) que se fija a un sitio diana en una región de interés (p. ej., un gen muy expresado, un gen asociado a la enfermedad o un gen de refugio) en un genoma, en donde la TALE comprende uno o varios dominios de fijación TALE manipulados genéticamente. En una realización, la TALE es una nucleasa (TALEN) que escinde una región genómica deseada de interés, en donde la TALEN comprende uno o varios dominios TALE de fijación al ADN manipulados genéticamente y un dominio de escisión, o semidominio de escisión, de nucleasa. Los dominios de escisión y los semidominios de escisión se pueden obtener, por ejemplo, de diferentes endonucleasas de restricción y/o endonucleasas de asentamiento (*homings*). En una realización, los semidominios de escisión proceden de una endonucleasa de restricción de tipo IIS (p. ej., *FokI*). En determinadas realizaciones, el dominio TALE de fijación al ADN reconoce un sitio diana en un gen muy expresado, asociado a la enfermedad o de refugio.

Se describe en la presente memoria un sistema CRISPR/Cas que se fija al sitio diana en una región de interés (p. ej., un gen muy expresado, un gen asociado a la enfermedad o un gen de refugio) en un genoma, en donde el sistema CRISPR/Cas comprende una nucleasa CRISPR/Cas y un crRNA/tracrRNA (o un ARN guía único) manipulado genéticamente. En determinadas realizaciones, el sistema CRISPR/Cas reconoce un sitio diana en un gen muy expresado, asociado a la enfermedad o de refugio.

La NDZn, la TALEN y/o el sistema CRISPR/Cas, tal y como se describen en la presente memoria, podría fijarse a, 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 región no transcrita, bien delante o bien detrás de la región codificante. En algunas realizaciones de la descripción, la NDZn, la TALEN y/o el sistema

CRISPR/Cas se fija a, y/o escinde, un gen muy expresado, por ejemplo, un gen de globina en los glóbulos rojos (RBC). Véase, p. ej., la solicitud de patente de los EE. UU. n.º 61/670.451 titulada «Methods and Compositions for Delivery of Biologics», registrada el 11 de julio de 2012. En otras realizaciones de la descripción, la NDZn, la TALEN y/o el sistema CRISPR/Cas se fija a, y/o escinde, un gen de refugio, por ejemplo, un gen de CCR5, un gen de PPP1R12C (también conocido como AAV S1), de albúmina, HPRT o un gen *Rosa*. Véanse, p. ej., las publicaciones de patentes de los EE. UU. n.ºs 20080299580, 20080159996 y 201000218264 y la solicitudes de patente de los EE. UU. n.ºs 61/537.349, 61/560.506, 13/660.821 y la solicitud de patente de los EE. UU. n.º 61/670.451 titulada «Methods and Compositions for Regulation of Transgene Expression» registrada el 11 de julio de 2012. Además, para ayudar a la selección, se puede utilizar el locus de HPRT (véanse las solicitudes de patente de los EE. UU. 13/660.821 y 13/660.843). En otras realizaciones de la descripción, la NDZn, la TALEN y/o el sistema CRISPR/Cas se puede fijar a, y/o escindir, un gen asociado a la enfermedad (p. ej., el gen que codifica la hidrolasa lisosómica α -galactosidasa A (AGA), relacionada con la enfermedad de Fabry). En otro aspecto, en la presente memoria se describen composiciones que comprenden una o varias de las nucleasas con dedos de cinc y/o TALE o el sistema CRISPR/Cas descritos en la presente memoria. También se describen composiciones que comprenden una o varias de estas nucleasas y el ácido nucleico donante. En algunos aspectos, se describen nucleasas o sistemas CRISPR/Cas manipulados genéticamente capaces de escindir los genes reguladores aberrantes asociados a la enfermedad y los métodos para utilizar estas nucleasas para tratar la enfermedad mediante la eliminación o reducción de la expresión del producto génico aberrante.

En un aspecto, la invención describe un método para tratar una enfermedad por almacenamiento lisosomal mediante la introducción de un transgén correctivo en una célula destinataria adecuada (p. ej., célula sanguínea, hepatocito, célula del cerebro, célula troncal, célula precursora, etc.) de tal manera que se expresa el producto codificado por ese transgén correctivo. En una realización, el transgén correctivo se inserta en una línea celular para producir *in vitro* la proteína sustitutiva. Las células que comprenden el transgén o la proteína producida por las células se pueden utilizar para tratar un paciente que lo necesita, por ejemplo, tras la purificación de la proteína producida. En otra realización, el transgén correctivo se inserta en un tejido destinatario en el cuerpo, de tal manera que se produce *in vivo* la proteína sustitutiva. En algunos aspectos, la proteína expresada se excreta de la célula para actuar sobre otras células que carecen del transgén (p. ej., a través de la exportación a la sangre), o ser tomada por ellas. En algunos casos, el tejido destinatario es el hígado. En otros casos, el tejido destinatario es el cerebro. En otros casos, el destino es la sangre (p. ej., vasculatura). En otros casos, el destino es el músculo esquelético. En una realización, el gen correctivo comprende la secuencia de tipo silvestre del gen funcional, mientras que en otras realizaciones, la secuencia del transgén correctivo está alterada de alguna manera para que se le mejore su actividad biológica. En algunos aspectos, el transgén corrector comprende codones optimizados para incrementar la actividad biológica, mientras que, en otros aspectos, la secuencia está alterada para darle a la proteína resultante otra función deseada (p. ej., mejorar su estabilidad, alterar la carga para alterar la fijación al sustrato, etc.). En algunas realizaciones, el transgén está alterado para reducir la inmunogenia. En otros casos, el transgén está alterado de tal manera que la proteína codificada se vuelve un sustrato para la introducción mediada por transportador en determinados tejidos, tales como el cerebro (véase Gabathuler et al. (2010) *Neurobiology of Disease* 37: 48-57).

En otro aspecto, la invención da a conocer una proteína nucleásica manipulada genéticamente que es capaz de escindir (editar) el genoma de una célula troncal o precursora (p. ej., célula hematopoyética precursora, célula troncal del hígado, etc.) mediante la introducción de un transgén deseado. En algunos aspectos, las células troncales o precursoras editadas se expanden después y se les puede inducir su diferenciación en las células editadas maduras *ex vivo*, y a continuación las células se le dan al paciente. En otros aspectos, las células precursoras editadas (p. ej., células troncales CD34+) se dan en un trasplante de la médula ósea que, después de la implantación satisfactoria, proliferan y producen células editadas que a continuación se diferencian y maduran *in vivo*, y contienen el agente biológico expresado desde el transgén. En otros aspectos, las células troncales editadas son células troncales del músculo que a continuación se introducen en el tejido muscular. En algunos aspectos, la nucleasa manipulada genéticamente es una nucleasa con dedos de cinc (NDZn) y, en otros, la nucleasa es una nucleasa con TALE (TALEN), y en otros aspectos, se utiliza un sistema CRISPR/Cas. Las nucleasas se manipulan genéticamente para que tengan especificidad por el locus de refugio endógeno de la albúmina. Solo a modo de ejemplo no limitante, el locus de refugio puede ser el gen de la albúmina, mientras que el gen asociado a la enfermedad puede ser el gen *GLA* que codifica la hidrolasa lisosomal α -galactosidasa A (véase la tabla 2). Un gen que se expresa mucho en los glóbulos rojos (RBC) es la β -globina. En otro aspecto, las células transgénicas se sintetizan *ex vivo* mediante la electrosensibilización para incrementar su propensión a la ruptura tras la exposición a una fuente de energía (p. ej., ultrasonidos) (véase la solicitud de patente internacional WO 2002007752).

En la presente memoria se describe un polinucleótido que codifica una o varias NDZn, TALEN y/o sistema CRISPR/Cas descritos en la presente memoria. 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).

Se describe en la presente memoria un vector de expresión de NDZn, TALEN y/o el sistema CRISPR/Cas que comprende un polinucleótido, que codifica una o varias de NDZn, TALEN y/o sistema CRISPR/Cas descritos en la presente memoria, unido operativamente a un promotor. En una realización, el vector de expresión es un vector vírico.

Se describe en la presente memoria una célula hospedadora que comprende uno o varios vectores de expresión con NDZn, TALEN y/o sistema CRISPR/Cas tal y como se describe en la presente memoria. La célula hospedadora puede

estar transformada de forma estable o transfectada de forma transitoria, o una combinación de las mismas, con uno o varios vectores de expresión con NDZn, TALEN y/o el sistema CRISPR/Cas. En algunas realizaciones, la célula hospedadora es un hepatocito.

- 5 Se describe en la presente memoria un método para escindir un locus muy expresado, asociado a la enfermedad y/o de refugio en una célula, en donde el método comprende: introducir en la célula uno o varios polinucleótidos que codifican una o varias NDZn, TALEN y/o sistema CRISPR/Cas que se fija o fijan a un sitio diana en uno o varios locus destinatarios en tales condiciones que las NDZn, TALEN o sistema CRISPR/Cas se expresa o expresan y se escinden uno o varios locus. Los ejemplos no limitantes de NDZn, TALEN y/o sistemas CRISPR/Cas que se fijan a los locus muy expresados y/o de refugio se describen en las publicaciones de patente de los EE. UU. n.^{os} 20080299580; 20080159996; 10 y 201000218264, y las solicitudes de patente de los EE. UU. 13/660.821, 13/660.843, 13/624.193 y 13/624.217 y la solicitud de patente de los EE. UU. n.^o 61/670.451 titulada «Methods and Compositions for Delivery of Biologics».

- En otras realizaciones de la descripción, una secuencia genómica en cualquier gen destinatario está sustituida por el 15 transgén terapéutico, por ejemplo, mediante el uso de una NDZn, una TALEN y/o un sistema CRISPR/Cas (o vector que codifica dicha NDZn, TALEN y/o sistema CRISPR/Cas) tal y como se describe en la presente memoria, y una secuencia «donante» o transgén que está insertada en el gen después de la escisión dirigida con la NDZn, la TALEN y/o el sistema CRISPR/Cas. La secuencia donante puede estar presente en el vector de NDZn o TALEN, presente en un vector distinto (p. ej., vector de Ad, AAV o LV) o, como alternativa, se puede introducir en la célula mediante un 20 mecanismo diferente de introducción de ácidos nucleicos. Tal inserción de una secuencia nucleotídica donante en el locus destinatario (p. ej., gen muy expresado, gen asociado a la enfermedad, otro gen de refugio, etc.) da lugar a la expresión del transgén bajo control de los elementos de control genético endógenos del locus destinatario (p. ej., albúmina, globina, etc.). En algunos aspectos, la inserción del transgén de interés, por ejemplo, en un gen deseado (p. ej., albúmina) da lugar a la expresión de una secuencia de la proteína exógena intacta y carece de cualquier aminoácido 25 codificado por la diana (p. ej., albúmina). En otros aspectos, la proteína exógena expresada es una proteína de fusión y comprende aminoácidos codificados por el transgén y por el locus endógeno en el cual está insertado el transgén (p. ej., el locus destinatario endógeno o, como alternativa, de secuencias del transgén que codifican secuencias del locus destinatario). La diana puede ser cualquier gen, por ejemplo, un gen de refugio, tal como un gen de la albúmina, un gen de AAV S1, un gen de HPRT; un gen de CCR5; o un gen muy expresado, tal como un gen de globina en un glóbulo rojo 30 (p. ej., β -globina o γ -globina). En algunos casos, las secuencias endógenas estarán presentes sobre la porción del extremo amino (N) de la proteína exógena, mientras que, en otros, las secuencias endógenas estarán presentes sobre la porción del extremo carboxilo (C) de la proteína exógena. En otros casos, las secuencias endógenas estarán presentes sobre las porciones tanto del extremo N como C de la proteína exógena. Las secuencias endógenas pueden incluir secuencias endógenas de tipo silvestre o mutantes de longitud completa o, como alternativa, pueden incluir 35 secuencias aminoacídicas endógenas parciales. En algunas realizaciones, la fusión gen endógeno y el transgén está localizada en el locus endógeno dentro de la célula, mientras que en otras realizaciones, la secuencia codificante de secuencia endógena y el transgén está insertada en otro locus dentro de un genoma (p. ej., una secuencia del transgén de IDUA insertada en un locus de albúmina, de HPRT o de CCR5). En algunos aspectos, el refugio se selecciona de los locus de AAV S1, *Rosa*, albúmina, HPRT o CCR5 (véanse las publicaciones de patente de los EE. UU. en copropiedad 40 n.^{os} 20080299580; 20080159996 y 201000218264 y las solicitudes de patente de los EE. UU. 13/660.821, 13/660.843, 13/624.193 y 13/624.217 y la solicitud de patente de los EE. UU. n.^o 61/670.451 titulada «Methods and Compositions for Regulation of Transgen Expression», registrada el 11 de julio de 2011). En otras realizaciones, el gen asociado a la enfermedad se selecciona de GLA (la hidrolasa lisosomal α -galactosidasa A) o de uno o varios genes recogidos en la 45 tabla 2.

- En algunas realizaciones, el transgén se expresa de tal manera que un producto de proteína terapéutica queda retenido dentro de la célula (p. ej., célula precursora o madura). En otras realizaciones, el transgén está fusionado al dominio extracelular de una proteína de membrana, de tal manera que, tras la expresión, una fusión del transgén hará que la proteína terapéutica aparezca localizada en la superficie. En algunos aspectos, el dominio extracelular se elige de las 50 proteínas recogidas en la tabla 1. En algunos aspectos, las células editadas también comprenden una proteína transmembranaria para transportar las células a un tipo concreto de tejido. En un aspecto, la proteína transmembranaria es un anticuerpo, mientras que, en otros, la proteína transmembranaria es un receptor. En determinadas realizaciones, la célula es un precursor (p. ej., CD34+ o célula troncal hematopoyética) o glóbulo rojo maduro. En algunos aspectos, el producto de la proteína terapéutica codificada sobre el transgén se exporta al exterior de la célula para afectar o para 55 ser tomado por las células que carecen del transgén. En algunas realizaciones, la célula es un hepatocito que libera la proteína terapéutica al torrente circulatorio para actuar sobre los tejidos distales (p. ej., cerebro).

- La invención también da a conocer métodos y composiciones para la producción de una célula (p. ej., glóbulo rojo) que lleva una proteína terapéutica para una EAL que se puede utilizar de forma universal con todos los pacientes como un 60 producto alógeno. Esto permitiría el desarrollo de un único producto para el tratamiento de los pacientes con una EAL concreta, por ejemplo. Estos vehículos pueden comprender proteínas transmembranarias que ayudan al transporte de la célula. En un aspecto, la proteína transmembranaria es un anticuerpo, mientras que, en otros, la proteína transmembranaria es un receptor.

- 65 En un aspecto, la descripción da a conocer métodos y composiciones para genosuprimir los genes asociados a la enfermedad. En algunas realizaciones, estos genes son aquellos cuyos productos pueden regular la expresión de un

5 gen en una célula madura o precursora. En algunos aspectos, la genosupresión es para el sitio diana regulador del ADN para tales proteínas. En algunos aspectos, el gen regulador es aberrante de tal manera que la genosupresión del gen restaura la función normal. En otros aspectos, el gen a genosuprimir es un alelo asociado a la enfermedad, de tal manera que la genosupresión de este alelo enfermo permite la expresión de un alelo de tipo silvestre y restaura la función normal.

10 En una realización, el transgén se expresa a partir del promotor de la albúmina tras la inserción en el locus de la albúmina. El producto biológico codificado por el transgén se puede liberar a continuación al torrente circulatorio si el transgén se inserta *in vivo* en un hepatocito. En algunos aspectos, el transgén se inserta *in vivo* en el hígado en un vector vírico mediante inyección intravenosa.

En otra realización de la descripción, el transgén codifica un ARN no codificante, p. ej., un shRNA. La expresión del transgén antes de la maduración de la célula dará lugar a una célula que contiene el ARN no codificante de interés.

15 En otra realización, la invención describe células precursoras (células troncales hematopoyéticas, células troncales del músculo o células troncales hematopoyéticas (CTH) CD34+) en las que se ha insertado un transgén de tal manera que las células maduras procedentes de estas células precursoras contienen una gran cantidad del producto codificado por el transgén. En algunas realizaciones, estas células precursoras son células troncales pluripotentes inducidas (CTPi).

20 En algunas realizaciones, los métodos de la invención se pueden utilizar *in vivo* en los sistemas de animales transgénicos. En algunos aspectos, el animal transgénico se puede utilizar en el desarrollo del modelo, en donde el transgén codifica un gen humano. En algunos casos, el animal transgénico puede tener genosuprimido el correspondiente locus endógeno, lo que permite el desarrollo de un sistema *in vivo* en donde la proteína humana se puede estudiar en aislamiento. Tales modelos transgénicos se pueden utilizar con el propósito de criba para identificar moléculas pequeñas, o biomoléculas grandes, u otras entidades que pudieran interactuar con la proteína humana de interés, o bien modificarla. En algunos aspectos, el transgén está integrado en el locus seleccionado (p. ej., muy expresado o de refugio) en una célula troncal (p. ej., una célula troncal embrionaria, una célula troncal pluripotente inducida, una célula troncal hepática, una célula troncal neural, etc.), o un embrión animal obtenido por cualquiera de los métodos descritos en la presente memoria, y a continuación el embrión se implanta de tal manera que nace un animal vivo. El animal se cría a continuación hasta la madurez sexual y se le deja producir descendencia, en donde al menos algunos de los descendientes comprenden el transgén integrado.

35 Se describe en la presente memoria un método para la integración específica de sitio de una secuencia de ácido nucleico en un locus endógeno (p. ej., gen asociado a la enfermedad, muy expresado tal como la globina en los glóbulos rojos, o un gen de refugio tal como un gen de albúmina, de CCR5, de HPRT o Rosa) de un cromosoma, por ejemplo, en el cromosoma de un embrión. En determinadas realizaciones, el método comprende: (a) inyectarle a un embrión (i) al menos un vector de ADN, en donde el vector de ADN comprende una secuencia delantera y una secuencia posterior que flanquean la secuencia de ácido nucleico a integrarse, y (ii) al menos una molécula de ARN que codifica un dedo de cinc, una nucleasa TALE o un sistema CRISPR/Cas que reconoce el sitio de integración en el locus destinatario, y (b) cultivar el embrión para permitir la expresión de la NDZn, la TALEN y/o el sistema CRISPR/Cas, en donde una rotura bicatenaria introducida por la NDZn, la TALEN y/o el sistema CRISPR/Cas en el sitio de integración se repara mediante recombinación homóloga con el vector de ADN, de manera que se integra la secuencia de ácido nucleico en el cromosoma.

45 En cualquiera de las realizaciones anteriores, los métodos y compuestos de la invención se pueden combinar con otros agentes terapéuticos para el tratamiento de los sujetos con enfermedades por almacenamiento lisosomal. En algunos aspectos, los métodos y las composiciones se utilizan en combinación con los métodos y composiciones que permiten atravesar la barrera hematoencefálica. En otros aspectos, los métodos y las composiciones se utilizan en combinación con compuestos que se sabe que deprimen la respuesta inmunitaria del sujeto.

50 También se describe un kit, que comprende la NDZn, la TALEN y/o el sistema CRISPR/Cas de la invención. El kit puede comprender ácidos nucleicos que codifican la NDZn, la TALEN y/o el sistema CRISPR/Cas (p. ej., moléculas de ARN o los genes que codifican la NDZn, la TALEN y/o el sistema CRISPR/Cas contenidos en un vector de expresión idóneo), moléculas donantes, vectores de expresión que codifican las líneas de las células hospedadoras idóneas para el ARN guía único, instrucciones para realizar los métodos de la invención, y similares.

Estos y otros aspectos les resultarán muy evidentes al experto en la técnica a la luz de la descripción como un conjunto.

60 BREVE DESCRIPCIÓN DE LOS DIBUJOS

En la **figura 1, paneles A y B**, se describe un conjunto de geles que muestran los resultados de un ensayo de discordancias con Cel-I (SurveyorTM, Transgenomic) que mide la escisión en una localización de interés mediante una pareja de nucleasas tras las que se ha producido una NHEJ. La NHEJ provoca la inserción o delección de bases nucleotídicas («indels»), lo que a continuación crea una discordancia cuando la hebra de ADN se hibrida con una hebra de ADN de tipo silvestre. La figura 1A muestra los resultados medidos cuando se transfectan a 37 °C las parejas de nucleasas específicas de la albúmina al interior de las células Neuro2A y la figura 1B muestra los

resultados cuando se transducen parejas de nucleasas sometidas a choque hipotérmico (30 °C). El porcentaje de discordancias o porcentaje de indels es una medida de la actividad nucleásica de cada pareja en cada condición.

La figura 2 es un esquema que describe la estructura de cuatro donantes en AAV diseñados para proporcionar los transgenes terapéuticos para el tratamiento de las enfermedades de Fabry, Gaucher, Hurler y Hunter. Cada construcción donante contiene las secuencias de AAV (5'ITR y 3'ITR), brazos de homología flanqueantes para la inserción de los donantes en el locus de albúmina mediante mecanismos dependientes de la homología, un sitio aceptor de ajuste, el ADN que codifica la enzima de sustitución, y una etiqueta MYC-Flag para permitir la identificación de los donantes integrados.

En la figura 3, paneles A y B, se muestra la actividad de las NDZn de la albúmina de ratón *in vivo*. A ratones macho normales ($n = 6$) se les administró una dosis única de 200 μ l de $1,0 \times 10^{11}$ genomas de vectores totales de AAV2/8 o bien AAV2/8.2 que codifican la pareja específica de la NDZn murina formada por SBS30724 y SBS30725 para evaluar la infectividad del hígado mediante la detección de las copias del genoma del vector de AAV y la actividad de NHEJ *in vivo* en los ratones normales. Los vectores se administraron tal y como está descrito mediante una inyección en la vena de la cola de los ratones y, a los 14 días de la administración, se sacrificaron los ratones, se les recogió el hígado y se procesaron para cuantificar el ADN o las proteínas totales. La detección de las copias del genoma del vector de AAV se realizó por PCR cuantitativa, y la actividad de escisión de la NDZn se midió mediante el ensayo Cel-1 (Surveyor, Transkaryotic). **La figura 3A** ilustra un gel con los resultados de Cel-1 de ratones que recibieron AAV2/8 que contenía un casete de expresión de GFP o AAV2/8 que comprende la NDZn, en donde los AAV se produjeron mediante un sistema de expresión en 293 o un sistema de baculovirus. **La figura 3B** ilustra la cuantificación de los carriles del gel y muestra que la infección de los ratones con el AAV que contiene la NDZn específica de la albúmina da lugar a una actividad de NHEJ cuantificable de casi el 30%.

En la figura 4, paneles A y B, se muestra la inserción de un donante del transgén huGLa (deficiente en los pacientes que padecen la enfermedad de Fabry) en el locus de la albúmina en los ratones. En la figura 4A se muestra una transferencia de tipo Western contra la proteína huGLa codificada por el transgén, en donde la flecha indica la supuesta proteína. La comparación de las muestras de ratón de los ratones que recibieron tanto la NDZn como el donante (muestras 1-1, 1-2 y 1-3) con las muestras que únicamente recibieron la NDZn (4-1, 4-2, 4-3) o bien las que recibieron únicamente el donante de huGLa («donante de Fabry humano»), muestras 5-1 y 5-2, conduce a identificar una banda que coincide con el control del lisado del hígado humano. En la figura 4B se ilustran los resultados de ELISA que usaron un kit de ELISA específico para la huGLa, en donde las muestras se analizaron de ratones a los 14 días o bien a los 30 días de la introducción del virus. Las barras de error representan las desviaciones estándares ($n = 3$). Los resultados demuestran que los ratones que recibieron la NDZn y el donante (círculos) tenían una cantidad de señal de huGLa mayor que los que solo recibieron la NDZn (cuadrados) o solo recibieron el donante (triángulos).

En la figura 5, paneles A-D, se ilustran las transferencias de tipo Western que muestran la expresión en homogeneizados hepáticos de los transgenes donantes para EAL insertados en el locus de la albúmina en los ratones. En la figura 5A se muestran los resultados con el uso del transgén que codifica la IDUA, en la figura 5B se muestran los resultados con el uso del transgén de la GLA, en la figura 5C se muestran los resultados con el uso del transgén de la IDS, y en la figura 5D se muestran los resultados con el uso del transgén de la GBA.

