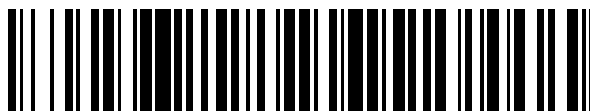


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 750 550**

51 Int. Cl.:

C12N 9/16 (2006.01)

A61K 31/7088 (2006.01)

A61K 31/713 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.02.2014** **PCT/US2014/019322**

87 Fecha y número de publicación internacional: **04.09.2014** **WO14134412**

96 Fecha de presentación y número de la solicitud europea: **28.02.2014** **E 14756449 (6)**

97 Fecha y número de publicación de la concesión europea: **24.07.2019** **EP 2961262**

54 Título: **Corrección de gen a base de TALEN**

30 Prioridad:

01.03.2013 US 201361771735 P

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

26.03.2020

73 Titular/es:

**REGENTS OF THE UNIVERSITY OF MINNESOTA
(100.0%)**

**600 McNamara Alumni Center 200 Oak Street S.E.
Minneapolis, MN 55455-2020, US**

72 Inventor/es:

**OSBORN, MARK, J.;
TOLAR, JAKUB;
BLAZAR, BRUCE y
VOYTAS, DANIEL, F.**

74 Agente/Representante:

RIZZO , Sergio

ES 2 750 550 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

Corrección de gen a base de TALEN

REFERENCIA CRUZADA A APLICACIÓN RELACIONADA

Antecedentes de la invención

- 5 **[0001]** La epidermólisis bullosa (EB) es un grupo de afecciones genéticas que provocan que la piel sea muy frágil y se formen ampollas fácilmente. Las ampollas y erosiones cutáneas se forman como respuesta a una lesión menor o fricción, como un roce o arañazo. La epidermólisis bullosa distrófica recesiva (EBDR o RDEB por sus siglas en inglés), la forma más grave y habitual de la enfermedad, se caracteriza por la formación de gran cantidad de ampollas y cicatrices en la piel y en las membranas mucosas. Las mutaciones de COL7A1 asociadas a la EBDR afectan a la capacidad del colágeno 7 para conectar la epidermis y la dermis; y la posterior separación de la epidermis y la dermis como resultado de la fricción o las lesiones leves provoca la formación de ampollas grave y la amplia formación de cicatrices en la piel asociadas a la EBDR. Las personas con EBDR muestran formación de ampollas cutáneas incurables, a menudo letales, y se enfrentan a un mayor riesgo de sufrir carcinoma agresivo de células escamosas. Las terapias de aumento de genes resultan prometedoras, aunque se corre el riesgo de sufrir mutagénesis insercional. Las herramientas actuales de terapia génica (p. ej., adición génica mediada por virus) dependen del aporte de copias funcionales de un gen terapéutico que se integran de manera aleatoria o parcialmente aleatoria en el genoma. Las consecuencias de la integración aleatoria son la alteración del locus en el que se sitúa la carga y una potencial inactivación o desregulación génica (efectos inespecíficos). Estas pueden derivar en efectos secundarios que amenacen la vida del paciente. Por consiguiente, se describen en el presente documento nucleasas modificadas efectoras de tipo activador de la transcripción (TALEN) para la edición precisa del genoma en células de pacientes con, por ejemplo, EBDR y otras afecciones genéticas.

Sumario de la invención

- 25 **[0002]** La invención da a conocer la composición de la reivindicación 1, la composición de la reivindicación 2 para el uso de las reivindicaciones 2-5, la proteína de las reivindicaciones 6 y 7, los vectores de la reivindicación 8 y 9, el ácido nucleico de la reivindicación 10 y la proteína de la reivindicación 11.

- 30 **[0003]** La presente invención supera los efectos inespecíficos proporcionando una corrección específica de sitio de la mutación. La corrección de la mutación se puede conseguir mediante la transformación o transfección de una célula. La célula se puede seleccionar de entre el grupo que consiste en un fibroblasto, queratinocito, célula madre pluripotente inducible, célula madre hematopoyética, célula madre mesenquimal, célula madre embrionaria, célula de progenie hematopoyética, célula T, célula B, célula glial, célula neural, célula progenitora neuroglial, célula madre neuroglial, célula muscular, célula pulmonar, célula pancreática, célula hepática y una célula del sistema reticuloendotelial.

- 35 **[0004]** Un método para tratar una enfermedad o trastorno genético provocado por una mutación genética puede comprender la puesta en contacto de una célula con uno o varios ácidos nucleicos que codifican una TALEN y una secuencia donadora de ácido nucleico, donde la proteína TALEN se expresa en la célula e induce una rotura de ADN bicatenario específica de sitio en un gen diana, donde la secuencia donadora es una plantilla para la reparación de ADN, que da como resultado una corrección de la mutación genética y proporciona una expresión génica correcta para tratar la enfermedad o trastorno genético. La célula puede ser un fibroblasto, queratinocito, célula madre pluripotente, hematopoyética, mesenquimal o embrionaria inducible, célula de progenie hematopoyética (como una célula T o una célula B), célula glial y neural, célula madre y progenitora neuroglial, célula muscular, célula pulmonar, célula pancreática y/o hepática y/o una célula del sistema reticuloendotelial. En el presente documento, se da a conocer el uso de uno o varios ácidos nucleicos para tratar una enfermedad o trastorno genético provocado por una mutación genética, donde dicho uno o varios ácidos nucleicos codifican una nucleasa efectora de tipo activador de la transcripción (TALEN) y una secuencia donadora de ácido nucleico, donde, cuando la proteína TALEN se expresa en una célula e induce una rotura de ADN bicatenario específica de sitio en un gen diana, y donde la secuencia donadora es una plantilla para la reparación de ADN, da como resultado una corrección de la mutación genética y proporciona una expresión génica correcta con el fin de tratar la enfermedad o trastorno genético.

- 50 **[0005]** La TALEN puede ser una TALEN izquierda y comprender, además, una TALEN derecha que coopere con la TALEN izquierda para provocar la rotura de doble cadena en el gen diana. El ácido nucleico que codifica la TALEN y/o la secuencia donadora de ácido nucleico puede formar parte de un vector o plásmido. La TALEN puede incluir un espaciador (p. ej., la secuencia espaciadora contiene de 12 a 30 nucleótidos de longitud).

- 55 **[0006]** El gen diana puede ser un gen con una alteración/mutación genética. El gen diana es COL7A1 (uno con una mutación que provoque, por ejemplo, la expresión aberrante de la proteína).

