

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 813 080**

51 Int. Cl.:

**C12N 9/22** (2006.01)

**A61K 48/00** (2006.01)

**C12N 15/90** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.03.2013 E 18198271 (1)**

97 Fecha y número de publicación de la concesión europea: **27.05.2020 EP 3444342**

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:

**22.03.2021**

73 Titular/es:

**SANGAMO THERAPEUTICS, 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 2 813 080 T3

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

## DESCRIPCIÓN

Métodos y composiciones para el tratamiento de enfermedades por almacenamiento lisosomal

5 **Campo técnico**

La presente divulgación está en el campo del tratamiento de las enfermedades por almacenamiento lisosomal (LSD) y la terapia génica.

10 **Antecedentes**

La terapia génica alberga un potencial enorme para una nueva era de agentes terapéuticos humanos. Estas metodologías permitirán el tratamiento para afecciones que hasta ahora no se han 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 esa célula exprese un producto que no se estaba produciendo previamente en esa célula. Los ejemplos de usos de esta tecnología incluyen la inserción de un gen que codifica una proteína terapéutica, la inserción de una secuencia codificadora que codifica una proteína que no está presente por alguna razón 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 pueden administrarse a una célula por una variedad de vías, de manera que el transgén se integra en el propio genoma de la célula y se mantiene ahí. En años recientes, se ha desarrollado una estrategia para la integración de transgenes que usa la escisión con nucleasas específicas de sitio para la inserción dirigida en un locus genómico elegido (véase, p. ej., la Patente U.S. en copropiedad 7.888.121). Las nucleasas, tales como nucleasas de dedo de cinc (ZFN), las nucleasas efectoras semejantes a activador de la transcripción (TALEN), o sistemas de nucleasas tales como el sistema CRISPR/Cas (que utiliza un ARN guía preparado por ingeniería), son específicas para genes diana y pueden utilizarse de manera que la construcción del transgén se inserta bien por reparación dirigida por homología (HDR) o por la captura de extremos durante procesos dirigidos por la unión de extremos no homólogos (NHEJ).

Los loci diana incluyen loci "seguros" tales como los genes AAVS1, HPRT y CCR5 en las células humanas, y *Rosa26* en las células murinas (véanse las Publicaciones de Patente de los Estados Unidos en copropiedad Nos. 20080299580; 20080159996 y 201000218264 y la Solicitud de Patente de los Estados Unidos 13/660.821). La integración mediada por nucleasas ofrece la perspectiva de una expresión transgénica mejorada, seguridad y durabilidad de la expresión incrementadas, comparado con las estrategias clásicas de integración que se basan en la integración aleatoria del transgén, ya que permite el posicionamiento exacto del transgén para un riesgo mínimo de silenciamiento génico o activación de oncogenes cercanos.

Mientras la administración del transgén a la célula diana es un obstáculo que debe superarse para llevar a cabo completamente esta tecnología, otro asunto que debe dominarse es asegurar que después de que el transgén se inserta en la célula y se expresa, el producto génico así codificado debe alcanzar la localización necesaria en el organismo, y producirse en concentraciones locales suficientes para ser eficaz. Para las enfermedades caracterizadas por la ausencia de una proteína o por la presencia de una aberrante no funcional, la administración de una proteína de tipo salvaje codificada por un transgén puede ser extremadamente útil.

Las enfermedades por almacenamiento lisosomal (LSD) son un grupo de enfermedades monogénicas metabólicas raras caracterizadas por la ausencia de proteínas lisosomales individuales funcionales implicadas normalmente en la degradación de lípidos, glicoproteínas y mucopolisacáridos de desecho. Estas enfermedades se caracterizan por una acumulación de estos compuestos en la célula ya que es incapaz de procesarlos para reciclado debido al mal funcionamiento de una enzima específica. Los ejemplos más comunes son las enfermedades de Gaucher (deficiencia de glucocerebrosidasa - nombre del gen: GBA), de Fabry (deficiencia de  $\alpha$  galactosidasa - GLA), de Hunter (deficiencia de iduronato-2-sulfatasa-IDS), de Hurler (deficiencia de alfa-L iduronidasa - IDUA), y de Niemann-Pick (deficiencia de esfingomielina fosfodiesterasa 1 - SMPD1). Cuando se agrupan conjuntamente, las LSD tienen una incidencia en la población de aproximadamente 1 en 7.000 nacimientos. Estas enfermedades tienen efectos devastadores en los que las padecen. Habitualmente, se diagnostican en primer lugar en bebés que pueden tener patrones de crecimiento faciales y corporales característicos y pueden tener un retraso mental moderado a severo. Las opciones de tratamiento incluyen terapia de reemplazo enzimático (ERT) en las que la enzima ausente se proporciona al paciente, habitualmente mediante inyección intravenosa en dosis altas. Dicho tratamiento solo es para tratar los síntomas y no es curativo, así se debe proporcionar al paciente una dosificación repetida de estas proteínas durante el resto de sus vidas y potencialmente pueden desarrollar anticuerpos neutralizantes frente a la proteína inyectada. Frecuentemente, estas proteínas tienen una semivida en suero corta, y así el paciente también debe soportar infusiones frecuentes de la proteína. Por ejemplo, los pacientes con la enfermedad de Gaucher que reciben el producto Cerezyme® (imiglucerasa) deben recibir infusiones tres veces a la semana. La producción y purificación de las enzimas también son problemáticas y así los tratamientos son muy costosos (>\$100.000 por año por paciente).

Así, permanece una necesidad de métodos y composiciones adicionales que puedan usarse para tratar una

enfermedad monogénica (p. ej., Enfermedades por almacenamiento lisosomal) a través de edición genómica, y de métodos para administrar un producto génico codificado por un transgén expresado a un nivel terapéuticamente relevante.

5 SUMARIO

La invención proporciona uno o más transgenes y una o más nucleasas con dedos de zinc (ZFN) para su uso en un método de tratamiento de una enfermedad por almacenamiento lisosomal, el método comprendiendo administrar el uno o más transgenes y el uno o más ZFN a un sujeto con necesidad de ello para generar una célula que produce una proteína que trata la enfermedad por almacenamiento lisosomal, en donde el transgén que codifica la proteína se inserta en un locus de albúmina endógena de la célula usando uno o más ZFN, y en donde la enfermedad por almacenamiento lisosomal es la enfermedad de Gaucher, Hunter, Hurler o Fabry.

La invención también proporciona uno o más transgenes y uno o más polinucleótidos que codifican una o más ZFN para su uso en un método de tratamiento de una enfermedad por almacenamiento lisosomal, el método comprendiendo administrar el uno o más transgenes y el uno o más polinucleótidos que codifican la una o más ZFN a un sujeto con necesidad de ello para generar una célula que produzca una proteína que trate la enfermedad por almacenamiento lisosomal, en donde el transgén que codifica la proteína se inserta en un locus de albúmina endógeno de la célula usando uno o más ZFN, y en donde la enfermedad por almacenamiento lisosomal es la enfermedad de Gaucher, Hunter, Hurler o Fabry.

En la presente se divulgan métodos y composiciones para tratar una enfermedad monogénica. Se describen métodos para la inserción de una secuencia transgénica en una célula diana adecuada en los que el transgén codifica una proteína que trata la enfermedad. La proteína terapéutica puede excretarse de la célula diana de manera que es capaz de afectar a o ser captada por otras células que no portan el transgén. También se divulgan métodos para la producción de una célula (p. ej., una célula madura o indiferenciada) que produce altos niveles de un agente terapéutico en el que la introducción de una población de estas células alteradas en un paciente suministrará la proteína necesaria para tratar una enfermedad o afección.

En un aspecto se describe una proteína de dedo de cinc (ZFP) que se une al sitio diana en una región de interés (p. ej., un gen asociado a enfermedad, un gen altamente expresado, un gen de albúmina u otro gen seguro) en un genoma, en el que la ZFP comprende uno o más dominios de unión de dedo de cinc preparados por ingeniería. En una realización, la ZFP es una nucleasa de dedo de cinc (ZFN) que escinde una región genómica diana de interés, en el que la ZFN comprende uno o más dominios de unión de dedo de cinc preparados por ingeniería y un dominio de escisión de nucleasa o semidominio de escisión. Los dominios de escisión y los semidominios de escisión pueden obtenerse, por ejemplo, a partir de varias endonucleasas de restricción y/o endonucleasas de dirección. En una realización, los semidominios de escisión derivan de una endonucleasa de restricción de Tipo IIS (p. ej., *Fok I*). En determinadas realizaciones, el dominio de dedo de cinc reconoce un sitio diana en un gen asociado a enfermedad o seguro tal como albúmina (p. ej., una proteína de dedo de cinc que tiene 5 o 6 dedos con las regiones de hélice de reconocimiento mostradas en una única fila de la Tabla 3).

En otro aspecto se describe una proteína TALE (semejante a activador de la transcripción) que se une al sitio diana en una región de interés (p. ej., un gen altamente expresado, un gen asociado a enfermedad o un gen seguro) en un genoma, en el que la TALE comprende uno o más dominios de unión de TALE preparados por ingeniería. La TALE es una nucleasa (TALEN) que escinde una región genómica diana de interés, en el que la TALEN comprende uno o más dominios de unión a ADN de TALE preparados por ingeniería y un dominio de escisión de nucleasa o semidominio de escisión. Los dominios de escisión y los semidominios de escisión pueden obtenerse, por ejemplo, a partir de varias endonucleasas de restricción y/o endonucleasas de dirección. Los semidominios de escisión pueden derivarse de una endonucleasa de restricción de Tipo IIS (p. ej., *Fok I*). El dominio de unión a ADN de TALE puede reconocer un sitio diana en un gen altamente expresado, asociado a enfermedad o seguro.

En otro aspecto se describe un sistema CRISPR/Cas que se une al sitio diana en una región de interés (p. ej., un gen altamente expresado, un gen asociado a enfermedad o un gen seguro) en un genoma, en el que el sistema CRISPR/Cas comprende una nucleasa CRISPR/Cas y un ARNcr/ARNtracr preparado por ingeniería (o único ARN guía). En determinadas realizaciones, el sistema CRISPR/Cas reconoce un sitio diana en un gen altamente expresado, asociado a enfermedad o seguro.

La ZFN, TALEN, y/o sistema CRISPR/Cas como se describen en la presente memoria pueden unirse a y/o escindir la región de interés en una región codificadora o no codificadora en o adyacente al gen, tal como, por ejemplo, una secuencia líder, secuencia remolque o intrón, o en una región no transcrita, bien secuencia arriba o secuencia abajo de la región codificadora. En ciertos aspectos de la descripción, la ZFN, TALEN, y/o sistema CRISPR/Cas se une a y/o escinde un gen altamente expresado, por ejemplo, un gen de globina en células sanguíneas rojas (RBC). Véase, p. ej., la Solicitud U.S. No. 61/670.451, titulada "Methods and Compositions for Delivery of Biologics", presentada el 11 de julio, 2012. En otras realizaciones, la ZFN, TALEN, y/o sistema CRISPR/Cas se une a y/o escinde un gen seguro, por ejemplo, un gen CCR5, un gen PPP1R12C (también conocido como AAV S1), albúmina, HPRT o un gen *Rosa*. Véanse, p. ej., las Publicaciones de Patente U.S. Nos. 20080299580; 20080159996 y 201000218264 y

Las Solicitudes U.S. Nos. 61/537.349, 61/560.506, 13/660.821 y la Solicitud U.S. no. 61/670.451 titulada "Methods and Compositions for Regulation of Transgene Expression", presentada el 11 de julio, 2012. Además, para ayudar en la selección, puede usarse el locus HPRT (véanse las solicitudes de patente U.S. 13/660.821 y 13/660.843). La ZFN, TALEN, y/o sistema CRISPR/Cas pueden unirse a y/o escindir un gen asociado a enfermedad (p. ej., el gen que codifica la  $\alpha$ -galactosidasa hidrolasa A (AGA) lisosomal, relacionada con la Enfermedad de Fabry). En otro aspecto, en la presente memoria se describen composiciones que comprenden una o más de las nucleasas de dedo de cinc y/o TALE o sistema CRISPR/Cas descritos en la presente memoria. También se describen composiciones que comprenden una o más de estas nucleasas y ácido nucleico donante. En algunos aspectos, se describen nucleasas o sistemas CRISPR/Cas preparados por ingeniería capaces de escindir genes reguladores aberrantes asociados a enfermedad y métodos para usar estas nucleasas para tratar la enfermedad mediante la reducción o eliminación de la expresión del producto génico aberrante.

En un aspecto, se un método para tratar una enfermedad por almacenamiento lisosomal mediante la inserción en un transgén correctivo en una célula diana adecuada (p. ej., célula sanguínea, célula hepática, célula cerebral, célula madre, célula precursora, etc.) de manera que el producto codificado por ese transgén correctivo se expresa. El transgén correctivo puede insertarse en una línea celular para la producción *in vitro* de la proteína de reemplazo. Las células que comprenden el transgén o la proteína producida por las células pueden usarse para tratar a un paciente que lo necesita, por ejemplo, después de la purificación de la proteína producida. El transgén correctivo puede insertarse en un tejido diana en el cuerpo de manera que la proteína de reemplazo se produce *in vivo*. En algunos aspectos, la proteína expresada se excreta de la célula para actuar en o ser captada por otras células (p. ej., mediante la exportación en la sangre) que carecen del transgén. En algunos casos, el tejido diana es el hígado. En otros casos, el tejido diana es el cerebro. En otros casos, la diana es la sangre (p. ej., vasculatura). En otros casos, la diana es músculo esquelético. El gen correctivo puede comprender la secuencia de tipo salvaje del gen que funciona, o la secuencia del transgén correctivo está alterada de alguna manera para proporcionar una actividad biológica aumentada. En algunos aspectos, el transgén correctivo comprende codones optimizados para incrementar la actividad biológica, mientras en otros aspectos, la secuencia está alterada para proporcionar a la proteína resultante una función más deseada (p. ej., mejora en la estabilidad, alteración de la carga para alterar la unión al sustrato etc.). El transgén puede alterarse para lograr una inmunogenicidad reducida. En otros casos, el transgén está alterado de manera que la proteína codificada se vuelve un sustrato para la administración mediada por transportador en tejidos específicos tales como el cerebro (véase Gabathuler et al. (2010) *Neurobiology of Disease* 37: 48-57).

En otro aspecto, se divulga una proteína nucleasa preparada por ingeniería capaz de escindir (editar) el genoma de una célula madre o precursora (p. ej., precursor de célula sanguínea, célula madre hepática, etc.) para la introducción de un transgén deseado. En algunos aspectos, las células madre o precursoras editadas se expanden entonces y pueden inducirse para que se diferencien en células editadas maduras *ex vivo*, y después las células se proporcionan al paciente. En otros aspectos, los precursores editados (p. ej., células madre CD34+) se proporcionan en un trasplante de médula ósea que, después de un implante exitoso, proliferan produciendo células editadas que entonces se diferencian y se maduran *in vivo* y contienen el producto biológico expresado del transgén. En otros aspectos, las células madre editadas son células madre musculares que se introducen entonces en tejido muscular. En algunos aspectos, la nucleasa preparada por ingeniería es una Nucleasa de Dedo de Cinc (ZFN) y en otros, la nucleasa es una nucleasa TALE (TALEN), y en otros aspectos, se usa un sistema CRISPR/Cas. Las nucleasas pueden prepararse por ingeniería para tener especificidad para el locus seguro de la albúmina. Como ejemplo no limitativo solo, el locus seguro puede ser el gen de la albúmina mientras el gen asociado a enfermedad puede ser el gen *GLA* que codifica la  $\alpha$ -galactosidasa hidrolasa A lisosomal (Véase la Tabla 2). Un gen que está altamente expresado en células sanguíneas rojas (RBC) es el de beta-globina. En otro aspecto, las células transgénicas se sensibilizan *ex vivo* mediante electrosensibilización para incrementar su susceptibilidad para la disrupción después de exposición a una fuente de energía (p. ej., ultrasonidos) (véase WO 2002007752).

En otro aspecto describe en la presente un polinucleótido que codifica una o más ZFN, TALEN, y/o sistema CRISPR/Cas descrito 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).

En otro aspecto se describe en la presente un vector de expresión de ZFN, TALEN, y/o sistema CRISPR/Cas que comprende un polinucleótido, que codifica una o más ZFN, TALEN, y/o sistema CRISPR/Cas descrito en la presente memoria, unido de forma operativa a un promotor. El vector de expresión puede ser un vector viral.

En otro aspecto se describe una célula huésped que comprende uno o más vectores de expresión de ZFN, TALEN, y/o sistema CRISPR/Cas como se describe en la presente memoria. La célula huésped puede transformarse de forma estable o transfectarse de forma transitoria o una combinación de estas con uno o más vectores de expresión de ZFN, TALEN, y/o sistema CRISPR/Cas. La célula huésped puede ser una célula hepática.

En otro aspecto se describe en la presente un método para escindir un locus altamente expresado, asociado a enfermedad y/o seguro en una célula, comprendiendo el método: introducir, en la célula, uno o más polinucleótidos que codifican una o más ZFN, TALEN, y/o sistema CRISPR/Cas que se une(n) a un sitio diana en los uno o más loci

diana bajo condiciones tales que ZFN(s),TALEN(s) o el sistema CRISPR/Cas se exprese(n) y el uno o más loci se escindan. Los ejemplos no limitativos de ZFN, TALEN, y/o sistemas CRISPR/Cas que se unen a loci altamente expresados y/o seguros se describen en las Publicaciones U.S. Nos. 20080299580; 20080159996 y 201000218264 y las solicitudes U.S. 13/660.821, 13/660.843, 13/624.193 y 13/624.217 y la Solicitud U.S. no. 61/670.451, titulada "Methods and Compositions for Delivery of Biologics".

En otros aspectos de la descripción, una secuencia genómica en cualquier gen diana se reemplaza con el transgén terapéutico, por ejemplo, usando una ZFN, TALEN, y/o sistema CRISPR/Cas (o vector que codifica dicha ZFN, TALEN, y/o sistema CRISPR/Cas) como se describe en la presente memoria y una secuencia o transgén "donante" que se inserta en el gen después de la escisión dirigida con la ZFN, TALEN, y/o sistema CRISPR/Cas. La secuencia donante puede estar presente en el vector de ZFN o TALEN, presente en un vector separado (p. ej., vector de Ad, AAV o LV) o, alternativamente, puede introducirse en la célula usando un mecanismo de administración de ácidos nucleicos diferente. Dicha inserción de una secuencia de nucleótidos donante en el locus diana (p. ej., gen altamente expresado, gen asociado a enfermedad, otro gen seguro, etc.) da como resultado la expresión del transgén bajo el control de los elementos de control genético endógenos del locus diana (p. ej., albúmina, globina, etc.). En algunos aspectos, la inserción del transgén de interés, por ejemplo, en un gen diana (p. ej., albúmina), da como resultado la expresión de una secuencia de proteína exógena intacta y carece de aminoácidos codificados 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 que se inserta el transgén (p. ej., del locus diana endógeno o, alternativamente, de secuencias en el transgén que codifican secuencias del locus diana). La diana puede ser cualquier gen, por ejemplo, un gen seguro tal como un gen de albúmina, un gen de AAVS1, un gen de HPRT; un gen de CCR5; o un gen altamente expresado tal como un gen de globina en una RBC (p. ej., beta globina o gamma globinas). En algunos casos, las secuencias endógenas estarán presentes en la parte amino (N)-terminal de la proteína exógena, mientras en otros, las secuencias endógenas estarán presentes en la parte carboxi (C)-terminal de la proteína exógena. En otros casos, las secuencias endógenas estarán presentes tanto en las partes N como C-terminal de la proteína exógena. Las secuencias endógenas pueden incluir secuencias endógenas de tipo salvaje de longitud completa o mutantes o, alternativamente, pueden incluir secuencias de aminoácidos parcialmente endógenas. La fusión gen endógeno-transgén puede estar localizada en el locus endógeno en la célula mientras en otras realizaciones, la secuencia endógena-secuencia codificadora del transgén se inserta en otro locus en un genoma (p. ej., una secuencia de IDUA-transgén insertada en un locus de la albúmina, HPRT o CCR5). En algunos aspectos, el seguro se selecciona del locus de AAVS1, Rosa, albúmina, HPRT o CCR5 (véanse las Publicaciones U.S. en copropiedad Nos. 20080299580; 20080159996; y 201000218264 y las solicitudes U.S. 13/660.821, 13/660.843, 13/624.193 y 13/624.217 y la Solicitud U.S. no 61/670.451, titulada "Methods and Compositions for Regulation of Transgene Expression", presentada el 11 de julio, 2011). En otras realizaciones, el gen asociado a enfermedad se selecciona de GLA ( $\alpha$ -galactosidasa hidrolasa A lisosomal), o de uno o más genes listados en la Tabla 2.

En algunas realizaciones, el transgén se expresa de manera que un producto de proteína terapéutica se retiene en la célula (p. ej., célula precursora o madura). En otras realizaciones, el transgén se fusiona con el dominio extracelular de una proteína de membrana de manera que después de la expresión, una fusión transgénica dará como resultado la localización en la superficie de la proteína terapéutica. En algunos aspectos, el dominio extracelular se elige de las proteínas listadas en la Tabla 1. En algunos aspectos, las células editadas también comprenden una proteína transmembrana para transportar a las células a un tipo de tejido particular. En un aspecto, la proteína transmembrana es un anticuerpo, mientras en otros, la proteína transmembrana es un receptor. En determinadas realizaciones, la célula es una RBC precursora (p. ej., CD34+ o célula madre hematopoyética) o madura. En algunos aspectos, el producto de proteína terapéutica codificado en el transgén se exporta fuera de la célula para afectar o ser captado por células que no tienen el transgén. En determinadas realizaciones, la célula es una célula hepática que libera la proteína terapéutica en la corriente sanguínea para actuar en tejidos distales (p. ej., cerebro).

También se divulgan métodos y composiciones para la producción de una célula (p. ej., RBC) que porta una proteína terapéutica para una LSD que puede usarse de forma universal para todos los pacientes como un producto alogénico. Esto permitiría el desarrollo de un único producto para el tratamiento de pacientes con una LSD particular, por ejemplo. Estos vehículos pueden comprender proteínas transmembrana para asistir en el transporte de la célula. En un aspecto, la proteína transmembrana es un anticuerpo, mientras en otros, la proteína transmembrana es un receptor.

En un aspecto, se divulgan métodos y composiciones para la inactivación de genes asociados a la enfermedad. En algunas realizaciones, estos genes son aquellos cuyos productos pueden regular la expresión de un gen en una célula precursora o madura. En algunos aspectos, la inactivación es en el sitio diana regulador en el ADN para dichas proteínas. En algunos aspectos, el gen regulador es aberrante de manera que la inactivación del gen restaura la función normal. En otros aspectos, el gen que se va a inactivar es un alelo asociado a enfermedad de manera que la inactivación de ese alelo enfermo permite la expresión a partir de un alelo de tipo salvaje y restaura la función normal.

En una realización, el transgén se expresa a partir del promotor de la albúmina después de la inserción en el locus de la albúmina. El agente biológico codificado por el transgén puede liberarse entonces en la corriente sanguínea

si el transgén se inserta en un hepatocito *in vivo*. En algunos aspectos, el transgén se administra a hígado *in vivo* en un vector viral a través de una inyección intravenosa.

En otra realización, el transgén codifica un ARN no codificador, p. ej., un ARNsh. La expresión del transgén antes de la maduración de la célula dará como resultado una célula que contiene el ARN no codificador de interés.