La figura 6 es un esquema que muestra los dos tipos de inserción de donante que se puede producir después de la escisión mediada por la NDZn. La inserción del donante mediada por la NHEJ dará lugar a la integración de toda la construcción EAL-donante, mientras que la inserción mediada por HDR hará que se incorpore solo el ADNc que incluye el sitio aceptor del ajuste de F9. En la figura 6 también se describe la ubicación de los dos cebadores de PCR («mALB-OOF1» y «Acc651-SA-rev-sh») utilizados para detectar el tipo de integración que se ha producido.

En la figura 7, paneles A-C, se muestran los resultados de la PCR radiomarcada con 32 P realizada sobre los homogeneizados de hígado de los ratones que contienen los transgenes para EAL integrados 30 días después del tratamiento. En la figura 7A se ilustran los ratones con el transgén IDUA, en la figura 7B se ilustran los que llevan el transgén GLA y en la figura 7C los que llevan el transgén IDS. En todos los casos, las bandas indican que la inserción de los transgenes se ha producido a través de la integración mediada por HDR y mediada por NHEJ.

La figura 8 es un esquema que ilustra el diseño de los donantes para EAL que contienen etiquetas de epítomos. Se indica la localización y las secuencias de las etiquetas Myc y Flag.

En la figura 9, paneles A y B, se ilustra un gel de PCR radiomarcada con 32 P realizada tal y como está descrito más arriba en homogeneizados de hígado de los ratones con donantes para EAL integrados que contienen las etiquetas de epítopo. En la figura 9A se muestra que la integración se produjo a través de la integración dirigida del transgén GLA, IDUA e IDS mediada por NHEJ y mediada por HDR. En la figura 9B se muestra lo mismo para el transgén GBA.

DESCRIPCIÓN DETALLADA

Se describen en la presente memoria métodos y composiciones para tratar o prevenir una enfermedad por almacenamiento lisosomal (EAL). La invención da a conocer métodos y composiciones para insertar un gen que codifica una proteína que está ausente o se expresa de manera insuficiente en el sujeto con la EAL, de tal manera que el gen se expresa en el hígado y se expresa la proteína terapéutica (de sustitución). La invención también describe la alteración de una célula (p. ej., glóbulo rojo precursor o maduro, CTPi o hepatocito) de tal manera que produce gran cantidad del agente terapéutico, y la introducción de una población de estas células alteradas en un paciente suministrará esa proteína necesaria. El transgén codificará una proteína o ARN estructural deseados que son beneficiosos desde el punto de vista terapéutico en un paciente que lo necesita.

Así pues, los métodos y las composiciones de la descripción se pueden utilizar para expresar, a partir de un transgén, proteínas beneficiosas desde el punto de vista terapéutico desde cualquier locus (p. ej., locus de la albúmina muy expresada) para remplazar las enzimas que son defectuosas en las enfermedades por almacenamiento lisosomal. Adicionalmente, la invención da a conocer métodos y composiciones para el tratamiento de estas enfermedades mediante la inserción de las secuencias en locus muy expresados en células tales como los hepatocitos.

Además, el transgén se puede introducir en células procedentes del paciente, p. ej., células troncales pluripotentes inducidas (CTPI) procedentes del paciente u otros tipos de células troncales (embrionarias o hematopoyéticas) para ser usado en la implantación final. Resulta particularmente útil la inserción del transgén asociado a la enfermedad en una célula troncal hematopoyética para su implantación en un paciente que lo necesita. A medida que las células troncales se diferencian en células maduras, contendrán gran cantidad de la proteína sustitutiva para suministrarla a los tejidos.

General

La puesta en práctica de los métodos, así como la preparación y el uso de las composiciones descritas en la presente memoria, emplean, a menos que se indique de otra manera, técnicas convencionales en biología molecular, bioquímica, estructura y análisis de la cromatina, química computacional, cultivo celular, ADN recombinante y los campos relacionados que están dentro de las capacidades de la técnica. Estas técnicas se explican con detalle en la bibliografía. Véanse, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, 1989 y su tercera edición, 2001; Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Nueva York, 1987 y las actualizaciones periódicas; la serie *Methods in Enzymology*, Academic Press, San Diego; Wolffe, *Chromatin Structure and Function*, tercera edición, Academic Press, San Diego, 1998; *Methods in Enzymology*, vol. 304, «Chromatin» (P. M. Wassarman y A. P. Wolffe, eds.), Academic Press, San Diego, 1999; y *Methods in Molecular Biology*, vol. 119, «Chromatin Protocols» (P. B. Becker, ed.) Humana Press, Totowa, 1999.

Definiciones

Los términos «ácido nucleico», «polinucleótido» y «oligonucleótido» se utilizan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en forma monocatenaria o bicatenaria. Para los propósitos de la presente descripción, estos términos no se deben interpretar como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar los análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en la base, en el azúcar o en los restos de fosfato (p. ej., esqueletos de fosforotioato). En general, un análogo de un nucleótido concreto tiene la misma especificidad de apareamiento de bases, a saber, un análogo de la A se emparejará con la T.

Los términos «polipéptido», «péptido» y «proteína» se utilizan indistintamente para referirse a un polímero de restos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de los correspondientes aminoácidos que aparecen en la naturaleza.

«Fijación» se refiere a una interacción no covalente entre macromoléculas que es específica de la secuencia (p. ej., entre una proteína y un ácido nucleico). No todos los componentes de una interacción de fijación necesitan ser específicos de la secuencia (p. ej., se pone en contacto con restos de fosfato en un esqueleto de ADN), siempre y cuando la interacción en su conjunto sea específica de la secuencia. Tales interacciones se caracterizan por lo general por una constante de disociación (K_d) de 10^{-5} M^{-1} o menor. «Afinidad» se refiere a la fuerza de la fijación: el incremento de la afinidad de fijación se correlaciona con una K_d menor.

Una «proteína que se fija» o «proteína de fijación» es una proteína que es capaz de fijarse de forma no covalente a otra molécula. Una proteína de fijación puede fijarse, por ejemplo, a una molécula de ADN (una proteína de fijación al ADN), a una molécula de ARN (una proteína de fijación al ARN) y/o a una molécula de proteína (una proteína que se fija a otra proteína). En el caso de una proteína que se fija a otra proteína, puede fijarse a sí misma (para formar homodímeros, homotrímeros, etc.) y/o se puede fijar a una o varias moléculas de una o varias proteínas diferentes. Una proteína de fijación puede tener más de un tipo de actividad de fijación. Por ejemplo, las proteínas con dedos de cinc tienen actividad de fijación al ADN, de fijación al ARN y de fijación a las proteínas.

Una «proteína de fijación al ADN con dedos de cinc» (o dominio de fijación) es una proteína, o un dominio dentro de una proteína más grande, que se fija al ADN de una manera específica de la secuencia a través de uno o varios dedos de cinc, que son regiones de la secuencia de aminoácidos dentro del dominio de fijación cuya estructura está estabilizada por la coordinación de un ion de cinc. La terminología proteína de fijación al ADN con dedos de cinc a menudo se abrevia como proteína con dedos de cinc o PDZn.

Un «dominio TALE de fijación al ADN» o «TALE» es un polipéptido que comprende uno o varios dominios o unidades con repeticiones de TALE. Los dominios con repeticiones intervienen en la fijación del TALE a su secuencia semejante en el ADN diana. Una sola «unidad de repetición» (también denominada «repetición») tiene típicamente de 33 a 35 aminoácidos de longitud y muestra al menos alguna homología de secuencia con otras secuencias repetidas de TALE dentro de una proteína TALE que se produce en la naturaleza.

Los dominios de fijación TALE y los dominios de fijación con dedos de cinc se pueden «manipular genéticamente» para que se fijen a una secuencia de nucleótidos predeterminada, por ejemplo, mediante manipulación genética (por la alteración de uno o varios aminoácidos) de la región de la hélice de reconocimiento de una proteína TALE o de los dedos de cinc que aparecen en la naturaleza. Por lo tanto, las proteínas de fijación al ADN (con dedos de cinc o las

5 TALE) manipuladas genéticamente son proteínas que no aparecen en la naturaleza. Los ejemplos no limitantes de los métodos para la manipulación genética de las proteínas de fijación al ADN son el diseño y la selección. Una proteína de fijación al ADN diseñada es una proteína que no aparece en la naturaleza cuyo diseño o composición es principalmente el resultado de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de normas de sustitución y algoritmos computadorizados para procesar la información en una base de datos que almacena la

10 información de los diseños y datos de fijación que existen para las PDZn y/o TALE. Véanse, por ejemplo, las patentes de los EE. UU. 6.140.081; 6.453.242; y 6.534.261; véanse también las solicitudes de patente internacional WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496, y la publicación de los EE. UU. n.º 20110301073.

15 Una proteína con dedos de cinc o TALE «seleccionada» es una proteína que no se encuentra en la naturaleza, cuya producción es principalmente el resultado de un proceso empírico, tal como la presentación en fagos, trampa de interacción o selección híbrida. Véanse, p. ej., las patentes de los EE. UU. n.ºs US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; las solicitudes de patente internacional WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084, y la publicación de los EE. UU. n.º 20110301073.

«Recombinación» se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, «recombinación homóloga (RH)» se refiere a la forma especializada de tal intercambio que tiene lugar, por ejemplo, durante la reparación de roturas bicatenarias en las células a través de los mecanismos de

25 reparación dirigidos por la homología. Este proceso requiere la homología de secuencia nucleotídica, utiliza una molécula «donante» como modelo para la reparación de una molécula «diana» (a saber, la que experimenta la rotura bicatenaria), y se conoce de maneras diferentes como «conversión génica sin entrecruzamiento» o «conversión génica de distancia corta», porque conduce a la transferencia de información genética desde el donante a la diana. Sin desear comprometerse con ninguna teoría concreta, tal transferencia puede implicar la corrección de discordancias en el ADN

30 del heterodúplex que se forma entre la diana rota y el donante, y/o «hibridación de las hebras dependiente de la síntesis», en la que se utiliza el donante para volver a sintetizar la información genética que se volverá parte de la diana, y/o los procesos relacionados. Tal RH especializada a menudo da lugar a una alteración de la secuencia de la molécula diana, de tal manera que parte o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

35 En los métodos de la descripción, una o varias nucleasas dirigidas selectivamente tal y como se describe en la presente memoria crean una rotura bicatenaria en la secuencia diana (p. ej., cromatina celular) en un sitio predeterminado, y se puede introducir en la célula un polinucleótido «donante», que tiene homología con la secuencia nucleotídica en la región de la rotura. La presencia de la rotura bicatenaria se ha demostrado que facilita la integración de la secuencia donante. La secuencia donante se puede integrar físicamente o, como alternativa, el polinucleótido donante se utiliza

40 como molde para reparar la rotura mediante la recombinación homóloga, lo que da lugar a la inserción en la cromatina celular de toda o parte de la secuencia de nucleótidos que está en el donante. Así pues, una primera secuencia en la cromatina celular se puede alterar y, en determinadas realizaciones, se puede convertir en una secuencia presente en un polinucleótido donante. Así pues, la utilización de los términos «sustituir» o «sustitución» se puede entender como que representa la sustitución de una secuencia nucleotídica por otra (a saber, la sustitución de una secuencia en el sentido de su información) y no requiere necesariamente una sustitución física o química de un polinucleótido por otro.

45

En cualquiera de los métodos descritos en la presente memoria, otras parejas de proteínas con dedos de cinc o TALEN se pueden utilizar para la escisión bicatenaria adicional de otros sitios diana dentro de la célula.

50 En determinadas realizaciones de los métodos para la recombinación y/o sustitución y/o alteración dirigidas de una secuencia en una región de interés en la cromatina celular, se altera una secuencia cromosómica mediante recombinación homóloga con una secuencia nucleotídica «donante» exógena. Tal recombinación homóloga se estimula por la presencia de una rotura bicatenaria en la cromatina celular, si las secuencias homólogas a la región de la rotura están presentes.

55 En cualquiera de los métodos descritos en la presente memoria, la primera secuencia nucleotídica (la «secuencia donante») puede contener secuencias que son homólogas, pero no idénticas, a las secuencias genómicas de la región de interés, con lo que se estimula la recombinación homóloga para insertar una secuencia no idéntica en la región de interés. Así pues, en determinadas realizaciones, porciones de la secuencia donante que son homólogas a la secuencia

60 de la región de interés muestran entre aproximadamente el 80 y el 99% (o cualquier entero entre ellas) de identidad de secuencia con la secuencia genómica que se sustituye. En otras realizaciones, la homología entre la secuencia donante y la genómica es de más del 99%, por ejemplo, si sólo difiere 1 nucleótido entre las secuencias donante y genómica entre más de 100 pares de bases contiguas. En algunos casos, una porción no homóloga de la secuencia donante puede contener secuencias que no están presentes en la región de interés, de tal manera que las nuevas secuencias se

65 insertan en la región de interés. En estos casos, la secuencia no homóloga está por lo general flanqueada por secuencias de 50 a 1.000 pares de bases (o cualquier valor entero entre ellos) o cualquier número de pares de bases

mayor que 1.000, que son homólogas o idénticas a la secuencia de la región de interés. En otras realizaciones, la secuencia donante es no homóloga a la primera secuencia y se inserta en el genoma mediante mecanismos de recombinación no homóloga.

5 Cualquiera de los métodos descritos en la presente memoria se pueden utilizar para la inactivación completa o parcial de una o varias secuencias diana en una célula mediante la integración dirigida de la secuencia donante que destruye la expresión del gen o genes de interés. También se dan a conocer líneas celulares con genes parcial o completamente inactivados.

10 Además, los métodos de la integración dirigida tal y como se describen en la presente memoria también se pueden utilizar para integrar una o varias secuencias exógenas. La secuencia del ácido nucleico exógeno puede comprender, por ejemplo, uno o varios genes o moléculas de ADNc, o cualquier tipo de secuencia codificante o no codificante, así como uno o varios elementos de control (p. ej., promotores). Además, la secuencia de ácido nucleico exógeno puede producir una o varias moléculas de ARN (p. ej., los ARN pequeños ahorquillados (shRNA), los ARN inhibidores (RNAi), los microARN (miRNA), etc.).

«Escisión» se refiere a la rotura del esqueleto covalente de una molécula de ADN. La escisión se puede iniciar mediante una serie de métodos que incluyen, pero sin limitarse a ellos, la hidrólisis enzimática o química de un enlace fosfodiéster. Se pueden producir tanto la escisión monocatenaria como la escisión bicatenaria, y la escisión bicatenaria se puede 20 producir de resultados de dos escisiones monocatenarias diferentes. La escisión del ADN dará lugar a la producción de extremos romos o de extremos escalonados. En determinadas realizaciones, los polipéptidos de fusión se utilizan para la escisión dirigida y selectiva de ADN bicatenario.

Un «semidominio de escisión» es una secuencia polipeptídica que, junto con un segundo polipéptido (idéntico o bien 25 diferente) forma un complejo que tiene actividad de escisión (preferiblemente actividad de escisión bicatenaria). Los términos «primer y segundo semidominio de escisión», «semidominios de escisión + y -» y «semidominios de escisión derecho e izquierdo» se utilizan indistintamente para referirse a las parejas de semidominios de escisión que se dimerizan.

30 Un «semidominio de escisión manipulado genéticamente» es un semidominio de escisión que se ha modificado para que forme obligatoriamente heterodímeros con otro semidominio de escisión (p. ej., otro semidominio de escisión manipulado genéticamente). Véanse también las publicaciones de patente de los EE. UU. n.^{os} 2005/0064474, 2007/0218528, 2008/0131962 y 2011/0201055.

35 El término «secuencia» se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser de ADN o de ARN; puede ser lineal, circular o ramificada, y puede ser monocatenaria o bicatenaria. La terminología «secuencia donante» se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede ser de cualquier longitud, por ejemplo, entre 2 y 10.000 nucleótidos de longitud (o cualquier valor entero entre ellos o por encima), preferiblemente entre aproximadamente 100 y 1.000 nucleótidos de longitud (o cualquier entero entre ellos), 40 más preferiblemente entre aproximadamente 200 y 500 nucleótidos de longitud.

Un «gen asociado a la enfermedad» es uno que es defectuoso de alguna manera en una enfermedad monogénica. Los ejemplos no limitantes de enfermedades monogénicas incluyen la inmunodeficiencia combinada intensa, la fibrosis quística, las enfermedades por almacenamiento lisosomal (p. ej., de Gaucher, de Hurler, de Hunter, de Fabry, de 45 Neimann-Pick, de Tay-Sach, etc.), anemia drepanocítica y talasemia.

«Cromatina» es la estructura de la nucleoproteína que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteína, que incluye las histonas y las proteínas cromosómicas no histónicas. La mayor parte de la cromatina celular eucariota existe en forma de nucleosomas, en donde la estructura central del 50 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 ADN conector (de longitud variable según el organismo) que se extiende entre las estructuras centrales de los nucleosomas. Una molécula de la histona H1 está asociada por lo general con el ADN conector. Para los propósitos de la presente descripción, el término «cromatina» pretende abarcar todos los tipos de nucleoproteínas celulares, tanto procarióticas como eucarióticas. La cromatina celular incluye tanto la 55 cromatina cromosómica como la episómica.

Un «cromosoma» es un complejo de cromatina que comprende todo el genoma de una célula, o una porción del mismo. El genoma de una célula se caracteriza a menudo por su cariotipo, que es el conjunto de los cromosomas que comprenden el genoma de la célula. El genoma de una célula puede comprender uno o varios cromosomas.

60 Un «episoma» es un ácido nucleico, un complejo nucleoproteico u otra estructura que comprende un ácido nucleico que no forma parte del cariotipo de una célula y que es capaz de replicarse. Los ejemplos de episomas incluyen plásmidos y determinados genomas víricos.

65 Un «sitio diana» o «secuencia diana» es una secuencia de ácido nucleico que define una porción de un ácido nucleico al cual se fijará una molécula de fijación, siempre y cuando se den las condiciones suficientes para la fijación.

Una molécula «exógena» es una molécula que no suele estar presente en una célula, pero que se puede introducir en una célula mediante uno o varios métodos genéticos, bioquímicos o de otro tipo. La «presencia normal en la célula» se determina con respecto a la etapa de desarrollo concreta y las condiciones del entorno de la célula. Así pues, por ejemplo, una molécula que está presente solo durante el desarrollo embrionario del músculo es una molécula exógena con respecto a un miocito de adulto. De igual forma, una molécula inducida por el choque térmico es una molécula exógena con respecto a una célula que no ha sufrido un choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión funcional de una molécula endógena disfuncional o una versión disfuncional de una molécula endógena que suele ser funcional.

Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal cual se genera en un proceso químico combinatorio, o una macromolécula, tal como una proteína, ácido nucleico, glúcido, lípido, glucoproteína, lipoproteína, polisacárido, cualquier derivado por modificación de las moléculas anteriores, o cualquier complejo que comprende una o varias de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden tener una o dos hebras; pueden ser lineales, ramificados o circulares; y pueden tener cualquier longitud. Los ácidos nucleicos incluyen los que son capaces de formar dobles hélices, así como ácidos nucleicos que forman hélices triples. Véanse, por ejemplo, las patentes de los EE. UU. n.^{os} 5.176.996 y 5.422.251. Las proteínas incluyen, pero sin limitarse a ellas, proteínas de fijación al ADN, factores de transcripción, factores de remodelación de la cromatina, proteínas de fijación al ADN metilado, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, cinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicinas.

Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, esto es, una proteína o ácido nucleico exógenos. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma vírico infectivo, un plásmido o episoma introducido en la célula, o un cromosoma que no suele estar presente en la célula. Los métodos para la introducción de moléculas exógenas dentro las células los conocen los expertos en la técnica e incluyen, aunque sin limitarse a ellos, la transferencia mediada por lípidos (esto es, liposomas, que incluyen lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo con partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores víricos. Una molécula exógena puede ser también del mismo tipo de molécula que una molécula endógena, pero procedente de una especie diferente a la de la célula de la que procede. Por ejemplo, una secuencia de ácido nucleico de humano se puede introducir en una línea celular que procede originalmente de ratón o hámster.

Por el contrario, una molécula «endógena» es una que suele estar presente en una célula concreta en una etapa de desarrollo concreta en unas condiciones ambientales concretas. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, de un cloroplasto o de otro orgánulo, o un ácido nucleico episómico que aparece en la naturaleza. Otras moléculas endógenas pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

Una molécula de «fusión» es una molécula en la que están conectadas, preferiblemente de forma covalente, dos o más moléculas que hacen de subunidades. Las moléculas de las subunidades pueden ser del mismo tipo químico de molécula, o pueden ser de diferentes tipos químicos de moléculas. Los ejemplos del primer tipo de molécula de fusión incluyen, pero sin limitarse a ellos, proteínas de fusión (por ejemplo, una fusión entre el dominio de fijación al ADN de PDZn o TALE, y uno o varios dominios de activación) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita más arriba). Los ejemplos del segundo tipo de molécula de fusión incluyen, pero sin limitarse a ellos, una fusión entre un ácido nucleico capaz de formar hélices triples y un polipéptido, y una fusión entre un fijador del surco menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede ser el resultado de introducir en la célula la proteína de fusión, o de la inserción en una célula de un polinucleótido que codifica la proteína de fusión, en donde se transcribe el polinucleótido, y se traduce el transcrito, para generar la proteína de fusión. El transayuste, la escisión de polipéptidos y la ligación polipeptídica también pueden estar implicadas en la expresión de una proteína en una célula. Los métodos para la introducción del polinucleótido y del polipéptido en las células se presentan en otro lugar de esta descripción.

Un «gen», para los propósitos de la presente descripción, incluye una región de ADN que codifica un producto génico (véase más adelante), así como todas las regiones del ADN que regulan la producción del producto génico, tanto si tales secuencias regulatorias son adyacentes a las secuencias codificantes y/o transcritas, como si no lo son. Por consiguiente, un gen incluye, pero sin estar necesariamente limitado a ello, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como los sitios de fijación del ribosoma y los sitios internos de entrada de ribosomas, potenciadores, silenciadores, aisladores, elementos delimitantes, orígenes de replicación, sitios de fijación a la matriz y regiones de control del locus.

«Expresión génica» hace referencia a la conversión de la información contenida en un gen en su correspondiente producto génico. Un producto génico puede ser el producto transcripcional directo de un gen (esto es, 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 los ARN que están modificados, por procesos tales como la adición de la caperuza, poliadenilación, metilación y riboedición, y las proteínas modificadas por, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristoilación y glucosilación.

«Modulación» de la expresión génica hace referencia a un cambio de la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitarse a ello, activación génica y represión génica. La edición del genoma (esto es, escisión, alteración, inactivación, mutación al azar) se puede utilizar para modular la expresión. La inactivación génica hace referencia a cualquier reducción de la expresión génica en comparación con una célula que no incluye una PDZn o TALEN tal y como se describe en la presente memoria. Así pues, la inactivación génica puede ser parcial o completa.

Una «región de interés» es cualquier región de la cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante dentro o adyacente a un gen, en la que se desea fijar una molécula exógena. La fijación puede tener el propósito de escisión dirigida del ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma organular (p. ej., mitocondria, cloroplasto) o un genoma vírico infectante, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas, tales como, por ejemplo, secuencias líder, secuencias de remolque o intrones, o dentro de regiones sin transcribir, bien delante o detrás de la región codificante. Una región de interés puede ser tan pequeña que tenga un solo par de nucleótidos o de hasta 2.000 pares de nucleótidos de longitud, o cualquier valor entero de pares de nucleótidos.

Las células «eucariotas» incluyen, pero sin limitarse a ellas, células fúngicas (tales como levadura), células vegetales, células animales, células de mamífero y células humanas (p. ej., linfocitos T).

Los «glóbulos rojos» o eritrocitos son células diferenciadas a término procedentes de las células troncales hematopoyéticas. Les falta el núcleo y la mayoría de los orgánulos celulares. Los glóbulos rojos contienen hemoglobina para transportar el oxígeno desde los pulmones a los tejidos periféricos. De hecho, de cada glóbulo rojo, el 33% es hemoglobina. También transportan el CO₂ producido por las células durante el metabolismo, al extraerlo de los tejidos, para devolverlo a los pulmones y liberarlo durante la espiración. Los glóbulos rojos se producen en la médula ósea en respuesta a la hipoxia sanguínea que está mediada por la liberación de eritropoyetina (EPO) desde el riñón. La EPO hace incrementar el número de proeritroblastos y acorta el tiempo requerido para la maduración completa de los glóbulos rojos. Al cabo de aproximadamente 120 días, ya que los glóbulos rojos no contienen un núcleo ni ninguna otra capacidad regenerativa, se retiran las células de la circulación mediante la actividad fagocítica de los macrófagos en hígado, bazo y ganglios linfáticos (~90%) o bien mediante hemólisis en el plasma (~10%). Después de ser engullidos por los macrófagos, los componentes químicos de los glóbulos rojos se degradan dentro de las vacuolas de los macrófagos debido a la acción de las enzimas lisosomales.