[0007] La enfermedad genética puede ser epidermólisis bullosa, osteogénesis imperfecta, disqueratosis congénita, las mucopolisacaridosis, distrofia muscular, fibrosis quística (CFTR), anemia de Fanconi, las

esfingolipidosis, las lipofuscinosis, adrenoleucodistrofia, inmunodeficiencia combinada grave, anemia de células falciformes o talasemia.

[0008] Un método para tratar una enfermedad o trastorno genético provocado por una mutación genética puede comprender a) introducir en una célula (i) un primer ácido nucleico que codifica un primer monómero de endonucleasa efectora de tipo activador de la transcripción (TAL), (ii) un segundo ácido nucleico que codifica un segundo monómero de endonucleasa efectora TAL, y (iii) una secuencia donadora, donde cada uno de dichos primer y segundo monómeros de endonucleasa efectora TAL comprende una pluralidad de secuencias de repetición efectoras TAL y un dominio de endonucleasa FokI, donde cada una de dicha pluralidad de secuencias de repetición efectoras TAL comprende un diresiduo variable de repetición, donde dicho primer monómero de endonucleasa efectora TAL comprende la capacidad para fijarse a una primera secuencia de medio sitio de un ADN diana en el interior de dicha célula y comprende la capacidad para escindir dicho ADN diana cuando dicho segundo monómero de endonucleasa efectora TAL está fijado a una segunda secuencia de medio sitio de dicho ADN diana, donde dicho ADN diana comprende dicha primera secuencia de medio sitio y dicha segunda secuencia de medio sitio separadas por una secuencia espaciadora, y donde dicho primer y segundo medio sitio poseen la misma secuencia de nucleótidos o distintas secuencias de nucleótidos, donde dicha secuencia donadora comprende homología con la diana al menos en los extremos 5' y 3' de la secuencia diana y la alteración genética preseleccionada y es una plantilla para la reparación de ADN, que da como resultado una corrección de la mutación genética; y (b) cultivar la célula en condiciones en las que se expresen el primer y el segundo monómero de endonucleasa efectora TAL, con el fin de corregir la mutación y restablecer la expresión génica correcta. Cada uno del primer y el segundo ácido nucleico puede comprender un espaciador (distinto de la secuencia espaciadora). La secuencia espaciadora se puede localizar entre la pluralidad de secuencias de repetición efectoras TAL y el dominio de endonucleasa FokI. La secuencia espaciadora puede contener de 12 a 30 nucleótidos. En el presente documento se da a conocer el uso de uno o varios ácidos nucleicos para tratar una enfermedad o trastorno genético provocado por una mutación genética, donde (i) un primer ácido nucleico codifica un primer monómero de endonucleasa efectora de tipo activador de la transcripción (TAL), (ii) un segundo ácido nucleico codifica un segundo monómero de endonucleasa efectora TAL, y (iii) una secuencia donadora, donde cada uno de dichos primer y segundo monómeros de endonucleasa efectora TAL comprende una pluralidad de secuencias de repetición efectoras TAL y un dominio de endonucleasa FokI, donde cada una de dicha pluralidad de secuencias de repetición efectoras TAL comprende un diresiduo variable de repetición, donde dicho primer monómero de endonucleasa efectora TAL comprende la capacidad para unirse a una primera secuencia de medio sitio de un ADN diana en el interior de dicha célula y comprende la capacidad para escindir dicho ADN diana cuando dicho segundo monómero de endonucleasa efectora TAL se fija a una segunda secuencia de medio sitio de dicho ADN diana, donde dicho ADN diana comprende dicha primera secuencia de medio sitio y dicha segunda secuencia de medio sitio separadas por una secuencia espaciadora, y donde dichos primer y segundo medios sitios poseen la misma secuencia de nucleótidos o distintas secuencias de nucleótidos, donde dicha secuencia donadora comprende homología con la diana al menos en los extremos 5' y 3' de la secuencia diana y la alteración genética preseleccionada y es una plantilla para la reparación de ADN, que da como resultado una corrección de la mutación genética; y donde (b) cultivar la célula en condiciones en las que se expresen el primer y el segundo monómero de endonucleasa efectora TAL, para corregir la mutación y restablecer la expresión génica correcta.

[0009] También se da a conocer un ácido nucleico que comprende una secuencia donadora, donde la secuencia donadora es una plantilla para la reparación de ADN específica de sitio, que da como resultado una corrección de una mutación genética, donde la secuencia donadora comprende homología con al menos los extremos 5' y 3' de la secuencia diana, donde una parte de la secuencia donadora comprende una secuencia de reparación para corregir la secuencia diana para su uso junto con una proteína TALEN. El donador puede comprender la SEQ ID NO:22. La diana es COL7A1 (un gen con una mutación). En una forma de realización, los extremos 5' y 3' del donador presentan, cada uno, al menos 100 bases de identidad de secuencia con la diana.

[0010] El ácido nucleico puede comprender la SEQ ID NO:29 o 30. También se dan a conocer las proteínas codificadas o expresadas por los ácidos nucleicos de TALEN.

[0011] Se da a conocer un vector o plásmido que comprende una secuencia donadora, donde la secuencia donadora es una plantilla para la reparación de ADN específica de sitio, que da como resultado una corrección de una mutación genética, donde la secuencia donadora comprende homología con al menos los extremos 5' y 3' de la secuencia diana, donde una parte de la secuencia donadora comprende una secuencia de reparación para corregir la secuencia diana para su uso junto con una proteína TALEN. El donador puede comprender la SEQ ID NO:22. La diana es COL7A1 (con una mutación). Los extremos 5' y 3' del donador pueden presentar, cada uno, al menos 100 bases de identidad de secuencia con la diana. Se da a conocer un vector o plásmido que comprende una o varias de las SEQ ID NO: 22, 31, 28, 29 o 30. También se da a conocer una célula huésped aislada que comprende una o varias de las SEQ ID NO: 22, 31, 28, 29 o 30 o las proteínas expresadas a partir de tales secuencias. En el presente documento se describe una línea celular transfectada que comprende las SEQ ID NO: 22, 31, 28, 29 o 30 o las proteínas expresadas a partir de tales secuencias.

[0012] Un método para tratar una enfermedad o trastorno genético provocado por una mutación genética puede comprender la puesta en contacto de una célula con un ácido nucleico que codifica una TALEN, donde la TALEN

corrige la mutación y, por ejemplo, restablece la expresión de genes correcta, o aumenta la expresión de genes. La célula puede ser un fibroblasto. La TALEN puede ser una TALEN izquierda y comprender, además, una TALEN derecha que coopere con la TALEN izquierda para realizar un corte de doble cadena en un ADN. La molécula de ácido nucleico puede ser un vector. La molécula de ácido nucleico puede ser un plásmido. La TALEN puede incluir un espaciador, por ejemplo, de 12 a 30 nucleótidos de longitud. En una forma de realización, la enfermedad genética es epidermólisis bullosa.