En la presente se describen células precursoras aisladas (células madre hematopoyéticas, células madre musculares o células madre hematopoyéticas (HSC) CD34+) en las que se ha insertado un transgén, de manera que las células maduras derivadas de estos precursores contienen altos niveles del producto codificado por el transgén. En algunas realizaciones, estos precursores son células madre pluripotentes inducidas (iPSC).

En algunos aspectos los métodos de la divulgación pueden usarse *in vivo* en sistemas de animales transgénicos. En algunos aspectos, el animal transgénico puede usarse en el desarrollo de un modelo en el que el transgén codifica un gen humano. En algunos casos, el animal transgénico puede inactivarse en el locus endógeno correspondiente, permitiendo el desarrollo de un sistema *in vivo*, en el que la proteína humana puede estudiarse aislada. Dichos modelos transgénicos pueden usarse para propósitos de cribado para identificar moléculas pequeñas, o biomoléculas grandes u otras entidades, que pueden interactuar con o modificar la proteína humana de interés. En algunos aspectos, el transgén se integra en el locus seleccionado (p. ej., altamente expresado o seguro) en una célula madre (p. ej., una célula madre embrionaria, una célula madre pluripotente inducida, una célula madre hepática, una célula madre neural, etc.) o embrión animal obtenido por cualquiera de los métodos descritos en la presente memoria, y después el embrión se implanta de manera que nace un animal vivo. El animal se cría entonces hasta la madurez sexual y se permite que produzca crías en las que al menos parte de las crías comprende el transgén integrado.

En un aspecto adicional más, se proporciona en la presente un método para la integración específica de sitio de una secuencia de ácido nucleico en un locus endógeno (p. ej., un gen asociado a enfermedad, altamente expresado tal como globina en RBC, o seguro, tal como el gen de la albúmina, CCR5, HPRT o *Rosa*) de un cromosoma, por ejemplo, en el cromosoma de un embrión. En ciertos aspectos, el método comprende: (a) inyectar a un embrión (i) al menos un vector de ADN, en el que el vector de ADN comprende una secuencia en 5' y una secuencia en 3' que flanquean la secuencia de ácido nucleico que se va a integrar, y (ii) al menos una molécula de ARN que codifica una nucleasa de dedo de cinc, TALE o sistema CRISPR/Cas, que reconoce el sitio de integración en el locus diana, y (b) cultivar el embrión para permitir la expresión de la ZFN, TALEN, y/o sistema CRISPR/Cas, en el que una rotura bicatenaria introducida en el sitio de integración por la ZFN, TALEN, y/o sistema CRISPR/Cas se repara, mediante recombinación homóloga con el vector de ADN, de manera que se integra la secuencia de ácido nucleico en el cromosoma.

Los métodos y compuestos descritos en la presente pueden combinarse con otros agentes terapéuticos para el tratamiento de sujetos con enfermedades por almacenamiento lisosomal. En algunos aspectos, los métodos y composiciones se usan en combinación con métodos y composiciones para permitir el paso a través de la barrera hematoencefálica. En otros aspectos, los métodos y composiciones se usan en combinación con compuestos que se sabe que suprimen la respuesta inmune del sujeto.

También se divulga un kit, que comprende la ZFN, TALEN, y/o sistema CRISPR/Cas. El kit puede comprender ácidos nucleicos que codifican la ZFN, TALEN, y/o sistema CRISPR/Cas (p. ej., moléculas de ARN o los genes que codifican ZFN, TALEN, y/o sistema CRISPR/Cas contenidos en un vector de expresión adecuado), moléculas donantes, vectores de expresión que codifican el único ARN guía, líneas celulares huéspedes adecuadas, instrucciones para realizar los métodos de la invención, y semejantes.

Estos y otros aspectos serán fácilmente evidentes para el experto en la técnica a la vista de la descripción como un todo.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **Figura 1, paneles A y B**, representan un conjunto compuesto de geles que demuestra los resultados de un ensayo de concordancia errónea Cel-I (Surveyor™, Transgenomic) que mide la escisión en una localización de interés por una pareja de nucleasas que ha sido seguida por un evento NHEJ. NHEJ causa la inserción o delección de bases de nucleótido ("indels") que entonces crea una concordancia errónea cuando la cadena de ADN se hibrida con una cadena de ADN de tipo salvaje. La Figura 1A muestra los resultados medidos cuando la transfección de las parejas de nucleasas específicas de albúmina en células Neuro2A se llevó a cabo a 37°C y la Figura 1B muestra los resultados cuando la transducción de las parejas de nucleasas se realizó bajo choque hipotérmico (30°C). El porcentaje de concordancia errónea, o % de indels, es una medida de la actividad nucleasa de cada pareja bajo cada condición.

La **Figura 2** es un esquema que representa la estructura de cuatro donantes AAV diseñados para proporcionar 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), que flanquean los brazos de homología

para la inserción de los donantes en el locus de la albúmina por mecanismos dependientes de homología, un sitio aceptor de corte y empalme, el ADN que codifica la enzima de reemplazo y una Etiqueta MYC-Flag para permitir la identificación de los donantes integrados.

La **Figura 3, paneles A y B**, demuestran la actividad de las ZFN de albúmina de ratón *in vivo*. Se administró a ratones machos normales (n=6) una única dosis de 200 microlitros de  $1,0 \times 10^{11}$  genomas de vector total bien de AAV2/8 o AAV2/8.2 que codifica la pareja de ZFN específica murina SBS30724 y SBS30725 para evaluar la infectividad hepática por la detección de las copias del genoma del vector AAV y la actividad NHEJ *in vivo* en ratones normales. Los vectores se proporcionaron por inyección en la vena de la cola en ratones como se describe, y 14 días después de la administración, los ratones se sacrificaron, los hígados se recogieron y procesaron para la cuantificación del ADN o proteínas totales. La detección de las copias del genoma del vector AAV se realizó por PCR cuantitativa y la actividad de escisión de la ZFN se midió usando el ensayo Cel-1 (Surveyor, Transkaryotic). La **Figura 3A** representa un gel con los resultados de Cel-1 de los ratones a los que se proporcionó AAV2/8 que contenía un casete de expresión de GFP o AAV2/8 que comprendía las ZFN, en el que los AAV se produjeron mediante un sistema de expresión 293 o un sistema de baculovirus. La **Figura 3B** representa la cuantificación de los carriles en el gel y muestra que la infección de los ratones con el AAV que contenía las ZFN específicas de albúmina da como resultado cerca de un 30% de actividad NHEJ cuantificable.

La **Figura 4, paneles A y B**, demuestran la inserción de un donante de transgén huGLa (deficiente en pacientes que padecen la enfermedad de Fabry) en el locus de la albúmina en ratones. La Figura 4A muestra una transferencia Western frente a la proteína huGLa codificada por el transgén, en el que la flecha indica la presunta proteína. La comparación de las muestras de ratón de los ratones que recibieron tanto ZFN como donante (muestras 1-1, 1-2 y 1-3) con las muestras que recibieron bien solo ZFN (4-1, 4-2, 4-3) o los que solo recibieron el donante huGLa ("donante hu Fabry"), muestras 5-1 y 5-2 da lugar a la identificación de una banda que coincide con el control de lisado hepático humano. La Figura 4B representa los resultados de ELISA usando un kit de ELISA específico de huGLa, en el que las muestras se analizaron de ratones bien 14 o 30 días después de la introducción del virus. Las barras de error representan las desviaciones estándar (n=3). Los resultados demuestran que los ratones que recibieron tanto la ZFN como el donante (círculos) tenían cantidades mayores de señal de huGLa que los que solo recibieron ZFN (cuadrados) o solo recibieron donante (triángulos).

La **Figura 5, paneles A-D**, representa transferencias Western que demuestran la expresión en homogenados de hígado de los transgenes del donante LSD insertados en el locus de la albúmina en ratones. La Figura 5A muestra los resultados usando el transgén que codifica IDUA, La Figura 5B muestra los resultados usando el transgén GLA, la Figura 5C muestra los resultados usando el transgén IDS, y la Figura 5D muestra los resultados usando el transgén GBA.

La **Figura 6** es un esquema que representa los dos tipos de inserción de donante que pueden ocurrir después de la escisión mediada por ZFN. La inserción de donante mediada por NHEJ dará como resultado que se integre la construcción entera LSD-Donante, mientras la inserción mediada por HDR solo causará la incorporación del ADNc incluyendo el sitio del aceptor de corte y empalme F9. La Figura 6 también representa la localización de los dos cebadores de PCR ("mALB-OOF1" y "Acc651-SA-inv-sh") usados para detectar el tipo de integración que se ha producido.

La **Figura 7, paneles A-C**, representan los resultados de PCR con radiomarcaje con  $^{32}\text{P}$  realizada en homogenados de hígado en ratones que contienen los transgenes LSD integrados 30 días después del tratamiento. La Figura 7A representa a los ratones con el transgén IDUA, la Figura 7B representa a aquellos con el transgén GLA, y la Figura 7C representa a aquellos con el transgén IDS. En todos los casos, las bandas indican que la inserción de los transgenes se ha producido a través de integración tanto mediada por NHEJ como mediada por HDR.

La **Figura 8** es un esquema que ilustra el diseño de los donantes LSD que contienen etiquetas de epítipo. Se indican la localización y secuencias de las etiquetas Myc y Flag.

La **Figura 9, paneles A y B**, representa un gel de PCR con radiomarcaje con  $^{32}\text{P}$  realizada como se ha descrito anteriormente en homogenados de hígado de ratones con donantes LSD integrados que contienen etiquetas de epítipo. La Figura 9A muestra que la integración se había producido a través de integración tanto mediada por NHEJ como mediada por HDR para el transgén diana GLA, IDUA e IDS. La Figura 9B muestra lo mismo para el transgén GBA.

### Descripción detallada

En la presente memoria se divulgan métodos y composiciones para tratar o prevenir una enfermedad por almacenamiento lisosomal (LSD). Se divulgan métodos y composiciones para la inserción de un gen que codifica una proteína que no está presente o se expresa de manera insuficiente en el sujeto con la LSD de manera que el gen se expresa en el hígado y se expresa la proteína terapéutica (de reemplazo). También se describe la alteración de una célula (p. ej., RBC precursora o madura, iPSC o célula hepática) de manera que produce altos niveles del agente terapéutico y la introducción de una población de estas células alteradas en un paciente suministrará la proteína necesaria. El transgén puede codificar una proteína deseada o ARN estructural que es beneficioso terapéuticamente en un paciente que lo necesita.

Así, los métodos y composiciones de la descripción pueden usarse para expresar a partir de un transgén,

proteínas terapéuticamente beneficiosas de locus de albúmina para reemplazar enzimas que son defectivas en enfermedades por almacenamiento lisosomal. Adicionalmente, se divulgan métodos y composiciones para el tratamiento de estas enfermedades mediante la inserción de las secuencias en loci altamente expresados en células tales como células hepáticas.

Además, el transgén puede introducirse en células derivadas de pacientes, p. ej., células madre pluripotentes inducidas (iPSC) derivadas de pacientes u otros tipos de células madre (embrionarias o hematopoyéticas) para uso en un implante eventual. Particularmente útil es la inserción del transgén asociado a enfermedad en una célula madre hematopoyética para el implante en un paciente que lo necesita. Al diferenciarse las células madre en células maduras, contendrán altos niveles de la proteína de reemplazo para administración a los tejidos.

## General

La práctica de los métodos, así como la preparación y uso de las composiciones descritas en la presente memoria emplean, a no ser que se indique otra cosa, técnicas convencionales de biología molecular, bioquímica, estructura y análisis de la cromatina, química computacional, cultivo celular, ADN recombinante y campos relacionados como están dentro de la técnica. Estas técnicas se explican totalmente en la bibliografía. Véase, por ejemplo, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edición, Cold Spring Harbor Laboratory Press, 1989 y Tercera edición, 2001; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, 1987 y actualizaciones periódicas; la serie METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Tercera Edición, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman 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 usan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en forma bien mono o bicatenaria. Para los propósitos de la presente descripción, estos términos no deben considerarse como limitantes respecto a la longitud del polímero. Los términos también engloban análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en los restos de base, azúcar y/o fosfato (p. ej., núcleos de fósforoato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo de A se emparejará por bases con T.

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

"Unión" se refiere a una interacción específica de secuencia no covalente entre macromoléculas (p. ej., entre una proteína y un ácido nucleico). No es necesario que todos los componentes de una interacción de unión sean específicos de secuencia (p. ej., los contactos con residuos fosfato en un núcleo de ADN), siempre que la interacción como un todo sea específica de secuencia. Dichas interacciones se caracterizan generalmente por una constante de disociación ( $K_d$ ) de  $10^{-6}$  M<sup>-1</sup> o menor. "Afinidad" se refiere a la fuerza de la unión: estando correlacionada una afinidad de unión incrementada con una  $K_d$  menor.

Una "proteína de unión" es una proteína que es capaz de unirse no covalentemente a otra molécula. Una proteína de unión puede unirse, por ejemplo, a una molécula de ADN (una proteína de unión a ADN), una molécula de ARN (una proteína de unión a ARN) y/o una molécula de proteína (una proteína de unión a proteína). En el caso de una proteína de unión a proteína, puede unirse a sí misma (para formar homodímeros, homotrímeros, *etc.*) y/o puede unirse a una o más moléculas de una proteína o proteínas diferentes. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas de dedo de cinc tienen actividad de unión a ADN, de unión a ARN y de unión a proteínas.

Una "proteína de unión a ADN de dedo de cinc" (o dominio de unión) es una proteína, o un dominio en una proteína mayor, que se une a ADN de una manera específica de secuencia a través de uno o más dedos de cinc, que son regiones de secuencia de aminoácidos en el dominio de unión cuya estructura se estabiliza a través de la coordinación con un ion cinc. El término proteína de unión a ADN de dedo de cinc se abrevia frecuentemente como proteína de dedo de cinc o ZFP.

Un "dominio de unión a ADN TALE" o "TALE" es un polipéptido que comprende uno o más dominios/unidades de repetición TALE. Los dominios de repetición están implicados en la unión de la TALE a su secuencia de ADN diana cognada. Una única "unidad de repetición" (también referida como una "repetición") tiene una longitud típicamente de 33-35 aminoácidos y presenta al menos alguna homología de secuencia con otras secuencias de repetición TALE en una proteína TALE natural.

Los dominios de unión de dedo de cinc y TALE pueden "prepararse por ingeniería" para unirse a una secuencia de nucleótidos predeterminada, por ejemplo, mediante la preparación por ingeniería (alteración de uno o más aminoácidos) de la región de la hélice de reconocimiento de una proteína de dedo de cinc o TALE natural. Por lo tanto, las proteínas de unión a ADN preparadas por ingeniería (de dedo de cinc o TALE) son proteínas que no son naturales. Los ejemplos no limitativos de métodos para preparar por ingeniería proteínas de unión a ADN son diseño y selección. Una proteína de unión a ADN diseñada es una proteína no natural cuyo diseño/composición resulta principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y algoritmos computerizados para procesar la información en una base de datos que almacena información sobre los diseños y datos de unión de ZFP y/o TALE existentes. Véanse, por ejemplo, las Patentes US 6.140.081; 6.453.242; y 6.534.261; véanse también WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496 y la Publicación U.S. No. 20110301073.

Una proteína de dedo de cinc o TALE "seleccionada" es una proteína que no encuentra en la naturaleza cuya producción resulta principalmente de un proceso empírico tal como exposición en fago, trampa de interacción o selección de híbridos. Véanse, p. ej., US 5.789.538; US 5.925.523, US 6.007.988; US 6.013.453; US 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084 y Publicación U.S. No. 20110301073.

"Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, "recombinación homóloga (HR)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble cadena en las células a través de mecanismos de reparación dirigidos por homología. Este proceso requiere homología en las secuencias de nucleótidos, usa una molécula "donante" para la reparación con molde de una molécula "diana" (es decir, la que experimenta la rotura de doble cadena), y se conoce de forma variada como "conversión génica sin entrecruzamiento" o "conversión génica de extensión corta", porque da lugar a la transferencia de información genética del donante a la diana. Sin pretender la vinculación a ninguna teoría particular, dicha transferencia puede implicar la corrección de concordancia errónea de ADN heterodúplex que se forma entre la diana rota y el donante, y/o "hibridación de cadenas dependiente de síntesis", en la que el donante se usa para resintetizar la información genética que formará parte de la diana, y/o procesos relacionados. Dicha HR especializada da como resultado frecuentemente una alteración de la secuencia de la molécula diana de manera que parte o toda de la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

En los métodos de la descripción, una o más nucleasas dirigidas como se describe en la presente memoria crean una rotura de doble cadena en la secuencia diana (p. ej., cromatina celular) en un sitio predeterminado, y un polinucleótido "donante", que tiene homología con la secuencia de nucleótidos en la región de la rotura, puede introducirse en la célula. Se ha mostrado que la presencia de la rotura de doble cadena facilita la integración de la secuencia donante. La secuencia donante puede integrarse físicamente o, alternativamente, el polinucleótido donante se usa como un molde para la reparación de la rotura mediante recombinación homóloga, produciendo la introducción de todo o parte de la secuencia de nucleótidos como en el donante en la cromatina celular. Así, una primera secuencia en la cromatina celular puede alterarse y, en determinadas realizaciones, puede convertirse en una secuencia presente en un polinucleótido donante. Así, el uso de los términos "reemplazar" o "reemplazo" puede entenderse que representa el reemplazo de una secuencia de nucleótidos por otra (es decir, el reemplazo de una secuencia en el sentido de información) y no requiere necesariamente el reemplazo físico o químico de un polinucleótido por otro.

En cualquiera de los métodos descritos en la presente memoria, pueden usarse parejas adicionales de proteínas de dedo de cinc o TALEN para la escisión de doble cadena adicional de sitios diana adicionales en la célula.

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

En cualquiera de los métodos descritos en la presente memoria, la primera secuencia de nucleótidos (la "secuencia donante") puede contener secuencias que son homólogas, pero no idénticas, a secuencias genómicas en la región de interés, estimulando de esta manera la recombinación homóloga para insertar una secuencia no idéntica en la región de interés. Así, en determinadas realizaciones, partes de la secuencia donante que son homólogas a secuencias en la región de interés presentan una identidad de secuencia entre aproximadamente el 80 al 99% (o cualquier número entero intermedio) con la secuencia genómica que se reemplaza. En otras realizaciones, la homología entre la secuencia donante y genómica es mayor del 99%, por ejemplo, si solo 1 nucleótido se diferencia entre las secuencias donante y genómica de más 100 pares de bases contiguos. En determinados casos, una parte no homóloga de la secuencia donante puede contener secuencias que no están presentes en la región de interés, de manera que se introducen nuevas secuencias en la región de interés. En estos casos, la secuencia no homóloga está flanqueada generalmente por secuencias de 50-1.000 pares de bases (o cualquier valor entero intermedio) o cualquier número de pares de bases mayor de 1.000, que son homólogas o idénticas a secuencias en la región de interés. En

otras realizaciones, la secuencia donante no es homóloga a la primera secuencia y se inserta en el genoma por mecanismos de recombinación no homóloga.

5 Cualquiera de los métodos descritos en la presente memoria puede usarse para la inactivación parcial o completa de una o más secuencias diana en una célula por la integración dirigida de una secuencia donante que  
 10 disrumpe la expresión del o de los genes de interés. También se proporcionan líneas celulares con genes inactivados parcialmente o completamente.

10 Además, los métodos de integración dirigida como se describen en la presente memoria también pueden usarse para integrar una o más secuencias exógenas. La secuencia de ácido nucleico exógena puede comprender,  
 15 por ejemplo, uno o más genes o moléculas de ADNc, o cualquier tipo de secuencia codificadora o no codificadora, así como uno o más elementos de control (p. ej., promotores). Además, la secuencia de ácido nucleico exógena puede producir una o más moléculas de ARN (p. ej., ARN de horquilla pequeños (ARNsh), ARN inhibidores (ARNi), microARN (ARNmi), etc.).

15 "Escisión" se refiere a la rotura del núcleo covalente de una molécula de ADN. La escisión puede iniciarse por una variedad de métodos incluyendo, pero no limitado a, hidrólisis enzimática o química de un enlace fosfodiéster.  
 20 Tanto la escisión monocatenaria como la escisión bicatenaria son posibles y la escisión bicatenaria puede ocurrir como resultado de dos eventos distintos de escisión monocatenaria. La escisión del ADN puede dar como resultado la producción de extremos romos o extremos escalonados. En determinadas realizaciones, se usan polipéptidos de fusión para la escisión dirigida de ADN bicatenario.

25 Un "semidominio de escisión" es una secuencia de polipéptido que, en conjunción con un segundo polipéptido (bien idéntico o diferente), forma un complejo que tiene actividad de escisión (preferiblemente, actividad de escisión de doble cadena). Los términos "primer y segundo semidominios de escisión", "semidominios de escisión + y -" y "semidominios de escisión derecho e izquierdo" se usan indistintamente para hacer referencia a parejas de semidominios de escisión que se dimerizan.

30 Un "semidominio de escisión preparado por ingeniería" es un semidominio de escisión que se ha modificado de manera que forma heterodímeros obligados con otro semidominio de escisión (p. ej., otro semidominio de escisión preparado por ingeniería). Véanse también las Publicaciones de Patente U.S. Nos. 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 ADN o ARN; puede ser lineal, circular o ramificada y puede ser bien monocatenaria o bicatenaria. El término "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede tener cualquier longitud, por ejemplo, una longitud entre 2 y 10.000 nucleótidos (o cualquier valor entero intermedio o superior), preferiblemente una longitud de entre aproximadamente 100 y 1.000 nucleótidos (o cualquier número entero intermedio), más preferiblemente una longitud entre aproximadamente 200 y 500 nucleótidos.

40 Un "gen asociado a enfermedad" es uno que es defectivo de alguna manera en una enfermedad monogénica. Los ejemplos no limitativos de enfermedades monogénicas incluyen inmunodeficiencia combinada grave, fibrosis quística, enfermedades por almacenamiento lisosomal (p. ej., de Gaucher, de Hurler, de Hunter, de Fabry, de Neimann-Pick, de Tay-Sach, etc), anemia de células falciformes y talasemia.

45 "Cromatina" es la estructura de nucleoproteínas que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteínas, incluyendo histonas y proteínas cromosómicas distintas de las histonas. La mayoría de la cromatina celular eucariota existe en la forma de nucleosomas, en los que un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociado con un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y ADN conector (de longitud variable dependiendo del organismo) que se extiende entre los núcleos de los nucleosomas. Una molécula de histona H1 está asociada generalmente con el ADN conector. Para los propósitos de la presente descripción, el término "cromatina" se pretende que englobe todos los tipos de nucleoproteínas celulares, tanto procariotas como eucariotas. La cromatina celular incluye tanto cromatina cromosómica como episomal.

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

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

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

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

Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como se genera por un proceso químico combinatorio, o una macromolécula tal como una proteína, ácido nucleico, carbohidrato, lípido, glicoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprende una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser mono o bicatenarios; pueden ser lineales, ramificados o circulares; y pueden tener cualquier longitud. Los ácidos nucleicos incluyen los capaces de formar dúplex, así como ácidos nucleicos que forman tríplex. Véanse, por ejemplo, las Patentes U.S. Nos. 5.176.996 y 5.422.251. Las proteínas incluyen, pero no están limitadas a, proteínas de unión a ADN, factores de transcripción, factores de remodelación de la cromatina, proteínas de unión a ADN metiladas, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, quinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicinas.

Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, p. ej., una proteína o ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma viral infeccioso, un plásmido o episoma introducido en una célula, o un cromosoma que normalmente no está presente en la célula. Los métodos para la introducción de moléculas exógenas en células son conocidos para los expertos en la técnica e incluyen, pero no están limitados a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores virales. Una molécula exógena también puede ser el mismo tipo de molécula que una molécula endógena pero derivada de una especie diferente de la célula de la que deriva. Por ejemplo, una secuencia de ácido nucleico humano puede introducirse en una línea celular derivada originariamente de un ratón o hámster.