Los «tejidos secretores» son los tejidos de un animal que secretan productos al exterior de cada una de las células en una luz de algún tipo que típicamente procede de un epitelio. Los ejemplos de tejidos secretores que están localizados en el tubo digestivo incluyen las células que revisten el intestino, el páncreas y la vesícula biliar. Otros tejidos secretores incluyen el hígado, los tejidos asociados al ojo, y las membranas mucosas, tales como las glándulas salivales, las glándulas mamarias, la glándula prostática, la hipófisis y otros miembros del sistema endocrino. Adicionalmente, los tejidos secretores incluyen cada una de las células de un tipo de tejido que tienen capacidad de secreción.

Los términos «unión operativa» y «operativamente unido» (o «unido operativamente») se utilizan indistintamente para hacer referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en donde los componentes se distribuyen de tal manera que ambos componentes funcionan normalmente y ofrecen la posibilidad de que al menos uno de los componentes pueda mediar en una función que ejerce al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unida operativamente 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 varios factores reguladores de la transcripción. Una secuencia reguladora de la transcripción suele estar por lo general unida operativamente 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, incluso aunque no sean contiguas.

Con respecto a los polipéptidos de fusión, la terminología «unido operativamente» puede hacer referencia al hecho de que cada uno de los componentes realiza la misma función cuando está unido al otro componente de la misma forma que lo haría si no estuviera así unido. Por ejemplo, con respecto a un polipéptido de fusión en el que un dominio de fijación al ADN de PDZn o TALE está fusionado a un dominio de activación, el dominio de fijación al ADN de PDZn o TALE y el dominio de activación están en una unión operativa si, en el polipéptido de fusión, la porción del dominio de fijación al ADN de PDZn o TALE es capaz de fijarse a su sitio diana y/o su sitio de fijación, mientras que el dominio de activación es capaz de inducir la expresión génica. Cuando un polipéptido de fusión en el que un dominio de fijación al ADN de PDZn o TALE está fusionado a un dominio de escisión, el dominio de fijación al ADN de PDZn o TALE y el dominio de escisión están en una unión operativa si, en el polipéptido de fusión, la porción del dominio de fijación al ADN de PDZn o TALE es capaz de fijarse a su sitio diana y/o su sitio de fijación, mientras que el dominio de escisión es capaz de escindir el ADN cercano al sitio diana.

Un «fragmento funcional» de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, aunque conserva 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 correspondiente molécula nativa y/o puede contener una o varias sustituciones de aminoácidos o nucleótidos. Los métodos para determinar la función de un ácido nucleico (p. ej., función codificante, capacidad para hibridarse a otro ácido nucleico) se conocen bien en la técnica. De igual forma, se conocen bien los métodos para determinar la función de la proteína. Por ejemplo, se puede determinar la función de fijación al ADN de un polipéptido, por ejemplo, mediante fijación a filtro, cambio de la movilidad electroforética o ensayos de inmunoprecipitación. La escisión del ADN se puede ensayar mediante electroforesis en gel. Véase Ausubel et al., más arriba. La capacidad que tiene una proteína para interactuar con otra proteína se puede determinar, por ejemplo, mediante coimmunoprecipitación, ensayos de doble híbrido o complementación, tanto genética como bioquímica. Véanse, por ejemplo, Fields et al. (1989) *Nature* 340: 245-246; patente de los EE. UU. n.º 5.585.245 y solicitud de internacional patente en el marco del PCT WO 98/44350.

Un «vector» es capaz de transferir secuencias génicas a las células destinatarias. Típicamente, «construcción en vector», «vector de expresión» y «vector de transferencia génica» hacen referencia a cualquier construcción de ácido nucleico capaz de dirigir la expresión de un gen de interés y que puede transferir secuencias génicas a las células destinatarias. Así pues, el término incluye la clonación y los vehículos de expresión, así como vectores integrativos.

Un «gen indicador» o «secuencia indicadora» se refiere a cualquier secuencia que produce un producto proteico que se mide con facilidad, preferiblemente, aunque no necesariamente, en un ensayo convencional. Los genes indicadores idóneos incluyen, pero sin limitarse a ellos, secuencias que codifican proteínas que intervienen en la resistencia a los antibióticos (p. ej., resistencia a la ampicilina, resistencia a la neomicina, resistencia al G418, resistencia a la puromicina), secuencias que codifican proteínas coloreadas o fluorescentes o luminiscentes (p. ej., proteína fluorescente verde, proteína fluorescente verde realzada, proteína fluorescente roja, luciferasa) y proteínas que intervienen en la mejora del crecimiento celular y/o de la amplificación génica (p. ej., dihidrofolato reductasa). Las etiquetas de epítomos incluyen, por ejemplo, una o varias copias de FLAG, His, myc, Tap, HA o cualquier secuencia de aminoácidos detectable. Las «etiquetas de expresión» incluyen secuencias que codifican indicadores que pueden estar unidos operativamente a una secuencia génica deseada para monitorizar la expresión del gen de interés.

Los términos «sujeto» y «paciente» se utilizan indistintamente y se refieren a los mamíferos, tales como los pacientes humanos y los primates no humanos, así como a animales experimentales, tales como conejos, perros, gatos, ratas, ratones y otros animales. Por consiguiente, el término «sujeto» o «paciente», tal y como se utilizan en la presente memoria, hacen referencia a cualquier paciente o sujeto mamífero al que se pueden administrar las células alteradas de la invención y/o las proteínas producidas por las células alteradas de la invención. Los sujetos de la presente invención incluyen los que tienen una EAL.

35 Nucleasas

Se describen en la presente memoria composiciones, en particular nucleasas, que son útiles para dirigir un gen para la inserción de un transgén, por ejemplo, nucleasas que son específicas de un gen de refugio, tal como la albúmina. En determinadas realizaciones, la nucleasa se produce en la naturaleza. En otras realizaciones, la nucleasa no se produce en la naturaleza, a saber, está manipulada genéticamente en el dominio de fijación al ADN y/o en el dominio de escisión. Por ejemplo, el dominio de fijación al ADN de una nucleasa o sistema con nucleasa que se produce en la naturaleza se puede alterar para que se fije a un sitio diana seleccionado (p. ej., una meganucleasa que se ha manipulado genéticamente para que se fije a un sitio diferente del sitio de fijación habitual, o un sistema CRISPR/Cas que utiliza un único ARN guía manipulado genéticamente). En otras realizaciones, la nucleasa comprende dominios heterólogos de escisión y de fijación al ADN (p. ej., nucleasas con dedos de cinc; nucleasas efectoras TAL; dominios de fijación al ADN de la meganucleasa con dominios de escisión heterólogos).

A. Dominios de fijación al ADN

50 En determinadas realizaciones, la nucleasa es una meganucleasa (endonucleasa de asentamiento). Las meganucleasas que se producen en la naturaleza reconocen sitios de escisión de 15 a 40 pares de bases y se agrupan habitualmente en cuatro familias: la familia LAGLIDADG, la familia GIY-YIG, la familia de la caja His-Cyst y la familia HNH. Las endonucleasas de asentamiento de ejemplo incluyen I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-Csml, I-PanI, I-Scell, I-PpoI, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII. Se conocen sus secuencias de reconocimiento. Véanse también la patente de los EE. UU. n.º 5.420.032; la patente de los EE. UU. n.º 6.833.252; Belfort et al. (1997) *Nucleic Acids Res.* 25: 3379-3388; Dujon et al. (1989) *Gene* 82: 115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; 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.

60 En determinadas realizaciones, la nucleasa comprende una endonucleasa de asentamiento (meganucleasa) manipulada genéticamente (que no aparece en la naturaleza). Se conocen las secuencias de reconocimiento de las endonucleasas de asentamiento y las meganucleasas tales como I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-Csml, I-PanI, I-Scell, I-PpoI, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII. Véanse también la patente de los EE. UU. n.º 5.420.032; la patente de los EE. UU. n.º 6.833.252; Belfort et al. (1997) *Nucleic Acids Res.* 25: 3379-3388; Dujon et al. (1989) *Gene* 82: 115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; 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. Además, la especificidad de fijación al ADN de las endonucleasas de asentamiento y de las meganucleasas se puede manipular genéticamente para que se fije a sitios diana que no son naturales. Véanse, por ejemplo, Chevalier et al. (2002) *Molec. Cell.* 10: 895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31: 2952-2962; Ashworth et al. (2006) *Nature* 441: 656-659; Paques et al. (2007) *Current Gene Therapy* 7: 49-66; publicación de patente de los EE. UU. n.º 20070117128. Los dominios de fijación al ADN de las endonucleasas de asentamiento y de las meganucleasas se puede alterar en el contexto de la nucleasa como un conjunto (a saber, de tal manera que la nucleasa incluya el dominio de escisión semejante) o se puede fusionar a un dominio de escisión heterólogo.

En otras realizaciones, el dominio de fijación al ADN comprende un dominio de fijación al ADN efector TAL que se produce en la naturaleza o manipulado genéticamente (que no se produce en la naturaleza). Véase, p. ej., la publicación de patente de los EE. UU. n.º 20110301073. Las bacterias patógenas de plantas del género *Xanthomonas* se sabe que ocasionan muchas enfermedades en plantas de cultivo importantes. La patogenia de *Xanthomonas* depende del sistema de secreción de tipo III conservado (T3S) que inyecta más de 25 proteínas efectoras diferentes en la célula vegetal. Entre estas proteínas inyectadas están los efectores de tipo activador de la transcripción (TALE) que imitan a los activadores de la transcripción vegetales y manipulan el transcriptoma de la planta (véase Kay et al. (2007) *Science* 318: 648-651). Estas proteínas contienen un dominio de fijación al ADN y un dominio de activación transcripcional. Una de las TALE mejor caracterizadas es AvrBs3 de *Xanthomonas campestris* pv. *Vesicatoria* (véase Bonas et al. (1989) *Mol Gen. Genet.* 218: 127-136 y la solicitud de patente internacional WO 2010079430). Las TALE contienen un dominio centralizado de repeticiones en tándem, cada repetición contiene aproximadamente 34 aminoácidos, que son la clave para la especificidad de fijación al ADN de estas proteínas. Además, contienen una secuencia de localización nuclear y un dominio de activación transcripcional ácido (para una revisión, véase Schornack S. et al. (2006) *J. Plant Physiol* 163 (3): 256-272). Además, en las bacterias fitopatógenas *Ralstonia solanacearum*, se han hallado dos genes, denominados brg11 y hpx17, que son homólogos a la familia de AvrBs3 de *Xanthomonas* en la cepa GMI1000 de *R. solanacearum* biovar 1 y en la cepa RS1000 del biovar 4 (véase Heuer et al. (2007) *Appl and Envir Micro* 73 (13): 4379-4384). Estos genes tienen una secuencia de nucleótidos idéntica al 98,9% entre sí, pero difieren en una delección de 1.575 pb en el dominio con repeticiones de hpx17. Sin embargo, ambos productos génicos tienen una identidad de secuencia de menos del 40% con las proteínas de la familia de AvrBs3 de *Xanthomonas*.

Así pues, en algunas realizaciones, el dominio de fijación al ADN que se fija a un sitio diana en un locus deseado (p. ej., albúmina u otro refugio) es un dominio manipulado genéticamente de un efector TAL similar a los procedentes de los patógenos de planta *Xanthomonas* (véanse Boch et al. (2009) *Science* 326: 1509-1512 y Moscou y Bogdanove, (2009) *Science* 326: 1501) y *Ralstonia* (véanse Heuer et al. (2007) *Applied and Environmental Microbiology* 73(13): 4379-4384); las publicaciones de patente de los EE. UU. n.ºs 20110301073 y 20110145940. Véanse, p. ej., las TALEN de albúmina en las solicitudes de patente de los EE. UU. n.ºs 13/624.193 y 13/624.217.

En determinadas realizaciones, el dominio de fijación al ADN comprende una proteína con dedos de cinc (p. ej., una proteína con dedos de cinc que se fija a un sitio diana en un gen de albúmina o en un gen de refugio). Preferiblemente, la proteína con dedos de cinc no aparece en la naturaleza, ya que está manipulada genéticamente para que se fije a un sitio diana de elección. Véanse, 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 los EE. UU. n.ºs 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 los EE. UU. n.ºs 2005/0064474; 2007/0218528; 2005/0267061.

La manipulación genética de un dominio TALE o de fijación con dedos de cinc puede proporcionar una nueva especificidad de fijación cuando se compara con una proteína con dedos de cinc que se produce en la naturaleza. Los métodos de ingeniería genética incluyen, pero sin limitarse a ellos, el diseño racional o diferentes tipos de selección. El diseño racional incluye, por ejemplo, la utilización de bases de datos que comprenden secuencias nucleotídicas de tripletes (o tetrapletes) y secuencias de aminoácidos con dedos de cinc independientes, en donde cada secuencia nucleotídica de triplete o tetraplete está asociada a una o varias secuencias aminoacídicas de dedos de cinc que se fijan a la secuencia concreta del triplete o del tetraplete. Véanse, por ejemplo, las patentes de los EE. UU. en copropiedad n.ºs 6.453.242 y 6.534.261.

Los métodos de selección de ejemplo, que incluyen los sistemas de exposición de fagos y de doble híbrido, se describen en las patentes de los EE. UU. n.ºs 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 las solicitudes de patente internacional n.ºs WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y la patente de Gran Bretaña GB 2.338.237. Además, se ha descrito la mejora de la especificidad de fijación para los dominios de fijación con dedos de cinc, por ejemplo, en la solicitud de patente internacional en copropiedad WO 02/077227.

Además, tal y como se describe en estas y otras referencias, los dominios de ADN (p. ej., dominios TALE o proteínas con numerosos dedos de cinc) se pueden unir entre sí con cualquier secuencia conectora idónea, que incluye, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véanse también las patentes de los EE. UU. n.ºs 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras de ejemplo de 6 o más aminoácidos de longitud. Las proteínas de fijación al ADN descritas en la presente memoria pueden incluir cualquier combinación de conectores idóneos entre cada uno de los dedos de cinc de la proteína. Además, la mejora de la especificidad de fijación para los dominios de

fijación con dedos de cinc se ha descrito, por ejemplo, en la solicitud de patente internacional en copropiedad WO 02/077227.

La selección de los sitios diana, los dominios de fijación al ADN y los métodos para el diseño y la construcción de las proteínas de fusión (y los polinucleótidos que las codifican) los conocen bien los expertos en la técnica y se describen con detalle en las patentes de los EE. UU. n.^{os} 6.140.0815; 789.538; 6.453.242; 6.534.261; 5.925.523; 6.007.988; 6.013.453; 6.200.759; las solicitudes de patente internacional 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/53059; WO 98/53060; WO 02/016536 y WO 03/016496, y la publicación de patente de los EE. UU. n.º 20110301073.

Además, tal y como se describe en estas y otras referencias, los dominios de fijación al ADN (p. ej., proteínas con numerosos dedos de cinc) se puede unir entre sí mediante una secuencia conectora idónea, que incluye, por ejemplo, conectores de 5 o más aminoácidos de longitud. Véanse también las patentes de los EE. UU. n.^{os} 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras de ejemplo de 6 o más aminoácidos de longitud. Las proteínas descritas en la presente memoria pueden incluir cualquier combinación de conectores idóneos entre cada uno de los dedos de cinc de la proteína.

B. Dominios de escisión

Cualquier dominio de escisión idóneo se puede unir operativamente a un dominio de fijación al ADN para formar una nucleasa. Por ejemplo, los dominios de fijación al ADN de las PDZn se han fusionado a dominios de nucleasa para crear las NDZn, una entidad funcional que es capaz de reconocer la diana de ácido nucleico para la que se diseñó a través de su dominio de fijación al ADN manipulado genéticamente (PDZn) y hacer que el ADN se corte cerca del sitio de fijación de la PDZn a través de la actividad de la nucleasa. Véase, p. ej., Kim et al. (1996) *Proc Natl Acad Sci USA* 93(3): 1156-1160. Más recientemente, las NDZn se han utilizado para la modificación del genoma de muchos organismos diferentes. Véanse, por ejemplo, las publicaciones de patente de los Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; y la publicación internacional WO 07/014275. Asimismo, los dominios TALE de fijación al ADN se han fusionado a dominios de nucleasa para crear las TALEN. Véase, p. ej., la publicación de patente de los EE. UU. n.º 20110301073.

Tal y como se observó más arriba, el dominio de escisión puede ser heterólogo al dominio de fijación al ADN, por ejemplo, un dominio de fijación al ADN con dedos de cinc y un dominio de escisión de una nucleasa o un dominio TALEN de fijación al ADN y un dominio de escisión, o un dominio de fijación al ADN de meganucleasa y un dominio de escisión de una nucleasa diferente. Los dominios de escisión heterólogos pueden proceder de cualquier endonucleasa o exonucleasa. Las endonucleasas de ejemplo de las cuales se puede obtener un dominio de escisión incluyen, pero sin limitarse a ellas, endonucleasas de restricción y endonucleasas de asentamiento. Véanse, por ejemplo, el catálogo de 2002-2003 de 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 de S1, nucleasa de la judía mungo; DNasa I pancreática; nucleasa microcócica, endonucleasa HO de la levadura; véase también Linn et al. (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Uno o varias de estas enzimas (o fragmentos funcionales las mismas) se pueden utilizar como fuente de dominios de escisión y semidominios de escisión.

De igual forma, un semidominio de escisión puede proceder de cualquier nucleasa o porción de la misma, tal y como se presenta más arriba, que requiera la dimerización para la actividad de escisión. En general, se necesitan dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Como alternativa, se puede utilizar una única proteína que comprenda dos semidominios de escisión. Los dos semidominios de escisión pueden proceder de la misma endonucleasa (o fragmentos funcionales de la misma) o cada semidominio de escisión puede proceder de una endonucleasa diferente (o fragmentos funcionales de las mismas). Además, los sitios diana para las dos proteínas de fusión se sitúan, preferiblemente, el uno con respecto al otro, de tal manera que la fijación de las dos proteínas de fusión a sus respectivos sitios diana coloca los semidominios de escisión en una orientación espacial del uno con respecto al otro que permite que los semidominios de escisión formen un dominio de escisión funcional, p. ej., mediante dimerización. Así pues, en determinadas realizaciones, los extremos cercanos a los sitios diana están separados por 5 a 8 nucleótidos o por 15 a 18 nucleótidos. No obstante, los dos sitios diana pueden estar separados por cualquier número entero de nucleótidos o pares de nucleótidos (p. ej., de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión está comprendido entre los sitios diana.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de fijarse al ADN en secuencias específicas (en un sitio de reconocimiento) y de escindir el ADN en el sitio de la fijación o cerca de él. Determinadas enzimas de restricción (p. ej., de tipo IIS) escinden el ADN en los sitios retirados del sitio de reconocimiento, y se pueden separar los dominios de fijación y de escisión. Por ejemplo, la enzima de tipo IIS *FokI* cataliza la escisión bicatenaria del ADN, a 9 nucleótidos de su sitio de reconocimiento en una hebra y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véanse, por ejemplo, las patentes de los EE. UU. 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: 31978-31982. Así pues, en una realización, 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 varios dominios con dedos de cinc, que pueden o no

estar manipulados genéticamente.

Una enzima de restricción de tipo IIS de ejemplo, cuyo dominio de escisión se puede separar del dominio de fijación, es *FokI*. Esta enzima en concreto es activa en forma de dímero. Bitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 10570-10575. Por consiguiente, para los propósitos de la presente descripción, la porción de la enzima *FokI* utilizada en las proteínas de fusión descritas se considera un semidominio de escisión. Así pues, para la escisión bicatenaria dirigida y/o el reemplazo dirigido de las secuencias celulares mediante el uso de fusiones de dedos de cinc con *FokI*, se pueden utilizar dos proteínas de fusión, cada una con un semidominio de escisión de *FokI*, para reconstituir un dominio de escisión catalíticamente activo. Como alternativa, también se puede utilizar una única molécula polipeptídica que contiene un dominio de fijación al ADN y dos semidominios de escisión de *FokI*.

Un dominio de escisión o semidominio de escisión puede ser cualquier porción de una proteína que conserva la actividad de escisión o que conserva la capacidad para multimerizarse (p. ej., dimerizarse) para formar un dominio de escisión funcional.

Las enzimas de restricción de tipo IIS de ejemplo se describen en la publicación de patente internacional WO 07/014275. Otras enzimas de restricción también contienen dominios de escisión y de fijación que se pueden separar, y están contempladas en la presente descripción. Véase, por ejemplo, Roberts et al. (2003) *Nucleic Acids. Res.* 31: 418-420.

En determinadas realizaciones, el dominio de escisión comprende uno o varios semidominios de escisión manipulados genéticamente (también denominados mutantes del dominio de dimerización) que disminuyen al mínimo o impiden la homodimerización, tal y como se describe, por ejemplo, en las patentes de los EE. UU. n.ºs 20050064474; 20060188987 y 20080131962. Los restos de aminoácidos de las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de *FokI* son todos ellos dianas que influyen en la dimerización de los semidominios de escisión de *FokI*.

Los semidominios de escisión manipulados genéticamente de ejemplo de *FokI* que forman heterodímeros obligatoriamente incluyen una pareja en la que un primer semidominio de escisión incluye mutaciones en los restos aminoácidos de las posiciones 490 y 538 de *FokI* y un segundo semidominio de escisión incluye mutaciones en los restos aminoácidos 486 y 499.

Así pues, en una realización, una mutación de la posición 490 reemplaza la Glu (E) por Lys (K); la mutación en 538 reemplaza la Iso (I) por Lys (K); la mutación en 486 reemplaza la Gln (Q) por Glu (E); y la mutación de la posición 499 reemplaza la Iso (I) por Lys (K). Específicamente, los semidominios de escisión manipulados genéticamente descritos en la presente memoria se prepararon mediante la mutación de las posiciones 490 (E→K) y 538 (I→K) en un semidominio de escisión para producir un semidominio de escisión manipulado genéticamente que se denomina «E409K:I538K» y mediante la mutación de las posiciones 486 (Q→E) y 499 (I→L) en otro semidominio de escisión para producir un semidominio de escisión manipulado genéticamente que se denomina «Q486E:I499L». Los semidominios de escisión manipulados genéticamente que se describen en la presente memoria son mutantes que forman heterodímeros obligatoriamente en los que la escisión aberrante está disminuida al mínimo o totalmente ausente. Véase, p. ej., la publicación de patente de los EE. UU. n.º 2008/0131962.

En determinadas realizaciones, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 486, 499 y 496 (número referido a la *FokI* de tipo silvestre), por ejemplo, mutaciones que reemplazan el resto de Gln (Q) de tipo silvestre de la posición 486 por un resto de Glu (E), el resto de Iso (I) de tipo silvestre de la posición 499 por un resto de Leu (L), y el resto de Asn (N) de tipo silvestre de la posición 496 por un resto de Asp (D) o Glu (E) (también denominados dominios «ELD» y «ELE», respectivamente). En otras realizaciones, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 490, 538 y 537 (números referidos a la *FokI* de tipo silvestre), por ejemplo, mutaciones que reemplazan el resto de Glu (E) de tipo silvestre de la posición 490 por un resto de Lys (K), el resto de Iso (I) de tipo silvestre de la posición 538 por un resto de Lys (K), y el resto de His (H) de tipo silvestre de la posición 537 por un resto de Lys (K) o un resto de Arg (R) (también denominados dominios «KKK» y «KKR», respectivamente). En otras realizaciones, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 490 y 537 (numeración referida a la *FokI* de tipo silvestre), por ejemplo, mutaciones que reemplazan el resto de Glu (E) de tipo silvestre de la posición 490 por un resto de Lys (K) y el resto de His (H) de tipo silvestre de la posición 537 por un resto de Lys (K) o un resto de Arg (R) (también denominados dominios «KIK» y «KIR», respectivamente) (véase la publicación de patente de los EE. UU. n.º 20110201055). Los semidominios de escisión manipulados genéticamente que se describen en la presente memoria se pueden preparar mediante cualquier método idóneo, por ejemplo, mediante mutagénesis específica de sitio de los semidominios de escisión de tipo silvestre (*FokI*) como se describe en las publicaciones de patente de los EE. UU. n.ºs 20050064474; 20080131962; y 20110201055.