[0013] Un método para tratar una enfermedad o trastorno genético provocado por una mutación genética puede comprender a) introducir en una célula (i) un primer ácido nucleico que codifica un primer monómero de endonucleasa efectora de tipo activador de la transcripción (TAL) y (ii) un segundo ácido nucleico que codifica un segundo monómero de endonucleasa efectora TAL, donde cada uno de dichos primer y segundo monómeros de endonucleasa efectora TAL comprende una pluralidad de secuencias de repetición efectoras TAL y un dominio de endonucleasa FokI, donde cada una de dicha pluralidad de secuencias de repetición efectoras TAL comprende un diresiduo variable de repetición, donde dicho primer monómero de endonucleasa efectora TAL comprende la capacidad para fijarse a una primera secuencia de medio sitio de un ADN diana en el interior de dicha célula y comprende la capacidad para escindir dicho ADN diana cuando dicho segundo monómero de endonucleasa efectora TAL se fija a una segunda secuencia de medio sitio de dicho ADN diana, donde dicho ADN diana comprende dicha primera secuencia de medio sitio y dicha segunda secuencia de medio sitio separadas por una secuencia espaciadora, y donde dicho primer y segundo medio sitio poseen la misma secuencia de nucleótidos o distintas secuencias de nucleótidos; y (b) cultivar la célula en condiciones en las que se expresen el primer y el segundo monómero de endonucleasa efectora TAL, con el fin de corregir la mutación y restablecer la expresión correcta de genes.

[0014] Se da a conocer un ácido nucleico que codifica una TALEN y una secuencia donadora de ácido nucleico, donde, cuando la proteína TALEN se expresa en una célula, induce una rotura de ADN bicatenario específica de sitio en un gen diana, y, además, donde la secuencia donadora es una plantilla para la reparación de ADN, que da como resultado una corrección de la mutación genética y proporciona una expresión génica correcta para tratar la enfermedad o trastorno genético. La descripción proporciona el ácido nucleico, donde la célula es un fibroblasto, queratinocito, célula madre pluripotente, hematopoyética, mesenquimal o embrionaria inducible, célula de progenie hematopoyética (como una célula T o una célula B), célula glial y neural, célula madre y progenitora neuroglial, célula muscular, célula pulmonar, célula pancreática y/o célula hepática y/o una célula del sistema reticuloendotelial. La descripción proporciona el ácido nucleico, donde la TALEN es una TALEN izquierda y que comprende, además, una TALEN derecha que coopera con la TALEN izquierda para provocar la rotura de doble cadena en el gen diana. La TALEN derecha puede ser codificada por el ácido nucleico o un segundo ácido nucleico. La TALEN izquierda y la TALEN derecha pueden comprender una pluralidad de secuencias de repetición efectoras TAL y un dominio de endonucleasa. Cada una de la TALEN derecha e izquierda puede comprender un espaciador (distinto de la secuencia espaciadora). La secuencia espaciadora se puede localizar entre la pluralidad de secuencias de repetición efectoras TAL y el dominio de endonucleasa. La secuencia espaciadora puede ser codificada por una secuencia de 12 a 30 nucleótidos. La descripción proporciona el ácido nucleico, donde dicho ácido nucleico que codifica la TALEN y/o la secuencia donadora de ácido nucleico forma parte de un vector o plásmido. La descripción proporciona el ácido nucleico, donde el gen diana es un gen con una alteración/mutación genética. La invención proporciona el ácido nucleico, donde el gen diana es COL7A1. La descripción proporciona el ácido nucleico, donde la TALEN incluye un espaciador. La descripción proporciona el ácido nucleico donde la secuencia espaciadora contiene de 12 a 30 nucleótidos de longitud. La descripción proporciona el ácido nucleico, donde la enfermedad genética es epidermólisis bullosa, osteogénesis imperfecta, disqueratosis congénita, las mucopolisacaridosis, distrofia muscular, fibrosis quística (CFTR), anemia de Fanconi, las esfingolipidosis, las lipofuscinosis, adrenoleucodistrofia, inmunodeficiencia combinada grave, anemia de células falciformes o talasemia. La descripción proporciona el ácido nucleico, donde la enfermedad genética es epidermólisis bullosa. La descripción proporciona al menos un ácido nucleico que comprende (i) un primer ácido nucleico que codifica un primer monómero de endonucleasa efectora de tipo activador de la transcripción (TAL), (ii) un segundo ácido nucleico que codifica un segundo monómero de endonucleasa efectora TAL, y (iii) una secuencia donadora, donde cada uno de dichos primer y segundo monómeros de endonucleasa efectora TAL comprende una pluralidad de secuencias de repetición efectoras TAL y un dominio de endonucleasa FokI, donde cada una de dicha pluralidad de secuencias de repetición efectoras TAL comprende un diresiduo variable de repetición, donde dicho primer monómero de endonucleasa efectora TAL comprende la capacidad para fijarse a una primera secuencia de medio sitio de un ADN diana en el interior de dicha célula y comprende la capacidad para escindir dicho ADN diana cuando dicho segundo monómero de endonucleasa efectora TAL está unida a una segunda secuencia de medio sitio de dicho ADN diana, donde dicho ADN diana comprende dicha primera secuencia de medio sitio y dicha segunda secuencia de medio sitio separadas por una secuencia espaciadora, y donde dicho primer y segundo medio sitio poseen la misma secuencia de nucleótidos o distintas secuencias de nucleótidos, donde dicha secuencia donadora comprende homología con la diana al menos en los extremos 5' y 3' de la secuencia diana y la alteración genética preseleccionada y es una plantilla para la reparación de ADN, que da como resultado una corrección de la mutación genética; y (b) cultivar la célula en condiciones en las que se expresen el primer y el segundo monómero de endonucleasa efectora TAL, con el fin de corregir la mutación y restablecer la expresión génica correcta. La descripción proporciona una proteína codificada o expresada por el ácido nucleico. La descripción

proporciona un vector o plásmido que comprende el ácido nucleico. La descripción proporciona una célula huésped aislada que comprende el ácido nucleico.

[0015] La descripción contempla el uso de los ácidos nucleicos, vectores, células huésped y proteínas de la descripción para tratar una enfermedad o trastorno genético provocado por una mutación genética.