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

Una molécula "de fusión" es una molécula en la que dos o más moléculas de subunidades están unidas, preferiblemente covalentemente. Las moléculas de subunidades pueden ser del mismo tipo químico de molécula, o pueden ser diferentes tipos químicos de moléculas. Los ejemplos del primer tipo de molécula de fusión incluyen, pero no están limitados a, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión a ADN de ZFP o TALE y uno o más dominios de activación) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita *supra*). Los ejemplos del segundo tipo de molécula de fusión incluyen, pero no están limitados a, una fusión entre un ácido nucleico que forma tríplex y un polipéptido, y una fusión entre un ligante del surco menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede resultar de la administración de la proteína de fusión a la célula o de la administración de un polinucleótido que codifica la proteína de fusión en una célula, en el que el polinucleótido se transcribe y el transcrito se traduce, para generar la proteína de fusión. El corte y empalme en trans, escisión del polipéptido y ligación del polipéptido también pueden estar implicados en la expresión de una proteína en una célula. Los métodos para la administración de polinucleótidos y polipéptidos a células se presentan en otro lugar en esta descripción.

Un "gen", para los propósitos de la presente descripción, incluye una región de ADN que codifica un producto génico (véase *infra*), así como todas las regiones de ADN que regulan la producción del producto génico, ya estén o no dichas secuencias reguladoras adyacentes a secuencias codificadoras y/o transcritas. De acuerdo con esto, un gen incluye, pero no está necesariamente limitado a, secuencias promotoras, terminadoras, secuencias que regulan la traducción tales como sitios de unión a ribosomas y sitios de entrada interna a ribosomas, potenciadores, silenciadores, aislantes, elementos de límite, orígenes de replicación, sitios de unión a la matriz y regiones de control de locus.

"Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto directo de la transcripción de un gen (p. ej., ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por la traducción de un

ARNm. Los productos génicos también incluyen ARN que están modificados, por procesos tales como recubrimiento, poliadenilación, metilación, y edición, y proteínas modificadas, por ejemplo, por metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristilación y glicosilación.

5 "Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero no está limitado a, activación génica y represión génica. La edición genómica (p. ej., escisión, alteración, inactivación, mutación aleatoria) puede usarse para modular la expresión. La inactivación génica se refiere a cualquier reducción en la expresión génica comparado con una célula que no incluye una ZFP o TALEN como se describe en la presente memoria. Así, la inactivación génica puede ser parcial o completa.

10 Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificadora en o adyacente a un gen, en la que es deseable unir una molécula exógena. La unión puede ser para los propósitos de la escisión dirigida del ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma de orgánulo (p. ej., mitocondrial, de cloroplasto), o un genoma viral infeccioso, por ejemplo. Una región de interés puede estar en la región codificadora de un gen, en regiones no codificadoras transcritas tales como, por ejemplo, secuencias líder, secuencias remolque o intrones, o en regiones no transcritas, bien en 5' o en 3' de la región codificadora. Una región de interés puede ser tan pequeña como una única pareja de nucleótidos o tener una longitud de hasta 2.000 parejas de nucleótidos, o cualquier valor entero de parejas de nucleótidos.

20 Las células "eucariotas" incluyen, pero no están limitadas a, células fúngicas (tales como de levaduras), células de plantas, células animales, células de mamíferos y células humanas (p. ej., células T).

25 Las "Células Sanguíneas Rojas" (RBC) o eritrocitos son células diferenciadas de forma terminal derivadas de células madre hematopoyéticas. Carecen de una nucleasa y de la mayor parte de los orgánulos celulares. Las RBC contienen hemoglobina para transportar el oxígeno desde los pulmones a los tejidos periféricos. De hecho, el 33% de las RBC de un individuo es hemoglobina. También transportan el CO<sub>2</sub> producido por las células durante el metabolismo fuera de los tejidos y de nuevo a los pulmones para su liberación durante la exhalación. Las RBC se producen en la médula ósea en respuesta a la hipoxia en la sangre que está mediada por la liberación de eritropoyetina (EPO) por el riñón. La EPO causa un incremento en el número de proeritroblastos y acorta el tiempo requerido para la maduración completa de las RBC. Después de aproximadamente 120 días, como las RBC no contienen un núcleo ni ningunas otras capacidades regenerativas, las células se eliminan de la circulación bien por las actividades fagocíticas de los macrófagos en el hígado, bazo y nodos linfáticos (~90%) o por hemólisis en el plasma (~10%). Después de ser engullidas por los macrófagos, los componentes químicos de las RBC se degradan en las vacuolas de los macrófagos debido a la acción de las enzimas lisosomales.

40 Los "tejidos secretores" son aquellos tejidos en un animal que secretan productos fuera de la célula del individuo en un lumen de algún tipo que deriva típicamente del epitelio. Los ejemplos de tejidos secretores que están localizados en el tracto gastrointestinal incluyen las células que revisten el intestino, el páncreas, y la vesícula biliar. Otros tejidos secretores incluyen el hígado, tejidos asociados con el ojo y las membranas mucosas tales como las glándulas salivares, glándulas mamarias, la glándula de la próstata, la glándula pituitaria y otros miembros del sistema endocrino. Adicionalmente, los tejidos secretores incluyen células individuales de un tipo de tejido que son capaces de secreción.

45 Los términos "unión operativa" y "unido operativamente" (o "unido de forma operativa") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en la que los componentes se organizan de manera que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los otros componentes. Como ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unida de forma operativa a una secuencia codificadora si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificadora en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción generalmente está unida de forma operativa en *cis* con una secuencia codificadora, pero no es necesario que esté directamente adyacente a esta. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está unido de forma operativa a una secuencia codificadora, incluso si no son contiguos.

60 Con respecto a los polipéptidos de fusión, el término "unido de forma operativa" puede referirse al hecho de que cada uno de los componentes realiza la misma función en unión al otro componente que la que haría si no estuvieran unidos. Por ejemplo, con respecto a un polipéptido de fusión en el que un dominio de unión a ADN de ZFP o TALE está fusionado con un dominio de activación, el dominio de unión a ADN de ZFP o TALE y el dominio de activación están en unión operativa si, en el polipéptido de fusión, la parte del dominio de unión a ADN de ZFP o TALE es capaz de unirse a su sitio diana y/o su sitio de unión, mientras el dominio de activación es capaz de regular al alza la expresión génica. Cuando un polipéptido de fusión en el que un dominio de unión a ADN de ZFP o TALE está fusionado con un dominio de escisión, el dominio de unión a ADN de ZFP o TALE y el dominio de escisión están en unión operativa si, en el polipéptido de fusión, la parte del dominio de unión a ADN de ZFP o TALE es capaz de unirse

a su sitio diana y/o su sitio de unión, mientras el dominio de escisión es capaz de escindir el ADN que se encuentra cerca del sitio diana.

5 Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero que todavía retiene 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 residuos que la molécula nativa correspondiente, y/o puede contener una o más sustituciones de aminoácidos o nucleótidos. Los métodos para determinar la función de un ácido nucleico (p. ej., función codificadora, capacidad para hibridar con otro ácido nucleico) son muy conocidos en la técnica. De forma similar, los métodos para determinar la función de proteínas son muy conocidos. Por ejemplo, la función de unión al ADN de un polipéptido puede determinarse, por ejemplo, por ensayos de unión a filtros, desplazamiento en la movilidad electroforética, o inmunoprecipitación. La escisión del ADN puede ensayarse por electroforesis en gel. Véase, Ausubel *et al.*, *supra*. La capacidad de una proteína de interactuar con otra proteína puede determinarse, por ejemplo, por coimmunoprecipitación, ensayos de dos híbridos o complementación, tanto genética como bioquímica. Véase, por ejemplo, Fields *et al.* (1989) *Nature* **340**:245-246; Patente U.S. No. 5.585.245 y PCT WO 98/44350.

Un "vector" es capaz de transferir secuencias génicas a células diana. Típicamente, "construcción de vector", "vector de expresión" y "vector de transferencia génica" significan cualquier construcción de ácido nucleico capaz de dirigir la expresión de un gen de interés y que puede transferir secuencias génicas a células diana. Así, el término incluye vehículos de clonación y expresión, así como vectores de integración.

Un "gen informador" o "secuencia informadora" se refiere a cualquier secuencia que produce un producto proteico que se mide fácilmente, preferiblemente, aunque no necesariamente, en un ensayo rutinario. Los genes informadores adecuados incluyen, pero no están limitados a, secuencias que codifican proteínas que median la resistencia a antibióticos (p. ej., resistencia a ampicilina, resistencia a neomicina, resistencia a G418, resistencia a puromicina), secuencias que codifican proteínas coloreadas o fluorescentes o luminiscentes (p. ej., proteína verde fluorescente, proteína verde fluorescente aumentada, proteína roja fluorescente, luciferasa), y proteínas que median el crecimiento celular aumentado y/o la amplificación génica (p. ej., dihidrofolato reductasa). Las etiquetas de epítipo incluyen, por ejemplo, una o más copias de FLAG, His, myc, Tap, HA o cualquier secuencia de aminoácidos detectable. Las "etiquetas de expresión" incluyen secuencias que codifican informadores que pueden estar unidas de forma operativa a una secuencia génica deseada con el fin de monitorizar la expresión del gen de interés.

Los términos "sujeto" y "paciente" se usan indistintamente y se refieren a mamíferos tales como pacientes humanos y primates no humanos, así como a animales experimentales tales como conejos, perros, gatos, ratas, ratones y otros animales. De acuerdo con esto, el término "sujeto" o "paciente" tal y como se usa en la presente memoria significa cualquier paciente o sujeto mamífero al que pueden administrarse 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 aquellos que tienen una LSD.

#### 40 **Nucleasas**

En la presente memoria se describen composiciones, particularmente nucleasas, que son útiles en tomar como diana un gen para la inserción de un transgén, por ejemplo, nucleasas que son específicas para un gen seguro tal como albúmina. En determinadas realizaciones, la nucleasa es natural. En otras realizaciones, la nucleasa no es natural, es decir, se prepara por ingeniería en el dominio de unión a ADN y/o dominio de escisión. Por ejemplo, el dominio de unión a ADN de una nucleasa o sistema de nucleasas natural puede alterarse para unirse a un sitio diana seleccionado (p. ej., una meganucleasa que se ha preparado por ingeniería para unirse a un sitio diferente del sitio de unión cognado o un sistema CRISPR/Cas que utiliza un ARN único de guía preparado por ingeniería). La nucleasa puede comprender dominios de unión a ADN y de escisión heterólogos (p. ej., nucleasas de dedo de cinc; nucleasas efectoras TAL; dominios de unión a ADN de meganucleasa con dominios de escisión heterólogos).

##### A. Dominios de unión a ADN

En ciertos aspectos, la nucleasa es una meganucleasa (endonucleasa homing). Las meganucleasas naturales reconocen sitios de escisión de 15-40 pares de bases y se agrupan comúnmente en cuatro familias: la familia LAGLIDADG, la familia GIY-YIG, la familia de la caja His-Cyst y la familia HNH. Las endonucleasas homing ejemplares incluyen I-Scel, I-CeuI, PI-PspI, PI-Sce, I-ScelV, I-Csml, I-PanI, I-Scell, I-Ppol, I-ScellIII, I-Crel, I-TeV1, I-TeVII e I-TeVIII. Sus secuencias de reconocimiento se conocen. Véase también la Patente U.S. No. 5.420.032; la Patente U.S. No. 6.833.252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25**:3379-3388; Dujon *et al.* (1989) *Gene* **82**:115-118; Perler *et al.* (1994) *Nucleic Acids Res.* **22**:1125-1127; Jasin (1996) *Trends Genet.* **12**:224-228; Gimble *et al.* (1996) *J. Mol. Biol.* **263**:163-180; Argast *et al.* (1998) *J. Mol. Biol.* **280**:345-353 y el catálogo de New England Biolabs.

En ciertos aspectos, la nucleasa comprende una endonucleasa homing preparada por ingeniería (no natural) (meganucleasa). Las secuencias de reconocimiento de las endonucleasas homing y meganucleasas tales como I-Scel, I-CeuI, PI-PspI, PI-Sce, I-ScelV, I-Csml, I-PanI, I-Scell, I-Ppol, I-ScellIII, I-Crel, I-TeV1, I-TeVII e I-TeVIII son

conocidas. Véase también la Patente U.S. No. 5.420.032; la Patente U.S. No. 6.833.252, Belfort *et al.* (1997) *Nucleic Acids Res.* **25**:3379-3388; Dujon *et al.* (1989) *Gene* **82**:115-118; Perler *et al.* (1994) *Nucleic Acids Res.* **22**, 1125-1127; Jasin (1996) *Trends Genet.* **12**:224-228; Gimble *et al.* (1996) *J. Mol. Biol.* **263**:163-180; Argast *et al.* (1998) *J. Mol. Biol.* **280**:345-353 y el catálogo de New England Biolabs. Además, la especificidad de la unión a ADN de las endonucleasas homing y meganucleasas puede prepararse por ingeniería para unirse a sitios diana no naturales. Véase, por ejemplo, Chevalier *et al.* (2002) *Molec. Cell* **10**:895-905; Epinat *et al.* (2003) *Nucleic Acids Res.* **31**:2952-2962; Ashworth *et al.* (2006) *Nature* **441**:656-659; Paques *et al.* (2007) *Current Gene Therapy* **7**:49-66; la Publicación de Patente U.S. No. 20070117128. Los dominios de unión a ADN de las endonucleasas homing y meganucleasas pueden alterarse en el contexto de la nucleasa como un todo (es decir, de manera que la nucleasa incluye el dominio de escisión cognado) o pueden fusionarse con un dominio de escisión heterólogo.

En otros aspectos, el dominio de unión a ADN comprende un dominio de unión a ADN efector de TAL natural o preparado por ingeniería (no natural). Véase, p. ej., la Publicación de Patente U.S. No. 20110301073. Se sabe que las bacterias patógenicas de plantas del género *Xanthomonas* causan muchas enfermedades en plantas cultivadas importantes. La patogenicidad de *Xanthomonas* depende de un sistema de secreción conservado de tipo III (T3S) que inyecta más de 25 proteínas efectoras diferentes en la célula de la planta. Entre estas proteínas inyectadas están efectores semejantes a activador de la transcripción (TALE) que mimetizan los activadores transcripcionales de las plantas y manipulan el transcriptoma de la planta (véase, Kay *et al.* (2007) *Science* **318**:648-651). Estas proteínas contienen un dominio de unión a ADN y un dominio de activación de la transcripción. 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 WO2010079430). Las TALE contienen un dominio centralizado de repeticiones en tándem, conteniendo cada repetición aproximadamente 34 aminoácidos, que son clave para la especificidad de la unión al ADN de estas proteínas. Además, contienen una secuencia de localización nuclear y un dominio de activación transcripcional ácido (para una revisión, véase Schornack S, *et al.* (2006) *J Plant Physiol* **163**(3): 256-272). Además, en las bacterias fitopatógenas *Ralstonia solanacearum* se ha encontrado que dos genes, designados brg11 y hpx17 son homólogos a la familia AvrBs3 de *Xanthomonas* en la cepa de *R. solanacearum* biovar 1 GM1000 y en la cepa biovar 4 RS1000 (Véase, Heuer *et al.* (2007) *Appl and Envir Micro* **73**(13): 4379-4384). Estos genes son idénticos en el 98,9% en la secuencia de nucleótidos entre sí, pero se diferencian por una delección de 1.575 pb en el dominio de repetición de hpx17. Sin embargo, ambos productos génicos tienen menos del 40% de identidad de secuencia con las proteínas de la familia AvrBs3 de *Xanthomonas*.

Así, en algunas realizaciones, el dominio de unión a ADN que se une a un sitio diana en un locus diana (p. ej., albúmina u otro seguro) es un dominio preparado por ingeniería de un efector TAL similar a los derivados de los patógenos de plantas de *Xanthomonas* (véase, Boch *et al.*, (2009) *Science* **326**: 1509-1512 y Moscou y Bogdanove, (2009) *Science* **326**: 1501) y *Ralstonia* (véase, Heuer *et al.* (2007) *Applied and Environmental Microbiology* **73**(13): 4379-4384); Publicaciones de Patente U.S. Nos. 20110301073 y 20110145940. Véase, p. ej., las TALEN de albúmina en la Solicitud U.S. No. 13/624.193 y 13/624.217.

El dominio de unión a ADN comprende una proteína de dedo de cinc (p. ej., una proteína de dedo de cinc que se une a un sitio diana en un gen de albúmina o seguro). Preferiblemente, la proteína de dedo de cinc no es natural ya que se ha preparado por ingeniería para unirse a un sitio diana elegido. Véase, por ejemplo, Beerli *et al.* (2002) *Nature Biotechnol.* **20**:135-141; Pabo *et al.* (2001) *Ann. Rev. Biochem.* **70**:313-340; Isalan *et al.* (2001) *Nature Biotechnol.* **19**:656-660; Segal *et al.* (2001) *Curr. Opin. Biotechnol.* **12**:632-637; Choo *et al.* (2000) *Curr. Opin. Struct. Biol.* **10**:411-416; Patentes U.S. Nos. 6.453.242, 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.030.215; 6.794.136; 7.067.317; 7.262.054; 7.070.934; 7.361.635; 7.253.273; y Publicaciones de Patente U.S. Nos. 2005/0064474; 2007/0218528; 2005/0267061.

Un dominio de unión de dedo de cinc o TALE preparado por ingeniería puede tener una nueva especificidad de unión, comparado con una proteína de dedo de cinc natural. Los métodos de preparación por ingeniería incluyen, pero no están limitados a, diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos en tripletes (o cudrupletes) y secuencias de aminoácidos de dedo de cinc individuales, en las que cada secuencia de nucleótidos en triplete o cudruplete se asocia con una o más secuencias de aminoácidos de los dedos de cinc que se unen a la secuencia en triplete o cudruplete. Véanse, por ejemplo, las Patentes U.S. en copropiedad 6.453.242 y 6.534.261.

Los métodos de selección ejemplares, incluyendo los sistemas de exposición en fago y de dos híbridos, se describen en las Patentes US 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como en WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, la potenciación de la especificidad de la unión para los dominios de unión de dedo de cinc se ha descrito, por ejemplo, en WO 02/077227 en copropiedad.

Además, como se describe en estas y otras referencias, los dominios de ADN (p. ej., los dominios de proteínas con múltiples dedos de cinc o TALE) pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, los conectores con una longitud de 5 o más aminoácidos. Véanse también las Patentes U.S. Nos. 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras ejemplares con una longitud de 6 o más

aminoácidos. Las proteínas de unión a ADN descritas en la presente memoria pueden incluir cualquier combinación de conectores adecuados entre los dedos de cinc individuales de la proteína. Además, la potenciación de la especificidad de la unión para los dominios de unión de dedo de cinc se ha descrito, por ejemplo, en WO 02/077227 en copropiedad.

La selección de sitios diana; dominios de unión a ADN y métodos para el diseño y construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos para los expertos en la técnica y se describen con detalle en las Patentes U.S. Nos. 6.140.0815; 789.538; 6.453.242; 6.534.261; 5.925.523; 6.007.988; 6.013.453; 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496 y en la Publicación U.S. No. 20110301073.

Además, como se describe en estas y otras referencias, los dominios de unión a ADN (p. ej., proteínas con múltiples dedos de cinc) pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores con una longitud de 5 o más aminoácidos. Véanse también las Patentes U.S. Nos. 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras ejemplares con una longitud de 6 o más aminoácidos. Las proteínas descritas en la presente memoria pueden incluir cualquier combinación de conectores adecuados entre los dedos de cinc individuales de la proteína.

## B. Dominios de escisión

Cualquier dominio de escisión adecuado puede unirse de forma operativa a un dominio de unión a ADN para formar una nucleasa. Por ejemplo, los dominios de unión a ADN de ZFP se han fusionado con dominios de nucleasa para crear ZFN - una entidad funcional que es capaz de reconocer su ácido nucleico diana pretendido a través de su dominio de unión a ADN preparado por ingeniería (ZFP) y causar que el ADN se corte cerca del sitio de unión de ZFP mediante la actividad de la nucleasa. Véase, p. ej., Kim *et al.* (1996) *Proc Nat'l Acad Sci USA* 93(3):1156-1160. Más recientemente, las ZFN se han usado para la modificación genómica en una variedad de organismos. 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 de unión a ADN de TALE se han fusionado con dominios de nucleasa para crear TALEN. Véase, p. ej., la Publicación U.S. No. 20110301073.

Como se ha indicado anteriormente, los dominios de escisión pueden ser heterólogos respecto al dominio de unión a ADN, por ejemplo, un dominio de unión a ADN de dedo de cinc y un dominio de escisión de una nucleasa o un dominio de unión a ADN de una TALEN y un dominio de escisión, o dominio de unión a ADN de meganucleasa y dominio de escisión de una nucleasa diferente. Los dominios de escisión heterólogos pueden obtenerse de cualquier endonucleasa o exonucleasa. Las endonucleasas ejemplares de las que puede derivarse un dominio de escisión incluyen, pero no están limitadas a, endonucleasas de restricción y endonucleasas homing. Véase, por ejemplo, el Catálogo 2002-2003 de, New England Biolabs, Beverly, MA; y Belfort *et al.* (1997) *Nucleic Acid Res.* **25**:3379-3388. Las enzimas adicionales que escinden ADN son conocidas (p. ej., Nucleasa S1; nucleasa de frijol mungo; ADNasa I pancreática I; nucleasa micrococcal; endonucleasa HO de levaduras; véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o fragmentos funcionales de estas) puede usarse como una fuente de dominios de escisión y semidominios de escisión.

De forma similar, un semidominio de escisión puede derivar de cualquier nucleasa o parte de esta, como se ha mostrado anteriormente, lo que requiere la dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Alternativamente, puede usarse una única proteína que comprende dos semidominios de escisión. Los dos semidominios de escisión pueden derivar de la misma endonucleasa (o fragmentos funcionales de esta), o cada semidominio de escisión puede derivar de una endonucleasa diferente (o fragmentos funcionales de esta). Además, los sitios diana para las dos proteínas de fusión se disponen preferiblemente, con respecto uno del otro, de manera que la unión de las dos proteínas de fusión a sus sitios diana respectivos coloca a los semidominios de escisión en una orientación espacial entre sí que permite que los semidominios de escisión formen un dominio de escisión funcional, p. ej., por dimerización. Así, en determinadas realizaciones, los bordes cercanos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número entero de nucleótidos o pares de nucleótidos puede interponerse entre dos sitios diana (p. ej., de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios diana.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unirse al ADN de manera específica de secuencia (en un sitio de reconocimiento), y escindir el ADN en o cerca del sitio de unión. Determinadas enzimas de restricción (p. ej., Tipo IIS) escinden el ADN en sitios retirados del sitio de reconocimiento y tienen dominios de unión y de escisión separables. Por ejemplo, la enzima de Tipo IIS *Fok I* cataliza la escisión bicatenaria del ADN, a 9 nucleótidos de su sitio de reconocimiento en una cadena y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véanse, por ejemplo, las Patentes US 5.356.802; 5.436.150 y 5.487.994; así como Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89**:4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* **90**:2764-

2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* **91**:883-887; Kim *et al.* (1994b) *J. Biol. Chem.* **269**:31,978-31,982. Así, 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 más dominios de unión de dedo de cinc, que pueden estar o no preparados por ingeniería.