Como alternativa, las nucleasas se pueden ensamblar *in vivo* en el sitio diana del ácido nucleico mediante la cuestionable denominación de tecnología de «enzima divisora» (véase, p. ej., la publicación de patente de los EE. UU. n.º 20090068164). Los componentes de tales enzimas divisoras se pueden expresar desde construcciones de expresión independientes o bien pueden estar conectadas en un marco abierto de lectura allí donde los componentes individuales están separados, por ejemplo, mediante una secuencia de IRES o de péptido 2A de autoescisión. Los componentes pueden ser dominios de fijación con dedos de cinc individuales o dominios de un dominio de fijación a ácido nucleico de

una meganucleasa.

Las nucleasas se pueden cribar por su actividad antes de usarlas, por ejemplo, en un sistema de cromosoma de levadura tal y como se describe en las solicitudes de patente internacional WO 2009/042163 y 20090068164. Las construcciones de expresión para nucleasa se pueden diseñar con facilidad con los métodos conocidos en la técnica. Véanse, por ejemplo, las publicaciones de patente de los Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; y la solicitud de patente internacional WO 07/014275. La expresión de la nucleasa puede ser bajo el control de un promotor constitutivo o de un promotor inducible, por ejemplo, el promotor de la galactocinasa, que se activa (deja de reprimirse) en presencia de rafinosa y/o galactosa, y se reprime en presencia de glucosa.

El sistema con nucleasa CRISPR (grupos de repeticiones palindrómicas cortas en intervalos regulares)/Cas (asociado a CRISPR) es un sistema con nucleasa recientemente manipulado genéticamente en un sistema bacteriano que se puede utilizar para la manipulación genética del genoma. Se basa en parte de la respuesta inmunitaria adaptativa de muchas bacterias y arqueas. Cuando un virus o un plásmido invade una bacteria, algunos segmentos del ADN del invasor se convierten en ARN de CRISPR (crRNA) mediante la respuesta 'inmunitaria'. Este crRNA se asocia entonces por una región de complementariedad parcial con otro tipo de ARN denominado tracrRNA para guiar a la nucleasa Cas9 a una región homóloga al crRNA en el ADN diana denominada «protoespaciador». La Cas9 escinde el ADN para generar extremos romos tras el corte bicatenario en los sitios especificados por una secuencia guía de 20 nucleótidos que está contenida dentro del transcrito de crRNA. La Cas9 necesita tanto el crRNA como el tracrRNA para el reconocimiento y escisión del ADN específico del sitio. Este sistema se acaba de manipular genéticamente de tal manera que el crRNA y el tracrRNA se pueden combinar en una única molécula (el «ARN guía único»), y la porción equivalente al crRNA del ARN guía único se puede manipular genéticamente para que guíe a la nucleasa Cas9 hacia cualquier secuencia diana deseada (véase Jinek et al. (2012) *Science* 337, 816-821, Jinek et al. (2013), *eLife* 2:e00471, y David Segal (2013) *eLife* 2:e00563). Así pues, el sistema CRISPR/Cas se puede manipular genéticamente para crear una rotura bicatenaria en una diana deseada de un genoma, y la reparación de la rotura bicatenaria puede verse influida por el uso de inhibidores de la reparación que ocasionarán un incremento de la reparación propensa a errores.

Sitios diana

Tal y como se describe en detalle más arriba, los dominios de ADN se pueden manipular genéticamente para que se fijen a cualquier secuencia elegida de un locus, por ejemplo, un gen de albúmina u otro gen de refugio. Un dominio de fijación al ADN manipulado genéticamente puede tener una nueva especificidad de fijación cuando se compara con el dominio de fijación al ADN que aparece en la naturaleza. Los métodos de manipulación (ingeniería) genética incluyen, pero sin limitarse a ellos, el diseño racional y diferentes tipos de selección. El diseño racional incluye, por ejemplo, la utilización de bases de datos que comprenden secuencias de nucleótidos de tripletes (o tetrapletes) y secuencias aminoacídicas individuales (por ejemplo, dedos de cinc), en los que cada secuencia nucleotídica triplete o tetraplete está asociada a una o varias secuencias de aminoácidos del dominio de fijación al ADN que se fija a la secuencia de triplete o tetraplete concreta. Véanse, por ejemplo, las patentes de los EE. UU. en copropiedad 6.453.242 y 6.534.261. También se puede llevar a cabo el diseño racional de los dominios efectores TAL. Véase, por ejemplo, la publicación n.º 20110301073.

Los métodos de selección de ejemplo que se pueden aplicar a los dominios de fijación al ADN, que incluyen la exposición en fagos y los sistemas de doble híbrido, se describen en las patentes de los EE. UU. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como las solicitudes de patente internacional WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y la patente de Gran Bretaña GB 2.338.237.

Selección de sitios diana; nucleasas y métodos para diseñar y construir proteínas de fusión (y los polinucleótidos que las codifican) los conocen bien los expertos en la técnica y se describen con detalle en las publicaciones de solicitud de patente de los EE. UU. n.ºs 20050064474 y 20060188987.

Además, tal y como se describe en estas y otras referencias, los dominios de fijación al ADN (p. ej., proteínas con varios dedos de cinc) se pueden unir entre sí mediante alguna secuencia conectora, que incluyen, por ejemplo, conectores de 5 o más aminoácidos. Véanse, p. ej., las patentes de los EE. UU. n.ºs 6.479.626; 6.903.185; y 7.153.949 para las secuencias conectoras de ejemplo de 6 o más aminoácidos de longitud. Las proteínas descritas en la presente memoria pueden incluir cualquier combinación de conectores adecuados entre cada uno de los dominios de fijación al ADN de la proteína. Véase, también, la publicación de los EE. UU. n.º 20110301073.

Donantes

Tal y como se observó más arriba, la inserción de una secuencia exógena (también denominada una «secuencia donante» o «donante»), por ejemplo, para la corrección de un gen mutante o para que se incremente la expresión de un gen de tipo silvestre. Se apreciará claramente con facilidad que la secuencia donante no es típicamente idéntica a la secuencia genómica en donde se coloca. Una secuencia donante puede contener una secuencia no homóloga flanqueada por dos regiones de homología para permitir la HDR eficaz en la localización de interés. Adicionalmente, las secuencias donantes pueden comprender una molécula de vector que contiene secuencias que no son homólogas a la

región de interés en la cromatina celular. Una molécula donante puede contener varias regiones discontinuas de homología con la cromatina celular. Por ejemplo, para la inserción dirigida de secuencias que no suelen estar presentes en una región de interés, dichas secuencias pueden estar presentes en una molécula de ácido nucleico donante y flanqueadas por regiones de homología a la secuencia de la región de interés.

5 El polinucleótido donante puede ser ADN o ARN, monocatenario o bicatenario, y se puede introducir en una célula en forma lineal o circular. Si se introduce en forma lineal, los extremos de la secuencia donante pueden estar protegidos (p. ej., de la degradación exonucleolítica) por los métodos conocidos por los expertos en la técnica. Por ejemplo, se añaden uno o varios restos de dideoxinucleótidos al extremo en 3' de una molécula lineal y/o se ligan oligonucleótidos autocomplementarios a uno o ambos extremos. Véanse, por ejemplo, Chang et al. (1987) *Proc. Natl. Acad. Sci. USA* 84: 4959-4963; Nehls et al. (1996) *Science* 272: 886-889. Otros métodos para proteger los polinucleótidos exógenos de la degradación incluyen, pero sin limitarse a ellos, la adición de uno o varios grupos amino en el extremo y el uso de enlaces modificados entre nucleótidos, tales como, por ejemplo, fosforotioatos, fosforamidatos y restos de O-metilribosa o desoxirribosa.

15 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 una resistencia a antibióticos. Además, los polinucleótidos donantes se pueden introducir como ácido nucleico desnudo, como ácido nucleico formando complejo con un agente tal como un liposoma o poloxámero, o se pueden introducir mediante virus (p. ej., adenovirus, AAV, virus del herpes, retrovirus, lentivirus y lentivirus carentes de integrasa (IDLV)).

25 El donante se inserta por lo general de tal modo que su expresión queda impulsada por el promotor endógeno del sitio de la integración, a saber, el promotor que impulsa la expresión del gen endógeno en el cual se inserta el donante (p. ej., muy expresado, albúmina, AAVS1, HPRT, etc.). Sin embargo, resultará evidente que el donante puede comprender un promotor y/o potenciador, por ejemplo, un promotor constitutivo o un promotor inducible o específico de tejido.

30 La molécula donante se puede insertar en un gen endógeno de tal manera que se exprese todo, parte o nada del gen endógeno. Por ejemplo, un transgén tal y como se describe en la presente memoria se puede insertar en un locus de albúmina u otro locus, de tal manera que se expresan parte (el extremo amino y/o el extremo carboxilo del transgén que codifica la enzima lisosomal) o nada de las secuencias de albúmina endógena, por ejemplo, como una fusión con el transgén que codifica las secuencias lisosomales. En otras realizaciones, el transgén (p. ej., con o sin otras secuencias codificantes tales como de la albúmina) se integra en cualquier locus endógeno, por ejemplo, un locus de refugio. Véanse, p. ej., las publicaciones de patente de los EE. UU. 20080299580; 20080159996 y 201000218264.

35 Cuando las secuencias endógenas (endógena o parte del transgén) se expresan con el transgén, las secuencias endógenas (p. ej., albúmina, etc.) pueden ser secuencias completas (de tipo silvestre o mutante) o secuencias parciales. Preferiblemente, las secuencias endógenas son funcionales. Los ejemplos no limitantes de la función de estas secuencias parciales o completas (p. ej. albúmina) incluyen el incremento de la semivida en el suero del polipéptido expresado por el transgén (p. ej., gen terapéutico) y/o actuar como un vehículo.

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

45 En determinadas realizaciones, la secuencia exógena (donante) comprende una fusión de una proteína de interés y, como su compañero de fusión, un dominio extracelular de una proteína de la membrana, lo que hace que la proteína de fusión esté localizada en la superficie de la célula. Esto permite que la proteína codificada por el transgén tenga la posibilidad de actuar en el suero. En el caso del tratamiento para una EAL, la enzima codificada por la fusión transgénica sería capaz de actuar sobre los productos metabólicos que se acumulan en el suero desde su localización en la superficie de la célula (p. ej., glóbulos rojos). Además, si el glóbulo rojo es engullido por un macrófago esplénico tal como ocurre en el transcurso normal de la degradación, el lisosoma formado cuando el macrófago engulle la célula expondría la proteína de fusión fijada a la membrana a concentraciones elevadas de productos metabólicos en el lisosoma al pH que favorezca de forma más natural a esa enzima. Los ejemplos no limitantes de los posibles compañeros de fusión se muestran a continuación en la tabla 1.

55

60

65

Tabla 1: Ejemplos de posibles compañeros de fusión

Nombre	Actividad
Banda 3	Transportador de aniones, constituye hasta el 25% de la proteína de la superficie de la membrana del glóbulo rojo
Acuaporina 1	Transportador de agua
Glut1	Transportador de glucosa y ácido L-deshidroascórbico
Proteína del antígeno Kidd	Transportador de urea
RhAG	Transportador de gas
ATP1A1, ATP1B1	Na ⁺ /K ⁺ -ATPasa
ATP2B1, ATP2B2, ATP2B3, ATP2B4	Ca ²⁺ -ATPasa
NKCC1, NKCC2	Cotransportador de Na ⁺ K ⁺ 2Cl ⁻
SLC12A3	Cotransportador de Na ⁺ Cl ⁻
SLC12A1, SLA12A2	Cotransportador de Na ⁺ K ⁺
KCC1	Cotransportador de K ⁺ Cl ⁻
KCNN4	Canal de Gardos

Las enfermedades por almacenamiento lisosomal se dividen típicamente en cinco clases. Estas clases se muestran a continuación en la tabla 2 junto con ejemplos específicos de las enfermedades. Así pues, las moléculas donantes descritas en la presente memoria pueden incluir secuencias que codifican una o varias enzimas ausentes o defectuosas en los sujetos con enfermedades por almacenamiento lisosomal, que incluyen, pero sin limitarse a ellas, las proteínas mostradas en la tabla 2.

Tabla 2: enfermedades por almacenamiento lisosomal

1. DEFECTOS EN LA DEGRADACIÓN DE GLUCANOS			
Proteína	Enfermedad	Gen asociado a la enfermedad	Producto acumulado
I. Defectos en la degradación de glucoproteínas			
α-Sialidasa (neuraminidasa)	Sialidosis	<i>NEU1</i>	Glucopéptidos y oligosacáridos sializados
Catepsina A	Galactosialidosis	<i>CTSA</i>	Polisacárido
α-Manosidasa lisosomal	α-Manosidosis	<i>MAN2B1</i>	Glucoproteínas y oligosacáridos ricos en manosa
β-Manosidasa lisosomal	β-Manosidosis	<i>MANBA</i>	
Glucosilasparaginasa	Aspartilglucosaminuria	<i>AGA</i>	Glucoasparaginas
α-L-Fucosidasa	Fucosidosis	<i>FUCA1</i>	Fucosa
α-N-Acetilglucosaminidasa	Síndrome de Sanfilippo B	<i>NAGLU</i>	Glucosaminoglucano
II. Defectos en la degradación de glucolípidos			
A. Gangliosido GM1			
β-Galactosidasa	Gangliosidosis GM1/MPS IVB	<i>GLB1</i>	Sulfato de queratano
Subunidad α de la β-hexosaminidasa	Gangliosidosis GM2 (Tay-Sachs)	<i>HEXA</i>	Gangliósido GM2
Subunidad β de la β-hexosaminidasa	Gangliosidosis GM2 (Sandhoff)	<i>HEXB</i>	Gangliósido GM2

(continuación)

1. DEFECTOS EN LA DEGRADACIÓN DE GLUCANOS			
Proteína	Enfermedad	Gen asociado a la enfermedad	Producto acumulado
Proteína activadora de GM2	Gangliosidosis GM2	GM2A	Gangliósido GM2
Glucocerebrosidasa	Enfermedad de Gaucher	GBA	Glucocerebrósido
Saposina C	Enfermedad de Gaucher (atípica)	PSAP	Glucocerebrósido
B. Defectos en la degradación de sulfátidos			
Arilsulfatasa A	Leucodistrofia metacromática	<i>ARSA</i>	Sulfátido
Saposina B	Leucodistrofia metacromática	<i>PSAP</i>	Sulfátido
Enzima generadora de formilglucina	Deficiencia de varias sulfatasas	<i>SUMF1</i>	Lípidos sulfatados
β -Galactosilceramidasa (Krabbe)	Leucodistrofia de células globoides	<i>GALC</i>	Galactocerebrósido
C. Defectos en la degradación de la globotriaosilceramida			
α -Galactosidasa A	Fabry	<i>GLA</i>	Globotriaosilceramida
III. Defectos en la degradación de glucosaminoglucanos (mucopolisacaridosis)			
A. Degradación del sulfato de heparano			
Iduronato sulfatasa	MPS II (Hunter)	<i>IDS</i>	Sulfato de dermatano, sulfato de heparano
Iduronidasa	MPS I (Hurler, Scheie)	<i>IDUA</i>	Sulfato de dermatano, sulfato de heparano
N-sulfatasa de heparano	MPS IIIa (Sanfilippo A)	<i>SGSH</i>	Sulfato de heparano
Acetil-CoA transferasa	MPS IIIc (Sanfilippo C)	<i>HGSNAT</i>	Sulfato de heparano
N-acetilglucosaminidasa	MPS IIIb (Sanfilippo B)	<i>NAGLU</i>	Sulfato de heparano
β -Glucuronidasa	MPS VII (Sly)	<i>GUSB</i>	
N-acetilglucosamina 6-sulfatasa	MPS IIId (Sanfilippo D)	<i>GNS</i>	Sulfato de heparano
B. Degradación de otros mucopolisacáridos			
B-Galactosidasa	MPS VIB (Morquio B)	<i>GLB1</i>	Sulfato de queratano
Galactosa 6-sulfatasa	MPS IVA (Morquio A)	<i>GALNS</i>	Sulfato de queratano, 6-sulfato de condroitina
Hialuronidasa	MPS IX	<i>HYAL1</i>	Ácido hialurónico
C. Defectos en la degradación del glucógeno			
α -Glucosidasa	Pompe	<i>GAA</i>	Glucógeno
2. DEFECTOS EN LA DEGRADACIÓN DE LÍPIDOS			
I. Defectos en la degradación de esfingomielinas			
Esfingomielinasa ácida	Niemann Pick de tipo A	<i>SMPD1</i>	Esfingomielina
Ceramidasa ácida	Lipogranulomatosis de Farber	<i>ASAHL1</i>	Mucopolisacárido ácido no sulfonatado

(continuación)

1. DEFECTOS EN LA DEGRADACIÓN DE GLUCANOS			
Proteína	Enfermedad	Gen asociado a la enfermedad	Producto acumulado
II. Defectos en la degradación de triglicéridos y de ésteres de colesterol			
Lipasa ácida	Enfermedad por almacenamiento de ésteres de colesterol y de Wolman	<i>LIPA</i>	Ésteres de colesterol
3. DEFECTOS EN LA DEGRADACIÓN DE PROTEÍNAS			
Catepsina K	Picnodistosis	<i>CTSK</i>	
Triptidilpeptidasa	Lipofuscinosis ceroides	<i>PPT2</i>	
Tioesterasa de proteínas palmitoiladas	Lipofuscinosis ceroides	<i>PPT1</i>	
4. DEFECTOS EN LOS TRANSPORTADORES LISOSOMALES			
Cistinosina (transporte de cistina)	Cistinosis	<i>CTNS</i>	
Sialina (transporte de ácido siálico)	Enfermedad de Salla	<i>SLC17A5</i>	Ácido N-acetilneuramínico
5. DEFECTOS EN LAS PROTEÍNAS DE TRANSPORTE LISOSOMAL			
Subunidad γ de la fosfotransferasa	Mucopolidosis III (célula I)	<i>GNPTG</i>	
Mucopolipina 1 (canal de cationes)	Mucopolidosis	<i>MCOLN1</i>	
Proteína membranaria 2 asociada al lisosoma	Danon	<i>LAMP2</i>	
Enfermedad de Niemann-Pick, de tipo C1	Niemann-Pick de tipo C	<i>NPC1</i>	Colesterol LDL
Tioesterasa 1 de proteínas palmitoiladas	Lipofuscinosis ceroides (enfermedad de Batten)	<i>CLN3</i>	Material de almacenamiento de lipopigmentos autofluorescentes
Lipofuscinosis 6 ceroides neuronal	Lipofuscinosis 6 ceroides neuronal	<i>CLN6</i>	
Lipofuscinosis 8 ceroides neuronal	Lipofuscinosis 8 ceroides neuronal	<i>CLN8</i>	
Regulador del transporte lisosomal	Chediak-Higashi	<i>LYST</i>	
Miocilina	Griscelli de tipo 1	<i>MYOC</i>	
Proteína 27A asociada a RAS	Griscelli de tipo 2	<i>RAB27A</i>	
Melanofilina	Griscelli de tipo 3	<i>MLPH o MYO5A</i>	
Subunidad β de AP3	Hermansky Pudlak	<i>AP3B1</i>	Ceroides

En algunos casos, el donante puede ser el gen endógeno que ha sido modificado. Aunque la respuesta de anticuerpos al tratamiento enzimático sustitutivo varía con respecto a la enzima terapéutica específica en cuestión y con cada uno de los pacientes, se ha visto una respuesta inmunitaria significativa en muchos pacientes con EAL que son tratados con el reemplazo enzimático. Además, la relevancia de estos anticuerpos para la eficacia del tratamiento también es variable (véase Katherine Ponder, (2008) *J Clin Invest* 118 (8): 2686). Así pues, los métodos y las composiciones de la presente invención pueden comprender el uso de las moléculas donantes cuya secuencia ha sido alterada mediante cambios de

aminoácidos funcionalmente silenciosos en sitios que se sabe que son epítomos sensibilizadores para las respuestas inmunitarias endógenas, de tal manera que el polipéptido producido por tal donante es menos inmunógeno.

Los pacientes con EAL a menudo tienen secuelas neurológicas debido a que el cerebro carece de la enzima ausente. Por desgracia, a menudo es difícil administrar tratamientos al cerebro a través de la sangre debido a la impermeabilidad de la barrera hematoencefálica. Así pues, los métodos y las composiciones de la invención se pueden utilizar junto con métodos que incrementan la entrada del agente terapéutico al cerebro. Hay algunos métodos que ocasionan una apertura transitoria de las uniones oclusivas entre las células de los capilares del cerebro. Los ejemplos incluyen la alteración osmótica transitoria mediante el uso de una administración intracarotídea de una solución hipertónica de manitol, el uso de ultrasonidos enfocados y la administración de un análogo de la bradicinina. Como alternativa, los agentes terapéuticos se pueden diseñar para que utilicen los receptores o mecanismos de transporte para el transporte específico al interior del cerebro. Los ejemplos de receptores específicos que se pueden utilizar incluyen el receptor de la transferrina, el receptor de la insulina o las proteínas 1 y 2 (LRP-1 y LRP-2) relacionadas con el receptor de las lipoproteínas de baja densidad. Se sabe que LRP interacciona con un abanico de proteínas secretadas, tales como apoE, tPA, PAI-1, etc., y, de esa forma, la fusión a una secuencia de reconocimiento de LRP de una de estas proteínas puede facilitar el transporte de la enzima al interior del cerebro, después de expresar en el hígado la proteína terapéutica y secretarla al torrente circulatorio (véase Gabathuler, (2010) mencionada más arriba).

Introducción

Las nucleasas, polinucleótidos que codifican estas nucleasas, polinucleótidos donantes y composiciones que comprenden las proteínas y/o polinucleótidos descritos en la presente memoria se pueden introducir *in vivo* o *ex vivo* mediante cualquier medio adecuado.

Los métodos para introducir nucleasas tal y como se describe en la presente memoria se describen, por ejemplo, en las patentes de los EE. UU. n.ºs 6.453.242; 6.503.717; 6.534.261; 6.599.692; 6.607.882; 6.689.558; 6.824.978; 6.933.113; 6.979.539; 7.013.219 y 7.163.824.

Las nucleasas y/o construcciones de donantes tal y como se describen en la presente memoria también se pueden introducir con vectores que contienen secuencias que codifican una o varias de las proteínas TALEN o con dedos de cinc. Se puede utilizar cualquier sistema de vector, entre ellos, pero sin limitarse a ellos, vectores plasmídicos, vectores retrovíricos, vectores lentivíricos, vectores adenovíricos, vectores del virus de la viruela; vectores del virus del herpes y vectores dependovíricos, etc. Véanse también las patentes de los EE. UU. n.ºs 6.534.261; 6.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824. Además, resultará evidente que cualquiera de estos vectores puede comprender una o varias de las secuencias necesarias para el tratamiento. Así pues, cuando una o más nucleasas y una construcción de donante se introducen en la célula, las nucleasas y/o el polinucleótido donante pueden ser llevados por el mismo vector o por vectores diferentes. Cuando se utilizan varios vectores, cada vector puede comprender una secuencia que codifica una o varias nucleasas y/o construcciones de donante.

Los métodos convencionales de transferencia génica basados o no en virus se pueden utilizar para introducir ácidos nucleicos que codifican nucleasas y construcciones de donantes en las células (p. ej., células de mamífero) y tejidos destinatarios. Los sistemas de introducción con vectores no víricos incluyen plásmidos de ADN, ácido nucleico desnudo y ácido nucleico formando complejo con un vehículo de introducción, tal como un liposoma o poloxámero. Los sistemas de introducción con vectores víricos incluyen virus de ADN y ARN, cuyos genomas quedan episómicos o bien integrados después de ser introducidos en la célula. Para una revisión de los procedimientos de genoterapia, véanse Anderson, *Science* 256: 808-813 (1992); Nabel y Felgener, *TIBTECH* 11: 211-217 (1993); Mitani y Caskey, *TIBTECH* 11: 162-166 (1993); Dillon, *TIBTECH* 11: 167-175 (1993); Miller, *Nature* 357: 455-460 (1992); Van Brunt, *Biotechnology* 6 (10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8: 35-36 (1995); Kremer y Perricaudet, *British Medical Bulletin* 51 (1): 31-44 (1995); Haddada et al., en *Current Topics in Microbiology and Immunology*, Doerfler y Böhm (eds.) (1995); y Yu et al., *Gene Therapy* 1: 13-26 (1994).