5 Breve descripción de los dibujos

[0016]

Figuras 1A-F. Especificidad de TALEN, estructura de nucleasa y modificación de gen COL7A1. (a) Sitio diana de COL7A1 en el cromosoma 3 y unión a matriz TALEN. Se muestra un esquema del cromosoma humano tres y la región del exón 13 específica. Las flechas indican los conjuntos de cebadores utilizados para posteriores análisis, y la línea con el cuadro gris moteado es el donador utilizado en (f). (b) Sitio diana de COL7A1 y los componentes básicos del complejo de nucleasa. La TALEN está constituida por una delección en el N-terminal de 152 residuos de TALE de *Xanthomonas*, seguida del dominio de repetición, y una subregión del C-terminal +63 fusionada con el dominio catalítico de la nucleasa FokI. (SEQ ID NO: 33; SEQ ID NO: 34) (c) Reconocimiento de base del residuo variable de repetición (RVD). Los RVD NN, NI, HD y NG (que se fijan, respectivamente, a guanina, adenina, citosina y timina) están codificados para la correspondiente matriz completa en 1b. (d) Esquema de rotura de ADN de doble cadena (DSB) generada por TALEN (rayo) y posibles mecanismos de reparación celular utilizados para la reparación de la rotura. (SEQ ID NO: 35; SEQ ID NO: 36). (e) Evaluación de unión de extremos no homólogos con probabilidad de error mediante secuenciación de Sanger de células tratadas con TALEN. Se llevó a cabo una PCR de ciclo limitante, seguida de una clonación al azar; se secuenciaron 75 clones, mostrando 64 de ellos un 100 % de alineamiento con la base de datos de genomas y 11 mostrando delecciones inducidas por unión de extremos no homólogos (NHEJ) que se representan con rayas. Los sitios diana de TALEN izquierdo y derecho se indican en mayúsculas y en negrita, y la secuencia espaciadora se representa en letras minúsculas. Las bases totales delecionadas se representan a la derecha y se indica «del» seguido del número de bases perdidas. (f) Reparación dirigida por homología (HDR). El donador de oligonucleótidos monocatenario (ssODN) contenía 65 pb de homología del gen COL7A1 en el brazo izquierdo y 101 pb en el derecho con una secuencia corta externa que sirve como un sitio de cebador único (cuadro gris moteado). Tres resultados de PCR de cebadores de amplificación con pares de cebadores endógenos (indicados con flechas marcadas con i. e iii.). Inserción de TALEN de los resultados de ODN en un segundo producto de PCR más pequeño generado mediante los pares de cebadores ii. e iii. El número al final de las células tratadas con TALEN indica el índice de HDR determinado por densitometría. (SEQ ID NO: 37 a (SEQ ID NO: 48).

Figura 2. Modificación con TALEN del gen COL7A1 evaluada mediante el ensayo de nucleasa Surveyor. Evaluación NHEJ mediante nucleasa Surveyor en fibroblastos EBDR. Se realizó una PCR de ciclo limitante de un fragmento de ~350 pb seguida de un ensayo de desemparejamiento Surveyor. La NHEJ inducida por TALEN queda demostrada por el patrón de bandas predecible de -200 y 300 pb (flechas). A la derecha se muestra el locus de COL7A1 no modificado en células de control.

Figuras 3A-C. Diseño de donador TALEN COL7A1 y reparación dirigida por homología. (a) Locus COL7A1 con mutación indicada mediante asterisco. Debajo se encuentra el donador, alineado con su relación con el locus endógeno que consta de secuencias genómicas de COL7A1 de un brazo izquierdo con 706 pb de largo y un 100 % de homología con respecto al locus genómico. Entre los brazos izquierdo y derecho, diseñado para insertarse en el intrón entre los exones 12 y 13, se encuentra un casete de puromicina PGK entre dos sitios loxp (cuadro, sitios loxp indicados mediante flechas flanqueantes). El brazo derecho era de 806 pb de largo y contenía 5 cambios de base. Cuatro de estos eran polimorfismos de mutación silenciosa (SPMP) (a los que se hace referencia en la dirección 5' y en la dirección 3') que servían como marcadores para la identificación de eventos basados en HDR; el último era la base normalizada que corrige el codón de terminación prematuro. El cuadro representa tres de los SPMP que se localizaban a 10 pb uno del otro. La base normal (es decir, con reversión de la mutación) se indica mediante el cuadro y el SPMP terminal (en la dirección 3') que suprime un sitio de enzima de restricción Apal se representa mediante un cuadro negro. El rayo indica el sitio diana de TALEN y los cebadores de PCR (flechas negras), diseñados de tal manera que uno se encuentre en el brazo donador y el otro fuera de este; se utilizaron para análisis conforme a lo representado. (SEQ ID NO: 49). Detección de SPMP en fibroblastos EBDR. El tratamiento con TALEN y la amplificación por PCR, seguidos por la digestión con Apal y la secuenciación de Sanger muestra (b) la presencia del SPMP resistente a Apal que deriva del donador y solo puede estar presente tras el corte con TALEN y la reparación dirigida por homología utilizando el donador exógeno como plantilla, (SEQ ID NO: 50) (c) mostrando la base no modificada (sensible a Apal) que había aparecido un evento heterocigoto de HDR (SEQ ID NO: 51).

Figura 4A-B. Escisión por Cre recombinasa de PGK-puromicina. (a) Esquema de donador con PGK puromicina entre dos sitios loxp. Introducción de un plásmido de Cre recombinasa en fibroblastos resistentes a la puromicina que dan como resultado la supresión del transgén de puromicina. (b) Unión genómica de loxp/COL7A1. Se empleó PCR para demostrar la presencia de una huella de loxP (triángulo/secuencia

inferior) en el intrón entre los exones 12 y 13 en las células de EBDR tratadas con TALEN/donador. (SEQ ID NO: 52).

Figura 5A-D. Análisis de secuencia con entrecruzamiento temprano. (a) leyenda de secuencias marcadoras introducidas en el donador. Flecha = SPMP en dirección 5'; línea = la base 1837 que provoca la EBDR, flecha = SPMP en dirección 3'. (SEQ ID NO: 53). Entrecruzamiento en la dirección 5'. Secuenciación de Sanger que muestra la incorporación de los SPMP en la dirección 5' (b) mantenimiento de la mutación en la base 1837 (SEQ ID NO: 54; (SEQ ID NO: 55) (c) y ausencia del SPMP en la dirección 3' (SEQ ID NO: 56; (SEQ ID NO: 57) (d) indicando que se dio HDR del donador, pero no consiguió corregir la mutación. La leyenda se ha fijado para incluir D (SEQ ID NO: 58; (SEQ ID NO: 59).