Una enzima de restricción de Tipo IIS ejemplar, cuyo dominio de escisión es separable del dominio de unión, es *Fok I*. Esta enzima particular es activa como un dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95**: 10,570-10,575. De acuerdo con esto, para los propósitos de la presente descripción, la parte de la enzima *Fok I* usada en las proteínas de fusión descritas se considera un semidominio de escisión. Así, para la escisión bicatenaria dirigida y/o el reemplazo dirigido de secuencias celulares usando las fusiones de dedo de cinc-*Fok I*, pueden usarse dos proteínas de fusión, comprendiendo cada una un semidominio de escisión de *Fok I*, para reconstituir un dominio de escisión catalíticamente activo. Alternativamente, también puede usarse una única molécula de polipéptido que contiene un dominio de unión a ADN y dos semidominios de escisión de *Fok I*.

Un dominio de escisión o semidominio de escisión puede ser cualquier parte de una proteína que retiene la actividad de escisión, o que retiene la capacidad de multimerizar (p. ej., dimerizar) para formar un dominio de escisión funcional.

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

En determinadas realizaciones, el dominio de escisión comprende un o más semidominios de escisión preparados por ingeniería (también referidos como mutantes del dominio de dimerización) que minimizan o previenen la homodimerización, como se describe, por ejemplo, en las Publicaciones de Patente U.S. Nos. 20050064474; 20060188987 y 20080131962. Los residuos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, y 538 de *Fok I* son todos dianas para influir en la dimerización de los semidominios de escisión de *Fok I*.

Los semidominios de escisión de *Fok I* preparados por ingeniería ejemplares que forman heterodímeros obligados incluyen una pareja en la que un primer semidominio de escisión incluye mutaciones en los residuos de aminoácidos en las posiciones 490 y 538 de *Fok I* y un segundo semidominio de escisión incluye mutaciones en los residuos de aminoácidos 486 y 499.

Así, en una realización, una mutación en 490 reemplaza Glu (E) con Lys (K); la mutación en 538 reemplaza Iso (I) con Lys (K); la mutación en 486 reemplaza Gln (Q) con Glu (E); y la mutación en la posición 499 reemplaza Iso (I) con Lys (K). Específicamente, los semidominios de escisión preparados por ingeniería descritos en la presente memoria se prepararon mutando las posiciones 490 (E→K) y 538 (I→K) en un semidominio de escisión para producir un semidominio de escisión preparado por ingeniería designado "E490K:I538K" y mutando las posiciones 486 (Q→E) y 499 (I→L) en otro semidominio de escisión para producir un semidominio de escisión preparado por ingeniería designado "Q486E:I499L". Los semidominios de escisión preparados por ingeniería descritos en la presente memoria son mutantes heterodímeros obligados en los que la escisión aberrante se minimiza o suprime. Véase, p. ej., la Publicación de Patente U.S. No. 2008/0131962.

En determinadas realizaciones, el semidominio de escisión preparado por ingeniería comprende mutaciones en las posiciones 486, 499 y 496 (numeración relativa a *Fok I* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo de tipo salvaje de Gln (Q) en la posición 486 con un residuo de Glu (E), el residuo de tipo salvaje de Iso (I) en la posición 499 con un residuo de Leu (L) y el residuo de tipo salvaje de Asn (N) en la posición 496 con un residuo de Asp (D) o Glu (E) (también referidos como dominios "ELD" y "ELE", respectivamente). En otras realizaciones, el semidominio de escisión preparado por ingeniería comprende mutaciones en las posiciones 490, 538 y 537 (numeración relativa a *Fok I* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo de tipo salvaje de Glu (E) en la posición 490 con un residuo de Lys (K), el residuo de tipo salvaje de Iso (I) en la posición 538 con un residuo de Lys (K), y el residuo de tipo salvaje de His (H) en la posición 537 con un residuo de Lys (K) o un residuo de Arg (R) (también referidos como dominios "KKK" y "KKR", respectivamente). En otras realizaciones, el semidominio de escisión preparado por ingeniería comprende mutaciones en las posiciones 490 y 537 (numeración relativa a *Fok I* de tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo de tipo salvaje de Glu (E) en la posición 490 con un residuo de Lys (K) y el residuo de tipo salvaje de His (H) en la posición 537 con un residuo de Lys (K) o un residuo de Arg (R) (también referidos como dominios "KIK" y "KIR", respectivamente). (Véase la Publicación de Patente US No. 20110201055). Los semidominios de escisión preparados por ingeniería descritos en la presente memoria pueden prepararse usando cualquier método adecuado, por ejemplo, por mutagénesis dirigida a sitio de semidominios de escisión de tipo salvaje (*Fok I*) como se describe en las Publicaciones de Patente U.S. Nos. 20050064474; 20080131962; y 20110201055.

Alternativamente, las nucleasas pueden ensamblarse *in vivo* en el sitio diana del ácido nucleico usando la

tecnología denominada "enzima dividida" ("split-enzyme") (véase, p. ej., la Publicación de Patente U.S. No. 20090068164). Los componentes de dichas enzimas divididas pueden expresarse bien en construcciones de expresión separadas o pueden unirse en un marco de lectura abierto donde los componentes individuales se separan, por ejemplo, por un péptido 2A de autoescisión o secuencia IRES. Los componentes pueden ser dominios de unión de dedo de cinc individuales o dominios de un dominio de unión a ácido nucleico de una meganucleasa.

Las nucleasas pueden cribarse para detectar la actividad antes del uso, por ejemplo, en un sistema cromosómico basado en levaduras como se describe en WO 2009/042163 y 20090068164. Las construcciones de expresión de nucleasas pueden diseñarse fácilmente usando métodos conocidos en la técnica. Véanse, p. ej., las Publicaciones de Patente de los Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; y la Publicación Internacional WO 07/014275. La expresión de la nucleasa puede estar bajo el control de un promotor constitutivo o un promotor inducible, por ejemplo, el promotor de la galactoquinasa que se activa (desreprime) en presencia de rafinosa y/o galactosa y se reprime en presencia de glucosa.

El sistema de nucleasa CRISPR (Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas Regularmente, del inglés Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (Asociado a CRISPR, del inglés Crispr Asociated) es un sistema de nucleasas recientemente preparado por ingeniería basado en un sistema bacteriano que puede usarse para la ingeniería genómica. Se basa en parte en la respuesta inmune adaptativa de muchas bacterias y archaea. Cuando un virus o plásmido invade una bacteria, segmentos del ADN del invasor se convierten en ARN de CRISPR (ARNcr) por la respuesta 'inmune'. Este ARNcr se asocia entonces, a través de una región de complementariedad parcial, con otro tipo de ARN denominado ARNtracr para guiar a la nucleasa Cas9 a una región homóloga al ARNcr en el ADN diana denominada un "protoespaciador". Cas9 escinde el ADN para generar extremos romos en el DSB en sitios especificados por una secuencia guía de 20 nucleótidos contenida en el transcrito de ARNcr. La Cas9 requiere tanto al ARNcr como al ARNtracr para el reconocimiento y escisión específicos de sitio en el ADN. Este sistema se ha preparado ahora por ingeniería de manera que el ARNcr y ARNtracr pueden combinarse en una molécula (el "único ARN guía"), y la parte equivalente de ARNcr del único ARN guía puede prepararse preparado por ingeniería para guiar a la nucleasa Cas9 para tomar como diana cualquier secuencia deseada (véase, Jinek *et al* (2012) *Science* 337, p. 816-821, Jinek *et al*, (2013), *eLife* 2:e00471, y David Segal, (2013) *eLife* 2:e00563). Así, el sistema CRISPR/Cas puede prepararse por ingeniería para crear un DSB en una diana deseada en un genoma y la reparación del DSB puede verse influida por el uso de inhibidores de la reparación para causar un incremento en la reparación tendente a error.

### Sitios diana

Como se ha descrito con detalle anteriormente, los dominios de ADN pueden prepararse por ingeniería para unirse a cualquier secuencia elegida en un locus, por ejemplo, un gen de albúmina u otro seguro. Un dominio de unión a ADN preparado por ingeniería puede tener una nueva especificidad de unión, comparado con un dominio de unión a ADN natural. Los métodos de ingeniería incluyen, pero no están limitados a, diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos en triplete (o cuadruplete) y secuencias de aminoácidos individuales (p. ej., de dedo de cinc), en las que cada secuencia de nucleótidos en triplete o cuadruplete está asociada con una o más secuencias de aminoácidos del dominio de unión a ADN que se une a la secuencia en triplete o cuadruplete particular. Véanse, por ejemplo, Las Patentes U.S. en copropiedad 6.453.242 y 6.534.261. El diseño racional de los dominios efectores TAL también puede realizarse. Véase, p. ej., la Publicación U.S. No. 20110301073.

Los métodos de selección ejemplares aplicables a los dominios de unión a ADN, incluyendo exposición en fago y sistemas de dos híbridos se describen en las Patentes US 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como en WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.

La selección de sitios diana; nucleasas y métodos para el diseño y construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos para los expertos en la técnica y se describen con detalle en las Publicaciones de Solicitud de Patente U.S. Nos. 20050064474 y 20060188987.

Además, como se describe en estas y otras referencias, los dominios de unión a ADN (p. ej., proteínas con múltiples dedos de cinc) pueden unirse entre sí usando cualquier secuencia conectora adecuada, incluyendo, por ejemplo, conectores de 5 o más aminoácidos. Véanse, p. ej., las Patentes U.S. Nos. 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras ejemplares con una longitud de 6 o más aminoácidos. Las proteínas descritas en la presente memoria pueden incluir cualquier combinación de conectores adecuados entre los dominios de unión a ADN individuales de la proteína. Véase también la Publicación U.S. No. 20110301073.

### Donantes

Como se ha indicado anteriormente, la inserción de una secuencia exógena (también denominada una "secuencia donante" o "donante"), por ejemplo, para la corrección de un gen mutante o para la expresión incrementada

de un gen de tipo salvaje. Será fácilmente evidente que la secuencia donante no es típicamente idéntica a la secuencia genómica en la que se pone. Una secuencia donante puede contener una secuencia no homóloga flanqueada por dos regiones de homología para permitir la HDR eficiente en la localización 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 respecto a la cromatina celular. Por ejemplo, para la inserción dirigida de secuencias que no están presentes normalmente en una región de interés, dichas secuencias pueden estar presentes en una molécula de ácido nucleico donante y estar flanqueadas por regiones de homología con la secuencia en la región de interés.

El polinucleótido donante puede ser ADN o ARN, monocatenario o bicatenario y puede introducirse en una célula en forma lineal o circular. Si se introduce en forma lineal, los extremos de la secuencia donante pueden protegerse (p. ej., de la degradación exonucleolítica) por métodos conocidos para los expertos en la técnica. Por ejemplo, uno o más residuos de didesoxinucleótidos se añaden al extremo 3' de una molécula lineal y/o se ligan oligonucleótidos autocomplementarios a uno o ambos extremos. Véase, por ejemplo, Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls *et al.* (1996) *Science* 272:886-889. Los métodos adicionales para proteger polinucleótidos exógenos de la degradación incluyen, pero no están limitados a, adición de grupo(s) amino terminales y el uso de uniones internucleotídicas modificadas tales como, por ejemplo, fósforotioatos, fósforoamidatos, y residuos de O-metil ribosa o desoxirribosa.

Un polinucleótido puede introducirse en una célula como parte de una molécula de vector que tiene secuencias adicionales tales como, por ejemplo, orígenes de replicación, promotores y genes que codifican resistencia a antibióticos. Además, los polinucleótidos donantes pueden introducirse como ácido nucleico desnudo, como ácido nucleico formando un complejo con un agente tal como un liposoma o poloxámero, o pueden administrarse por virus (p. ej., adenovirus, AAV, herpesvirus, retrovirus, lentivirus y lentivirus defectivos en integrasa (IDLV)).

El donante se inserta generalmente de manera que su expresión está dirigida por el promotor endógeno en el sitio de la integración, concretamente el promotor que dirige la expresión del gen endógeno en el que se inserta el donante (p. ej., altamente expresado, albúmina, AAVS1, HPRT, etc.). Sin embargo, será evidente que el donante puede comprender un promotor y/o potenciador, por ejemplo, un promotor constitutivo o un promotor inducible o específico de tejido.

La molécula donante puede insertarse en un gen endógeno de manera que todo, parte o nada del gen endógeno se expresa. Por ejemplo, un transgén como se describe en la presente memoria puede insertarse en un locus de albúmina u otro de manera que parte (N-terminal y/o C-terminal al transgén que codifica la enzima lisosomal) o nada de las secuencias de albúmina endógenas se expresan, por ejemplo, como una fusión con el transgén que codifica las secuencias lisosomales. En otras realizaciones, el transgén (p. ej., con o sin secuencias codificadoras adicionales tales como para albúmina) se integra en cualquier locus endógeno, por ejemplo, un locus seguro. Véanse, p. ej., las publicaciones de patente U.S. 20080299580, 20080159996 y 201000218264.

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

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

En determinadas realizaciones, la secuencia exógena (donante) comprende una fusión de una proteína de interés y, como su pareja de fusión, un dominio extracelular de una proteína de membrana, lo que causa que la proteína de fusión esté localizada en la superficie de la célula. Esto permite que la proteína codificada por el transgén actúe potencialmente en el suero. En el caso del tratamiento para una LSD, la enzima codificada por la fusión transgénica será capaz de actuar en los productos metabólicos que se acumulan en el suero desde su localización en la superficie de la célula (p. ej., RBC). Además, si la RBC es engullida por un macrófago esplénico como es el curso normal de degradación, el lisosoma formado cuando el macrófago engulle la célula expondrá la proteína de fusión unida a membrana a las altas concentraciones de productos metabólicos en el lisosoma al pH más favorable naturalmente para esa enzima. Los ejemplos no limitativos de parejas de fusión potenciales se muestran a continuación en la Tabla 1.

**Tabla 1: Ejemplos de parejas de fusión potenciales**

Nombre	Actividad
Band 3	Transportador de aniones, constituye en 25% de la proteína de la superficie de la membrana de RBC
Acuaporina 1	transportador de agua

Glut1	transportador de glucosa y ácido L-deshidroascórbico
proteína antigénica Kidd	transportador de urea
RhAG	transportador de gas
ATP1A1, ATP1B1	Na <sup>+</sup> /K <sup>+</sup> - ATPasa
ATP2B1, ATP2B2, ATP2B3, ATP2B4	Ca <sup>2+</sup> - ATPasa
NKCC1, NKCC2	cotransportador de Na <sup>+</sup> K <sup>+</sup> 2Cl <sup>-</sup>
SLC12A3	cotransportador de Na <sup>+</sup> Cl <sup>-</sup>
SLC12A1, SLA12A2	cotransportador de Na-K
KCC1	cotransportador de K-Cl
KCNN4	Canal de Gardos

Las enfermedades por almacenamiento lisosomal se encuentran típicamente en cinco clases. Estas clases se muestran a continuación en la Tabla 2 junto con ejemplos específicos de las enfermedades. Así, las moléculas donantes descritas en la presente memoria pueden incluir secuencias que codifican una o más enzimas ausentes o deficientes en sujetos con enfermedades por almacenamiento lisosomal, incluyendo, pero no limitado a, las proteínas mostradas en la Tabla 2.

**Tabla 2: Enfermedades por almacenamiento lisosomal**

<b>1. DEFECTOS EN LA DEGRADACIÓN DE GLICANOS</b>			
<b>Proteína</b>	<b>Enfermedad</b>	<b>Gen asociado a la enfermedad</b>	<b>Producto acumulado</b>
<b>I. Defectos en la degradación de glicoproteínas</b>			
α-Sialidasa (neuraminidasa)	Sialidosis	<i>NEU1</i>	glicopéptidos y oligosacáridos sializados
Catepsina A	Galactosialidosis	<i>CTSA</i>	polisacárido
alfa-manosidasa lisosomal	α-Manosidosis	<i>MAN2B1</i>	glicoproteínas y oligosacáridos ricos en manosa
beta-manosidasa lisosomal	β-Manosidosis	<i>MANBA</i>	
Glicosilparaginasa	Aspartilglucosaminuria	<i>AGA</i>	glicoparaginas
Alfa L Fucosidasa	Fucosidosis	<i>FUCA1</i>	fucosa
α-N-Acetilglucosaminidasa	síndrome de Sanfilippo B	<i>NAGL U</i>	glicosaminoglicano
<b>II. Defectos en la degradación de glicolípidos</b>			
<b>A. Gangliósido GM1</b>			
β-Galactosidasa	gangliosidosis GM1 / MPS IVB	<i>GLB1</i>	queratán sulfato
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
proteína activadora de GM2	gangliosidosis GM2	<i>GM2A</i>	gangliósido GM2
Glucocerebrosidasa	enfermedad de Gaucher	<i>GBA</i>	glucocerebrósido
Saposina C	enfermedad de Gaucher (atípica)	<i>PSAP</i>	glucocerebrósido
<b>B. Defectos en la degradación de sulfátido</b>			
Arilsulfatasa A	Leucodistrofia metacromática	<i>ARSA</i>	sulfátido
Saposina B	Leucodistrofia metacromática	<i>PSAP</i>	sulfátido
Enzima generadora de formil-glicina	Deficiencia de sulfatasa múltiple	<i>SUMF1</i>	lípidos sulfatados
β-Galactosilceramidasa (Krabbe)	Leucodistrofia de células globoides	<i>GALC</i>	galactocerebrósido
<b>C. Defectos en la degradación de globotriaosilceramida</b>			
α-Galactosidasa A	Fabry	<i>GLA</i>	globotriaosilceramida
<b>III. Defectos en la degradación de Glicosaminoglicano (Mucopolisacáridos)</b>			
<b>A. Degradación de heparán sulfato</b>			
Iduronato sulfatasa	MPS II (Hunter)	<i>IDS</i>	Dermatán sulfato, Heparán sulfato
Iduronidasa	MPS I (Hurler, Scheie)	<i>IDUA</i>	Dermatán sulfato, Heparán sulfato
Heparan N-sulfatasa	MPS IIIa (Sanfilippo A)	<i>SGSH</i>	Heparán sulfato
Acetil-CoA transferasa	MPS IIIc (Sanfilippo C)	<i>HGSNAT</i>	Heparán sulfato
N-acetil glucosaminidasa	MPS IIIb (Sanfilippo B)	<i>NAGL U</i>	Heparán sulfato
β-glucuronidasa	MPS VII (Sly)	<i>GUSB</i>	

N-acetil glucosamina 6-sulfatasa	MPS III d (Sanfilippo D)	GNS	Heparán sulfato
<b>B. Degradación de otros mucopolisacáridos</b>			
B-Galactosidasa	MPS VIB (Morquio B)	GLB1	Queratán sulfato,
Galactosa 6-sulfatasa	MPS IVA (Morquio A)	GALNS	Queratán sulfato, Condroitin 6-sulfato
Hialuronidasa	MPS IX	HYAL1	Ácido hialurónico
<b>C. Defectos en la degradación de glucógeno</b>			
α-Glucosidasa	Pompe	GAA	Glucógeno
<b>2. DEFECTOS EN LA DEGRADACIÓN DE LÍPIDOS</b>			
<b>I. Defectos en la degradación de esfingomielina</b>			
Esfingomielinasa ácida	Niemann Pick tipo A	SMPD1	esfingomielina
Ceramidasa ácida	Lipogranulomatosis de Farber	ASAH1	mucopolisacárido ácido no sulfonatado
<b>II. Defectos en la degradación de triglicéridos y éster de colesterol</b>			
Lipasa ácida	Enfermedad de Wolman y de almacenamiento de éster de colesterol	LIPA	ésteres de colesterol
<b>3. DEFECTOS EN LA DEGRADACIÓN DE PROTEÍNAS</b>			
Catepsina K	Picnodistostosis	CTSK	
Tripeptidil peptidasa	Lipofuscinosis ceroide	PPT2	
Palmitoil-proteína tioesterasa	Lipofuscinosis ceroide	PPT1	
<b>4. DEFECTOS EN LOS TRANSPORTADORES LISOSOMALES</b>			
Cistinosina (transporte de cistina)	Cistinosis	CTNS	
Sialina (transporte del ácido siálico)	Enfermedad de Salla	SLC17A5	Ácido N-acetilneuramínico
<b>5. DEFECTOS EN LAS PROTEÍNAS DE TRANSPORTE LISOSOMALES</b>			
Subunidad γ de fosfotransferasa	Mucopolipidosis III (célula I)	GNPTG	
Mucopolipina-1 (canal de cationes)	Mucopolipidosis	MCOLN1	
PROTEÍNA DE MEMBRANA 2 ASOCIADA A LISOSOMAS	Danon	LAMP2	
Enfermedad de Niemann-Pick, tipo C1	Niemann Pick tipo C	NPC1	colesterol LDL
palmitoil-proteína tioesterasa-1	Lipofuscinosis ceroide (Enfermedad de Batten)	CLN3	material de almacenamiento de lipopigmentos autofluorescente
lipofuscinosis neuronal 6	Lipofuscinosis ceroide 6	CLN 6	
lipofuscinosis neuronal 8	Lipofuscinosis ceroide 8	CLN 8	
REGULADOR DEL TRANSPORTE LISOSOMAL	Chediak-Higashi	LYST	
Miocilina	Griscelli Tipo 1	MYOC	
proteína 27A asociada a RAS	Griscelli Tipo 2	RAB27A	
Melanofilina	Griscelli Tipo 3	MLPH o MYO5A	
Subunidad β de AP3	Hermansky Pudliak	AP3B1	ceroide

En algunos casos, el donante puede ser un gen endógeno que se ha modificado. Aunque la respuesta de anticuerpo a la terapia de reemplazo enzimático varía respecto a la enzima terapéutica específica en cuestión y con el paciente individual, se ha observado una respuesta inmune significativa en muchos pacientes con LSD que se han tratado con reemplazo enzimático. Además, la relevancia de estos anticuerpos respecto a la eficacia del tratamiento también es variable (véase Katherine Ponder, (2008) *J Clin Invest* 118(8):2686). Así, los métodos y composiciones descritos en la presente pueden comprender el uso de moléculas donantes cuya secuencia se ha alterado por cambios de aminoácidos funcionalmente silenciosos en sitios que se sabe que son epítomos de cebado para las respuestas inmunes endógenas, de manera que el polipéptido producido por dicho donante es menos inmunogénico.

Los pacientes con LSD tienen frecuentemente secuelas neurológicas debido a la carencia de la enzima ausente en el cerebro. Desafortunadamente, frecuentemente es difícil administrar agentes terapéuticos al cerebro a través de la barrera hematoencefálica. Así, los métodos y composiciones de la invención pueden usarse conjuntamente con métodos para incrementar la administración del agente terapéutico en el cerebro. Hay algunos métodos que causan una apertura transitoria de las uniones firmes entre las células de los capilares del cerebro. Los ejemplos incluyen la disrupción osmótica transitoria a través del uso de una administración intracarótida de una disolución hipertónica de manitol, el uso de ultrasonidos focalizados y la administración de un análogo de bradiquinina. Alternativamente, los agentes terapéuticos pueden diseñarse para utilizar receptores o mecanismos de transporte para el transporte específico en el cerebro. Los ejemplos de receptores específicos que pueden usarse incluyen el receptor de la transferrina, el receptor de la insulina o las proteínas relacionadas con el receptor de la lipoproteína de baja densidad 1 y 2 (LRP-1 y LRP-2). Se sabe que LRP interacciona con un rango de proteínas secretadas tales como apoE, tPA, PAI-1 etc, y así la fusión de una secuencia de reconocimiento de una de estas proteínas para LRP puede facilitar el transporte de la enzima en el cerebro, después de la expresión en el hígado de la proteína terapéutica y la secreción en la corriente sanguínea (véase Gabathuler, (2010) *ibid*).