Los métodos para la introducción no vírica de ácidos nucleicos incluyen la electroporación, lipofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados ácido nucleico:lipido o policación, ADN desnudo, viriones artificiales y captación del ADN potenciada por el agente. También se puede utilizar la sonicoporación que utiliza, p. ej., el sistema Sonitron 2000 (Rich-Mar) para la introducción de ácidos nucleicos.

Otros sistemas de introducción de ácidos nucleicos de ejemplo incluyen los proporcionados por Amaxa Biosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) y Copernicus Therapeutics Inc. (véase, por ejemplo, la patente de los EE. UU. US 6.008.336). La lipofección se describe en, p. ej., las patentes de los EE. UU. n.ºs 5.049.386; 4.946.787; y 4.897.355) y los reactivos de lipofección están disponibles en el mercado (p. ej., TransfectamTM y LipofectinTM). Los lípidos catiónicos y neutros que son idóneos para la lipofección eficiente de los polinucleótidos por reconocimiento del receptor incluyen los de Felgner, solicitudes de patente internacional WO 91/17424 y WO 91/16024.

La preparación de los complejos lípido:ácido nucleico, entre ellos los liposomas dirigidos tales como los complejos inmunolipídicos, los conocen bien los expertos en la técnica (véanse, 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 los EE. UU. n.ºs 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).

5

Otros métodos de introducción incluyen el uso del empaquetamiento de los ácidos nucleicos a insertar en vehículos de introducción EnGeneIC (VIE). Estos VIE se administran específicamente a los tejidos destinatarios mediante anticuerpos biespecíficos, en donde un brazo del anticuerpo tiene especificidad por el tejido destinatario y el otro tiene especificidad por el VIE. El anticuerpo lleva el VIE a la superficie de la célula destinataria y a continuación el VIE se introduce en la célula por endocitosis. Una vez en la célula, se libera el contenido (véase MacDiarmid et al. (2009) *Nature Biotechnology* 27 (7): 643).

10

El uso de los sistemas basados en virus de ARN o ADN para la introducción de los ácidos nucleicos que codifican las PDZn manipuladas genéticamente tiene la ventaja de que son procesos muy evolucionados para dirigir un virus a las células específicas del cuerpo y transportar la carga útil vírica al núcleo. Los vectores víricos se pueden administrar directamente a los sujetos (*in vivo*) o se pueden utilizar para tratar células *in vitro* y entonces administrar las células modificadas a los sujetos (*ex vivo*). Los sistemas convencionales basados en virus para la introducción de las PDZn incluyen, pero sin limitarse a ellos, vectores para la transferencia génica retrovíricos, lentivíricos, adenovíricos, dependovíricos, del virus de la variolovacuna y del virus del herpes simple. Se puede integrar en el genoma del hospedador con los métodos de transferencia génica basados en retrovirus, lentivirus y dependovirus, que a menudo da lugar a la expresión a largo plazo del transgén insertado. Adicionalmente, se han observado altas eficacias de transducción en muchos tipos diferentes de células y de tejidos destinatarios.

15

20

El tropismo de un retrovirus se puede alterar con la incorporación de proteínas foráneas en la envoltura, lo que expande la posible población destinataria de las células diana. Los vectores lentivíricos son vectores retrovíricos que son capaces de transducir o infectar células que no se dividen y producir típicamente altos títulos víricos. La selección de un sistema retrovírico de transferencia génica depende del tejido destinatario. Los vectores retrovíricos están comprendidos por repeticiones terminales largas que actúan en *cis* con capacidad de empaquetamiento de hasta 6-10 kb de la secuencia foránea. Las LTR mínimas que actúan en *cis* son suficientes para la replicación y empaquetamiento de los vectores, que a continuación se utilizan para integrar el gen terapéutico en la célula destinataria para proporcionar la expresión permanente del transgén. Los vectores retrovíricos utilizados ampliamente incluyen los basados en el virus de la leucemia murina (MuLV), virus de la leucemia de los simios gibones (GaLV), virus de inmunodeficiencia de los simios (SIV), virus de inmunodeficiencia humana (HIV) y combinaciones de los mismos (véanse, 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); solicitud internacional de patente en el marco del PCT US94/05700).

25

30

35

En las aplicaciones en donde se prefiere la expresión transitoria, se pueden utilizar los sistemas basados en adenovirus. Los vectores basados en adenovirus son capaces transducir con una eficiencia muy alta en muchos tipos de células y no requieren la división celular. Con tales vectores, se han obtenido títulos elevados y niveles altos de expresión. Este vector se puede producir en grandes cantidades en un sistema relativamente simple. Los vectores de dependovirus («AAV») también se utilizan para transducir células con los ácidos nucleicos deseados, p. ej., para la producción *in vitro* de ácidos nucleicos y péptidos, y para los procedimientos de genoterapia *in vivo* y *ex vivo* (véanse, p. ej., West et al., *Virology* 160: 38-47 (1987); patente de los EE. UU. n.º 4.797.368; solicitud de patente internacional 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 una serie de publicaciones, entre ellas, la patente de los EE. UU. n.º 5.173.414; Tratschin et al., *Mol. Cell. Biol.* 5: 3251-3260 (1985); Tratschin et al., *Mol. Cell. Biol.* 4: 2072-2081 (1984); Hermonat y Muzyczka, *PNAS* 81: 6466-6470 (1984); y Samulski et al., *J. Virol.* 63: 3822-3828 (1989).

40

45

50

Al menos seis estrategias de vectores víricos están actualmente disponibles para la transferencia génica en los ensayos clínicos, las cuales utilizan estrategias que implican la complementación de vectores defectuosos por genes insertados en líneas de células cooperadoras para generar el agente transductor.

pLASN y MFG-S son ejemplos de los vectores retrovíricos que se han utilizado en los 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 utilizado en un ensayo de genoterapia (Blaese et al., *Science* 270: 475-480 (1995)). Se han observado unas eficiencias de transducción del 50% o más para los vectores de MFG-S empaquetados (Ellem et al., *Immunol Immunother.* 44 (1): 10-20 (1997); Dranoff et al., *Hum. Gene Ther.* 1: 111-2 (1997)). Los vectores dependovíricos recombinantes (rAAV) son sistemas alternativos y prometedores de inserción de genes basados en el dependovirus de tipo 2 de parvovirus no patógeno y defectuoso. Todos los vectores proceden de un plásmido que conserva solo las repeticiones terminales invertidas de 145 pb del AAV flanqueando el casete de expresión del transgén. La transferencia génica eficiente y la administración estable del transgén debido a la integración en los genomas de la célula transducida son características clave para este sistema de vector (Wagner et al., *Lancet* 351 (9117): 1702-3 (1998); Kearns et al., *Gene Ther.* 9: 748-55 (1996)). Otros serotipos de AAV, entre ellos, pero sin limitarse a ellos, por ejemplo, AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV8.2, AAV9 y AAV rh10, y un AAV pseudotipado tal como AAV2/8, AAV2/5 y AAV2/6 también se pueden utilizar de acuerdo con la presente invención.

60

65

Los vectores adenovíricos (Ad) recombinantes incapaces de replicarse se pueden producir en títulos altos y pueden infectar con facilidad una serie de tipos de células diferentes. La mayoría de vectores adenovíricos están manipulados genéticamente de tal manera que un transgén reemplaza los genes de Ad E1a, E1b y/o E3; posteriormente, el vector incapaz de replicarse se propaga en las células 293 humanas que suministran en *trans* la función del gen eliminado. Los vectores de Ad pueden transducir varios tipos de tejidos *in vivo*, entre ellos las células diferenciadas que no se dividen, tales como las encontradas en hígado, riñón y músculo. Los vectores de Ad convencionales tienen una gran capacidad de transporte. Un ejemplo del uso de un vector de Ad en un ensayo clínico implicó la terapia con polinucleótidos para la inmunización antitumoral con inyección intramuscular (Sternan et al., *Hum. Gene Ther.* 7: 1083-9 (1998)). Otros ejemplos del uso de los vectores adenovíricos para la transferencia génica en los 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 utilizan para formar partículas de virus que son capaces de infectar una célula hospedadora. Tales células incluyen las células 293, que empaquetan adenovirus, y las células ψ 2 o las células PA317, que empaquetan retrovirus. Los vectores víricos utilizados en la genoterapia se suelen generar mediante una línea celular productora que empaqueta un vector de ácido nucleico en una partícula vírica. Los vectores típicamente contienen las secuencias víricas mínimas requeridas para el empaquetamiento y la posterior integración en un hospedador (si es aplicable), otras secuencias víricas se reemplazan por un casete de expresión que codifica la proteína a expresar. Las funciones víricas ausentes están suministradas en *trans* por la línea de células de empaquetamiento. Por ejemplo, los vectores de AAV utilizados en la genoterapia típicamente poseen sólo secuencias de repeticiones terminales invertidas (ITR) del genoma del AAV que son necesarias para el empaquetamiento y la integración en el genoma hospedador. El ADN vírico se empaqueta en una línea celular, que contiene un plásmido cooperador que codifica los otros genes del AAV, a saber, *rep* y *cap*, pero que carecen de las secuencias ITR. La línea celular también está infectada con el adenovirus a modo de cooperador. El virus cooperador favorece la replicación del vector de AAV y la expresión de los genes del AAV desde el plásmido cooperador. El plásmido cooperador no se empaqueta en cantidades significativas debido a la ausencia de secuencias ITR. La contaminación con el adenovirus se puede reducir mediante, p. ej., el tratamiento con calor al cual el adenovirus es más sensible que el AAV.

En muchas aplicaciones genoterápicas es deseable que el vector genoterápico se inserte con un alto grado de especificidad por un tipo de tejido concreto. En concordancia, un vector vírico se puede modificar para que sea específico de un tipo celular dado mediante la expresión de un ligando tal como una proteína de fusión con una proteína de la envoltura vírica sobre la superficie externa del virus. El ligando se escoge por tener afinidad por un receptor que se sabe que está presente en un tipo de célula de interés. Por ejemplo, Han et al. *Proc. Natl. Acad. Sci. USA* 92: 9747-9751 (1995) describieron que el virus de la leucemia murina de Moloney se puede modificar para que exprese la herregulina humana fusionada a la gp70, y los virus recombinantes infectan determinadas células de cáncer de mama de humano que expresan el receptor del factor de crecimiento epidérmico de humanos. Este principio se puede extender a otras parejas de virus y células destinatarias, en las que la célula destinataria expresa un receptor y el virus expresa una proteína de fusión que comprende un ligando del receptor de la superficie celular. Por ejemplo, se puede manipular genéticamente un fago filamentoso para que presente fragmentos de anticuerpo (p. ej., FAb o Fv) que tienen una afinidad de fijación específica por prácticamente cualquier receptor celular escogido. Aunque la descripción anterior se aplica principalmente a los vectores víricos, los mismos principios se pueden aplicar a los vectores que no sean víricos. Tales vectores se pueden manipular genéticamente para que contengan secuencias de captación específicas que favorecen la captación por las células destinatarias específicas.

Los vectores genoterápicos se pueden insertar *in vivo* mediante la administración a un paciente determinado, típicamente por administración sistémica (p. ej., infusión intravenosa, intraperitoneal, intramuscular, subdérmica o intracraneal), o por aplicación tópica, tal y como se describe más adelante. Como alternativa, los vectores se pueden introducir en las células *ex vivo*, tales como las células explantadas de un paciente determinado (p. ej., linfocitos, aspirados de médula ósea, biopsia de tejido) o las células troncales hematopoyéticas donantes universales, y luego se reimplantan las células en un paciente, normalmente después de seleccionar las células que han incorporado el vector. Los vectores (p. ej., retrovirus, adenovirus, liposomas, etc.) que contienen nucleasas y/o construcciones donantes también se pueden administrar directamente a un organismo para la transducción de las células *in vivo*. Como alternativa, se puede administrar el ADN desnudo. La administración se realiza por cualquiera de las vías que normalmente se usan para introducir una molécula en contacto final con la sangre o las células del tejido, que incluyen, pero sin limitarse a ellas, inyección, infusión, aplicación tópica y electroporación. Están disponibles los métodos idóneos para la administración de tales ácidos nucleicos y los conocen bien los expertos en la técnica, y, aunque se puede usar más de una vía para administrar una composición concreta, una vía particular puede ofrecer a menudo una reacción más inmediata y más eficaz que otra vía.

Los vectores idóneos para la introducción de los polinucleótidos descritos en la presente memoria incluyen los vectores lentivíricos no integrativos (IDLV). Véanse, 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 los EE. UU. 2009/054985.

Los vehículos farmacéuticamente aceptables están determinados en parte por la composición concreta que se ha de

administrar, así como por el método concreto que se utiliza para administrar la composición. En consecuencia, está a disposición una amplia gama de formulaciones idóneas de composiciones farmacéuticas, tal y como se describe más adelante (véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 17.^a edición, 1989).

- 5 Queda claro que las secuencias que codifican nucleasas y construcciones donantes se puede insertar mediante el mismo sistema o con uno diferente. Por ejemplo, un polinucleótido donante puede venir en un plásmido, mientras que una o varias nucleasas pueden venir en un vector de AAV. Además, los diferentes vectores se pueden administrar por la misma vía o por vías diferentes (inyección intramuscular, inyección en la vena de la cola, otra inyección venosa, administración intraperitoneal y/o inyección intramuscular). Los vectores se pueden introducir simultáneamente o en cualquier orden secuencial.

Las formulaciones para las administraciones tanto *ex vivo* como *in vivo* incluyen suspensiones en líquido o líquidos emulsionados. Los ingredientes activos a menudo están mezclados con excipientes que son farmacéuticamente aceptables y compatibles con el ingrediente activo. Los excipientes idóneos incluyen, por ejemplo, agua, solución salina, dextrosa, glicerol, etanol y similares, y combinaciones de los mismos. Además, la composición podría contener cantidades menores de sustancias auxiliares, tales como humectantes o emulsionantes, tamponantes del pH, estabilizantes y otros reactivos que realzan la eficacia de la composición farmacéutica.

Aplicaciones

20 Los métodos de esta invención contemplan el tratamiento de una enfermedad monogénica (p. ej., enfermedad por almacenamiento lisosomal). El tratamiento puede comprender la inserción del gen corregido asociado a la enfermedad en el locus de refugio (p. ej., albúmina) para que se exprese la enzima necesaria y se libere al torrente circulatorio. Resulta particularmente útil la inserción en una célula secretora, tal como un hepatocito, para liberar el producto al torrente circulatorio. Los métodos y composiciones de la invención también se pueden usar en cualquier circunstancia, en la que se desea el suministro de un transgén que codifica uno o varios agentes terapéuticos en una célula troncal hematopoyética, tal como células maduras (por ejemplo, glóbulos rojos) procedentes de las células que contienen el agente terapéutico. Estas células troncales se pueden diferenciar *in vivo* e *in vitro* y pueden proceder de un tipo de célula donante universal que se puede usar con todos los pacientes. Además, las células pueden contener una proteína transmembranaria para transportar las células por el cuerpo. El tratamiento también puede comprender el uso de células del paciente que contienen el transgén terapéutico, en donde las células se desarrollan *ex vivo* y luego se devuelven al paciente. Por ejemplo, las CTM que contienen un transgén idóneo se podrían insertar en un paciente a través del trasplante de médula ósea. Como alternativa, las células troncales tales como las células troncales del músculo o las CTPi, que se han editado con el transgén terapéutico, se podrían inyectar también en el tejido muscular.

35 Así pues, esta tecnología podría ser muy útil para una afección en la que un paciente carece de alguna proteína debido a problemas (p. ej., problemas de nivel de expresión o problemas con la proteína expresada al ser subfuncional o afuncional). Resulta particularmente útil con esta invención la expresión de transgenes para corregir o restaurar la funcionalidad de las enfermedades por almacenamiento lisosomal.

40 A modo de ejemplos no limitantes, se consiguió la producción de las proteínas defectuosas o ausentes, y se usó para tratar la enfermedad por almacenamiento lisosomal. Los ácidos nucleicos donantes que codifican las proteínas se pueden insertar en un locus de refugio (por ejemplo, albúmina o HPRT) y se pueden expresar gracias a un promotor exógeno o bien gracias al promotor presente en el refugio. Como alternativa, los donantes se pueden usar para corregir el gen defectuoso *in situ*. El transgén deseado se puede insertar en una célula troncal CD34+ y devolverse al paciente durante un trasplante de médula ósea. Finalmente, el ácido nucleico donante se podría insertar en una célula troncal CD34+ en un locus de β -globina, de tal manera que los glóbulos rojos maduros procedentes de esta célula tienen una concentración elevada del agente biológico codificado por el ácido nucleico donante. El agente biológico que contienen los glóbulos rojos puede ser entonces dirigido al tejido correcto a través de las proteínas transmembranarias (p. ej., receptor o anticuerpo). Además, los glóbulos rojos se pueden sensibilizar *ex vivo* a través de la electrosensibilización para hacerlos más propensos a romperse tras la exposición a una fuente de energía (véase la solicitud de patente internacional WO 2002007752).

55 En algunas aplicaciones, un gen endógeno puede ser desactivado mediante el uso de métodos y composiciones de la invención. Ejemplos de este aspecto incluyen la desactivación de un regulador génico aberrante o de un gen aberrante asociado a la enfermedad. En algunas aplicaciones, un gen endógeno aberrante se podría reemplazar, bien funcionalmente o bien *in situ*, por una versión de tipo silvestre del gen. El gen insertado se podría alterar también para mejorar la funcionalidad de la proteína expresada o reducirle la inmunogenia. En algunas aplicaciones, el gen insertado es una proteína de fusión que incrementa su transporte a un tejido seleccionado, tal como el cerebro.

60 Los siguientes ejemplos se refieren a realizaciones de ejemplo de la presente descripción en donde la nucleasa comprende una nucleasa con dedos de cinc (NDZn) o TALEN. Se apreciará que esto tiene el propósito de ilustrar solo a modo de ejemplo y que se pueden usar otras nucleasas o sistemas de nucleasas, por ejemplo, endonucleasas de asentamiento (meganucleasas) con dominios de fijación al ADN manipulados genéticamente y dominios de escisión heterólogos y/o un sistema CRISPR/Cas que comprende un ARN guía único manipulado genéticamente.

EJEMPLOS**Ejemplo 1: Diseño, construcción y caracterización general de las nucleasas específicas de la albúmina**

- 5 Se diseñaron proteínas con dedos de cinc y se incorporaron en plásmidos, vectores de AAV o adenovíricos esencialmente como se describe en Urnov et al. (2005) *Nature* 435 (7042): 646-651, Perez et al. (2008) *Nature Biotechnology* 26 (7): 808-816 y como se describe en la patente de los EE. UU. n.º 6.534.261. La tabla 3 muestra las hélices de reconocimiento dentro de cada dominio de fijación al ADN de las PDZn específicas de la albúmina de ejemplo, mientras que la tabla 4 muestra los sitios diana para estas PDZn (véanse los documentos provisionales de los EE. UU. en copropiedad 61/537.349 y 61/560.506). Los nucleótidos del sitio diana que entran en contacto a través de las hélices de reconocimiento de la PDZn están indicados en letras mayúsculas; los nucleótidos que no entran en contacto están indicados en letras minúsculas. También se diseñaron TALEN específicas de la albúmina y se presentan en las solicitudes de los EE. UU. n.ºs 13/624.193 y 13/624.217).

Tabla 3: Diseños de hélices para nucleasas con dedos de cinc específicas de la albúmina murina

Diana	SBS #	Diseño					
		F1	F2	F3	F4	F5	F6
Intrón 1	30724	TSGSLTR (SEQ ID n.º 1)	RSDALST (SEQ ID n.º 2)	QSATRTK (SEQ ID n.º 3)	TSGHLNR (SEQ ID n.º 4)	QSGNLAR (SEQ ID n.º 5)	NA
Intrón 1	30725	RSDHLA (SEQ ID n.º 6)	TKSNRTK (SEQ ID n.º 7)	DRSNLNR (SEQ ID n.º 8)	WRSSLRA (SEQ ID n.º 9)	DSSDRKK (SEQ ID n.º 10)	NA
Intrón 1	30732	TSGNLTR (SEQ ID n.º 11)	DRSTRRQ (SEQ ID n.º 12)	TSGSLTR (SEQ ID n.º 1)	ERGTLAR (SEQ ID n.º 13)	TSANLNR (SEQ ID n.º 14)	NA
Intrón 1	30733	DRSALAR (SEQ ID n.º 15)	RSDHLSE (SEQ ID n.º 16)	HRSDRTR (SEQ ID n.º 17)	QSGALAR (SEQ ID n.º 18)	QSGHLNR (SEQ ID n.º 19)	NS
Intrón 13	30759	RSDNLST (SEQ ID n.º 20)	DRSALAR (SEQ ID n.º 15)	DRSNLNR (SEQ ID n.º 8)	DGRNLRH (SEQ ID n.º 21)	RSDNLAR (SEQ ID n.º 22)	QSNALNR (SEQ ID n.º 23)
Intrón 13	30761	DRSNLNR (SEQ ID n.º 8)	LKQVLVR (SEQ ID n.º 24)	QSGNLAR (SEQ ID n.º 5)	QSTPLFA (SEQ ID n.º 25)	QSGALAR (SEQ ID n.º 18)	NA
Intrón 13	30760	DRSNLNR (SEQ ID n.º 8)	DGRNLRH (SEQ ID n.º 21)	RSDNLAR (SEQ ID n.º 22)	QSNALNR (SEQ ID n.º 23)	NA	NA
Intrón 13	30767	RSDNLNV (SEQ ID n.º 26)	HSNARKT (SEQ ID n.º 27)	RSDSLA (SEQ ID n.º 28)	QSGNLAR (SEQ ID n.º 5)	RSDSLNV (SEQ ID n.º 29)	QSGHLNR (SEQ ID n.º 19)
Intrón 13	30768	RSDNLSE (SEQ ID n.º 30)	ERANRNS (SEQ ID n.º 31)	QSANRTK (SEQ ID n.º 32)	ERGTLAR (SEQ ID n.º 13)	RSDALTQ (SEQ ID n.º 33)	NA
Intrón 13	30769	TSGSLTR (SEQ ID n.º 1)	DRSNLNR (SEQ ID n.º 8)	DGRNLRH (SEQ ID n.º 21)	ERGTLAR (SEQ ID n.º 13)	RSDALTQ (SEQ ID n.º 33)	NA
Intrón 12	30872	QSGHLAR (SEQ ID n.º 34)	RSDHLTQ (SEQ ID n.º 35)	RSDHLSQ (SEQ ID n.º 36)	WRSSLVA (SEQ ID n.º 37)	RSDVLSE (SEQ ID n.º 38)	RNQRKT (SEQ ID n.º 39)
Intrón 12	30873	QSGDLTR (SEQ ID n.º 40)	RSDALAR (SEQ ID n.º 41)	QSGDLTR (SEQ ID n.º 40)	RRDPLIN (SEQ ID n.º 42)	RSDNLNV (SEQ ID n.º 26)	IRSTLRD (SEQ ID n.º 43)

(continua)

5	Intrón 12	30876	RSDNLSV (SEQ ID n.º 26)	YSSTRNS (SEQ ID n.º 44)	RSDHLSA (SEQ ID n.º 6)	SYWSRTV (SEQ ID n.º 45)	QSSDLSR (SEQ ID n.º 46)	RTDALRG (SEQ ID n.º 47)
	Intrón 12	30877	RSDNLST (SEQ ID n.º 20)	QKSPLNT (SEQ ID n.º 48)	TSGNLTR (SEQ ID n.º 11)	QAENLKS (SEQ ID n.º 49)	QSSDLSR (SEQ ID n.º 46)	RTDALRG (SEQ ID n.º 47)
10	Intrón 12	30882	RSDNLSV (SEQ ID n.º 26)	RRAHLNQ (SEQ ID n.º 50)	TSGNLTR (SEQ ID n.º 11)	SDTNRFK (SEQ ID n.º 51)	RSDNLST (SEQ ID n.º 20)	QSGHLSR (SEQ ID n.º 19)
15	Intrón 12	30883	DSSDRKK (SEQ ID n.º 10)	DRSALSR (SEQ ID n.º 52)	TSSNRKT (SEQ ID n.º 53)	QSGALAR (SEQ ID n.º 18)	RSDHLSR (SEQ ID n.º 54)	NA

Tabla 4: Sitios diana de las NDZn específicas de la albúmina murina

20	Diana	SBS#	Sitio diana
	Intrón 1	30724	ctGAAGGTgGCAATGGTTcctctctgct_ (SEQ ID n.º 55)
25	Intrón 1	30725	ttTCTGTAAACGATCGGgaactggcatc_ (SEQ ID n.º 56)
	Intrón 1	30732	aaGATGCCaGTTCCCGATcggttacagga_ (SEQ ID n.º 57)
	Intrón 1	30733	agGGAGTAGCTTAGGTCagtgaagagaa_ (SEQ ID n.º 58)
30	Intrón 13	30759	acGTAGAGAACAACATCTAGattggtgg_ (SEQ ID n.º 59)
	Intrón 13	30761	ctGTAATAGAACTGACttacgttagctt_ (SEQ ID n.º 60)
	Intrón 13	30760	acGTAGAGAACAACatctagattggtgg_ (SEQ ID n.º 59)
35	Intrón 13	30767	agGGAATGtGAAATGATTcAGatatata_ (SEQ ID n.º 61)
	Intrón 13	30768	ccATGGCCTAACCAACAGtttatcttctt_ (SEQ ID n.º 62)
40	Intrón 13	30769	ccATGGCCTAACCAACAGtttatcttctt_ (SEQ ID n.º 62)
	Intrón 12	30872	ctTGGCTGTGTAGGAGGGGAgtagcagt_ (SEQ ID n.º 63)
	Intrón 12	30873	ttCCTAAGTTGGCAGTGGCAtgcttaat_ (SEQ ID n.º 64)
45	Intrón 12	30876	ctTTGGCTTTGAGGATTAAgcatgccac_ (SEQ ID n.º 65)
	Intrón 12	30877	acTTGGCTcCAAGATTTATAGccttaaa_ (SEQ ID n.º 66)
	Intrón 12	30882	caGGAAAGTAAGATAGGAAGgaatgtga_ (SEQ ID n.º 67)
50	Intrón 12	30883	ctGGGGTAAATGTCTCcttgctcttctt_ (SEQ ID n.º 68)

Ejemplo 2: Actividad de las nucleasas específicas de la albúmina murina

55 Se utilizaron parejass de NDZn que actúan selectivamente sobre el gen de la albúmina murina para ensayar la capacidad de estas NDZn para inducir roturas bicatenarias en un sitio diana específico. La secuencia de aminoácidos de la región de la hélice de reconocimiento de las NDZn indicadas se muestran a continuación en la tabla 3 y sus sitios diana se muestran en la tabla 4 (los sitios diana en el ADN están indicados en letras mayúsculas; los nucleótidos que no entran en contacto se indican en minúscula).