Figura 6A-D. Esquema del supuesto entrecruzamiento temprano. (a) Se muestran matrices de TALEN que se fijan a la secuencia diana y el donador se muestra a continuación. (b) La fijación al sitio diana y la dimerización con TALEN median una rotura de ADN de doble cadena (rayo) y estimulación de HDR utilizando el donador como plantilla de reparación. (c) Entrecruzamientos teóricos. El alineamiento del ADN endógeno y el donador da como resultado un entrecruzamiento (entrecruzamiento n.º 1), donde el material genético se intercambia de manera que los SPMP en la dirección 5' (cuadro) se incorporan mientras el segundo entrecruzamiento (flecha/entrecruzamiento n.º 2) se da en la dirección 5' de la base correctora y el SPMP en la dirección 3'. (d) Secuencia genómica resuelta que contiene secuencias donadoras parciales (líneas y cuadro) con mantenimiento de la base mutada (cuadro).

Figura 7A-C. Esquema de HDR y producción normal de ARNm. (a) Locus de COL7A1 endógeno mutado con sitio diana de TALEN indicado mediante un rayo. Se muestra la base mutada y, debajo, se encuentra el donador que da como resultado (b) la reparación del locus con presencia permanente de secuencias derivadas del donador desde el exón 12 a través del intrón entre los exones 15 y 16. (c) Análisis de ARNm. Los cebadores indicados amplificaron un producto que contiene la base correctora (cuadro y el cuadro negro de SPMP de *Apal*).

Figura 8A. Análisis de secuencia del corte con TALEN del donador. (SEQ ID NO: 60). (a) El ADNc de fibroblastos de EBDR tratados con TALEN se analizó mediante secuenciación directa de Sanger. El sitio de TALEN se traza en un cuadro rojo (cabe destacar que se trata de una secuencia de TALEN parcial, ya que el resto del sitio se encuentra dentro del intrón adyacente. La flecha muestra un límite exón/exón). La mutación de EBDR aparece subrayada y mostraba un retorno al estado natural (mutante = T, normal = C). El SPMP de *Apal* en dirección 3' está presente y se representa. El alineamiento de secuencia es de la secuencia de ADNc que se prevé que se codifica mediante el donador en la parte superior y la secuencia recuperada en la parte inferior. Los guiones/huecos muestran las deleciones probablemente debidas al corte con TALEN tras la HDR que inducía la posterior NHEJ (unión de extremos no homólogos). (SEQ ID NO: 61; SEQ ID NO: 62).

Figuras 9A-F. Edición genética mediada por TALEN de COL7A1 con HDR y expresión genética y proteica normalizada resultante. (a) Células corregidas con TALEN con conversión de la mutación a estado natural, (SEQ ID NO: 64) y (b) restablecimiento de la producción de colágeno tipo VII evaluado mediante inmunofluorescencia. (c) Secuenciación de ADNc de codón de terminación prematuro de EBDR homocigoto, (SEQ ID NO: 65) y (d) ausencia de producción proteica de colágeno tipo VII. (e) Secuenciación de Sanger del locus de COL7A1 de tipo natural, (SEQ ID NO: 66) y (f) expresión de colágeno tipo VII. Las células se tiñeron simultáneamente y los tiempos de exposición a microscopía confocal y de configuración del instrumento fueron idénticos. Los núcleos se tiñeron con DAPI y se muestran en azul.

Figura 10A-B. Secuenciación de Sanger de ARNm procedente de fibroblastos corregidos con TALEN. (a) El clon de fibroblasto 1-19 (SEQ ID NO: 67; SEQ ID NO: 68) y (b) 1-21 mostraba la presencia de la base corregida (línea) y el SPMP en dirección 3' (flecha). (SEQ ID NO: 69; SEQ ID NO: 70).

Figuras 11A-D. Perfil de mapas de integración de TALEN. (a) Esquema de rotura de ADN inducida por TALEN que acepta la carga de GFP, marcando permanentemente el locus genómico. (b) La coexpresión de TALEN e IDLV en 293 células dio como resultado células de GFP estables (análisis de citometría de flujo realizado 6 semanas después de la administración de TALEN e IDLV). (c) Esquema de la PCR lineal mediada por amplificación. La flecha azul indica el cebador LAM PCR, y las líneas discontinuas representan los productos de amplificación lineal que se clonaron y cartografiaron posteriormente para determinar el fragmento de fusión genómica de IDLV inducido por TALEN. (d) Partes integrantes identificadas (nr)LAM PCR/PCR. La recuperación de la secuencia LAM PCR y la búsqueda en la base de datos de genomas revelaron cinco sitios en los que se integró el IDLV. Las secuencias cartografiadas para la región espaciadora del sitio diana COL7A1 y cuatro sitios inespecíficos en los cromosomas 7, 16, 1 y 5 (ninguna de las últimas secuencias derivó de un exón codificante). (SEQ ID NO: 71-75).

Figura 12A-B. Lentivirus deficiente en integrasa. (a) Esquema de casete viral de GFP que se ha producido con una integrasa defectuosa. (b) Evolución temporal de la expresión de GFP de 293 IDLV en ausencia de TALEN mediante análisis secuencial durante 9 días, mostrando una rápida pérdida de GFP.

Las figuras 13 y 14 muestran constructos.

Descripción detallada de la invención

[0017] La descripción se dirige a la edición de ADN mediada por nucleasa efectora de tipo activador de la transcripción (TALEN) de mutaciones que provocan enfermedades en el contexto del genoma humano y células humanas para tratar pacientes afectados con trastornos genéticos. Esto supone un avance con respecto a los ensayos/herramientas de terapia génica anteriores que se basan en el aporte de copias funcionales de un gen terapéutico que se integran de manera aleatoria o parcialmente aleatoria en el genoma. Las consecuencias de los métodos de terapia génica anteriores son la alteración del locus en el que se sitúa la carga y una potencial inactivación o desregulación génica. Estas pueden derivar en efectos secundarios potencialmente mortales. El enfoque descrito en el presente documento maximiza la seguridad y la eficacia mediante el empleo de una TALEN personalizada para, por ejemplo, los genes humanos, que corrige únicamente el lugar de mutación, mientras que se conserva el resto del genoma en perfectas condiciones; en otras palabras, no existe una alteración del genoma restante, eliminando así los efectos inespecíficos asociados a la tecnología existente (p. ej., adición de genes mediada por virus). Se trata de un enfoque novedoso y es la primera terapia génica personalizada con corrección sin transgenes y mediada por TALEN de enfermedad que provoca la mutación en células, por ejemplo, células humanas. Por lo tanto, la tecnología se puede utilizar en células, tales como células humanas, de manera que se pueda corregir una mutación de pérdida de función sin dificultades con el restablecimiento de la función celular normal. En otras formas de realización, se puede mejorar la expresión génica.