### Administración

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

Los métodos para la administración de nucleasas como se describe en la presente memoria se describen, por ejemplo, en las Patentes U.S. Nos. 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 donantes como se describen en la presente memoria también pueden administrarse usando vectores que contienen secuencias que codifican una o más proteínas de dedo de cinc o TALEN. Puede usarse cualquier sistema de vector incluyendo, pero no limitado a, vectores plasmídicos, vectores retrovirales, vectores lentivirales, vectores de adenovirus, vectores de poxvirus; vectores de herpesvirus y vectores de virus adeno-asociados, etc. Véanse también las Patentes U.S. Nos. 6.534.261; 6.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824. Además, será evidente que cualquiera de estos vectores puede comprender una o más de las secuencias necesarias para el tratamiento. Así, cuando una o más nucleasas y una construcción donante se introducen en la célula, las nucleasas y/o polinucleótido donante pueden estar en el mismo vector o en vectores diferentes. Cuando se usan múltiples vectores, cada vector puede comprender una secuencia que codifica una o múltiples nucleasas y/o construcciones donantes.

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

Los métodos de administración no virales de ácidos nucleicos incluyen electroporación, lipofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados de polimerización o lípido:ácido nucleico, ADN desnudo, viriones artificiales, y captación de ADN aumentada por agentes. La sonoporación usando, p. ej., el sistema de Sonitron 2000 (Rich-Mar) también puede usarse para la administración de ácidos nucleicos.

Los sistemas de administración de ácidos nucleicos adicionales ejemplares incluyen aquellos proporcionados por Amaxa Biosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) y Copernicus Therapeutics Inc, (véase, por ejemplo, US6008336). La lipofección se describe, p. ej., en las Patentes U.S. Nos. 5.049.386; 4.946.787; y 4.897.355) y los reactivos de lipofección se venden comercialmente (p. ej., Transfectam™ y Lipofectin™). Los lípidos catiónicos y neutros que son adecuados para la lipofección eficiente por reconocimiento de receptor de polinucleótidos incluyen aquellos de Felgner, WO 91/17424, WO 91/16024.

La preparación de complejos de lípidos:ácidos nucleicos, incluyendo liposomas dirigidos tales como complejos de inmunolípidos, es muy conocida para un experto en la técnica (véase, p. ej., Crystal, *Science* 270:404-410 (1995); Blaese *et al.*, *Cancer Gene Ther.* 2:291-297 (1995); Behr *et al.*, *Bioconjugate Chem.* 5:382-389 (1994); Remy *et al.*, *Bioconjugate Chem.* 5:647-654 (1994); Gao *et al.*, *Gene Therapy* 2:710-722 (1995); Ahmad *et al.*, *Cancer*

Res. 52:4817-4820 (1992); Pat. U.S. Nos. 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028, y 4.946.787).

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

El uso de sistemas con base viral con ARN o ADN para la administración de ácidos nucleicos que codifican ZFP preparadas por ingeniería tiene la ventaja de procesos altamente evolucionados para dirigir a un virus a células específicas en el cuerpo y transportar la carga viral al núcleo. Los vectores virales pueden administrarse directamente a sujetos (*in vivo*) o pueden usarse para tratar células *in vitro* y las células modificadas se administran a sujetos (*ex vivo*). Los sistemas con base viral convencionales para la administración de ZFP incluyen, pero no están limitados a, vectores de virus retrovirales, lentivirus, adenovirales, adeno-asociados, de vaccinia y herpes simple para la transferencia génica. La integración en el genoma huésped es posible con los métodos de transferencia génica de virus retrovirus, lentivirus, y adeno-asociados, que frecuentemente dan como resultado la expresión a largo plazo del transgén insertado. Adicionalmente, se han observado eficiencias de transducción altas en muchos tipos de células diferentes y tejidos diana.

El tropismo de un retrovirus puede alterarse mediante la incorporación de proteínas de cubierta extrañas, expandiendo la población diana potencial de células diana. Los vectores lentivirales son vectores retrovirales que son capaces de transducir o infectar células que no se están dividiendo y típicamente producen altas titulaciones virales. La selección de un sistema de transferencia génica retroviral depende del tejido diana. Los vectores retrovirales están comprendidos por repeticiones terminales largas que actúan en *cis* con capacidad de empaquetamiento para hasta 6-10 kb de secuencia extraña. Las LTR mínimas que actúan en *cis* son suficientes para la replicación y empaquetamiento de los vectores, que se usan entonces para integrar el gen terapéutico en la célula diana para proporcionar la expresión permanente del transgén. Los vectores retrovirales usados ampliamente incluyen aquellos basados en el virus de la leucemia murina (MuLV), virus de la leucemia del mono gibón (GaLV), el virus de la inmunodeficiencia de simios (SIV), el virus de la inmunodeficiencia humana (VIH), y combinaciones de estos (véase, p. ej., Buchscher *et al.*, *J. Virol.* 66:2731-2739 (1992); Johann *et al.*, *J. Virol.* 66:1635-1640 (1992); Sommerfelt *et al.*, *Virology* 176:58-59 (1990); Wilson *et al.*, *J. Virol.* 63:2374-2378 (1989); Miller *et al.*, *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700).

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

Al menos seis estrategias de vectores virales están disponibles actualmente para transferencia génica en estudios clínicos, que utilizan las estrategias que implican complementación de vectores defectivos por genes insertados en líneas celulares auxiliares para generar el agente de transducción.

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

Los vectores de virus adeno-asociados recombinantes (rAAV) son sistemas de administración génica alternativos prometedores basados en el virus parvovirus adeno-asociado tipo 2 defectivo y no patógeno. Todos los vectores derivan de un plásmido que solo retiene las repeticiones terminales invertidas de AAV de 145 pb que flanquean 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, incluyendo como ejemplo no limitativo, AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV 8.2, AAV9, y AAV rh10 y AAV pseudotipado tal como AAV2/8, AAV2/5 y AAV2/6 también pueden usarse según la presente invención.

Los vectores adenovirales (Ad) recombinantes con replicación deficiente pueden producirse con una alta titulación y pueden infectar fácilmente varios tipos celulares diferentes. La mayor parte de los vectores de adenovirus se preparan por ingeniería de manera que un transgén reemplaza los genes E1a, E1b, y/o E3 de Ad; posteriormente, el vector con replicación defectiva se propaga en células 293 humanas que suministran la función del gen deleciónado en *trans*. Los vectores Ad pueden transducir múltiples tipos de tejidos *in vivo*, incluyendo células diferenciadas que no se dividen tales como las encontradas en el hígado y músculo. Los vectores Ad convencionales tienen una gran capacidad de transporte. Un ejemplo del uso de un vector Ad en un ensayo clínico implicó terapia con polinucleótidos para la inmunización antitumoral con inyección intramuscular (Sterman *et al.*, *Hum. Gene Ther.* 7:1083-9 (1998)). Los ejemplos adicionales del uso de vectores de adenovirus para la transferencia génica en ensayos clínicos incluyen Rosenecker *et al.*, *Infection* 24:1 5-10 (1996); Sterman *et al.*, *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh *et al.*, *Hum. Gene Ther.* 2:205-18 (1995); Alvarez *et al.*, *Hum. Gene Ther.* 5:597-613 (1997); Topf *et al.*, *Gene Ther.* 5:507-513 (1998); Sterman *et al.*, *Hum. Gene Ther.* 7:1083-1089 (1998).

Las células de empaquetamiento se usan para formar partículas de virus que son capaces de infectar una célula huésped. Dichas células incluyen células 293, que empaquetan adenovirus, y células  $\psi$ 2 o células PA317, que empaquetan retrovirus. Los vectores virales usados en terapia génica se generan habitualmente por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula viral. Los vectores contienen típicamente las secuencias virales mínimas requeridas para el empaquetamiento e integración posterior en un huésped (si es aplicable), reemplazándose otras secuencias virales por un casete de expresión que codifica la proteína que se va a expresar. Las funciones virales ausentes se suministran en *trans* por la línea celular de empaquetamiento. Por ejemplo, los vectores de AAV usados en terapia génica solo poseen típicamente secuencias de repetición terminal invertida (ITR) del genoma de AAV que se requieren para el empaquetamiento e integración en el genoma del huésped. El ADN viral se empaqueta en una línea celular, que contiene un plásmido auxiliar que codifica los otros genes de AAV, concretamente *rep* y *cap*, pero carece de secuencias ITR. La línea celular también se infecta con adenovirus como un auxiliar. El virus auxiliar estimula la replicación del vector de AAV y la expresión de los genes de AAV a partir del plásmido auxiliar. El plásmido auxiliar no se empaqueta en cantidades significativas debido a la ausencia de secuencias ITR. La contaminación con adenovirus puede reducirse, p. ej., por tratamiento con calor frente al que los adenovirus son más sensibles que los AAV.

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

Los vectores de terapia génica pueden administrarse *in vivo* por la administración a un paciente individual, típicamente por administración sistémica (p. ej., infusión intravenosa, intraperitoneal, intramuscular, subdérmica, o intracraneal) o aplicación tópica, como se describe más adelante. Alternativamente, los vectores pueden administrarse a células *ex vivo*, tales como a células extraídas de un paciente individual (p. ej., linfocitos, aspirados de médula ósea, biopsia de tejido) o células madre hematopoyéticas de donante universal, seguido de la reimplantación de las células en un paciente, habitualmente después de la selección de las células que han incorporado el vector.

Los vectores (p. ej., retrovirus, adenovirus, liposomas, etc.) que contienen nucleasas y/o construcciones donantes también pueden administrarse directamente en un organismo para la transducción de células *in vivo*. Alternativamente, puede administrarse ADN desnudo. La administración es por cualquiera de las rutas usadas normalmente para introducir una molécula en contacto final con las células de la sangre o tejido incluyendo, pero no limitado a, inyección, infusión, aplicación tópica y electroporación. Los métodos adecuados para administrar dichos ácidos nucleicos están disponibles y son muy conocidos para los expertos en la técnica, y, aunque puede usarse más de una ruta para administrar una composición particular, una ruta particular puede proporcionar frecuentemente una reacción más inmediata y más efectiva que otra ruta.

Los vectores adecuados para la introducción de polinucleótidos descritos en la presente memoria incluyen vectores de lentivirus que no se integran (IDLV). Véase, por ejemplo, Ory *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:11382-11388; Dull *et al.* (1998) *J. Virol.* 72:8463-8471; Zuffery *et al.* (1998) *J. Virol.* 72:9873-9880; Follenzi *et al.* (2000) *Nature Genetics* 25:217-222; Publicación de Patente U.S. No 2009/054985.

Los vehículos farmacéuticamente aceptables están determinados, en parte, por la composición particular que se está administrando, así como por el método particular usado para administrar la composición. De acuerdo con esto, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas disponible, como se describe más adelante (véase, p. ej., *Remington's Pharmaceutical Sciences*, 17ª ed., 1989).

Será evidente que las secuencias que codifican la nucleasa y las construcciones donantes pueden administrarse usando el mismo sistema o sistemas diferentes. Por ejemplo, un polinucleótido donante puede ser portado por un plásmido, mientras la una o más nucleasas pueden ser portadas por un vector de AAV. Además, los diferentes vectores pueden administrarse por la misma ruta o por rutas diferentes (inyección intramuscular, inyección en la vena de la cola, otra inyección intravenosa, administración intraperitoneal y/o inyección intramuscular. Los vectores pueden administrarse simultáneamente o en cualquier orden secuencial.

Las formulaciones tanto para las administraciones *ex vivo* como *in vivo* incluyen suspensiones en líquido o líquidos emulsionados. Los ingredientes activos se mezclan frecuentemente con excipientes que son farmacéuticamente aceptables y compatibles con el ingrediente activo. Los excipientes adecuados incluyen, por ejemplo, agua, disolución salina, dextrosa, glicerol, etanol o semejantes, y combinaciones de estos. Además, la composición puede contener cantidades menores de sustancias auxiliares, tales como, agentes humectantes o emulsionantes, agentes tamponadores de pH, agentes estabilizantes u otros reactivos que potencian la efectividad de la composición farmacéutica.

### Aplicaciones

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 asociado a enfermedad corregido en un locus seguro (p. ej., albúmina) para la expresión de la enzima necesaria y la liberación en la corriente sanguínea. La inserción en una célula secretora, tal como una célula de hígado para la liberación del producto en la corriente sanguínea, es particularmente útil. Los métodos y composiciones de la invención también pueden usarse en cualquier circunstancia en la que se desea suministrar un transgén que codifica uno o más agentes terapéuticos en una célula madre hematopoyética de manera que las células maduras (p. ej., RBC) derivadas de estas células contienen el agente terapéutico. Estas células madre pueden diferenciarse *in vitro* o *in vivo* y pueden derivar de un tipo de célula donante universal que puede usarse para todos los pacientes. Adicionalmente, las células pueden contener una proteína transmembrana para transportar las células en el cuerpo. El tratamiento también puede comprender el uso de células de pacientes que contienen el transgén terapéutico en el que las células se desarrollan *ex vivo* y después se introducen de nuevo en el paciente. Por ejemplo, las HSC que contienen un transgén adecuado pueden insertarse en un paciente a través de un trasplante de médula ósea. Alternativamente, las células madre tales como células madre de músculo o iPSC que se han editado usando el transgén terapéutico también pueden inyectarse en tejido muscular.

Así, esta tecnología puede ser útil en una afección en la que un paciente es deficiente en alguna proteína debido a problemas (p. ej., problemas en el nivel de expresión o problemas con la proteína expresada con un funcionamiento inferior o ausente). Particularmente útil con esta invención es la expresión de transgenes para corregir o restaurar la funcionalidad en trastornos de almacenamiento lisosomal.

Como ejemplos no limitativos, la producción de las proteínas defectivas o ausentes conseguida y usada para tratar la enfermedad por almacenamiento lisosomal. Los donantes de ácidos nucleicos que codifican las proteínas pueden insertarse en un locus seguro (p. ej., albúmina o HPRT) y expresarse bien usando un promotor exógeno o usando el promotor presente en el locus seguro. Alternativamente, pueden usarse donantes para corregir el gene defectivo *in situ*. El transgén deseado puede insertarse en una célula madre CD34+ y devolverse a un paciente durante un trasplante de médula ósea. Finalmente, el donante de ácido nucleico puede insertarse en una célula madre CD34+ en un locus de beta globina de manera que la célula sanguínea roja madura derivada de esta célula tiene una alta concentración del agente biológico codificado por el donante de ácido nucleico. Las RBC que contienen el agente biológico pueden dirigirse al tejido correcto mediante proteínas transmembrana (p. ej., receptor o anticuerpo). Adicionalmente, las RBC pueden sensibilizarse *ex vivo* mediante electrosensibilización para hacerlas más susceptibles a la disrupción después de la exposición a una fuente de energía (véase WO2002007752).

En algunas aplicaciones, un gen endógeno puede inactivarse por el uso de los métodos y composiciones de la invención. Los ejemplos de este aspecto incluyen inactivar un regulador génico aberrante o un gen asociado a enfermedad aberrante. En algunas aplicaciones, un gen endógeno aberrante puede reemplazarse, bien funcionalmente o *in situ*, con una versión de tipo salvaje del gen. El gen insertado también puede alterarse para mejorar la funcionalidad de la proteína expresada o para reducir su inmunogenicidad. En algunas aplicaciones, el gen insertado es una proteína de fusión para incrementar su transporte en un tejido seleccionado tal como el cerebro.

Los siguientes Ejemplos se refieren a realizaciones ejemplares de la presente descripción en las que la nucleasa comprende una nucleasa de dedo de cinc (ZFN) o TALEN. Se apreciará que esto solo es para propósitos de

ejemplificación y que otras nucleasas o sistemas de nucleasas pueden usarse, por ejemplo, endonucleasas homing (meganucleasas) con dominios de unión a ADN preparados por ingeniería y/o fusiones de dominios de unión a ADN de endonucleasas homing naturales o preparadas por ingeniería (meganucleasas) y dominios de escisión heterólogos y/o un sistema CRISPR/Cas que comprende un único ARN guía preparado por ingeniería.

5

### Ejemplos

#### **Ejemplo 1: Diseño, construcción y caracterización general de nucleasas específicas de albúmina**

10 Las proteínas de dedo de cinc se diseñaron e incorporaron en plásmidos, vectores de AAV o adenovirales 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 U.S. No. 6.534.261. La Tabla 3 muestra las hélices de reconocimiento en el dominio de unión a ADN de ZFP específicas de albúmina ejemplares mientras la Tabla 4 muestra los sitios diana para estas ZFP (véanse los Provisionales en copropiedad US 61/537.349 y 61/560.506). Los nucleótidos en el sitio diana que entran en contacto con las hélices de reconocimiento de las ZFP se indican en letras en mayúsculas; los nucleótidos no contactados se indican en minúsculas. Las TALEN específicas de albúmina también se diseñaron y se muestran en las Solicitudes U.S. Nos. 13/624.193 y 13/624.217.

15

**Tabla 3: Diseños de hélice de nucleasas de dedo de cinc específicas para albúmina murina**

Diana	No. SBS	Diseño					
		F1	F2	F3	F4	F5	F6
Intrón 1	30724	TSGSLTR (SEQ ID NO: 1)	RSDALST (SEQ ID NO:2)	QSATRTK (SEQ ID NO:3)	TSGHLSR (SEQ ID NO:4)	QSGNLR (SEQ ID NO:5)	NA
Intrón 1	30725	RSDHLSA (SEQ ID NO:6)	TKSNRTK (SEQ ID NO:7)	DRSNLSR (SEQ ID NO:8)	WRSSLRA (SEQ ID NO:9)	DSSDRKK (SEQ ID NO:10)	NA
Intrón 1	30732	TSGNLTR (SEQ ID NO:11)	DRSTRRQ (SEQ ID NO:12)	TSGSLTR (SEQ ID NO:1)	ERGLAR (SEQ ID NO:13)	TSANLSR (SEQ ID NO:14)	NA
Intrón 1	30733	DRSALAR (SEQ ID NO:15)	RSDHLSA (SEQ ID NO:16)	HRSDRTR (SEQ ID NO:17)	QSGALAR (SEQ ID NO:18)	QSGHLSR (SEQ ID NO:19)	NS
Intrón 13	30759	RSDNLST (SEQ ID NO:20)	DRSALAR (SEQ ID NO:15)	DRSNLSR (SEQ ID NO:8)	DGRNLRH (SEQ ID NO:21)	RSDNLAR (SEQ ID NO:22)	QSNALNR (SEQ ID NO:23)
Intrón 13	30761	DRSNLSR (SEQ ID NO:8)	LKQVLVR (SEQ ID NO:24)	QSGNLR (SEQ ID NO:5)	QSTPLFA (SEQ ID NO:25)	QSGALAR (SEQ ID NO:18)	NA
Intrón 13	30760	DRSNLSR (SEQ ID NO:8)	DGRNLRH (SEQ ID NO:21)	RSDNLAR (SEQ ID NO:22)	QSNALNR (SEQ ID NO:23)	NA	NA
Intrón 13	30767	RSDNLSV (SEQ ID NO:26)	HSNARKT (SEQ ID NO:27)	RSDLSA (SEQ ID NO:28)	QSGNLR (SEQ ID NO:5)	RSDLSV (SEQ ID NO:29)	QSGHLSR (SEQ ID NO:19)
Intrón 13	30768	RSDNLSE (SEQ ID NO:30)	ERANRNS (SEQ ID NO:31)	QSANRTK (SEQ ID NO:32)	ERGLAR (SEQ ID NO:13)	RSDALTQ (SEQ ID NO:33)	NA
Intrón 13	30769	TSGSLTR (SEQ ID NO:1)	DRSNLSR (SEQ ID NO:8)	DGRNLRH (SEQ ID NO:21)	ERGLAR (SEQ ID NO:13)	RSDALTQ (SEQ ID NO:33)	NA
Intrón 12	30872	QSGHLAR (SEQ ID NO:34)	RSDHLTQ (SEQ ID NO:35)	RSDHLSQ (SEQ ID NO:36)	WRSSLVA (SEQ ID NO:37)	RSDVLSE (SEQ ID NO:38)	RNQHRKT (SEQ ID NO:39)
Intrón 12	30873	QSGDLTR (SEQ ID NO:40)	RSDALAR (SEQ ID NO:41)	QSGDLTR (SEQ ID NO:40)	RRDPLIN (SEQ ID NO:42)	RSDNLSV (SEQ ID NO:26)	IRSTLRD (SEQ ID NO:43)
Intrón 12	30876	RSDNLSV (SEQ ID NO:26)	YSSTRNS (SEQ ID NO:44)	RSDHLSA (SEQ ID NO:6)	SYWSRTV (SEQ ID NO:45)	QSSDLR (SEQ ID NO:46)	RTDALRG (SEQ ID NO:47)
Intrón 12	30877	RSDNLST (SEQ ID NO:20)	QKSPINT (SEQ ID NO:48)	TSGNLTR (SEQ ID NO:11)	QAENLKS (SEQ ID NO:49)	QSSDLR (SEQ ID NO:46)	RTDALRG (SEQ ID NO:47)

<b>Intrón 12</b>	30882	RSDNLSV (SEQ ID NO:26)	RRAHLNQ (SEQ ID NO:50)	TSGNLTR (SEQ ID NO:11)	SDTNRFK (SEQ ID NO:51)	RSDNLST (SEQ ID NO:20)	QSGHLRS (SEQ ID NO:19)
<b>Intrón 12</b>	30883	DSSDRKK (SEQ ID NO:10)	DRSALSR (SEQ ID NO:52)	TSSNRKT (SEQ ID NO:53)	QSGALAR (SEQ ID NO:18)	RSDHLRS (SEQ ID NO:54)	NA

**Tabla 4: Sitios diana de ZFN específicas de albúmina murina**

Diana	No. SBS	Sitio diana
<b>Intrón 1</b>	30724	ctGAAGGTgGCAATGGTtctctctgct (SEQ ID NO:55)
<b>Intrón 1</b>	30725	ttTCCTGTAAACGATCGGgaactggcatc (SEQ ID NO:56)
<b>Intrón 1</b>	30732	aaGATGCCaGTTCCCGATcgttacagga (SEQ ID NO:57)
<b>Intrón 1</b>	30733	agGGAGTAGCTTAGGTcagtgaagagaa (SEQ ID NO:58)
<b>Intrón 13</b>	30759	acGTAGAGAACAACATCTAGattggtgg (SEQ ID NO:59)
<b>Intrón 13</b>	30761	ctGTAATAGAAACTGACttacgtagctt (SEQ ID NO:60)
<b>Intrón 13</b>	30760	acGTAGAGAACAACatctagattggtgg (SEQ ID NO:59)
<b>Intrón 13</b>	30767	agGGAATGtGAAATGATTcAGatata (SEQ ID NO:61)
<b>Intrón 13</b>	30768	ccATGGCCTAACAAACAGttatcttctt (SEQ ID NO:62)
<b>Intrón 13</b>	30769	ccATGGCCTAACAAACaGTTatcttctt (SEQ ID NO:62)
<b>Intrón 12</b>	30872	ctTGGCTGTGTAGGAGGGGAgttagcagt (SEQ ID NO:63)
<b>Intrón 12</b>	30873	ttCCTAAGTTGGCAGTGGCAtgcttaat (SEQ ID NO:64)
<b>Intrón 12</b>	30876	ctTTGGCTTTGAGGATTAAGcatgccac (SEQ ID NO:65)
<b>Intrón 12</b>	30877	acTTGGCTcCAAGATTTATAGccttaa (SEQ ID NO:66)
<b>Intrón 12</b>	30882	caGGAAAGTAAGATAGGAAGgaatgtga (SEQ ID NO:67)
<b>Intrón 12</b>	30883	ctGGGGTAAATGTCTCctgcttctt (SEQ ID NO:68)

**Ejemplo 2: Actividad de nucleasas específicas de albúmina murina**

5

Se usaron parejas de ZFN dirigidas al gen de la albúmina murina para ensayar la capacidad de estas ZFN de inducir DSB en un sitio diana específico. La secuencia de aminoácidos de la región de la hélice de reconocimiento de las ZFN indicadas se muestran más adelante en la Tabla 3 y sus sitios diana se muestran en la Tabla 4 (los sitios diana en el ADN se indican en letras mayúsculas; los nucleótidos no contactados se indican en minúsculas).