60 El ensayo Cel-I (Surveyor™, Transgenomics) tal y como se describe en Perez et al., (2008) *Nat. Biotechnol.* 26: 808-816 y Guschin et al. (2010) *Methods Mol. Biol.* 649: 247-56, se utilizó para detectar las modificaciones inducidas por las NDZn. En este ensayo, la amplificación por PCR del sitio diana fue seguida de la cuantificación de las inserciones y deleciones (indels) mediante el uso de la enzima detectora de discordancias Cel-I (Yang et al., (2000) *Biochemistry* 39, 3533-3541) que proporcionó una estimación del límite más bajo de la frecuencia de roturas bicatenarias. Tres días después de la transfección del vector de expresión con NDZn en las condiciones estándares (37 °C) o bien mediante un

choque hipotérmico (30 °C, véase la publicación de los EE. UU. en copropiedad n.º 20110041195), se aisló el ADN genómico de las células Neuro2A transfectadas con la o las NDZn con el kit DNeasy™ (Qiagen). En estos experimentos, todas las parejas de NDZn eran parejas de mutaciones de *FokI* ELD/KKR (descritas más arriba).

- 5 En la figura 1 se muestra una composición de los resultados del ensayo con Cel-I y demuestra que las NDZn mostradas a continuación son capaces de inducir la escisión en sus correspondientes sitios diana. El porcentaje de indels mostrados por debajo de los carriles indica la cantidad de genomas que se alteraron mediante NHEJ después de la escisión. Los datos también demuestran que se incrementó la actividad cuando el procedimiento de la transducción incorpora el choque hipotérmico.

10

Ejemplo 3: Escisión *in vivo* del locus de la albúmina

- Se ensayaron *in vivo* las NDZn específicas de la albúmina de ratón SBS30724 y SBS30725 que actúan selectivamente sobre una secuencia en el intrón 1. Las NDZn estaban insertadas en un vector de AAV2/8 tal y como se describe más arriba (Li et al. (2011) *Nature* 475 (7355): 217). Para facilitar la producción en el sistema del baculovirus, se utilizó el vector AAV2/8.2 para las preparaciones destinadas a la producción del baculovirus. El vector AAV2/8.2 se diferencia del vector AAV2/8 en que se le ha retirado una porción de la cápsida de AAV y ha sido reemplazada por la misma región de la cápsida de AAV2, con lo que se crea una cápsida quimérica. La región es el dominio de la fosfolipasa A2 en VP1. La producción de la NDZn que contiene partículas víricas se realizó mediante la preparación con el sistema HEK293, o bien un sistema de baculovirus que utiliza los métodos estándares de la técnica (véase Li et al., véase p. ej., la patente de los EE. UU. US 6.723.551). A continuación, las partículas víricas se administraron a ratones macho normales ($n = 6$) en una dosis única de 200 μ l con los $1,0 \times 10^{11}$ genomas de vector totales de AAV2/8 o bien de AAV2/8.2 que codifican la NDZn específica de la albúmina de ratón. A los 14 días de la administración de los vectores rAAV, se sacrificaron los ratones, se les recogió el hígado y se procesaron para obtener el ADN o las proteínas totales con los métodos estándares conocidos en la técnica. La detección de las copias del genoma del vector de AAV se realizó por PCR cuantitativa. Brevemente, los cebadores de qPCR se hicieron específicos de las secuencias de bGHPa contenida en el AAV como sigue:

- 30 Oligo200 (Directo) 5'-GTTGCCAGCCATCTGTTGTTT-3' (SEQ ID n.º 69)
Oligo201 (Inverso) 5'-GACAGTGGGAGTGGCACCTT-3' (SEQ ID n.º 70)
Oligo202 (Sonda) 5'-CTCCCCCGTGCCTTCCTTGACC-3' (SEQ ID n.º 71)

- La actividad de escisión de la NDZn se midió con un ensayo de Cel-I realizado con un aparato LC-GX (Perkin Elmer) de acuerdo con el protocolo del fabricante. La expresión de las NDZn *in vivo* se midió con un sistema de etiqueta FLAG de acuerdo con los métodos estándares. Los resultados (figura 3) demuestran que se expresaron las NDZn y que eran activas a la hora de escindir la diana en el gen en el hígado del ratón. En la figura se muestran los resultados de NHEJ de Cel-I para cada ratón en el estudio. El tipo de vector y su contenido se muestran en la parte superior de los carriles. La reparación de las discordancias después de la escisión de NDZn (indicada en % de indels) se detectó en aproximadamente el 16% en algunos de los ratones.

40

Las TALEN específicas de la albúmina también se ensayaron como se presenta en las solicitudes de patente de los EE. UU. n.ºs 13/624.193 y 13/624.217).

Ejemplo 4: Inserción *in vivo* de un gen asociado a la enfermedad corregido

45

- Las NDZn o TALEN específicas de la albúmina murina se utilizan a continuación para introducir un transgén que codifica un producto génico terapéutico para que se exprese en el locus de la albúmina. Los donantes se diseñaron para insertar el gen correcto para la enfermedad de Fabry (GLA), enfermedad de Gaucher (GBA), enfermedad de Hurler (IDUA) y enfermedad de Hunter en el locus de la albúmina. En estas construcciones de donante, el gen terapéutico estaba flanqueado por secuencias homólogas al gen de la albúmina. En 5' del transgén, todas las construcciones donantes contienen secuencias homólogas al intrón 1 de la albúmina murina, mientras que en 3' del gen, las construcciones contienen secuencias homólogas a la unión del intrón 1 con el exón 2 de la albúmina murina.

- 55 Luego, las construcciones de donantes se incorporan en un genoma de AAV y las partículas de AAV resultantes que contienen los donantes se purifican después con los métodos conocidos en la técnica. El material se utiliza para producir virus de AAV que contienen los genomas de AAV-donante con el método de transfección triple en las células HEK 293T y se purificaron con gradientes de densidad de CsCl, tal y como está descrito (véase Ayuso et al. (2010) *Gene Ther* 17(4), 503-510). El vector de AAV se diluirá en PBS antes de la inyección. Un abanico de partículas del vector que va de 5×10^9 a 5×10^{13} genomas víricos de AAV-donante se utilizarán junto con 1×10^9 a 1×10^{12} genomas víricos de partículas del vector AAV-NDZn a través de la vena de la cola o inyecciones intraperitoneales de los virus en los ratones del modelo de enfermedad o de tipo silvestre. Los genomas de AAV-NDZn, descritos más arriba, que contienen las NDZn específicas de la albúmina de ratón, se utilizarán, en combinación con los donantes GLA, GBA, IDUA e IDS-AAV. Los ensayos de Cel-I y PCR se realizarán sobre el ADN de hígado aislado en diferentes momentos para determinar la frecuencia de la inserción del donante inducido por NHEJ y por NDZn. También se pueden utilizar transferencias de tipo Southern. De acuerdo con el protocolo estándar, la cuantificación de GLA, GBA, IDUA e IDS

65

humanos en el plasma se realizará con un kit de ELISA para GLA, GBA, IDUA e IDS humanos o con un kit de ELISA para la etiqueta FLAG. También se realizan transferencias de tipo Western estándares. Los resultados demuestran que estos transgenes correctivos incrementan la expresión de la proteína terapéutica *in vivo*.

- 5 Por ejemplo, el gen que codifica la α -galactosidasa humana (deficiente en los pacientes con la enfermedad de Fabry) se introdujo en el locus de la albúmina de ratón. La pareja de NDZn 30724/30725 se utilizó tal y como se describe más arriba con el transgén de la α -galactosidasa. En este experimento, se trataron 3 ratones con un virus de AAV2/8 que contiene la pareja de NDZn a una dosis de $3,0 \times 10^{11}$ genomas víricos por ratón y un virus de AAV2/8 que contiene el donante huGLA a $1,5 \times 10^{12}$ genomas víricos por ratón. A los animales de control se les dio únicamente el virus que
- 10 contenía la NDZn, o bien solo el virus del donante huGLA. Las transferencias de tipo Western realizadas sobre los homogeneizados de hígado mostraron un incremento de la señal específica de la α -galactosidasa, lo que indica que el gen de la α -galactosidasa se había integrado y se estaba expresando (figura 4A). Además, se realizó un ELISA sobre el lisado de hígado con un kit de ensayo de la α -galactosidasa humana (Sino) de acuerdo con el protocolo del fabricante. Los resultados, mostrados en la figura 4B, mostraban un incremento de la señal en los ratones que se habían tratado
- 15 con las NDZn y el donante de huGLA.

Ejemplo 5: Diseño de las NDZn específicas de la albúmina humana

- Para diseñar las NDZn con especificidad por el gen de la albúmina humana, la secuencia de ADN del intrón 1 de la
- 20 albúmina humana se analizó con los métodos descritos más arriba para identificar las secuencias diana con el mejor potencial para la fijación de la NDZn (para una información más detallada, véanse las solicitudes de patente de los Estados Unidos en copropiedad n.ºs 13/624.193 y 13/624.217). La diana y las hélices se muestran en las tablas 5 y 6.

Tabla 5: Diseños de hélices para las nucleasas con dedos de cinc específicas de la albúmina humana

Diana	SBS #	Diseño					
		F1	F2	F3	F4	F5	F6
Intrón 1	35393	QSSDLSR (SEQ ID n.º 46)	LRHNLRA (SEQ ID n.º 72)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35394	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35396	QSSDLSR (SEQ ID n.º 46)	LKWNLR (SEQ ID n.º 76)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35398	QSSDLSR (SEQ ID n.º 46)	LRHNLRA (SEQ ID n.º 72)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35399	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35405	QSSDLSR (SEQ ID n.º 46)	WKWNLR A (SEQ ID n.º 77)	DQSNLRA (SEQ ID n.º 73)	RPYTLRL (SEQ ID n.º 74)	QSSDLSR (SEQ ID n.º 46)	HRSNLNK (SEQ ID n.º 75)
Intrón 1	35361	QSGNLAR (SEQ ID n.º 5)	LMQNRN Q (SEQ ID n.º 78)	LKQHLNE (SEQ ID n.º 79)	TSGNLTR (SEQ ID n.º 11)	RRYYLRL (SEQ ID n.º 80)	N/A
Intrón 1	35364	QSGNLAR (SEQ ID n.º 5)	HLGNLKT (SEQ ID n.º 81)	LKQHLNE (SEQ ID n.º 79)	TSGNLTR (SEQ ID n.º 11)	RRDWRR D (SEQ ID n.º 82)	N/A
Intrón 1	35370	QSGNLAR (SEQ ID n.º 5)	LMQNRN Q (SEQ ID n.º 78)	LKQHLNE (SEQ ID n.º 79)	TSGNLTR (SEQ ID n.º 11)	RRDWRR D (SEQ ID n.º 82)	N/A

Diana	SBS #	Diseño					
Intrón 1	35379	QRSNLVR (SEQ ID n.º 83)	TSSNRKT (SEQ ID n.º 53)	LKHHLTD (SEQ ID n.º 84)	TSGNLTR (SEQ ID n.º 11)	RRDWRR D (SEQ ID n.º 82)	N/A
Intrón 1	35458	DKSYLRP (SEQ ID n.º 85)	TSGNLTR (SEQ ID n.º 11)	HRSARKR (SEQ ID n.º 86)	QSSDLSR (SEQ ID n.º 46)	WRSSLKT (SEQ ID n.º 87)	N/A
Intrón 1	35480	TSGNLTR (SEQ ID n.º 11)	HRSARKR (SEQ ID n.º 86)	QSGDLTR (SEQ ID n.º 40)	NRHHLKS (SEQ ID n.º 88)	N/A	N/A
Intrón 1	35426	QSGDLTR (SEQ ID n.º 40)	QSGNLHV (SEQ ID n.º 89)	QSAHRKN (SEQ ID n.º 90)	STAALSY (SEQ ID n.º 91)	TSGSLSR (SEQ ID n.º 92)	RSDALAR (SEQ ID n.º 41)
Intrón 1	35428	QSGDLTR (SEQ ID n.º 40)	QRSNLNI (SEQ ID n.º 93)	QSAHRKN (SEQ ID n.º 90)	STAALSY (SEQ ID n.º 91)	DRSALSR (SEQ ID n.º 52)	RSDALAR (SEQ ID n.º 41)
Intrón 1	34931	QRTHLTQ (SEQ ID n.º 94)	DRSNLTR (SEQ ID n.º 95)	QSGNLAR (SEQ ID n.º 5)	QKVNRRAG (SEQ ID n.º 96)	N/A	N/A
Intrón 1	33940	RSDNLSV (SEQ ID n.º 26)	QANANRIT (SEQ ID n.º 97)	DQSNLRA (SEQ ID n.º 73)	QSAHRIT (SEQ ID n.º 98)	TSGNLTR (SEQ ID n.º 11)	HRSARKR (SEQ ID n.º 86)

Tabla 6: Sitios diana de las NDZn específicas de la albúmina humana

Diana	SBS #	Sitio diana
Intrón 1 (locus 2)	35393	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35394	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35396	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35398	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35399	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35405	<u>ccTATCCATTGCACTATGCTTtatttta</u> (SEQ ID n.º 99)
Intrón 1 (locus 2)	35361	<u>ttTGGGATAGTTATGAAttcaatcttca</u> (SEQ ID n.º 100)
Intrón 1 (locus 2)	35364	<u>ttTGGGATAGTTATGAAttcaatcttca</u> (SEQ ID n.º 100)
Intrón 1 (locus 2)	35370	<u>ttTGGGATAGTTATGAAttcaatcttca</u> (SEQ ID n.º 100)
Intrón 1 (locus 2)	35379	<u>ttTGGGATAGTTATGAAttcaatcttca</u> (SEQ ID n.º 100)
Intrón 1 (locus 3)	35458	<u>ccTGTGCTGTTGATCTCataaatagaac</u> (SEQ ID n.º 101)
Intrón 1 (locus 3)	35480	<u>ccTGTGCTGTTGATCTCataaatagaac</u> (SEQ ID n.º 101)
Intrón 1 (locus 3)	35426	<u>ttGTGGTTTTTAAAtAAAGCAtagtgc</u> (SEQ ID n.º 102)
Intrón 1 (locus 3)	35428	<u>ttGTGGTTTTTAAAtAAAGCAtagtgc</u> (SEQ ID n.º 102)
Intrón 1 (locus 4)	34931	<u>acCAAGAAGACAGActaaaatgaaaata</u> (SEQ ID n.º 103)
Intrón 1 (locus 4)	33940	<u>ctGTTGATAGACACTAAAGagtattag</u> (SEQ ID n.º 104)

Estas nucleasas se ensayaron por parejas para determinar la pareja con la mayor actividad. Las matrices resultantes de las parejas analizadas se muestran en las tablas 7 y 8 que vienen a continuación, en donde la NDZn utilizada para el lado derecho del dímero se muestra por la parte superior de cada matriz, y la NDZn utilizada para el lado izquierdo del dímero se recoge a la izquierda de cada matriz. La actividad resultante, según se determinó por el porcentaje de discordancias detectadas con el ensayo Cel-I, se muestra en el cuerpo de ambas matrices:

Tabla 7: Actividad de las NDZn específicas de la albúmina humana (% de dianas mutadas)

		35393	35394	35396	35398	35399	35405	Media
10	35361	18	19	25	22	23	21	21
	35364	n.d.	24	23	19	21	21	22
	35370	21	19	22	n.d.	22	23	21
15	35379	21	21	n.d.	19	19	21	20

Tabla 8: Actividad de las NDZn específicas de la albúmina humana (% de dianas mutadas)

20		35458	35480	Media
	35426	4,5	7	3
	35428	4,9	6	3,6

25 (nota: «n.d.» significa que el ensayo en esta pareja no se realizó).

Así pues, se han desarrollado nucleasas muy activas que reconocen las secuencias diana del intrón 1 de la albúmina humana.

30 Ejemplo 6: Diseño de las TALEN específicas de la albúmina

Se diseñaron TALEN que actúan selectivamente sobre secuencias dentro del intrón 1 de la albúmina humana. El reconocimiento de las bases se consiguió con las correspondencias canónicas de bases y RVD (el código de TALE: NI para A, HD para C, NN para G (NK en semirrepetición), NG para T). Se construyeron las TALEN tal y como se describió previamente (véase la publicación de patente de los EE. UU. en copropiedad n.º 20110301073). Las dianas para un subconjunto de TALEN estaban conservadas en los genes de la albúmina del macaco cangrejero y del macaco rhesus (para conocer más detalles, véanse las solicitudes de patente de los EE. UU. en copropiedad 13/624.193 y 13/624.217). Las TALEN se construyeron en los esqueletos de TALEN «+17» y «+63» tal y como se describe en la patente de los EE. UU. US 20110301073. Las dianas y los identificadores numéricos para las TALEN analizadas se muestran a continuación en la tabla 9.

Tabla 9: TALEN específicas de la albúmina

45	SBS #	Sitio	# de RVD	SEQ ID n.º
	102249	gtTGAAGATTGAATTCaTa	15	105
	102250	gtTGAAGATTGAATTCATAac	17	106
50	102251	gtGCAATGGATAGGTCTtt	15	107
	102252	atAGTGCAATGGATAGGtc	15	108
	102253	atTGAATTCATAACTATcc	15	109
	102254	atTGAATTCATAACTATCCca	17	110
55	102255	atAAAGCATAGTGCAATGGat	17	111
	102256	atAAAGCATAGTGCAATgg	15	112
	102257	ctATGCTTTATTTAAAAac	15	113
60	102258	ctATGCTTTATTTAAAAACca	17	114
	102259	atTTATGAGATCAACAGCaca	17	115
	102260	ctATTTATGAGATCAACAGca	17	116

(continuación)

SBS #	Sitio	# de RVD	SEQ ID n.º
5 102261	ttCATTTTAGTCTGTCTTctt	17	117
102262	atTTTAGTCTGTCTTctg	15	118
102263	ctAATACTCTTTTAGTGTct	16	119
10 102264	atCTAATACTCTTTTAGTGtc	17	120
102265	atAATTGAACATCATCctg	15	121
102266	atAATTGAACATCATCCTGag	17	122
15 102267	atATTGGGCTCTGATTCTac	17	123
102268	atATTGGGCTCTGATTCTct	15	124
102269	ttTTTCTGTAGGAATCAga	15	125
20 102270	ttTTTCTGTAGGAATCAGag	16	126
102271	ttATGCATTTGTTTCAaaa	15	127
102272	atTATGCATTTGTTTCAaa	15	128

25 A continuación, las TALEN se analizaron por parejas en las células HepG2 en función de su capacidad para inducir modificaciones en sus dianas cromosómicas endógenas, y los resultados mostraron que eran activas muchas proteínas que llevan el punto de truncamiento +17. De igual forma, muchas TALEN que llevan el punto de truncamiento +63 también eran activas (véase la tabla 10). El estudio comparativo directo con tres series de NDZn de albúmina sin optimizar (véase la tabla 10) mostraron que las TALEN y las NDZn tienen actividades que están aproximadamente en el mismo margen.

Tabla 10: Modificación de la diana inducida por las TALEN en las células HepG2-C3a

Pareja de muestras	TALEN C17	% de modificación, C17	TALEN C63	% de modificación, C63	Hueco
35 1	102251:102249	15	102251:102249	0	12
2	102251:102250	0	102251:102250	0	10
40 3	102252:102249	0	102252:102249	8,3	15
4	102252:102250	32	102252:102250	8,0	13
5	102255:102253	38	102255:102253	21	13
45 6	102255:102254	43	102255:102254	0	11
7	102256:102253	0	102256:102253	23	15
8	102256:102254	28	102256:102254	16	13
9	102259:102257	18	102259:102257	15	13
50 10	102259:102258	15	102259:102258	0	11
11	102260:102257	15	102260:102257	13	15
12	102260:102258	24	102260:102258	11	13
55 13	102263:102261	0	102263:102261	16	17
14	102263:102262	0	102263:102262	15	16
15	102264:102261	0	102264:102261	22	18
60 16	102264:102262	0	102264:102262	17	17
20	102267:102265	47	102267:102265	9,8	13
21	102267:102266	4,7	102267:102266	0	11
65 22	102268:102265	4,2	102268:102265	7,9	15

Pareja de muestras	TALEN C17	% de modificación, C17	TALEN C63	% de modificación, C63	Huevo
23	102268:102266	10	102268:102266	0	13
24	102271:102269	14	102271:102269	0	12
25	102271:102270	0	102271:102270	0	11
26	102272:102269	0	102272:102269	0	13
27	102272:102270	0	102272:102270	0	12
NDZn					
17	35361:35396	31	35361:35396	29	6
18	35426:35458	10	35426:35458	7	6
19	34931:33940	7,3	34931:33940	7	6

Tal y como se señaló con anterioridad (véase la publicación de patente de los EE. UU. en copropiedad n.º 20110301073), las TALEN C17 tienen mayor actividad cuando el tamaño del hueco entre los dos sitios diana de la TALEN es de aproximadamente 11 a 15 pb, mientras que las TALEN C63 mantienen la actividad con un tamaño de hueco de hasta 18 pb (véase la tabla 10).

Ejemplo 7: Detección *in vivo* de los transgenes de donantes para EAL

Se construyeron los donantes para cuatro transgenes de la enfermedad por almacenamiento lisosomal con el propósito de integrarlos en el intrón 1 del gen de la albúmina de ratón. Los transgenes eran α -galactosidasa A (GLA), β -glucosidasa ácida (GBA), α -L-iduronidasa (IDUA) e iduronato-2-sulfatasa (IDS), genes ausentes en las enfermedades de Fabry, Gaucher, Hurler y Hunter, respectivamente. Véase, p. ej., la figura 8.