Definiciones

[0018] Al describir y reivindicar la invención, se utilizará la siguiente terminología de acuerdo con las definiciones expuestas más adelante. A no ser que se definan de otro modo, todos los términos técnicos y científicos utilizados en el presente documento poseen el mismo significado que el que entiende normalmente un experto en la materia a la que pertenece la presente invención. Cualquier método y material similar o equivalente a los descritos en el presente documento se puede utilizar en la práctica o en los ensayos de la presente invención. Los valores específicos y preferidos listados más adelante para radicales, sustituyentes e intervalos se destinan únicamente a fines ilustrativos; no excluyen otros valores definidos ni otros valores dentro de intervalos definidos para los radicales y sustituyentes.

[0019] Según se utiliza en el presente documento, los artículos «un», «una», «unos» y «unas» se refieren a uno o más de uno, esto es, al menos uno, del objeto gramatical del artículo. A modo de ejemplo, «un elemento» se refiere a un elemento o a más de un elemento.

[0020] El término «aproximadamente», según se utiliza en el presente documento, significa alrededor de, en torno a, más o menos o circa. Cuando se utiliza el término «aproximadamente» junto con un intervalo numérico, este modifica ese intervalo ampliando los límites por encima y por debajo de los valores numéricos expuestos. Por lo general, el término «aproximadamente» se utiliza en el presente documento para modificar un valor numérico por encima y por debajo del valor indicado con una variación del 20 %.

[0021] El término «aislado(s)» se refiere a un(os) factor(es), célula o células que no estén asociadas a uno o varios factores, células o uno o varios componentes celulares que están asociados al/a los factor(es), célula o células *in vivo*.

[0022] Las «células» incluyen células de un vertebrado, o el «sujeto» es un vertebrado, tal como un mamífero, incluyendo un humano. Entre los mamíferos se incluyen, sin carácter limitativo, humanos, animales de granja, animales para deporte y animales de compañía. En el término «animal» se incluyen perros, gatos, peces, jerbos, cobayas, hámsteres, caballos, conejos, cerdos, ratones, monos (p. ej., homínidos, gorilas, chimpancés u orangutanes), ratas, ovejas, cabras, vacas y aves.

[0023] Un sujeto «de control» es un sujeto que presenta las mismas características que un sujeto de ensayo, tal como un tipo similar de enfermedad, etc. El sujeto de control puede examinarse, por ejemplo, durante el tiempo exacto o prácticamente el mismo tiempo durante el que se trata o se examina el sujeto de ensayo. El sujeto de control se puede examinar también, por ejemplo, en un momento lejano al momento en el que se examina el sujeto de ensayo, y los resultados del análisis del sujeto de control se pueden registrar para que los resultados registrados se puedan comparar con resultados obtenidos mediante el análisis de un sujeto de ensayo.

[0024] Un sujeto «de ensayo» es un sujeto que está siendo tratado.

[0025] Una «enfermedad» es un estado de salud de un sujeto, en el que el sujeto no puede mantener la homeostasis, y en el que, si la enfermedad no se mejora, entonces la salud del sujeto continúa deteriorándose. En cambio, un «trastorno» en un sujeto es un estado de salud en el que el sujeto es capaz de mantener la homeostasis, pero en el que el estado de salud del sujeto es menos favorable de lo que sería en ausencia del trastorno. Sin embargo, las definiciones de «enfermedad» y «trastorno», según se ha descrito anteriormente, no pretenden reemplazar las definiciones o el uso habitual relacionado con enfermedades o trastornos adictivos específicos.

[0026] Una enfermedad, afección o trastorno se «alivia», por ejemplo, si se reduce la gravedad de un síntoma de la enfermedad o trastorno, la frecuencia con la que un paciente experimenta dicho síntoma, o ambas.

[0027] Según se utiliza en el presente documento, una «cantidad efectiva» hace referencia, por ejemplo, a una cantidad suficiente para producir un efecto seleccionado, tal como el alivio de síntomas de una enfermedad o trastorno.

[0028] El término «medición del nivel de expresión» o «determinación del nivel de expresión», según se utiliza en el presente documento, se refiere, por ejemplo, a cualquier medida o ensayo que se pueda utilizar para correlacionar los resultados del ensayo con el nivel de expresión de un gen o proteína de interés. Dichos ensayos incluyen la medición del nivel de ARNm, niveles proteicos, etc., y se puede llevar a cabo mediante ensayos, como análisis de transferencia *northern* y *western*, ensayos de unión, inmunoelectrotransferencia, etc. El nivel de expresión puede incluir intervalos de expresión y se puede medir en relación con la cantidad real de un ARNm o proteína presente.

[0029] Según se utiliza en el presente documento, el término «vehículo farmacéuticamente aceptable» incluye, por ejemplo, cualquiera de los vehículos farmacéuticos habituales, como una solución salina tamponada con fosfato, agua, emulsiones, como una emulsión aceite/agua o agua/aceite, y diversos tipos de agentes humectantes. El término abarca también cualquiera de los agentes aprobados por una agencia reguladora del gobierno federal estadounidense o listados en la Farmacopea de EE. UU., para su uso en animales, incluyendo humanos.

[0030] El término «sal farmacéuticamente aceptable» se refiere, por ejemplo, a sales que conservan la eficacia biológica y las propiedades de los compuestos de la presente invención y que no resultan inapropiadas desde el punto de vista biológico ni de otro modo. En muchos casos, los compuestos de la presente invención son capaces de formar sales ácidas y/o básicas en virtud de la presencia de grupos amino y/o carboxilo o grupos similares a estos.

[0031] Según se utiliza en el presente documento, el término «fijarse específicamente» hace referencia, por ejemplo, a una molécula que reconoce y se fija a una molécula específica, pero no reconoce considerablemente ni se fija a otras moléculas en una muestra.

[0032] Según se utiliza en el presente documento, el término «síntoma» se refiere, por ejemplo, a cualquier fenómeno mórbido o desviación de la normalidad en cuanto a su estructura, función o sensación que experimenta el paciente e indica la presencia de la enfermedad.

[0033] Según se utiliza en el presente documento, el término «tratar» puede incluir la profilaxis de la enfermedad, trastorno o afección específica, o el alivio de los síntomas asociados a una enfermedad, trastorno o afección específica y/o a la prevención o eliminación de los síntomas. Un tratamiento «profiláctico» es, por ejemplo, un tratamiento administrado a un sujeto que no muestra signos de enfermedad o solo muestra signos tempranos de la enfermedad con el fin de reducir el riesgo de desarrollar patologías asociadas a la enfermedad. En el presente documento, «tratar» y «tratamiento» se utilizan indistintamente.

[0034] Un tratamiento «terapéutico» es, por ejemplo, un tratamiento administrado a un sujeto que muestra síntomas patológicos con el fin de disminuir o eliminar esos síntomas.