10

Se usó el ensayo Cel-I (Surveyor™, Transgenomics) como se describe en Perez *et al.*, (2008) *Nat. Biotechnol.* 26: 808-816 y Guschin *et al.*, (2010) *Methods Mol Biol.* 649:247-56), para detectar las modificaciones inducidas por ZFN. En este ensayo, la amplificación por PCR del sitio diana fue seguida por la cuantificación de inserciones y deleciones (indels) usando la enzima de detección de concordancia errónea Cel-I (Yang *et al.*, (2000) *Biochemistry* 39, 3533-3541) que proporcionó una estimación del límite inferior de la frecuencia de DSB. Tres días después de la transfección del vector de expresión de ZFN bien en condiciones estándar (37°C) o usando un choque hipotérmico (30°C, véase la Publicación U.S. en copropiedad No. 20110041195), el ADN genómico se aisló de células Neuro2A transfectadas con la o las ZFN usando el kit DNeasy™ (Qiagen). En estos experimentos, todas las parejas de ZFN fueron parejas con mutación ELD/KKR *FokI* (descrito anteriormente).

20

Una combinación de los resultados del ensayo Cel-I se muestra en la Figura 1, y demuestra que las ZFN mostradas más adelante son capaces de inducir la escisión en sus sitios diana respectivos. El porcentaje de indels mostrado por debajo de los carriles indica la cantidad de genomas que se alteraron por NHEJ después de la escisión. Los datos también demuestran una actividad incrementada cuando el procedimiento de transducción incorpora el choque hipotérmico.

25

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

30

Las ZFN específicas de albúmina de ratón, SBS30724 y SBS30725, que están dirigidas a una secuencia en el intrón 1 se ensayaron *in vivo*. Las ZFN se introdujeron en un vector AAV2/8 como se ha descrito previamente (Li *et al.* (2011) *Nature* 475 (7355): 217). Para facilitar la producción en el sistema de baculovirus, el vector AAV2/8.2 se usó para las preparaciones destinadas a la producción baculoviral. AAV2/8.2 se diferencia del vector AAV2/8 en que una parte de la cápside de AAV8 se ha eliminado y reemplazado por una misma región de la cápside de AAV2 creando una cápside quimérica. La región es el dominio de fosfolipasa A2 en VP1. La producción de las partículas virales que contienen ZFN se hizo bien por preparación usando un sistema HEK293 o un sistema de baculovirus usando métodos estándar en la técnica (Véase Li *et al.*, *ibid.*, véase, p. ej., US6723551). Las partículas de virus se administraron entonces a ratones machos normales (n=6) usando una única dosis de 200 microlitros de 1,0e11 genomas de vector total bien de AAV2/8 o AAV2/8.2 que codifica la ZFN específica de albúmina de ratón. 14 días después de la administración de vectores rAAV, los ratones se sacrificaron, los hígados se recogieron y procesaron para ADN o proteínas totales usando métodos estándar conocidos en la técnica. La detección de las copias del genoma del vector AAV se realizó por PCR cuantitativa. Brevemente, se hicieron cebadores de qPCR específicos para las secuencias bGHPa en el AAV

40

como sigue:

Oligo200 (**Directo**) 5'-GTTGCCAGCCATCTGTTGTTT-3' (SEQ ID NO:69)

5 Oligo201 (**Inverso**) 5'-GACAGTGGGAGTGGCACCTT-3' (SEQ ID NO:70)

Oligo202 (**Sonda**) 5'-CTCCCCGTCCTTCCTTGACC-3'(SEQ ID NO:71)

10 La actividad de escisión de la ZFN se midió usando un ensayo Cel-1 realizado usando un aparato LC-GX (Perkin Elmer), según el protocolo del fabricante. La expresión de las ZFN *in vivo* se midió usando un sistema FLAG-Tag según métodos estándar. Los resultados (Figura 3) demostraron que las ZFN se expresaron y que son activas en la escisión de la diana en el gen de hígado de ratón. En la Figura se muestran los resultados de Cel-I NHEJ para cada ratón en el estudio. El tipo de vector y sus contenidos se muestran por encima de los carriles. La reparación con concordancia errónea después de la escisión por ZFN (% indels indicado) se detectó en casi el 16% en algunos de los ratones.

Las TALEN específicas de albúmina también se ensayaron como se muestra en la Solicitud U.S. No. 13/624.193 y 13/624.217.

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

Las ZFN o TALEN específicas de albúmina murina se usan entonces para introducir el transgén que codifica un producto génico terapéutico en el locus de albúmina para su expresión. 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 donantes, el gen terapéutico estaba flanqueado por secuencias homólogas al gen de la albúmina. En 5' del transgén, las construcciones donantes contenían todas secuencias homólogas al intrón 1 de la albúmina murina, mientras en 3' del gen, las construcciones contienen secuencias homólogas al límite del intrón 1-exón 2 de la albúmina murina.

30 Las construcciones donantes se incorporan entonces en el genoma de un AAV y las partículas de AAV resultantes que contienen los donantes se purifican entonces usando métodos conocidos en la técnica. El material se usa para producir virus AAV que contienen genomas donantes de AAV usando el método de la triple transfección en células HEK 293T y se purifica en gradientes de densidad de CsCl como se ha descrito (véase Ayuso *et al.* (2010) *Gene Ther* 17(4), 503-510). El vector de AAV se diluirá en PBS antes de la inyección. Se usará un rango de 5e9 a 5e13 v.g. partículas del vector donante AAV conjuntamente con 1e9 a 1e12 vg de partículas del vector AAV-ZFN a través de inyecciones en la vena de la cola o intraperitoneal de los virus en ratones de tipo salvaje o con modelo de enfermedad. Los genomas de AAV-ZFN, descritos previamente, que contienen las ZFN específicas de albúmina de ratón se usarán, en combinación con los donantes AAV GLA, GBA, IDUA e IDS. Los ensayos de Cel-I y PCR se realizarán en ADN hepático aislado a varios puntos de tiempo para determinar la frecuencia de inserción de donante inducida por NHEJ y ZFN. También pueden usarse transferencias Southern. Según un protocolo estándar, la cuantificación de GLA, GBA, IDUA e IDS humanos en plasma se realizará usando un kit de ELISA de GLA, GBA, IDUA e IDS humanos o usando un kit de ELISA de Etiqueta FLAG. También se realizan transferencias Western estándar. Los resultados demuestran que estos transgenes correctivos pueden incrementar la expresión de la proteína terapéutica *in vivo*.

45 Por ejemplo, el gen que codifica alfa galactosidasa humana (deficiente en pacientes con la enfermedad de Fabry) se insertó en el locus de la albúmina de ratón. La pareja de ZFN 30724/30725 se usó como se ha descrito anteriormente usando un transgén de alfa galactosidasa. En este experimento, se trataron 3 ratones con un virus AAV2/8 que contenía la pareja de ZFN a una dosis de 3,0e11 genomas virales por ratón y un virus AAV2/8 que contenía el donante huGLa a 1,5e12 genomas virales por ratón. A los animales control se les proporcionó los virus que contenían ZFN solo o el virus donante huGLa solo. Las transferencias Western realizadas en los homogenados de hígado mostraron un incremento en la señal específica de la alfa galactosidasa, indicando que el gen de la alfa galactosidasa se había integrado y se estaba expresando (Figura 4A). Además, se realizó un ELISA en el lisado hepático usando un kit de ensayo para alfa galactosidasa humana (Sino) según el protocolo del fabricante. Los resultados, mostrados en la Figura 4B, demostraron un incremento en la señal en los ratones que se habían tratado tanto con las ZFN como con el donante huGLa.

#### **Ejemplo 5: Diseño de ZFN específicas de albúmina humana.**

60 Para diseñar ZFN con especificidad para el gen de la albúmina humana, la secuencia de ADN del intrón 1 de la albúmina humana se analizó usando métodos descritos previamente para identificar las secuencias diana con el mejor potencial para la unión de ZFN (para detalles, véase la Solicitud de Patente de los Estados Unidos en copropiedad No. 13/624.193 y 13/624.217). La diana y las hélices se muestran en las Tablas 5 y 6.

65 **Tabla 5: Diseños de hélices de nucleasas de dedo de cinc específicas de albúmina humana**

ES 2 813 080 T3

Diana	No. SBS	Diseño					
		F1	F2	F3	F4	F5	F6
Intrón 1	35393	QSSDLSR (SEQ ID NO:46)	LRHNLRA (SEQ ID NO:72)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35394	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35396	QSSDLSR (SEQ ID NO:46)	LKWNLRT (SEQ ID NO:76)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35398	QSSDLSR (SEQ ID NO:46)	LRHNLRA (SEQ ID NO:72)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35399	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35405	QSSDLSR (SEQ ID NO:46)	WKWNLRA (SEQ ID NO:77)	DQSNLRA (SEQ ID NO:73)	RPYTLRL (SEQ ID NO:74)	QSSDLSR (SEQ ID NO:46)	HRSNLNK (SEQ ID NO:75)
Intrón 1	35361	QSGNLAR (SEQ ID NO:5)	LMQNRN Q (SEQ ID NO:78)	LKQHLNE (SEQ ID NO:79)	TSGNLTR (SEQ ID NO:11)	RRYYLRL (SEQ ID NO:80)	N/A
Intrón 1	35364	QSGNLAR (SEQ ID NO:5)	HLGNLKT (SEQ ID NO:81)	LKQHLNE (SEQ ID NO:79)	TSGNLTR (SEQ ID NO:11)	RRDWRR D (SEQ ID NO:82)	N/A
Intrón 1	35370	QSGNLAR (SEQ ID NO:5)	LMQNRN Q (SEQ ID NO:78)	LKQHLNE (SEQ ID NO:79)	TSGNLTR (SEQ ID NO:11)	RRDWRR D (SEQ ID NO:82)	N/A
Intrón 1	35379	QRSNLVR (SEQ ID NO:83)	TSSNRKT (SEQ ID NO:53)	LKHHLTD (SEQ ID NO:84)	TSGNLTR (SEQ ID NO:11)	RRDWRR D (SEQ ID NO:82)	N/A
Intrón 1	35458	DKSYLRP (SEQ ID NO:85)	TSGNLTR (SEQ ID NO:11)	HRSARKR (SEQ ID NO:86)	QSSDLSR (SEQ ID NO:46)	WRSSLKT (SEQ ID NO:87)	N/A
Intrón 1	35480	TSGNLTR (SEQ ID NO:11)	HRSARKR (SEQ ID NO:86)	QSGDLTR (SEQ ID NO:40)	NRHHLKS (SEQ ID NO:88)	N/A	N/A
Intrón 1	35426	QSGDLTR (SEQ ID NO:40)	QSGNLHV (SEQ ID NO:89)	QSAHRKN (SEQ ID NO:90)	STAALSY (SEQ ID NO:91)	TSGLSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:41)
Intrón 1	35428	QSGDLTR (SEQ ID NO:40)	QRSNLNI (SEQ ID NO:93)	QSAHRKN (SEQ ID NO:90)	STAALSY (SEQ ID NO:91)	DRSALSR (SEQ ID NO:52)	RSDALAR (SEQ ID NO:41)
Intrón 1	34931	QRTHLTQ (SEQ ID NO:94)	DRSNLTR (SEQ ID NO:95)	QSGNLAR (SEQ ID NO:5)	QKVNRRAG (SEQ ID NO:96)	N/A	N/A
Intrón 1	33940	RSDNLSV (SEQ ID NO:26)	QNANRIT (SEQ ID NO:97)	DQSNLRA (SEQ ID NO:73)	QSAHRIT (SEQ ID NO:98)	TSGNLTR (SEQ ID NO:11)	HRSARKR (SEQ ID NO:86)

Tabla 6: Sitios diana de ZFN específicas de albúmina humana

Diana	No. SBS	Sitio diana
Intrón 1 (locus 2)	35393	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35394	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35396	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35398	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35399	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35405	ccTATCCATTGCACTATGCTttatttaa (SEQ ID NO:99)
Intrón 1 (locus 2)	35361	ttTGGGATAGTTATGAAAttcaatctca (SEQ ID NO:100)
Intrón 1 (locus 2)	35364	ttTGGGATAGTTATGAAAttcaatctca (SEQ ID NO:100)
Intrón 1 (locus 2)	35370	ttTGGGATAGTTATGAAAttcaatctca (SEQ ID NO:100)
Intrón 1 (locus 2)	35379	ttTGGGATAGTTATGAAAttcaatctca (SEQ ID NO:100)

<b>Intrón 1 (locus 3)</b>	35458	ccTGTGCTGTTGATCTCataaataqaac (SEQ ID NO:101)
<b>Intrón 1 (locus 3)</b>	35480	ccTGTGCTGTTGATctcataaataqaac (SEQ ID NO:101)
<b>Intrón 1 (locus 3)</b>	35426	ttGTGGTTTTTAAAtAAAGCAtagtcca (SEQ ID NO:102)
<b>Intrón 1 (locus 3)</b>	35428	ttGTGGTTTTTAAAtAAAGCAtagtcca (SEQ ID NO:102)
<b>Intrón 1 (locus 4)</b>	34931	acCAAGAAGACAGActaaaatgaaaata (SEQ ID NO:103)
<b>Intrón 1 (locus 4)</b>	33940	ctGTTGATAGACACTAAAAGagtattag (SEQ ID NO:104)

Estas nucleasas se ensayaron en parejas para determinar la pareja con la mayor actividad. Las matrices resultantes de parejas ensayadas se muestran en las Tablas 7 y 8, a continuación, donde la ZFN usada para el lado derecho del dímero se muestra a lo largo de la parte superior de cada matriz, y la ZFN usada para el lado izquierdo del dímero se lista en el lado izquierdo de cada matriz. La actividad resultante, según se determina por el porcentaje de concordancia errónea detectada usando el ensayo Cel-I, se muestra en el cuerpo de ambas matrices:

**Tabla 7: Actividad de ZFN específicas de albúmina humana (% de dianas mutadas)**

	35393	35394	35396	35398	35399	35405	prom.
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
35379	21	21	n.d.	19	19	21	20

**Tabla 8: Actividad de ZFN específicas de albúmina humana (% de dianas mutadas)**

	35458	35480	prom.
35426	4,5	7	3
35428	4,9	6	3,6

(nota: 'n.d.' significa que el ensayo en esa pareja no se hizo)

Así, se han desarrollado nucleasas altamente activas que reconocen secuencias diana en el intrón 1 de la albúmina humana.

#### **Ejemplo 6: Diseño de TALEN específicas de albúmina**

Las TALEN se diseñaron para tomar como diana secuencias en el intrón 1 de la albúmina humana. El reconocimiento de bases se consiguió usando las correspondencias de bases RVD canónicas (el "código TALE: NI para A, HD para C, NN para G (NK en la semirepetición), NG para T). Las TALEN se construyeron como se ha descrito previamente (véase la Publicación de Patente U.S. en copropiedad No. 20110301073). Las dianas para un subconjunto de TALEN estaban conservadas en los genes de la albúmina del mono cynomolgus y el macaco rhesus (para detalles, véase la Solicitud de Patente de los Estados Unidos en copropiedad 13/624.193 y 13/624.217). Las TALEN se construyeron en los núcleos de TALEN "+17" y "+63" como se describe en US20110301073. Las dianas e identificadores numéricos para las TALEN ensayadas se muestran a continuación en la Tabla 9.

**Tabla 9: TALEN específicas de albúmina**

No.SBS	sitio	No. de RVD	SEQ ID NO:
102249	gtTGAAGATTGAATTCAta	15	105
102250	gtTGAAGATTGAATTCATAac	17	106
102251	gtGCAATGGATAGGTCTtt	15	107
102252	atAGTGAATGGATAGGtc	15	108
102253	atTGAATTCATAACTATcc	15	109
102254	atTGAATTCATAACTATCCca	17	110
102255	atAAAGCATAGTGCAATGGat	17	111
102256	atAAAGCATAGTGCAATgg	15	112
102257	ctATGCTTTATTTAAAAac	15	113
102258	ctATGCTTTATTTAAAAcCa	17	114
102259	atTTATGAGATCAACAGCAca	17	115
102260	ctATTTATGAGATCAACAGca	17	116
102261	ttCATTTTAGTCTGTCTTctt	17	117
102262	atTTTAGTCTGTCTTcttg	15	118
102263	ctAATACTCTTTTAGTGTct	16	119
102264	atCTAATACTCTTTTAGTGtc	17	120
102265	atAATTGAACATCATCctg	15	121
102266	atAATTGAACATCATCCTGag	17	122
102267	atATTGGGCTCTGATTCTac	17	123
102268	atATTGGGCTCTGATTCTct	15	124
102269	ttTTTCTGTAGGAATCAga	15	125

102270	ttTTTCTGTAGGAATCAGag	16	126
102271	ttATGCATTTGTTTCAAaa	15	127
102272	atTATGCATTTGTTTCAAaa	15	128

5 Las TALEN se ensayaron entonces en parejas en las células HepG2 para determinar su capacidad de inducir modificaciones en sus dianas cromosómicas endógenas y los resultados mostraron que muchas proteínas que portaban el punto de truncamiento +17 eran activas. De forma similar, muchas TALEN que portaban el punto de truncamiento +63 también eran activas (véase la Tabla 10). Las comparaciones en paralelo con tres conjuntos de ZFN de albúmina no optimizadas (véase la Tabla 10) mostraron que las TALEN y las ZFN tienen actividades que están en el mismo rango aproximado.

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

Pareja de muestras	TALEN C17	% de modificación, C17	TALEN C63	% de modificación, C63	Hueco
1	102251:102249	15	102251:102249	0	12
2	102251:102250	0	102251:102250	0	10
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
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
10	102259:102258	15	102259:102258	0	11
11	102260:102257	15	102260:102257	13	15
12	102260:102258	24	102260:102258	11	13
13	102263:102261	0	102263:102261	16	17
14	102263:102262	0	102263:102262	15	16
15	102264:102261	0	102264:102261	22	18
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
22	102268:102265	4,2	102268:102265	7,9	15
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
ZFN					
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

10 Como se ha indicado previamente (véase la Publicación de Patente U.S. en copropiedad No. 20110301073), las TALEN C17 tienen una mayor actividad cuando el tamaño del hueco entre los dos sitios diana de TALEN es aproximadamente 11-15 pb, mientras las TALEN C63 mantienen la actividad a tamaños de hueco de hasta 18 pb (véase la Tabla 10).

15 **Ejemplo 7: Detección de transgenes donantes de LSD *in vivo***

20 Se construyeron donantes para cuatro transgenes de enfermedad por almacenamiento lisosomal para el propósito de integración en el gen de la albúmina de ratón en el intrón 1. Los transgenes fueron los genes de  $\alpha$ -galactosidasa A (GLA),  $\beta$ -glucosidasa ácida (GBA),  $\alpha$ -L-iduronidasa (IDUA) e Iduronato-2-sulfatasa (IDS), ausentes en las enfermedades de Fabry, de Gaucher, de Hurler y de Hunter, respectivamente. Véase, p. ej., la Figura 8. Los donantes se usaron entonces en estudios *in vivo* para observar la integración de los transgenes. Las ZFN específicas de albúmina murina y los donantes se insertaron todos en virus AAV2/8 como se describe en el Ejemplo 4, y entonces se inyectaron en ratones. En estos experimentos, el virus se formuló para inyección en D-PBS + 35 mM NaCl, 5% glicerol y se congeló antes de la inyección. Los virus que contenían donante y nucleasa se mezclaron conjuntamente antes de la congelación. En el Día 0, la preparación de virus se descongeló y se administró a los ratones por inyección en la vena de la cola donde el volumen total de la inyección fue 200  $\mu$ L. A los tiempos indicados, los ratones se sacrificaron y entonces se retiró el suero, hígado y bazo para análisis de proteínas y ADN por protocolos estándar. Los grupos de dosis se muestran a continuación en la Tabla 11.

30

Tabla 11: Grupos de tratamiento para la integración del transgén LSD

Grupo	Tratamiento	N/grupo/punto de tiempo
1	intrón 1 de Alb murina+ Fabry a relación 1:5; ZFN a 3,0e11, Donante a 1.5e12	3
2	intrón 1 de Alb murina+ donante Hunters a relación 1:5; ZFN a 3,0e11, Donante a 1.5e12	3
3	intrón 1 de Alb murina+ donante Hurlers a relación 1:5; ZFN a 3,0e11, Donante a 1.5e12	3
4	intrón 1 de Alb murina	3
5	Donante de Fabry solo	2
6	Donante de Hunter solo	2
7	Donante de Hurler solo	2

En el día 30, se analizaron los homogenados de hígado por análisis de transferencia Western para determinar la presencia de las proteínas de LSD codificadas por los donantes. Los homogenados de hígado se analizaron por transferencia Western usando metodologías estándar, usando los siguientes anticuerpos primarios: el anticuerpo monoclonal de conejo específico de  $\alpha$ -galactosidasa A (Fabry) se adquirió en Sino Biological, Inc.; el anticuerpo monoclonal de ratón específico de  $\alpha$ -L-Iduronidasa humana (Hurler) se adquirió en R&D Systems; el monoclonal de ratón específico de iduronato 2-Sulfatasa humana (Hunter) se adquirió en R&D Systems. Los resultados (Véase la Figura 5) demuestran expresión, especialmente en los ratones que recibieron tanto el virus que contenía ZFN como el virus que contenía el donante del transgén.

También se investigó la manera de la integración del ADN donante en el locus de la albúmina. Después de la escisión en el locus de la albúmina, el transgén donante podría integrarse potencialmente mediante recombinación dirigida por homología (HDR), utilizando las regiones de homología que flanquean en donante de transgén (Figura 2), o el transgén puede capturarse durante el proceso de unión de extremos no homólogos tendente a error (NHEJ). Los resultados de estas dos alternativas rendirán inserciones de diferentes tamaños cuando se someten a PCR, usando bien el cebador Acc651-SA-inv-sh (5'AAG AAT AAT TCT TTA GTG GTA 3', SEQ ID NO:129) que se une al sitio aceptor de corte y empalme F9 en todas las construcciones donantes de LSD y el cebador mALB-OOF1 (5' ATGAAGTGGGTAACCTTTCTC 3', SEQ ID NO:130) que se une al exón 1 de la albúmina de ratón en 5' del sitio de escisión de ZFN (véase la Figura 6), donde la integración del transgén por HDR dará como resultado la inserción de aproximadamente 680 pb mientras la integración mediante NHEJ dará como resultado la integración de aproximadamente 1.488 pb. Así, el ADN genómico aislado de los homogenados de hígado de los ratones tratados se sometió a PCR en presencia de nucleótidos radiomarcados con  $^{32}$ P y se corrieron en un gel. En las tres integraciones del transgén, se observó la integración a través de ambos mecanismos (véase la Figura 7).

También se diseñaron ADN donantes para incluir una secuencia de etiqueta para el reconocimiento posterior de la proteína usando anticuerpos específicos de la etiqueta. La etiqueta se diseñó para incorporar dos secuencias diferentes que codifican las etiquetas Myc (EQKLISEEDL, SEQ ID NO:131) y Flag (DYKDDDD SEQ ID NO:132). Los esquemas de los diseños de los donantes se muestran en la Figura 8. Los donantes se integraron como se ha descrito anteriormente y todos fueron capaces de integración como se demuestra por PCR (véase la Figura 9).