A continuación, los donantes se utilizaron en estudios *in vivo* para observar la integración de los transgenes. Las NDZn específicas de la albúmina murina y los donantes se insertaron todos en el virus AAV2/8 tal y como se describe en el ejemplo 4, y a continuación se inyectaron en los ratones. En estos experimentos, los virus se formularon para la inyección en D-PBS + NaCl a 35 mM, glicerol al 5% y se congelaron antes de la inyección. Los virus que contienen la nucleasa y el donante se mezclaron antes de la congelación. A día 0, se descongeló la preparación de los virus y se administró a los ratones por una inyección en la vena de la cola, en donde el volumen de inyección total fue de 200 μ l. En los tiempos indicados, se sacrificaron los ratones y a continuación se les recogieron suero, hígado y bazo para el análisis de proteínas y del ADN mediante los protocolos estándares. Los grupos de dosis se muestran a continuación en la tabla 11.

Tabla 11: Grupos de tratamiento para la integración de los transgenes para la EAL

Grupo	Tratamiento	N/grupo/punto temporal
1	Intrón 1 de albúmina murina + Fabry a una relación 1:5; NDZn a $3,0 \times 10^{11}$, donante a $1,5 \times 10^{12}$	3
2	Intrón 1 de albúmina murina + donante de Hunter a una relación 1:5; NDZn a $3,0 \times 10^{11}$, donante a $1,5 \times 10^{12}$	3
3	Intrón 1 de albúmina murina + donante de Hurler a una relación 1:5; NDZn $3,0 \times 10^{11}$, donante a $1,5 \times 10^{12}$	3
4	Intrón 1 de albúmina murina	3
5	Sólo donante de Fabry	2
6	Sólo donante de Hunter	2
7	Sólo donante de Hurler	2

El día 30, los homogeneizados de hígado se analizaron mediante análisis de transferencia de tipo Western en busca de la presencia de las proteínas para EAL codificadas por los donantes. Los homogeneizados de hígado se analizaron mediante transferencia de tipo Western con metodologías estándares, mediante el uso de los siguientes anticuerpos primarios: el anticuerpo monoclonal de conejo específico contra la α -galactosidasa A (enfermedad de Fabry) se compró a Sino Biological, Inc.; el anticuerpo monoclonal de ratón específico contra la α -L-iduronidasa humana (enfermedad de Hurler) se compró a R&D Systems; el anticuerpo monoclonal de ratón específico contra la iduronato 2-sulfatasa humana (enfermedad de Hunter) se compró a R&D Systems. Los resultados (véase la figura 5) muestran la expresión,

especialmente en los ratones que recibieron tanto el virus que contenía la NDZn como el virus que contenía el transgén donante.

También se investigó la manera de integrar el ADN donante en el locus de la albúmina. Después de la escisión en el locus de la albúmina, el transgén donante posiblemente se pudo integrar a través de la recombinación dirigida por la homología (HDR), gracias a las regiones de homología que flanquean el transgén donante (figura 2), o el transgén se pudo capturar durante el proceso de unión de extremos no homólogos (NHEJ) propenso a errores. Los resultados de estas dos alternativas producirán inserciones de diferente tamaño cuando se analizan por PCR, con el uso del cebador Acc651-SA-rev-sh (5' AAG AAT AAT TCT TTA GTG GTA 3', SEQ ID n.º 129) que se fija al sitio aceptor del ajuste de F9 en todas las construcciones de donantes para EAL y el cebador mALB-OOF1 (5' ATGAAGTGGGTAACCTTTCTC 3', SEQ ID n.º 130) que se fija al exón 1 de la albúmina de ratón por delante del sitio de escisión de la NDZn (véase la figura 6), donde la integración del transgén mediante la HDR dará lugar a la inserción de aproximadamente 680 pb, mientras que la integración a través de la NHEJ dará lugar a la integración de aproximadamente 1488 pb. Así pues, el ADN genómico aislado de los homogeneizados de hígado de los ratones tratados se analizó por PCR en presencia de nucleótidos radiomarcados con ³²P y se migró en un gel. En las tres integraciones de los transgenes se observó la integración a través de ambos mecanismos (véase la figura 7).

Los ADN de los donantes también se diseñaron para incluir una secuencia a modo de etiqueta para el reconocimiento posterior de la proteína con los anticuerpos específicos contra etiqueta. La etiqueta se diseñó para que incorpore dos secuencias diferentes que codifican las etiquetas Myc (EQKLISEEDL, SEQ ID n.º 131) y Flag (DYKDDDD, SEQ ID n.º 132). El esquema de los diseños de los donantes se muestra en la figura 8. Los donantes se integraron como se describe más arriba y todos fueron capaces de integrarse tal y como se demostró por PCR (véase la figura 9).

<110> Sangamo Biosciences, Inc.
 <120> Métodos y composiciones para el tratamiento de enfermedades por almacenamiento lisososomal
 <130> 8325-0097.40
 <140> PCT/US2013/032381
 <141> 15-03-2013
 <150> 61/704.072
 <151> 21-09-2012
 <150> 61/670.463
 <151> 11-07-2012
 <160> 134
 <170> PatentIn versión 3.5
 <210> 1
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

<400> 1
Thr Ser Gly Ser Leu Thr Arg
1 5

<210> 2
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

<400> 2
Arg Ser Asp Ala Leu Ser Thr
1 5

<210> 3
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

5 **<400> 3**
 Gln Ser Ala Thr Arg Thr Lys
 1 5

10 <210> 4
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

15 **<400> 4**
 Thr Ser Gly His Leu Ser Arg
 1 5

20 <210> 5
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

25 **<400> 5**
 Gln Ser Gly Asn Leu Ala Arg
 1 5

30 <210> 6
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>

35 <223> Descripción de secuencia artificial: péptido sintético

40 **<400> 6**
 Arg Ser Asp His Leu Ser Ala
 1 5

45 <210> 7
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

50 **<400> 7**
 Thr Lys Ser Asn Arg Thr Lys
 1 5

55 <210> 8
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

60 **<400> 8**
 Asp Arg Ser Asn Leu Ser Arg
 1 5

65 <210> 9
 <211> 7
 <212> PRT

<213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 5 **<400> 9**
 Trp Arg Ser Ser Leu Arg Ala
 1 5

 10 <210> 10
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 15 **<400> 10**
 Asp Ser Ser Asp Arg Lys Lys
 1 5

 20 <210> 11
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 11
 Thr Ser Gly Asn Leu Thr Arg
 30 1 5

 <210> 12
 <211> 7
 <212> PRT
 35 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 12
 40 **Asp Arg Ser Thr Arg Arg Gln**
 1 5

 <210> 13
 <211> 7
 45 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 50 **<400> 13**
 Glu Arg Gly Thr Leu Ala Arg
 1 5

 55 <210> 14
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 60 **<400> 14**
 Thr Ser Ala Asn Leu Ser Arg
 1 5

 65

5
 <210> 15
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

10
<400> 15
Asp Arg Ser Ala Leu Ala Arg
1 5

15
 <210> 16
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

20
<400> 16
Arg Ser Asp His Leu Ser Glu
1 5

25
 <210> 17
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

30
<400> 17
His Arg Ser Asp Arg Thr Arg
1 5

35
 <210> 18
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

40
<400> 18
Gln Ser Gly Ala Leu Ala Arg
1 5

45
 <210> 19
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

50
<400> 19
Gln Ser Gly His Leu Ser Arg
1 5

55
 <210> 20
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

60
 <210> 20
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

65

5 **<400> 20**
 Arg Ser Asp Asn Leu Ser Thr
 1 5

 10 <210> 21
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 15 **<400> 21**
 Asp Gly Arg Asn Leu Arg His
 1 5

 20 <210> 22
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 25 **<400> 22**
 Arg Ser Asp Asn Leu Ala Arg
 1 5

 30 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 35 <223> Descripción de secuencia artificial: péptido sintético
 <400> 23
 Gln Ser Asn Ala Leu Asn Arg
 1 5

 40 <210> 24
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 45 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 24
 Leu Lys Gln Val Leu Val Arg
 50 1 5

 55 <210> 25
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 60 **<400> 25**
 Gln Ser Thr Pro Leu Phe Ala
 1 5

 65 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>

<223> Descripción de secuencia artificial: péptido sintético
 <400> 26
 Arg Ser Asp Asn Leu Ser Val
 1 5
 <210> 27
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 27
 His Ser Asn Ala Arg Lys Thr
 1 5
 <210> 28
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 28
 Arg Ser Asp Ser Leu Ser Ala
 1 5
 <210> 29
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 29
 Arg Ser Asp Ser Leu Ser Val
 1 5
 <210> 30
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 30
 Arg Ser Asp Asn Leu Ser Glu
 1 5
 <210> 31
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 31
 Glu Arg Ala Asn Arg Asn Ser
 1 5
 <210> 32
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>

5 <223> Descripción de secuencia artificial: péptido sintético
 <400> 32
 Gln Ser Ala Asn Arg Thr Lys
 1 5

10 <210> 33
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 33
 15 Arg Ser Asp Ala Leu Thr Gln
 1 5

20 <210> 34
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 25 <400> 34
 Gln Ser Gly His Leu Ala Arg
 1 5

30 <210> 35
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 35 <223> Descripción de secuencia artificial: péptido sintético
 <400> 35
 Arg Ser Asp His Leu Thr Gln
 1 5

40 <210> 36
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 45 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 36
 Arg Ser Asp His Leu Ser Gln
 50 1 5

55 <210> 37
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 37
 60 Trp Arg Ser Ser Leu Val Ala
 1 5

65 <210> 38
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 38
 Arg Ser Asp Val Leu Ser Glu
 1 5
 <210> 39
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 39
 Arg Asn Gln His Arg Lys Thr
 1 5
 <210> 40
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 40
 Gln Ser Gly Asp Leu Thr Arg
 1 5
 <210> 41
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 41
 Arg Ser Asp Ala Leu Ala Arg
 1 5
 <210> 42
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 42
 Arg Arg Asp Pro Leu Ile Asn
 1 5
 <210> 43
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 43
 <400> 43
 Ile Arg Ser Thr Leu Arg Asp
 1 5
 <210> 44
 <211> 7

<212> PRT
<213> Secuencia artificial
<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 44

Tyr Ser Ser Thr Arg Asn Ser
1 5

<210> 45
<211> 7
<212> PRT
<213> Secuencia artificial
<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 45

Ser Tyr Trp Ser Arg Thr Val
1 5

<210> 46
<211> 7
<212> PRT
<213> Secuencia artificial
<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 46

Gln Ser Ser Asp Leu Ser Arg
1 5

<210> 47
<211> 7
<212> PRT
<213> Secuencia artificial
<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 47

Arg Thr Asp Ala Leu Arg Gly
1 5

<210> 48
<211> 7
<212> PRT
<213> Secuencia artificial
<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 48

Gln Lys Ser Pro Leu Asn Thr
1 5

<210> 49
<211> 7
<212> PRT
<213> Secuencia artificial
<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 49

Gln Ala Glu Asn Leu Lys Ser
1 5

5 <210> 50
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 50
 Arg Arg Ala His Leu Asn Gln
 1 5
 15 <210> 51
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 51
 Ser Asp Thr Asn Arg Phe Lys
 1 5
 25 <210> 52
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 52
 Asp Arg Ser Ala Leu Ser Arg
 1 5
 35 <210> 53
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 53
 Thr Ser Ser Asn Arg Lys Thr
 1 5
 50 <210> 54
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 54
 Arg Ser Asp His Leu Ser Arg
 1 5
 60 <210> 55
 <211> 28
 <212> ADN
 <213> Mus sp.
 <400> 55
 ctgaagggtgg caatggttcc tctctgct

5	<210> 56 <211> 28 <212> ADN <213> Mus sp. <400> 56 tttcctgtaa cgatcgggaa ctggcatc	28
10		
15	<210> 57 <211> 28 <212> ADN <213> Mus sp. <400> 57 aagatgccag ttcccgatcg ttacagga	28
20		
25	<210> 58 <211> 28 <212> ADN <213> Mus sp. <400> 58 agggagtagc ttaggtcagt gaagagaa	28
30		
35	<210> 59 <211> 28 <212> ADN <213> Mus sp. <400> 59 acgtagagaa caacatctag attggtgg	28
40		
45		
50	<210> 60 <211> 28 <212> ADN <213> Mus sp. <400> 60 ctgtaataga aactgactta cgtagctt	28
55		
60	<210> 61 <211> 28 <212> ADN <213> Mus sp. <400> 61 agggaatgtg aaatgattca gatata	28
65		
	<210> 62 <211> 28 <212> ADN <213> Mus sp. <400> 62 ccatggccta acaacagttt atcttctt	28
65	<210> 63 <211> 28 <212> ADN	

	<213> Mus sp. <400> 63 cttggctgtg taggagggga gtagcagt	28
5	<210> 64 <211> 28 <212> ADN <213> Mus sp.	
10	<400> 64 ttcctaagtt ggcagtggca tgcttaat	28
15	<210> 65 <211> 28 <212> ADN <213> Mus sp.	
20	<400> 65 ctttggcttt gaggattaag catgccac	28
25	<210> 66 <211> 28 <212> ADN <213> Mus sp.	
	<400> 66 acttggctcc aagatttata gccttaaa	28
30	<210> 67 <211> 28 <212> ADN <213> Mus sp.	
35	<400> 67 caggaaagta agataggaag gaatgtga	28
40	<210> 68 <211> 28 <212> ADN <213> Mus sp.	
45	<400> 68 ctggggtaaa tgtctccttg ctcttctt	28
50	<210> 69 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: cebador sintético	
55	<400> 69 gttgccagcc atctgttggt t	21
60	<210> 70 <211> 20 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: cebador sintético	
65	<400> 70 gacagtggga gtggcacctt	20

5 <210> 71
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 <220>
 10 <223> Descripción de secuencia artificial: sonda sintética
 <400> 71
 ctcccccgtag ccttccttga cc 22

15 <210> 72
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 20 <223> Descripción de secuencia artificial: péptido sintético
 <400> 72
 Leu Arg His Asn Leu Arg Ala
 1 5

25 <210> 73
 <211> 7
 <212> PRT
 30 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 73
 Asp Gln Ser Asn Leu Arg Ala
 1 5

35 <210> 74
 <211> 7
 <212> PRT
 40 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 74
 Arg Pro Tyr Thr Leu Arg Leu
 1 5

45 <210> 75
 <211> 7
 <212> PRT
 50 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 <400> 75
 His Arg Ser Asn Leu Asn Lys
 1 5

55 <210> 76
 <211> 7
 <212> PRT
 60 <213> Secuencia artificial
 <220>
 65 <223> Descripción de secuencia artificial: péptido sintético

<400> 76

Leu Lys Trp Asn Leu Arg Thr
1 5

<210> 77

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 77

Trp Lys Trp Asn Leu Arg Ala
1 5

<210> 78

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 78

Leu Met Gln Asn Arg Asn Gln
1 5

<210> 79

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 79

Leu Lys Gln His Leu Asn Glu
1 5

<210> 80

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 80

Arg Arg Tyr Tyr Leu Arg Leu
1 5

<210> 81

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 81

His Leu Gly Asn Leu Lys Thr
1 5

<210> 82

<211> 7

<212> PRT
<213> Secuencia artificial
<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 82

Arg Arg Asp Trp Arg Arg Asp

1 5

<210> 83
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 83

Gln Arg Ser Asn Leu Val Arg

1 5

<210> 84
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 84

Leu Lys His His Leu Thr Asp

1 5

<210> 85
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 85

Asp Lys Ser Tyr Leu Arg Pro

1 1 1 1 5

<210> 86
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 86

His Arg Ser Ala Arg Lys Arg

1 5

<210> 87
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

<400> 87**Trp Arg Ser Ser Leu Lys Thr****1 5**

<210> 88

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 88**Asn Arg His His Leu Lys Ser****1 5**

<210> 89

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 89**Gln Ser Gly Asn Leu His Val****1 5**

<210> 90

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 90**Gln Ser Ala His Arg Lys Asn****1 5**

<210> 91

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 91**Ser Thr Ala Ala Leu Ser Tyr****1 5**

<210> 92

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: péptido sintético

<400> 92**Thr Ser Gly Ser Leu Ser Arg****1 5**

<210> 93

<211> 7

<212> PRT

<213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 5 **<400> 93**
 Gln Arg Ser Asn Leu Asn Ile
 1 5

 10 <210> 94
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético
 15 **<400> 94**
 Gln Arg Thr His Leu Thr Gln
 1 5

 20 <210> 95
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 25 <223> Descripción de secuencia artificial: péptido sintético

 <400> 95
 Asp Arg Ser Asn Leu Thr Arg
 1 5
 30 <210> 96
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 35 <220>
 <223> Descripción de secuencia artificial: péptido sintético

 <400> 96
 Gln Lys Val Asn Arg Ala Gly
 40 1 5

 <210> 97
 <211> 7
 <212> PRT
 45 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

 <400> 97
 50 **Gln Asn Ala Asn Arg Ile Thr**
 1 5

 <210> 98
 <211> 7
 55 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: péptido sintético

 <400> 98
 60 **Gln Ser Ala His Arg Ile Thr**
 1 5

 65 <210> 99
 <211> 28

	<212> ADN		
	<213> Homo sapiens		
	<400> 99		
5	cctatccatt gcactatgct ttatttaa	28	
	<210> 100		
	<211> 28		
	<212> ADN		
10	<213> Homo sapiens		
	<400> 100		
	tttgggatag ttatgaattc aatcttca	28	
	<210> 101		
	<211> 28		
15	<212> ADN		
	<213> Homo sapiens		
	<400> 101		
	cctgtgctgt tgatctcata aatagaac	28	
20	<210> 102		
	<211> 28		
	<212> ADN		
	<213> Homo sapiens		
	<400> 102		
25	ttgtgggtttt taaataaagc atagtgc	28	
	<210> 103		
	<211> 28		
	<212> ADN		
30	<213> Homo sapiens		
	<400> 103		
	accaagaaga cagactaaaa tgaaaata	28	
	<210> 104		
35	<211> 28		
	<212> ADN		
	<213> Homo sapiens		
	<400> 104		
40	ctgttgatag acactaaaag agtattag	28	
	<210> 105		
	<211> 19		
	<212> ADN		
45	<213> Secuencia artificial		
	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 105		
50	gttgaagatt gaattcata	19	
	<210> 106		
	<211> 21		
	<212> ADN		
55	<213> Secuencia artificial		
	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 106		
	gttgaagatt gaattcataa c	21	
60	<210> 107		
	<211> 19		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
65	<223> Descripción de secuencia artificial: oligonucleótido sintético		

5	<400> 107 gtgcaatgga taggtcttt <210> 108 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	19
10		
15	<400> 108 atagtgcgaat ggataggtc <210> 109 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	19
20		
25	<400> 109 attgaattca taactatcc <210> 110 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	19
30		
35	<400> 110 attgaattca taactatccc a <210> 111 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	21
40		
45	<400> 111 ataaagcata gtgcaatgga t <210> 112 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	21
50		
55	<400> 112 ataaagcata gtgcaatgg <210> 113 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	19
60		
65	<400> 113 ctatgcttta tttaaaaac <210> 114 <211> 21 <212> ADN <213> Secuencia artificial	19

	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 114		
5	ctatgcttta tttaaaaacc a		21
	<210> 115		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
10	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 115		
	atttatgaga tcaacagcac a		21
15	<210> 116		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
20	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 116		
	ctatttatga gatcaacagc a		21
	<210> 117		
25	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
30	<400> 117		
	ttcatttttag tctgtcttct t		21
	<210> 118		
	<211> 19		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 118		
40	attttagtct gtcttcttg		19
	<210> 119		
	<211> 20		
	<212> ADN		
45	<213> Secuencia artificial		
	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 119		
50	ctaatactct tttagtgtct		20
	<210> 120		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
55	<220>		
	<223> Descripción de secuencia artificial: oligonucleótido sintético		
	<400> 120		
	atctaatact ctttttagtgt c		21
60	<210> 121		
	<211> 19		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
65	<223> Descripción de secuencia artificial: oligonucleótido sintético		

	<400> 121	
	ataattgaac atcatcctg	19
5	<210> 122 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
10	<400> 122	
	ataattgaac atcatcctga g	21
15	<210> 123 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
20	<400> 123	
	atattgggct ctgattccta c	21
25	<210> 124 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
30	<400> 124	
	atattgggct ctgattcct	19
35	<210> 125 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
40	<400> 125	
	tttttctgta ggaatcaga	19
45	<210> 126 <211> 20 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
50	<400> 126	
	tttttctgta ggaatcagag	20
55	<210> 127 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
	<400> 127	
	ttatgcattt gtttcaaaa	19
60	<210> 128 <211> 19 <212> ADN <213> Secuencia artificial <220> <223> Descripción de secuencia artificial: oligonucleótido sintético	
65	<400> 128	
	attatgcatt tgtttcaaa	19

5 <210> 129
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: cebador sintético
 <400> 129
 10 aagaataatt ctttagtggt a 21
 <210> 130
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <223> Descripción de secuencia artificial: cebador sintético
 <400> 130
 20 atgaagtggg taacctttct c 21
 <210> 131
 <211> 10
 <212> PRT
 <213> Secuencia artificial
 <220>
 25 <223> Descripción de secuencia artificial: péptido sintético
 <400> 131
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
 1 5 10
 30 <210> 132
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 35 <223> Descripción de secuencia artificial: péptido sintético
 <400> 132
 Asp Tyr Lys Asp Asp Asp Asp
 1 5
 40 <210> 133
 <211> 9
 <212> PRT
 <213> Desconocida
 <220>
 45 <223> Descripción de desconocida: secuencia desconocida de la familia LAGLIDADG
 <400> 133
 Leu Ala Gly Leu Ile Asp Ala Asp Gly
 1 5
 50 <210> 134
 <211> 29
 <212> PRT
 <213> Secuencia artificial
 <220>
 55 <223> Descripción de secuencia artificial: péptido sintético
 <400> 134
 Arg Thr Arg Pro Leu Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Ala
 1 5 10 15
 60
 Ala Asn Asp Ile Leu Asp Tyr Lys Asp Asp Asp Asp Lys
 20 25
 65

REIVINDICACIONES

1. Un método *in vitro* para generar una célula que produce una proteína que está ausente en un sujeto con una enfermedad por almacenamiento lisosomal, en donde el método comprende:
 - integrar un transgén que codifica la proteína en un locus endógeno de albúmina de la célula mediante el uso de una nucleasa que no se produce en la naturaleza, de tal manera que la célula produce la proteína que está ausente en la enfermedad por almacenamiento lisosomal.
2. El método *in vitro* de acuerdo con la reivindicación 1, en donde la enfermedad por almacenamiento lisosomal se selecciona del grupo que consiste en la enfermedad de Gaucher, de Fabry, de Hunter, de Hurler y de Niemann-Pick.
3. El método *in vitro* de acuerdo con la reivindicación 1 o la reivindicación 2, en donde el transgén codifica una proteína seleccionada del grupo que consiste en glucocerebrosidasa, α -galactosidasa, iduronato-2-sulfatasa, α -L-iduronidasa y esfingomielina fosfodiesterasa.
4. El método *in vitro* de acuerdo con cualquiera de las reivindicaciones 1 a 3, en donde la expresión del transgén está impulsada por un promotor endógeno.
5. El método *in vitro* de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde el transgén es una proteína de fusión que comprende aminoácidos codificados por el transgén y por el locus endógeno en el cual está integrado el transgén.
6. El método *in vitro* de acuerdo con cualquiera de las reivindicaciones 1 a 5, en donde el transgén codifica una proteína de fusión que comprende:
 - (a) un dominio extracelular de una proteína membranaria de tal manera que, tras la expresión, la proteína de fusión está localizada en la superficie de la célula; y/o
 - (b) un ligando para un receptor de tal manera que, tras la expresión, la proteína de fusión atraviesa la barrera hematoencefálica.
7. El método *in vitro* de acuerdo con cualquiera de las reivindicaciones 1 a 6, en donde la célula se selecciona del grupo que consiste en un glóbulo rojo, un hepatocito, un miocito y una célula troncal.
8. El método *in vitro* de acuerdo con la reivindicación 7, en donde la célula troncal es una célula troncal hematopoyética o una célula troncal pluripotente inducida.
9. El método *in vitro* de acuerdo con cualquiera de las reivindicaciones 1 a 8, en donde la nucleasa es una nucleasa con dedos de cinc (NDZn), una nucleasa TALE (TALEN) y/o un sistema CRISPR/Cas.
10. Una célula generada por el método de acuerdo con cualquiera de las reivindicaciones 1 a 9 para ser usada en un método para tratar una enfermedad por almacenamiento lisosomal, en donde el método comprende administrar la célula a un sujeto que lo necesita.
11. Una o varias nucleasas y uno o varios transgenes para ser usados en un método para tratar una enfermedad por almacenamiento lisosomal, en donde el método comprende administrar la una o varias nucleasas y el uno o varios transgenes a un sujeto que lo necesita para generar una célula que produce una proteína que está ausente en el sujeto, en donde el transgén que codifica la proteína está integrado en un locus endógeno de albúmina de la célula mediante el uso de una nucleasa que no se produce en la naturaleza, de tal manera que la célula produce la proteína que está ausente en la enfermedad por almacenamiento lisosomal.
12. La célula para ser usada de acuerdo con la reivindicación 10, en donde la célula es una célula troncal o precursora aislada, y el método comprende además la expansión y la diferenciación de la célula troncal o precursora antes de la administración.
13. El método de acuerdo con cualquiera de las reivindicaciones 1 a 9, la célula para ser usada de acuerdo con la reivindicación 10 o 12, o la una o varias nucleasas y el uno o varios transgenes para ser usados de acuerdo con la reivindicación 11, en donde el transgén se introduce en la célula mediante un vector vírico tal como un vector de AAV.