[0035] Una «cantidad terapéuticamente efectiva» de un compuesto es, por ejemplo, aquella cantidad de compuesto que sea suficiente para proporcionar un efecto beneficioso al sujeto al que se le administra el compuesto.

[0036] Según se utiliza en el presente documento, los «aminoácidos» se representan con el nombre completo de estos, mediante el código de tres letras correspondiente a este, o mediante el código de una letra correspondiente a este, según se indica en la siguiente tabla:

Nombre completo	Código de tres letras	Código de una letra
Ácido aspártico	Asp	D
Ácido glutámico	Glu	E
Lisina	Lys	K
Arginina	Arg	R
Histidina	His	H
Tirosina	Tyr	Y

Nombre completo	Código de tres letras	Código de una letra
(continuación)		
Cisteína	Cys	C
Asparagina	Asn	N
Glutamina	Gln	Q
Serina	Ser	S
Treonina	Thr	T
Glicina	Gly	G
Alanina	Ala	A
Valina	Val	V
Leucina	Leu	L
Isoleucina	Ile	I
Metionina	Met	M
Prolina	Pro	P
Fenilalanina	Phe	F
Triptófano	Trp	W

[0037] La expresión «aminoácido», según se utiliza en el presente documento, pretende incluir tanto aminoácidos naturales como sintéticos, y tanto D-aminoácidos como L-aminoácidos. «Aminoácido patrón» se refiere a cualquiera de los veinte L-aminoácidos patrón que se encuentran con frecuencia en péptidos de origen natural. «Residuo de aminoácido no patrón» se refiere a cualquier aminoácido, distinto de los aminoácidos patrón, independientemente de si se prepara de manera sintética o se deriva de una fuente natural. Según se utiliza en el presente documento, «aminoácido sintético» también abarca aminoácidos modificados químicamente, incluyendo, aunque sin carácter limitativo, sales, derivados de aminoácidos (como amidas) y sustituciones. Los aminoácidos contenidos en el interior de los péptidos de la presente invención, y especialmente en el terminal carboxilo o en el terminal amino, se pueden modificar mediante metilación, amidación, acetilación o sustitución con otros grupos químicos que pueden cambiar la semivida circulante del péptido sin afectar negativamente a su actividad. Asimismo, un enlace disulfuro puede estar presente o ausente en los péptidos de la invención.

[0038] El término «aminoácido» se utiliza indistintamente con «residuo de aminoácido», y puede hacer referencia a un aminoácido libre y a un residuo aminoacídico de un péptido. Se podrá observar, a partir del contexto en el que se utiliza el término si este se refiere a un aminoácido libre o a un residuo de un péptido.

[0039] Los aminoácidos se pueden clasificar en siete grupos en función de la cadena lateral R: (1) cadenas laterales alifáticas; (2) cadenas laterales que contienen un grupo hidroxilo (OH); (3) cadenas laterales que contienen átomos de azufre; (4) cadenas laterales que contienen un grupo ácido o amida; (5) cadenas laterales que contienen un grupo básico; (6) cadenas laterales que contienen un anillo aromático; y (7) prolina, un iminoácido en el que la cadena lateral se fusiona con el grupo amino.

[0040] Según se utiliza en el presente documento, el término «sustitución conservadora de aminoácido» se define en el presente documento como intercambios entre uno de los siguientes cinco grupos:

I. Residuos pequeños alifáticos, no polares o ligeramente polares:

Ala, Ser, Thr, Pro, Gly;

II. Residuos polares, con carga negativa y sus amidas:

Asp, Asn, Glu, Gln;

III. Residuos polares, con carga positiva:

His, Arg, Lys;

IV. Residuos grandes, no polares y alifáticos:

Met Leu, Ile, Val, Cys

5 V. Residuos grandes aromáticos:

Phe, Tyr, Trp

[0041] Según se utiliza en el presente documento, el término «ácido nucleico» abarca ARN, así como ADN y ADNc de cadena simple, doble y triple. Además, los términos «ácido nucleico», «ADN», «ARN» y términos similares también incluyen análogos de ácido nucleico, es decir, análogos que presenten una cadena principal distinta de una cadena principal de fosfodiéster. Por ejemplo, los denominados «ácidos nucleicos peptídicos», que se conocen en la técnica y presentan enlaces peptídicos en lugar de enlaces fosfodiéster en la cadena principal, se consideran dentro del alcance de la presente invención. Por «ácido nucleico» se hace referencia también a cualquier ácido nucleico, bien compuesto por desoxirribonucleósidos o ribonucleósidos, bien compuestos por enlaces fosfodiéster o enlaces modificados, tales como fosfotriéster, fosforamidoato, siloxano, carbonato, carboximetilester, acetamidoato, carbamato, tioéter, fosforamidoato puenteado, metilfosfonato puenteado, fosforamidoato puenteado, metilfosfonato puenteado, fosforotioato, metilfosfonato, fosforoditioato, fosforotioato puenteado o enlaces sulfona, y combinaciones de dichos enlaces. El término ácido nucleico también incluye específicamente ácidos nucleicos compuestos por bases distintas de las cinco bases de origen biológico (adenina, guanina, timina, citosina y uracilo). La notación convencional se utiliza en el presente documento para describir secuencias de polinucleótidos: el extremo izquierdo de una secuencia de polinucleótidos de cadena simple es el extremo 5'; la dirección izquierda de una secuencia de polinucleótidos de cadena doble se denomina dirección 5'. La dirección de la adición de nucleótidos 5' a 3' en los transcritos de ARN incipientes se denomina dirección de la transcripción. La cadena de ADN que presenta la misma secuencia que un ARNm se denomina «cadena codificante»; las secuencias en la cadena de ADN que se localizan 5' con respecto a un punto de referencia en el ADN se denominan «secuencias en dirección 5'; las secuencias en la cadena de ADN que se localizan 3' con respecto a un punto de referencia en el ADN se denominan «secuencias en dirección 3'».

[0042] A no ser que se especifique de otro modo, una «secuencia de nucleótidos que codifica una secuencia de aminoácidos» incluye todas las secuencias de nucleótidos que son versiones degeneradas entre sí y que codifican la misma secuencia de aminoácidos. Las secuencias de nucleótidos que codifican proteínas y ARN pueden incluir intrones.