#### LISTADO DE SECUENCIAS

<110> SANGAMO BIOSCIENCES, INC.

<120> MÉTODOS Y COMPOSICIONES PARA EL TRATAMIENTO DE ENFERMEDADES POR ALMACENAMIENTO LISOSOMAL

<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



**Gln Ser Gly Asn Leu Ala Arg**  
**1 5**

<210> 6

<211> 7

5 <212> PRT

<213> Secuencia Artificial

<220>

10 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 6

**Arg Ser Asp His Leu Ser Ala**  
**1 5**

15 <210> 7

<211> 7

<212> PRT

<213> Secuencia Artificial

20 <220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 7

**Thr Lys Ser Asn Arg Thr Lys**  
**1 5**

25 <210> 8

<211> 7

<212> PRT

30 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

35 <400> 8

**Asp Arg Ser Asn Leu Ser Arg**  
**1 5**

40 <210> 9

<211> 7

<212> PRT

<213> Secuencia Artificial

<220>

45 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 9

**Trp Arg Ser Ser Leu Arg Ala**  
**1 5**

50 <210> 10

<211> 7

<212> PRT

<213> Secuencia Artificial

55 <220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 10

**Asp Ser Ser Asp Arg Lys Lys**  
**1 5**

5 <210> 11  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 11

**Thr Ser Gly Asn Leu Thr Arg**  
**1 5**

15 <210> 12  
 <211> 7  
 <212> PRT  
 20 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

25 <400> 12

**Asp Arg Ser Thr Arg Arg Gln**  
**1 5**

30 <210> 13  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 35 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 13

**Glu Arg Gly Thr Leu Ala Arg**  
**1 5**

40 <210> 14  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 14

50 **Thr Ser Ala Asn Leu Ser Arg**  
**1 5**

<210> 15  
 <211> 7  
 55 <212> PRT  
 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 15

**Asp Arg Ser Ala Leu Ala Arg**

5     **1**                                 **5**

<210> 16

<211> 7

<212> PRT

10   <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

15   <400> 16

**Arg Ser Asp His Leu Ser Glu**

**1**   **5**

<210> 17

20   <211> 7

<212> PRT

<213> Secuencia Artificial

<220>

25   <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 17

**His Arg Ser Asp Arg Thr Arg**

**1**   **5**

30   <210> 18

<211> 7

<212> PRT

35   <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 18

40   **Gln Ser Gly Ala Leu Ala Arg**

**1**   **5**

<210> 19

45   <211> 7

<212> PRT

<213> Secuencia Artificial

<220>

50   <223> Descripción de Secuencia Artificial: Péptido sintético

<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

5 <400> 20

**Arg Ser Asp Asn Leu Ser Thr**  
**1 5**

<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

25 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

30 <400> 22

**Arg Ser Asp Asn Leu Ala Arg**  
**1 5**

35 <210> 23  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

40 <400> 23

**Gln Ser Asn Ala Leu Asn Arg**  
**1 5**

45 <210> 24  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

50 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 24

**Leu Lys Gln Val Leu Val Arg**  
**1 5**

55 <210> 25  
 <211> 7

<212> PRT  
 <213> Secuencia Artificial

<220>  
 5 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 25

**Gln Ser Thr Pro Leu Phe Ala**  
 1 5

10 <210> 26  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 26

20 **Arg Ser Asp Asn Leu Ser Val**  
 1 5

<210> 27  
 <211> 7  
 25 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

30 <400> 27

**His Ser Asn Ala Arg Lys Thr**  
 1 5

35 <210> 28  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 28

**Arg Ser Asp Ser Leu Ser Ala**  
 1 5

45 <210> 29  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

50 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 29

55 **Arg Ser Asp Ser Leu Ser Val**  
 1 5

<210> 30

<211> 7  
 <212> PRT  
 <213> Secuencia Artificial

5 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 30

**Arg Ser Asp Asn Leu Ser Glu**

10 **1 5**

<210> 31  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 31

**Glu Arg Ala Asn Arg Asn Ser**

20 **1 5**

<210> 32  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 32

**Gln Ser Ala Asn Arg Thr Lys**

30 **1 5**

<210> 33  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 33

**Arg Ser Asp Ala Leu Thr Gln**

40 **1 5**

<210> 34  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 34

**Gln Ser Gly His Leu Ala Arg**

50 **1 5**

5 <210> 35  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

10 <400> 35  
**Arg Ser Asp His Leu Thr Gln**  
 1 5

15 <210> 36  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 20 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 36  
**Arg Ser Asp His Leu Ser Gln**  
 1 5

25 <210> 37  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 37  
 35 **Trp Arg Ser Ser Leu Val Ala**  
 1 5

40 <210> 38  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 45 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 38  
**Arg Ser Asp Val Leu Ser Glu**  
 1 5

50 <210> 39  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 55 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 39

**Arg Asn Gln His Arg Lys Thr**  
**1 5**

5 <210> 40  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 40

**Gln Ser Gly Asp Leu Thr Arg**  
**1 5**

15 <210> 41  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 41

**Arg Ser Asp Ala Leu Ala Arg**  
**1 5**

25 <210> 42  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

35 <400> 42

**Arg Arg Asp Pro Leu Ile Asn**  
**1 5**

40 <210> 43  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 43

**Ile Arg Ser Thr Leu Arg Asp**  
**1 5**

50 <210> 44  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

55 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 44

**Tyr Ser Ser Thr Arg Asn Ser**  
**1 5**

5 <210> 45  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 45

**Ser Tyr Trp Ser Arg Thr Val**  
**1 5**

15 <210> 46  
 <211> 7  
 <212> PRT  
 20 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

25 <400> 46

**Gln Ser Ser Asp Leu Ser Arg**  
**1 5**

30 <210> 47  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 47

**Arg Thr Asp Ala Leu Arg Gly**  
**1 5**

40 <210> 48  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 48

50 **Gln Lys Ser Pro Leu Asn Thr**  
**1 5**

55 <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

10 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

15 <400> 50

**Arg Arg Ala His Leu Asn Gln**  
**1 5**

<210> 51

20 <211> 7

<212> PRT

<213> Secuencia Artificial

<220>

25 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 51

**Ser Asp Thr Asn Arg Phe Lys**  
**1 5**

30

<210> 52

<211> 7

<212> PRT

<213> Secuencia Artificial

35

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 52

40

**Asp Arg Ser Ala Leu Ser Arg**  
**1 5**

<210> 53

<211> 7

45 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

50

<400> 53

**Thr Ser Ser Asn Arg Lys Thr**  
**1 5**

55 <210> 54

<211> 7

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido sintético

5

<400> 54

**Arg Ser Asp His Leu Ser Arg**

**1 5**

10 <210> 55  
<211> 28  
<212> ADN  
<213> Mus sp.

15 <400> 55  
**ctgaaggtgg caatggttcc totctgct** 28

<210> 56  
<211> 28  
<212> ADN  
<213> Mus sp.

20 <400> 56  
**tttctgtaa cgatcgggaa ctggcatc** 28

<210> 57  
<211> 28  
<212> ADN  
<213> Mus sp.

25

30 <400> 57  
**aagatgccag ttcccgatcg ttacagga** 28

<210> 58  
<211> 28  
<212> ADN  
<213> Mus sp

35

40 <400> 58  
**agggagttagc ttaggtcagt gaagagaa** 28

<210> 59  
<211> 28  
<212> ADN  
<213> Mus sp.

45 <400> 59  
**acgtagagaa caacatctag attggtgg** 28

<210> 60  
<211> 28  
<212> ADN  
<213> Mus sp.

50

55 <400> 60  
**ctgtaataga aactgactta cgtagctt** 28

<210> 61  
<211> 28  
<212> ADN  
<213> Mus sp.

60

	<400> 61		
	<b>agggaatgtg aatgattca gatata</b>		<b>28</b>
5	<210> 62 <211> 28 <212> ADN <213> Mus sp.		
10	<400> 62		
	<b>ccatggccta acaacagttt atcttctt</b>		<b>28</b>
15	<210> 63 <211> 28 <212> ADN <213> Mus sp.		
20	<400> 63		
	<b>cttggctgtg taggagggga gtagcagt</b>		<b>28</b>
25	<210> 64 <211> 28 <212> ADN <213> Mus sp.		
30	<400> 64		
	<b>ttcctaagtt ggcagtggca tgcttaat</b>		<b>28</b>
35	<210> 65 <211> 28 <212> ADN <213> Mus sp.		
40	<400> 65		
	<b>ctttggcttt gaggattaag catgccac</b>		<b>28</b>
45	<210> 66 <211> 28 <212> ADN <213> Mus sp.		
50	<400> 66		
	<b>acttggctcc aagatttata gccttaaa</b>		<b>28</b>
55	<210> 67 <211> 28 <212> ADN <213> Mus sp.		
60	<400> 67		
	<b>caggaaagta agataggaag gaatgtga</b>		<b>28</b>
65	<210> 68 <211> 28 <212> ADN <213> Mus sp.		
70	<400> 68		
	<b>ctggggtaaa tgtctccttg ctcttctt</b>		<b>28</b>
75	<210> 69 <211> 21 <212> ADN <213> Secuencia Artificial		



**Arg Pro Tyr Thr Leu Arg Leu**  
**1 5**

5 <210> 75  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 75  
**His Arg Ser Asn Leu Asn Lys**  
**1 5**

15 <210> 76  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 76  
**Leu Lys Trp Asn Leu Arg Thr**  
**1 5**

25 <210> 77  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 77  
**Trp Lys Trp Asn Leu Arg Ala**  
**1 5**

35 <210> 78  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 78  
**Leu Met Gln Asn Arg Asn Gln**  
**1 5**

45 <210> 79  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

50 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

55 <400> 79  
**Leu Lys Gln His Leu Asn Glu**  
**1 5**

<210> 80  
 <211> 7  
 <212> PRT  
 5 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 10 <400> 80  
**Arg Arg Tyr Tyr Leu Arg Leu**  
**1 5**

<210> 81  
 <211> 7  
 15 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 20 <400> 81  
**His Leu Gly Asn Leu Lys Thr**  
**1 5**

<210> 82  
 <211> 7  
 25 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 30 <400> 82  
**Arg Arg Asp Trp Arg Arg Asp**  
**1 5**

<210> 83  
 <211> 7  
 35 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 40 <400> 83  
**Gln Arg Ser Asn Leu Val Arg**  
**1 5**

<210> 84  
 <211> 7  
 45 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 50 <400> 84  
**Leu Lys His His Leu Thr Asp**  
**1 5**

55 <210> 85  
 <211> 7

<212> PRT  
 <213> Secuencia Artificial

<220>  
 5 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 85  
**Asp Lys Ser Tyr Leu Arg Pro**  
 1 5

10 <210> 86  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 86  
**His Arg Ser Ala Arg Lys Arg**  
 1 5

20 <210> 87  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

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

30 <210> 88  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

40 <400> 88  
**Asn Arg His His Leu Lys Ser**  
 1 5

<210> 89  
 <211> 7  
 45 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

50 <400> 89  
**Gln Ser Gly Asn Leu His Val**  
 1 5

<210> 90  
 <211> 7  
 55 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

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

<210> 91  
 <211> 7  
 <212> PRT  
 10 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

15 <400> 91  
**Ser Thr Ala Ala Leu Ser Tyr**  
       **1**                             **5**

<210> 92  
 <211> 7  
 20 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

25 <400> 92  
**Thr Ser Gly Ser Leu Ser Arg**  
       **1**                             **5**

<210> 93  
 <211> 7  
 30 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

35 <400> 93  
**Gln Arg Ser Asn Leu Asn Ile**  
       **1**                             **5**

<210> 94  
 <211> 7  
 40 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 94  
**Gln Arg Thr His Leu Thr Gln**  
       **1**                             **5**

50 <210> 95  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

55 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 95  
**Asp Arg Ser Asn Leu Thr Arg**  
**1 5**

5 <210> 96  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

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

15 <210> 97  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

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

25 <210> 98  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

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

35 <210> 99  
 <211> 28  
 <212> ADN  
 <213> Homo sapiens

40 <400> 99  
**cctatccatt gcactatgct ttatttaa** **28**

45 <210> 100  
 <211> 28  
 <212> ADN  
 <213> Homo sapiens

50 <400> 100  
**tttgggatag ttatgaattc aatcttca** **28**

55 <210> 101  
 <211> 28  
 <212> ADN  
 <213> Homo sapiens

<400> 101  
**cctgtgctgt tgatctcata aatagaac** **28**

5	<210> 102 <211> 28 <212> ADN <213> Homo sapiens	
	<400> 102 <b>ttgtggtttt taaataaagc atagtgca</b>	28
10	<210> 103 <211> 28 <212> ADN <213> Homo sapiens	
15	<400> 103 <b>accaagaaga cagactaaaa tgaaaata</b>	28
20	<210> 104 <211> 28 <212> ADN <213> Homo sapiens	
25	<400> 104 <b>ctggtgatag aactataaag agtattag</b>	28
30	<210> 105 <211> 19 <212> ADN <213> Secuencia Artificial	
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
35	<400> 105 <b>gttgaagatt gaattcata</b>	19
40	<210> 106 <211> 21 <212> ADN <213> Secuencia Artificial	
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
45	<400> 106 <b>gttgaagatt gaattcataa c</b>	21
50	<210> 107 <211> 19 <212> ADN <213> Secuencia Artificial	
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
55	<400> 107 <b>gtgcaatgga taggtcttt</b>	19
60	<210> 108 <211> 19 <212> ADN <213> Secuencia Artificial	
	<220>	

	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 108	
5	<b>atagtgc</b> aat ggatagg <b>tc</b>	<b>19</b>
	<210> 109	
	<211> 19	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 109	
15	<b>attgaattca</b> taactat <b>cc</b>	<b>19</b>
	<210> 110	
	<211> 21	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
25	<400> 110	
	<b>attgaattca</b> taactat <b>ccc a</b>	<b>21</b>
	<210> 111	
	<211> 21	
30	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
35	<400> 111	
	<b>ataaagcata</b> gtgcaat <b>gga t</b>	<b>21</b>
	<210> 112	
40	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
45	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 112	
	<b>ataaagcata</b> gtgcaat <b>gg</b>	<b>19</b>
50	<210> 113	
	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
55	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 113	
60	<b>ctatgcttta</b> ttt <b>aaaaac</b>	<b>19</b>
	<210> 114	
	<211> 21	
	<212> ADN	
	<213> Secuencia Artificial	

	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
5	<400> 114 <b>ctatgcttta tttaaaaacc a</b>	<b>21</b>
	<210> 115 <211> 21 <212> ADN <213> Secuencia Artificial	
10		
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
15	<400> 115 <b>atztatgaga tcaacagcac a</b>	<b>21</b>
	<210> 116 <211> 21 <212> ADN <213> Secuencia Artificial	
20		
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
25	<400> 116 <b>ctatttatga gatcaacagc a</b>	<b>21</b>
	<210> 117 <211> 21 <212> ADN <213> Secuencia Artificial	
30		
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
35	<400> 117 <b>ttcatttttag tctgtcttct t</b>	<b>21</b>
40		
	<210> 118 <211> 19 <212> ADN <213> Secuencia Artificial	
45		
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
50	<400> 118 <b>attttagtct gtcttcttg</b>	<b>19</b>
	<210> 119 <211> 20 <212> ADN <213> Secuencia Artificial	
55		
	<220> <223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
60	<400> 119 <b>ctaatactct tttagtgtct</b>	<b>20</b>
	<210> 120 <211> 21	

	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
5	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 120	
	<b>atctaatact cttttagtgt c</b>	<b>21</b>
10	<210> 121	
	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 121	
	<b>ataattgaac atcatcctg</b>	<b>19</b>
20	<210> 122	
	<211> 21	
	<212> ADN	
	<213> Secuencia Artificial	
25	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 122	
30	<b>ataattgaac atcatcctga g</b>	<b>21</b>
	<210> 123	
	<211> 21	
	<212> ADN	
35	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
40	<400> 123	
	<b>atattgggct ctgattccta c</b>	<b>21</b>
	<210> 124	
	<211> 19	
45	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
50	<400> 124	
	<b>atattgggct ctgattcct</b>	<b>19</b>
	<210> 125	
55	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
60	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 125	
	<b>tttttctgta ggaatcaga</b>	<b>19</b>

	<210> 126	
	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
	<400> 126	
10	<b>tttttctgta ggaatcagag</b>	<b>20</b>
	<210> 127	
	<211> 19	
	<212> ADN	
15	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
20	<400> 127	
	<b>ttatgcattt gtttcaaaa</b>	<b>19</b>
	<210> 128	
	<211> 19	
25	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
30	<400> 128	
	<b>attatgcatt tgtttcaaa</b>	<b>19</b>
	<210> 129	
35	<211> 21	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
40	<223> Descripción de Secuencia Artificial: Cebador sintético	
	<400> 129	
	<b>aagaataatt ctttagtggt a</b>	<b>21</b>
45	<210> 130	
	<211> 21	
	<212> ADN	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de Secuencia Artificial: Cebador sintético	
	<400> 130	
55	<b>atgaagtggg taacctttct c</b>	<b>21</b>
	<210> 131	
	<211> 10	
	<212> PRT	
60	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Péptido sintético	
	<400> 131	

**Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu**  
**1 5 10**

5 <210> 132  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

**Asp Tyr Lys Asp Asp Asp Asp**  
**1 5**

15 <210> 133  
 <211> 9  
 <212> PRT  
 <213> Desconocido

20 <220>  
 <223> Descripción de Desconocido: Secuencia de la familia LAGLIDADG Desconocida

<400> 133  
**Leu Ala Gly Leu Ile Asp Ala Asp Gly**  
**1 5**

25 <210> 134  
 <211> 29  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <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**

40 **Ala Asn Asp Ile Leu Asp Tyr Lys Asp Asp Asp Asp Lys**  
**20 25**

45  
 50  
 55  
 60

## REIVINDICACIONES

- 5 **1.** Uno o más transgenes y una o más nucleasas con dedos de zinc (ZFN) para su uso en un método de tratamiento de una enfermedad por almacenamiento lisosomal, el método comprendiendo administrar el uno o más transgenes y el uno o más ZFN a un sujeto con necesidad de ello para generar un célula que produce una proteína que trata la enfermedad por almacenamiento lisosomal, en donde el transgén que codifica la proteína se inserta en un locus de albúmina endógeno de la célula usando el uno o más ZFN, y en donde la enfermedad por almacenamiento lisosomal es la enfermedad de Gaucher, Hunter, Hurler o Fabry.
- 10 **2.** Uno o más transgenes y uno o más polinucleótidos que codifican uno o más ZFN para su uso en un método de tratamiento de una enfermedad por almacenamiento lisosomal, el método comprendiendo administrar el uno o más transgenes y uno o más polinucleótidos que codifican el uno o más ZFN, a un sujeto con necesidad de ello para generar una célula que produce una proteína que trata la enfermedad por almacenamiento lisosomal, en donde el transgén que codifica la proteína se inserta en un locus de albúmina endógeno de la célula usando el uno o más ZFN, y en donde la enfermedad por almacenamiento lisosomal es la enfermedad de Gaucher, Hunter, Hurler o Fabry.
- 15 **3.** El uno o más transgenes y uno o más ZFN para su uso de acuerdo con la reivindicación 1, o el uno o más transgenes y uno o más polinucleótidos para su uso de acuerdo con la reivindicación 2, donde el uno o más transgenes y/o uno o más polinucleótidos que codifican uno o más ZFN se administran usando uno o más vectores.
- 20 **4.** El uno o más transgenes y uno o más ZFN para su uso o el uno o más transgenes y uno o más polinucleótidos para su uso de acuerdo con la reivindicación 3, en donde el uno o más vectores son vectores virales.
- 25 **5.** El uno o más transgenes y uno o más ZFN para su uso o el uno o más transgenes y uno o más polinucleótidos para su uso de acuerdo con la reivindicación 3 o la reivindicación 4, en donde el uno o más transgenes se administran en un vector separado al vector que codifica el ZFN.
- 30 **6.** El uno o más transgenes y uno o más ZFN para el uso o el uno o más transgenes y uno o más polinucleótidos para el uso de acuerdo con la reivindicación 4 o la reivindicación 5, en donde el vector viral es un vector AAV.
- 35 **7.** El uno o más transgenes o uno o más ZFN para el uso o el uno o más transgenes y uno o más polinucleótidos para el uso de acuerdo con cualquiera de las reivindicaciones anteriores, en donde la proteína es una  $\alpha$ -galactosidasa A (GLA), un ácido  $\beta$ -glucosidasa (GBA), una  $\alpha$ -L-iduronidasa (IDUA) y/o una iduronato-2-sulfatasa (IDS).
- 40 **8.** El uno o más transgenes y uno o más ZFN para el uso o el uno o más transgenes y uno o más polinucleótidos para el uso de acuerdo con la reivindicación 7, en donde la proteína es una proteína IDS humana (hIDS).
- 45
- 50
- 55
- 60
- 65

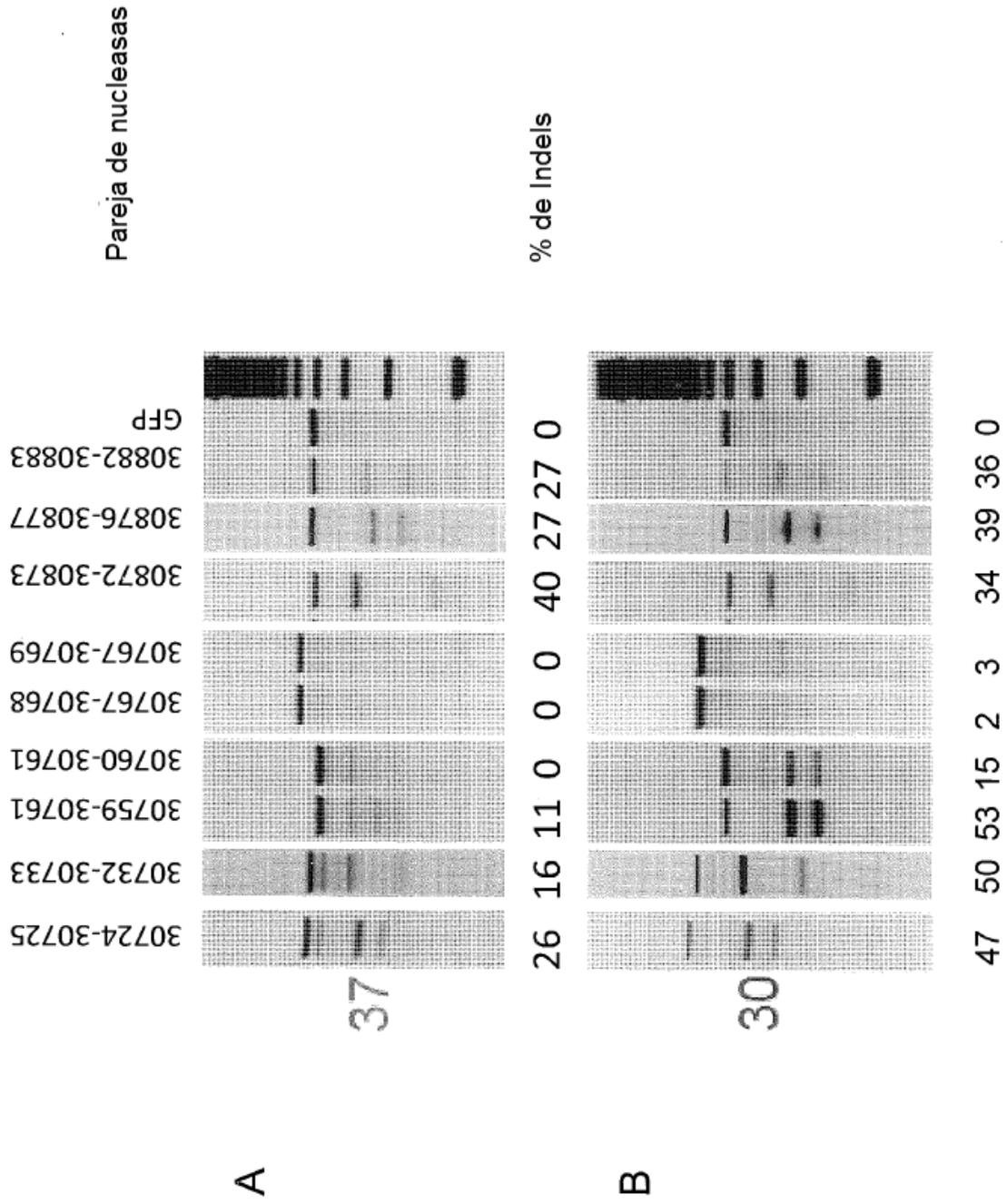
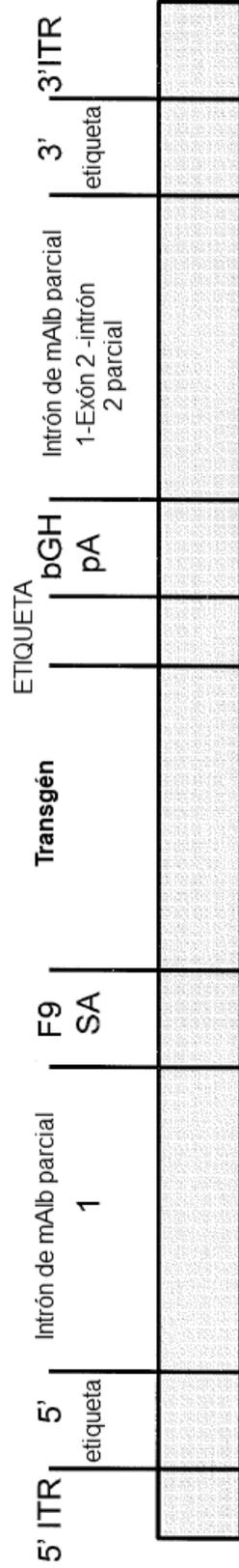