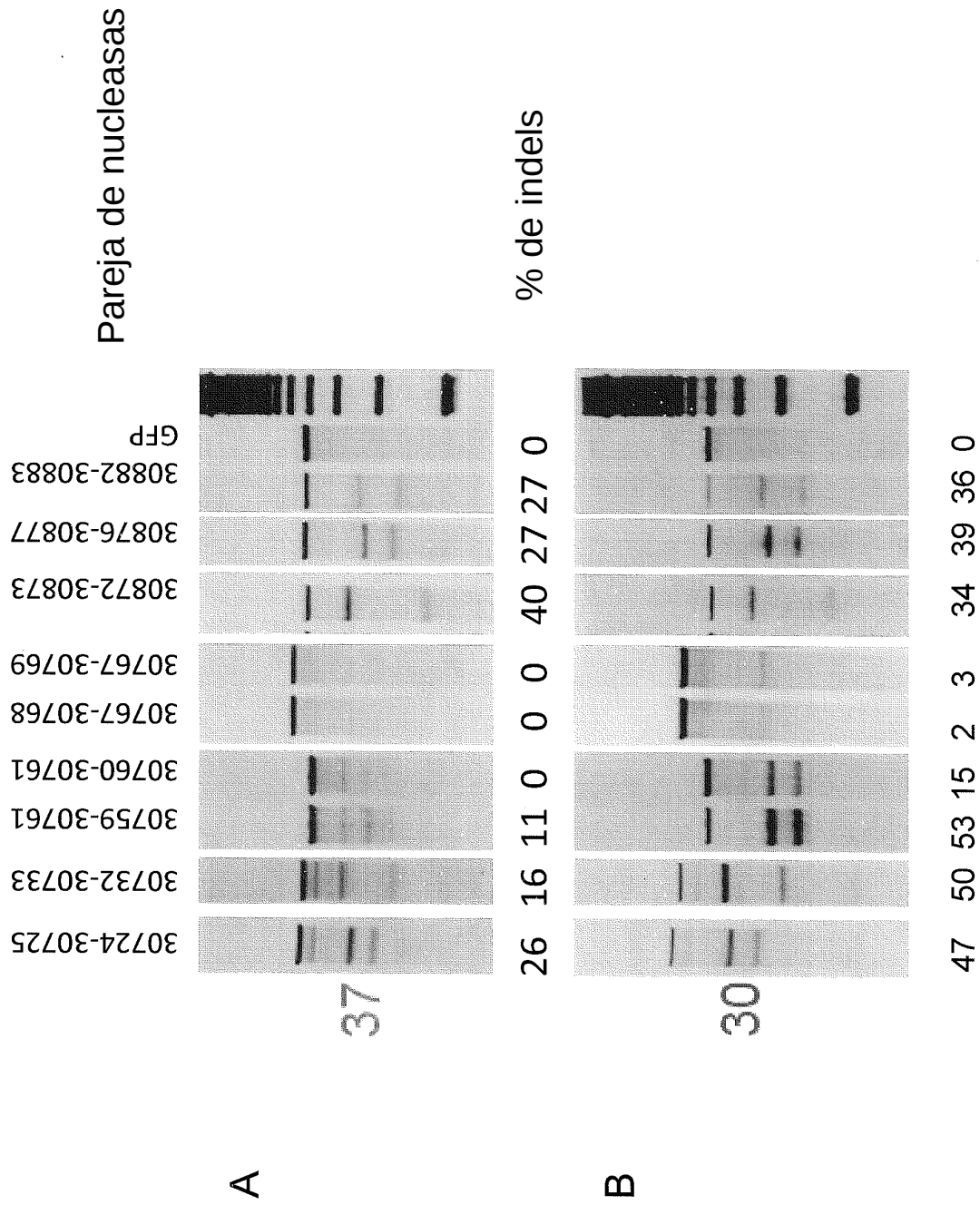


Figura 1

Donantes de etiquetas para la enfermedad por almacenamiento lisosomal

5' ITR	Etiqu. 5'	Intrón 1 parcial de Alb	F9 SA	Transgén	Etiqu.	bGH pA	Intrón 1 parcial de Albm-Exón 2-Intrón 2 parcial	Etiqu. 3'	3' ITR

Transgenes:

- Enfermedad de Fabry: GLA
- Enfermedad de Gaucher: GBA
- Enfermedad de Hurler: IDUA
- Enfermedad de Hunter: IDS

Figura 2

Los rAAV-2/8 y 2/8.2 generados en 293 o baculovirus y dirigidos al intrón 1 del gen de la albúmina son infecciosos y activos en los ratones normales

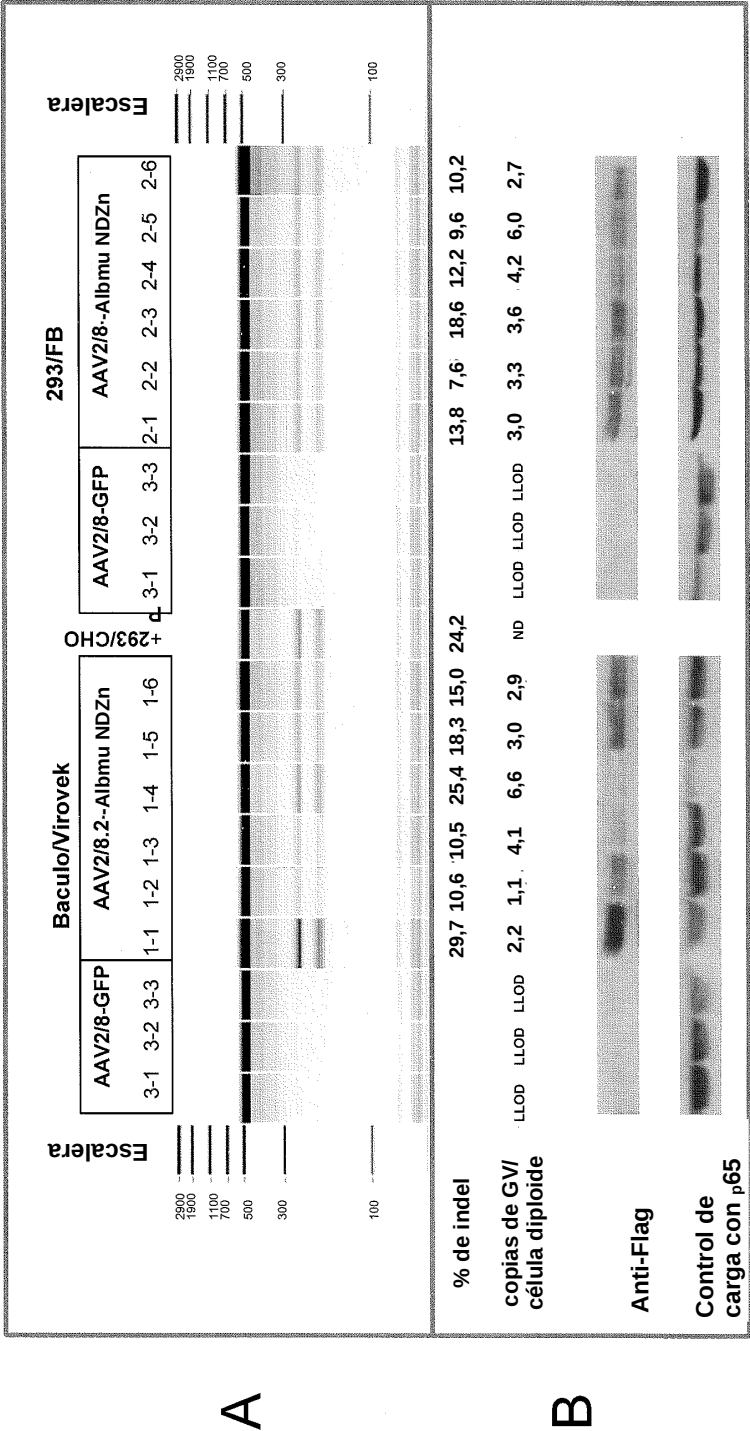


Figura 3

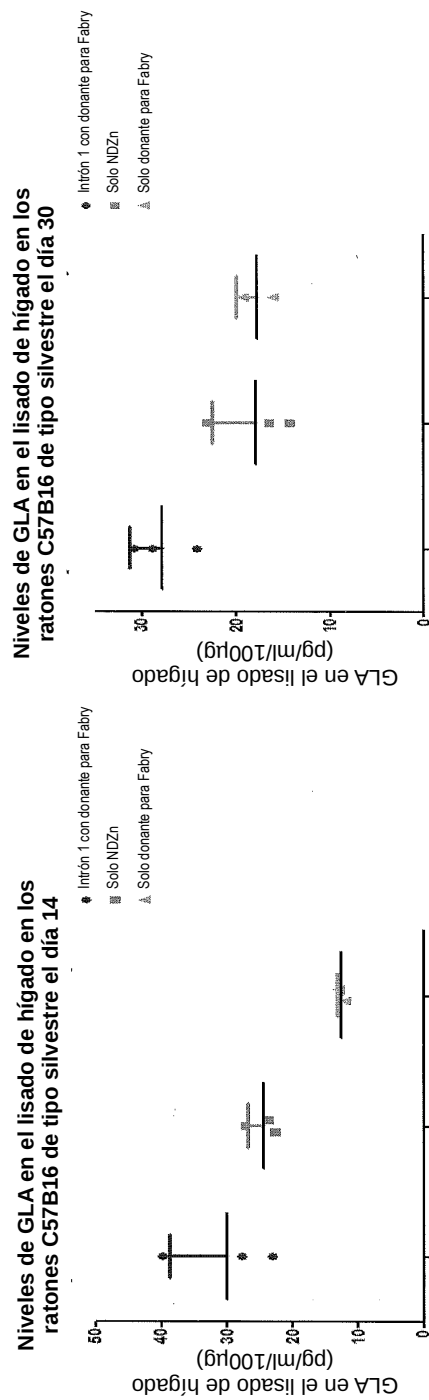
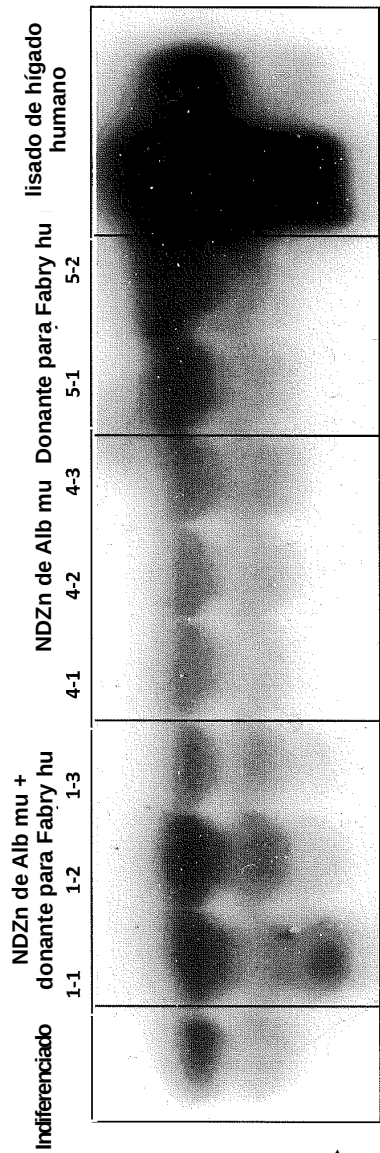


Figura 4

Figura 5

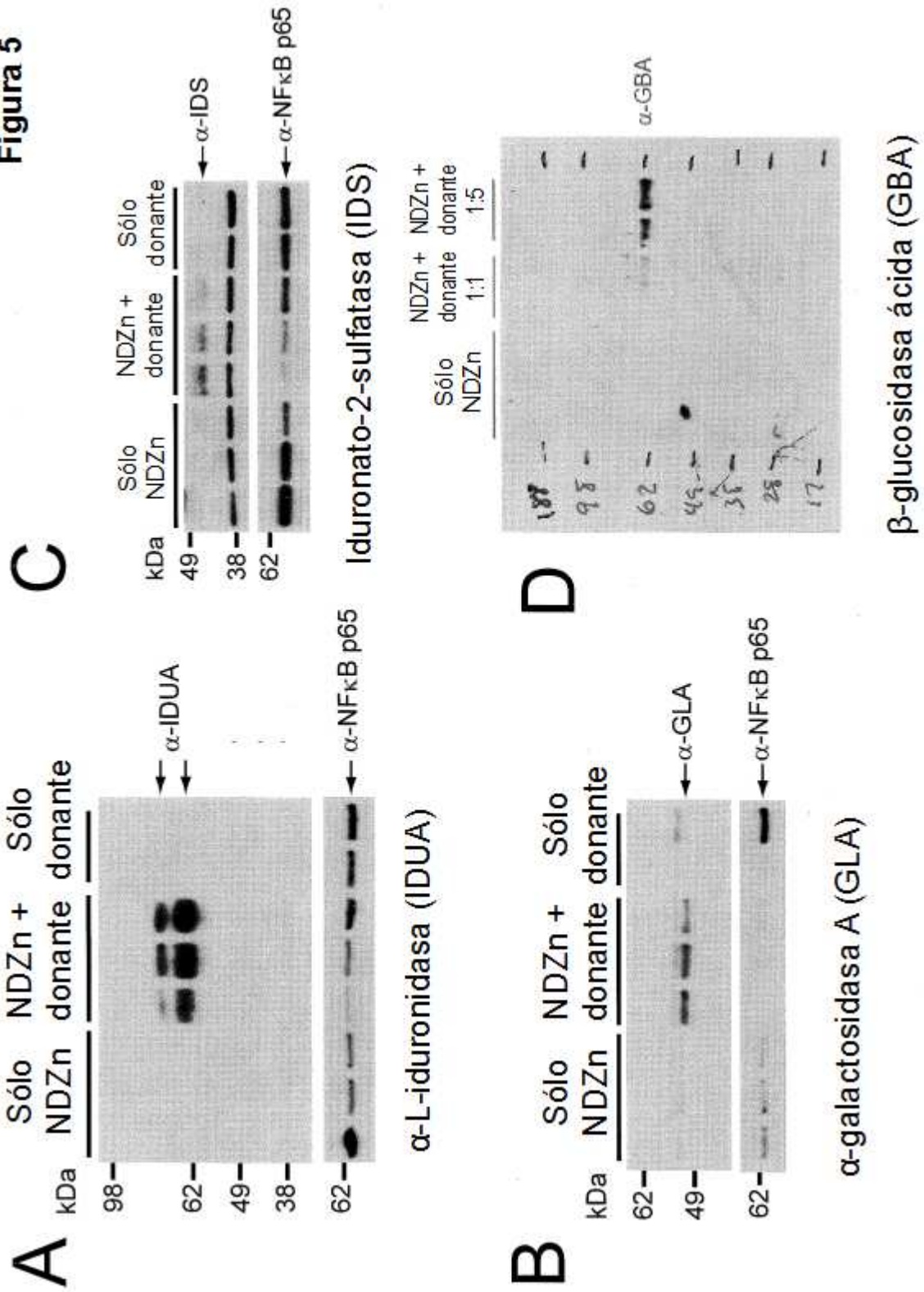
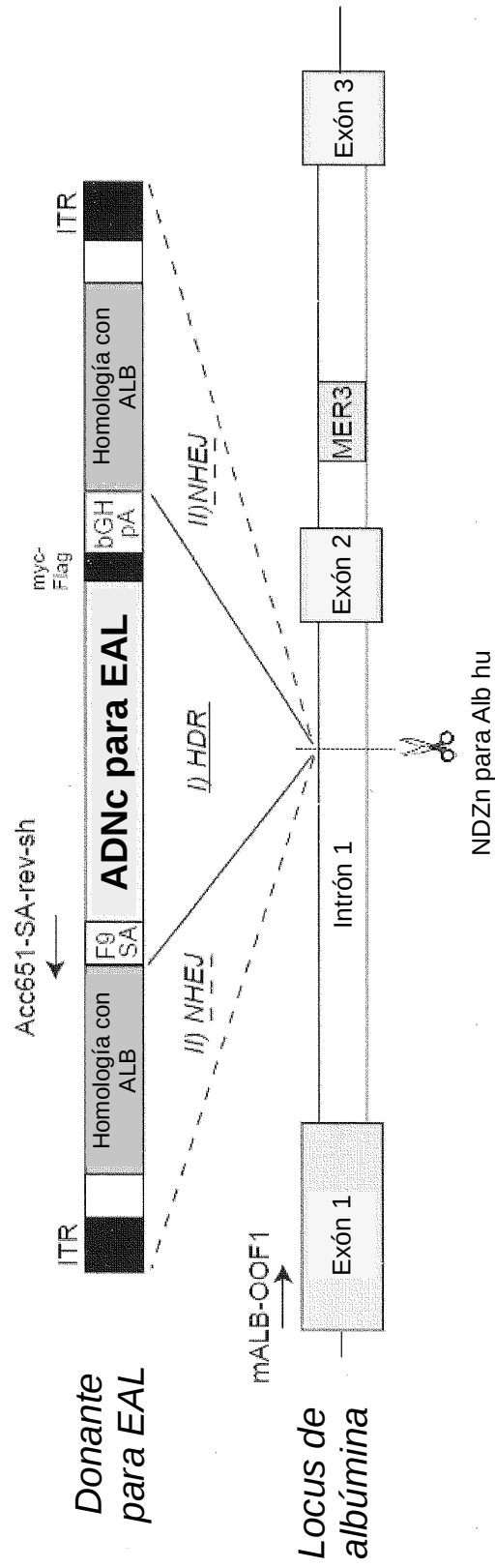


Figura 6



1) Integración del donante por HDR (gracias a los brazos de homología)

2) Integración del donante por NHEJ (captura de extremos)

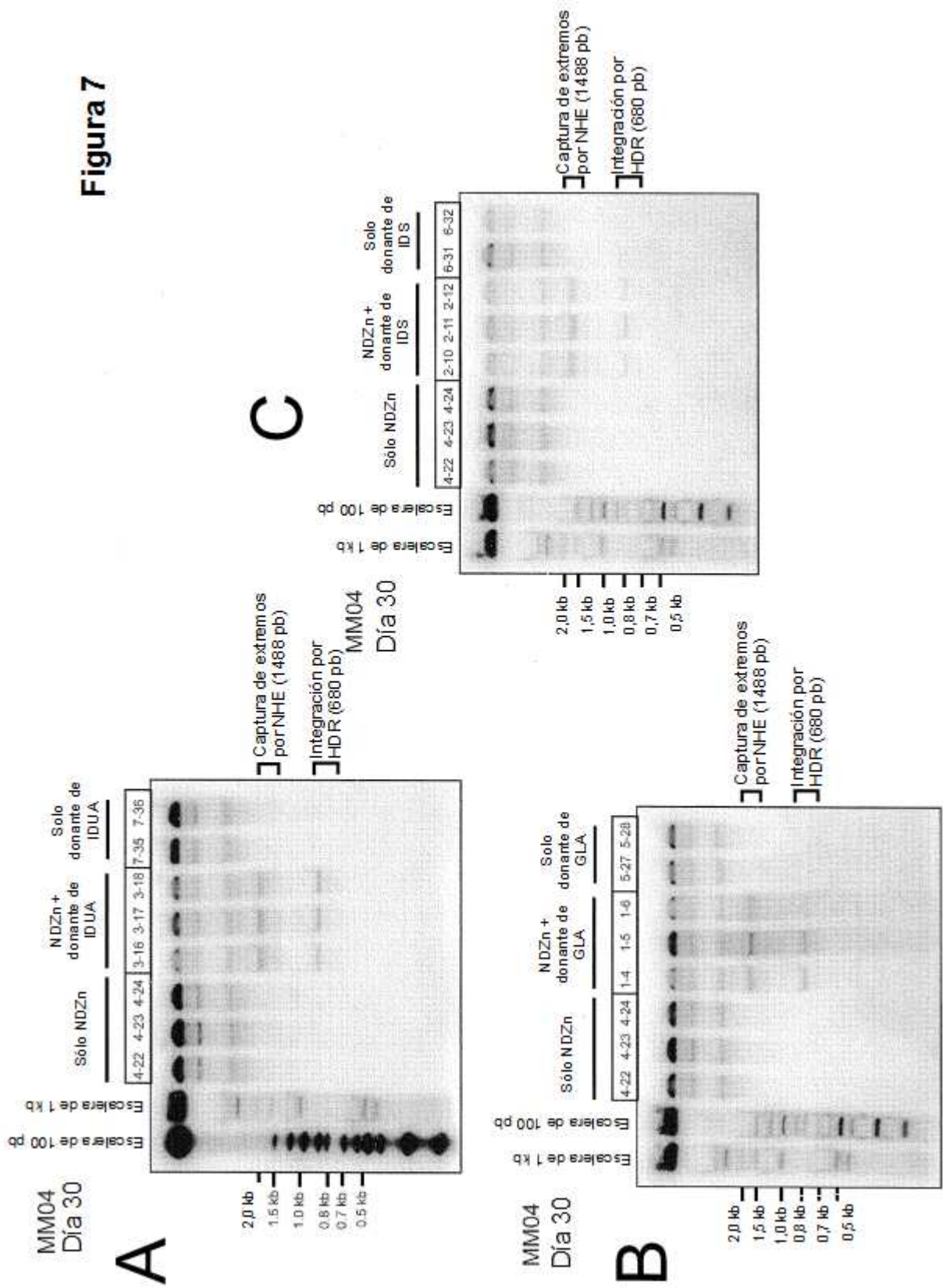


Figura 8

Donantes en AlbM

1. pAAV_Fabry_MF_DNA2.0_mALB_Intron1_D4									
5'ITR	Etiqu. 5'	Intrón 1 parcial de Alb-mu	F9 SA	α-galactosidasa A (GLA)	E TI Q	bGH pA	Intrón 1 parcial de Alb mu-Exón 2-Intrón 2 parcial	Etiqu. 3'	3'ITR
2. pAAV_Gaucher_MF_DNA2.0_mALB_Intron1_D4									
5'ITR	Etiqu. 5'	Intrón 1 parcial de Alb-mu	F9 SA	β-Glucosidasa ácida (GBA)	E TI Q	bGH pA	Intrón 1 parcial de Alb mu-Exón 2-Intrón 2 parcial	Etiqu. 3'	3'ITR
3. pAAV_Hurler_MF_DNA2.0_mALB_Intron1_D4									
5'ITR	Etiqu. 5'	Intrón 1 parcial de Alb-mu	F9 SA	α-L-Iduronidasa (DUA) [MPS-I]	E TI Q	bGH pA	Intrón 1 parcial de Alb mu-Exón 2-Intrón 2 parcial	Etiqu. 3'	3'ITR
4. pAAV_Hunter_MF_DNA2.0_mALB_Intron1_D4									
5'ITR	Etiqu. 5'	Intrón 1 parcial de Alb-mu	F9 SA	Iduronato-2-sulfatasa (IDS) [MPS-II]	E TI Q	bGH pA	Intrón 1 parcial de Alb mu-Exón 2-Intrón 2 parcial	Etiqu. 3'	3'ITR

Etiqueta: MYC-FLAG (MF)

Último AA del ADNc-RTRPLEQKLISEEDLAANDILDYKDDDDK-PARADA
MYC FLAG

Figura 9