Figura 1

**Donantes "TAG" de Enfermedad de Almacenamiento Lisosomal**



**Trangenes:**

- 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 por baculovirus o 293 dirigidos al intrón 1 del gen de la albúmina son infecciosos y activos en ratones normales

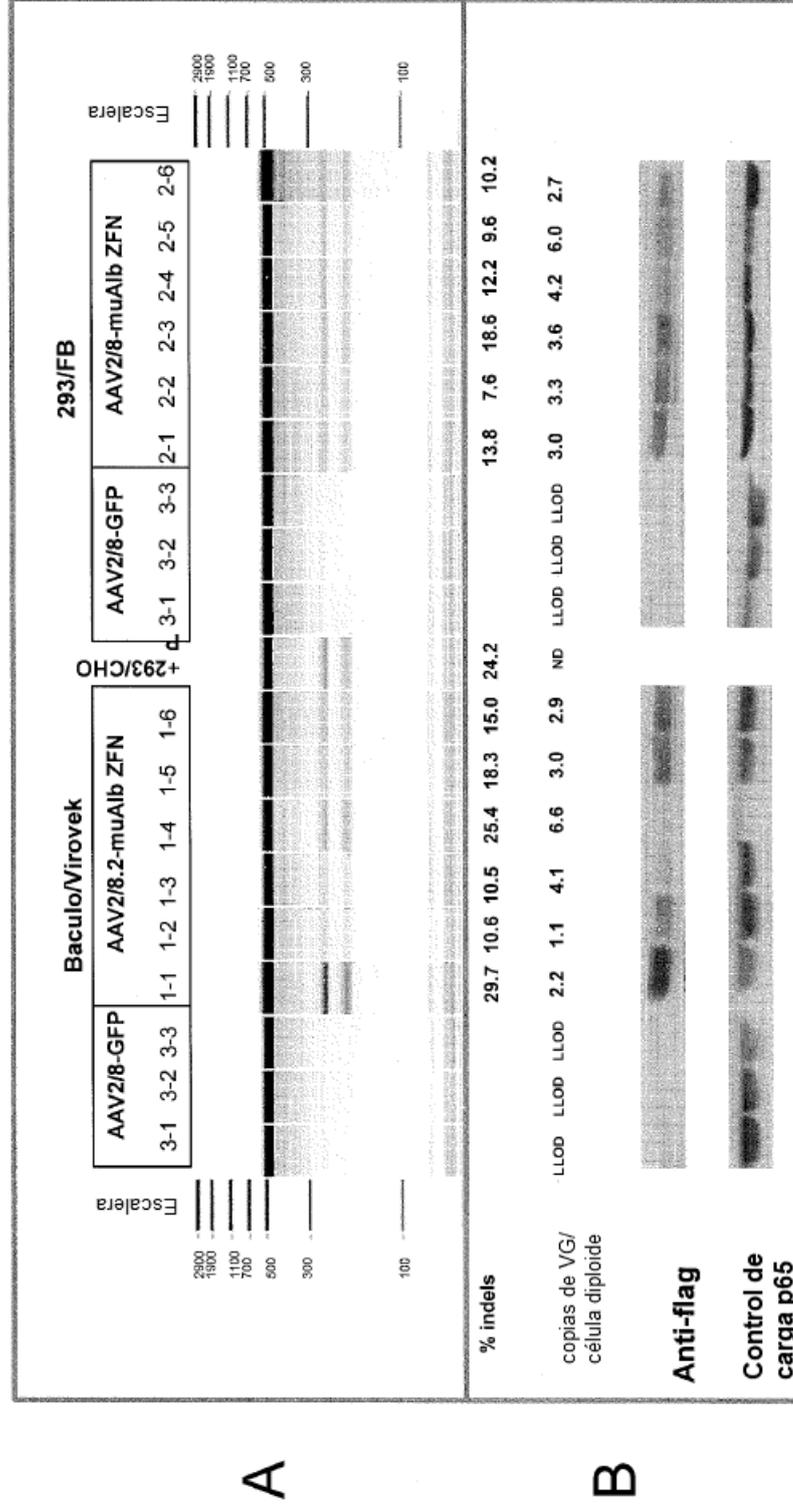
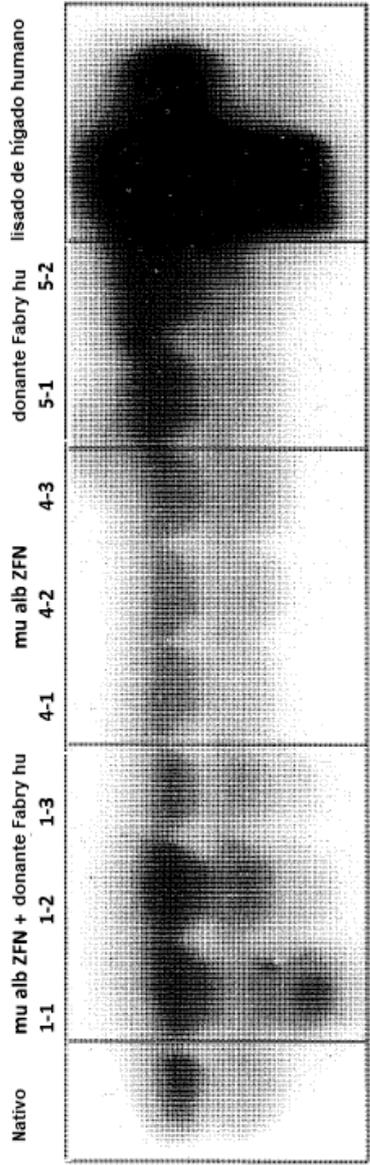
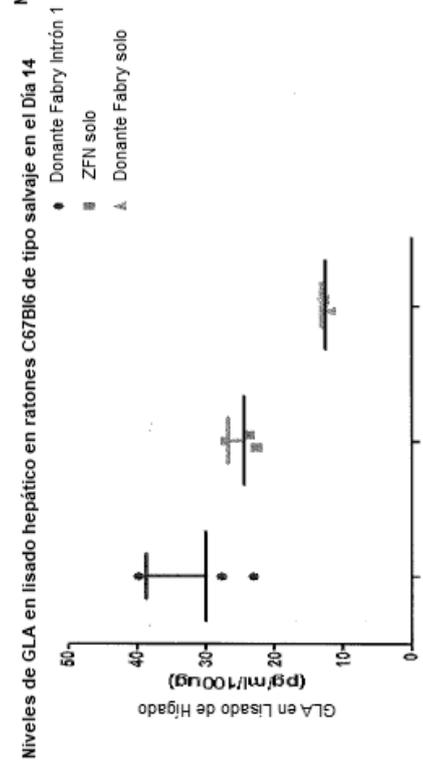
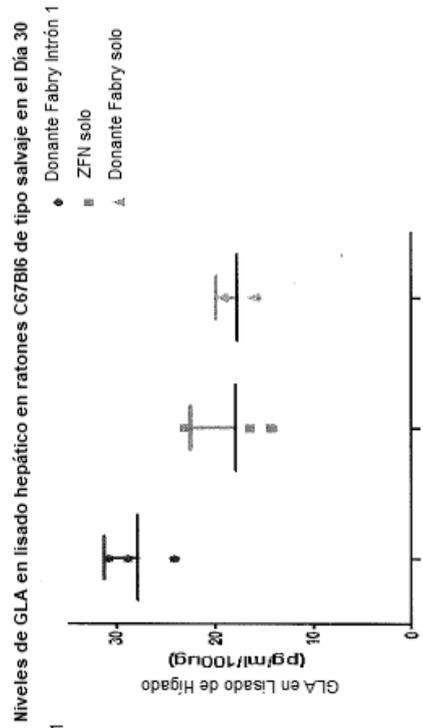


Figura 3



A



B

Figura 4

Figura 5

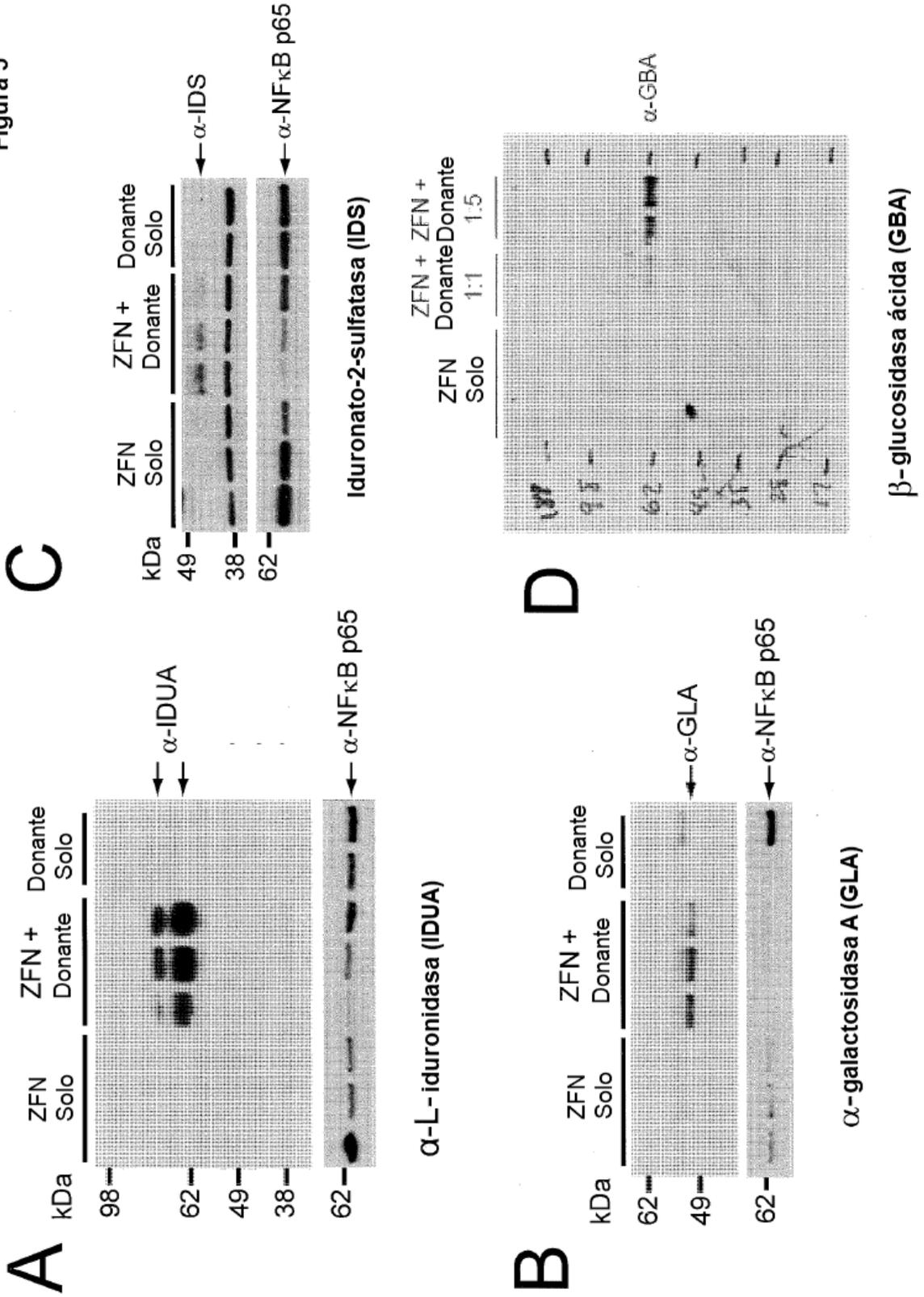
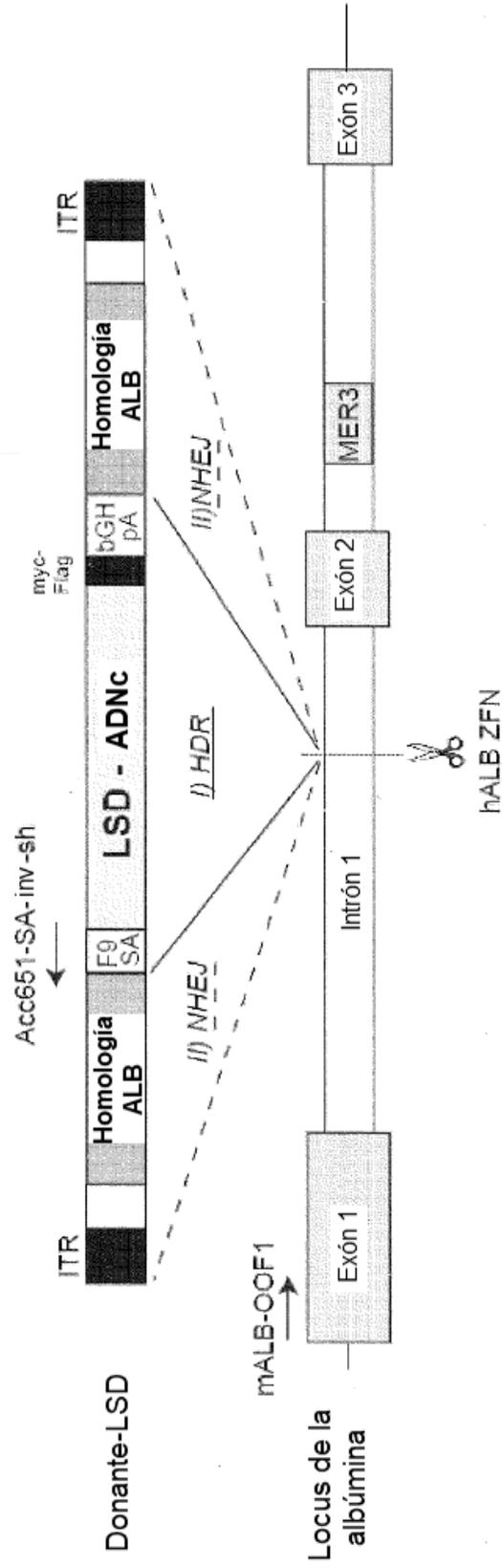


Figura 6



- 1) Integración del donante por HDR (usando brazos de homología)
- 2) Integración del donante por NHEJ (captura de extremos)

Figura 7

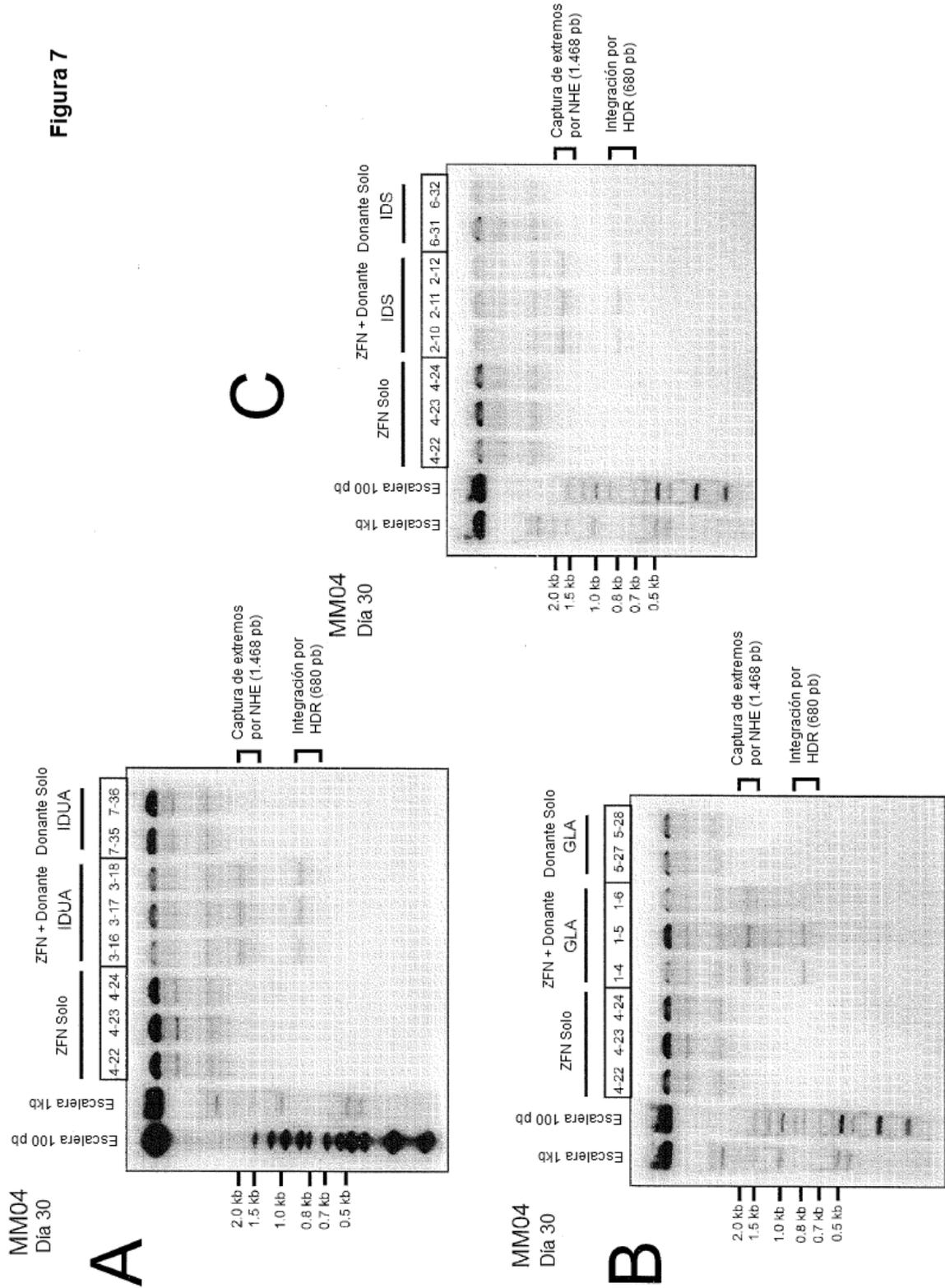
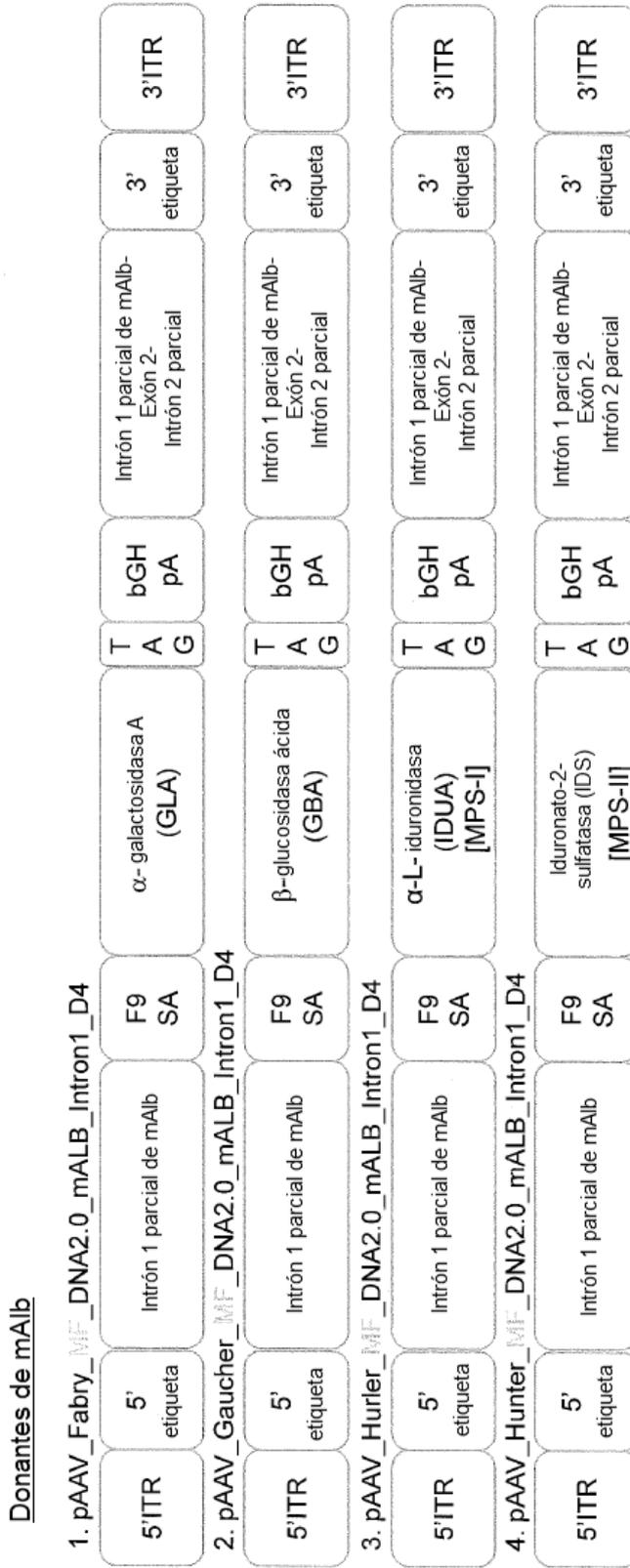


Figura 8



**TAG: MYC-FLAG (MF)**

**ADNc Último AA-RTRPLEQKLISEEDLAANDILDYKDDDK-PARADA**

**MYC**

**FLAG**

Figura 9