[0043] Según se utiliza en el presente documento, el término «homólogo/a(s)» hace referencia a la similitud de secuencias de subunidad entre dos moléculas poliméricas, p. ej., entre dos moléculas de ácido nucleico, p. ej., dos moléculas de ADN o dos moléculas de ARN, o entre dos moléculas polipeptídicas. Cuando una posición de subunidad en ambas moléculas es ocupada por la misma subunidad monomérica, p. ej., si una posición en cada una de las dos moléculas de ADN es ocupada por adenina, entonces son homólogas en esa posición. La homología entre dos secuencias es una función directa del número de posiciones coincidentes u homólogas, por ejemplo, si la mitad (p. ej., cinco posiciones en un polímero con diez subunidades de longitud) de las posiciones en dos secuencias de compuestos son homólogas, entonces las dos secuencias son homólogas en un 50 %, si el 90 % de las posiciones, p. ej., de 9 a 10 son coincidentes u homólogas, las dos secuencias comparten un 90 % de homología. A modo de ejemplo, las secuencias de ADN 3'ATTGCC5' y 3'TATGGC comparten un 50 % de homología.

[0044] Según se utiliza en el presente documento, «homología» se utiliza como sinónimo de «identidad».

[0045] La determinación del porcentaje de identidad entre dos secuencias nucleotídicas o aminoácidas se puede lograr utilizando un algoritmo matemático. Por ejemplo, un algoritmo matemático útil para comparar dos secuencias es el algoritmo de Karlin y Altschul (1990, Proc. Natl. Acad. Sci. USA 87:2264-2268), modificado como en Karlin y Altschul (1993, Proc. Natl. Acad. Sci. USA 90:5873-5877). Este algoritmo se incorpora en los programas NBLAST y XBLAST de Altschul, et al. (1990, J. Mol. Biol. 215:403-410), y es accesible, por ejemplo, a través del sitio web del Centro Nacional para la Información Biotecnológica (NCBI). Las búsquedas de nucleótidos BLAST se pueden llevar a cabo con el programa NBLAST (denominado «blastn» en el sitio web del NCBI), utilizando, por ejemplo, los siguientes parámetros: penalización por hueco = 5; penalización por extensión de hueco = 2; penalización por desemparejamiento = 3; recompensa por emparejamiento = 1; valor esperado 10,0; y tamaño de palabra = 11 para obtener secuencias de nucleótidos homólogas a un ácido nucleico descrito en el presente documento. Las búsquedas de proteínas BLAST se pueden realizar con el programa XBLAST (denominado «blastp» en el sitio web del NCBI) o el programa del NCBI «blastp», utilizando los siguientes parámetros: valor esperado 10,0, matriz de puntuación BLOSUM62 para obtener secuencias de aminoácidos homólogas a una molécula proteica descrita en el presente documento. Para obtener alineamientos con hueco para fines comparativos, se puede utilizar Gapped BLAST («BLAST con huecos»), según se describe en Altschul et al. (1997, Nucleic Acids Res. 25:3389-3402). De manera alternativa, se puede utilizar PSI-Blast o PHI-Blast para realizar una búsqueda iterativa que detecte relaciones distantes entre moléculas (Id.) y relaciones entre

moléculas que comparten un patrón común. Al utilizar los programas BLAST, Gapped BLAST, PSI-Blast y PHI-Blast, se pueden utilizar los parámetros por defecto de los respectivos programas (p. ej., XBLAST y NBLAST).

5 **[0046]** El porcentaje de identidad entre dos secuencias se puede determinar utilizando técnicas similares a las anteriormente descritas, con o sin permitir huecos. Al calcular el porcentaje de identidad, se cuentan normalmente emparejamientos exactos.

[0047] Los términos «comprender», «comprendiendo» y similares pueden presentar el significado que se les atribuye en la ley de patentes estadounidense y pueden significar «incluir», «incluyendo» y similares. Según se utiliza en el presente documento, «incluyendo» o «incluir» o similares significan incluyendo, sin carácter limitativo.

10 TALEN

[0048] Las nucleasas efectoras de tipo activador de la transcripción (TALEN) son enzimas de restricción artificiales generadas mediante la fusión del dominio de unión al ADN del efector TAL con un dominio de escisión del ADN. Estos reactivos permiten una escisión del ADN eficiente, programable y específica, y representan potentes herramientas para la edición del genoma *in situ*. Los efectores de tipo activador de la transcripción (TALE) pueden ser modificados rápidamente para fijarse prácticamente a cualquier secuencia de ADN. El término TALEN, según se utiliza en el presente documento, es amplio e incluye una TALEN monomérica que puede escindir ADN bicatenario sin la ayuda de otra TALEN. El término TALEN se utiliza también para referirse a uno o a ambos miembros de un par de TALEN que hayan sido modificadas para trabajar juntas en la escisión de ADN en el mismo sitio. Se puede hacer referencia a las TALEN que trabajan juntas como una TALEN izquierda y una TALEN derecha, que hacen alusión a la lateralidad del ADN. Véase USSN 12/965,590; USSN 13/426,991 (US 8,450,471); USSN 13/427,040 (US 8,440,431); USSN 13/427,137 (US 8,440,432); y USSN 13/738,381.

25 **[0049]** Los efectores TAL son proteínas secretadas por bacterias *Xanthomonas*. El dominio de unión al ADN contiene una secuencia de 33-34 aminoácidos muy conservada, con la excepción de los aminoácidos 12.^o y 13.^o. Estas dos ubicaciones son muy variables (residuo variable de repetición (RVD)) y muestran una fuerte correlación con el reconocimiento de nucleótidos concreto. Esta relación simple entre secuencia de aminoácidos y reconocimiento de ADN ha permitido la modificación de dominios de unión específicos al ADN seleccionando una combinación de segmentos de repetición que contienen los RVD apropiados.

30 **[0050]** El dominio de escisión no específico del ADN desde el extremo de la endonucleasa FokI se puede utilizar para construir nucleasas híbridas que estén activas en un ensayo con levadura. Estos reactivos están también activos en células vegetales y en células animales. Los estudios de TALEN iniciales utilizaban el dominio de escisión natural de FokI, aunque algunos estudios de TALEN posteriores también utilizaban variantes del dominio de escisión de FokI con mutaciones diseñadas para mejorar la especificidad de escisión y la actividad de escisión. El dominio FokI funciona como un dímero, que necesita dos constructos con dominios de unión al ADN únicos para los sitios en el genoma diana con una orientación y separación adecuadas. Tanto el número de residuos de aminoácidos entre el dominio de unión al ADN de TALEN y el dominio de escisión de FokI y el número de bases entre los dos sitios de unión individuales de TALEN son parámetros para lograr altos niveles de actividad. El número de residuos de aminoácidos entre el dominio de unión al ADN de TALEN y el dominio de escisión de FokI se puede modificar mediante la introducción de un espaciador (distinto de la secuencia espaciadora) entre la pluralidad de secuencias de repetición del efector TAL y el dominio de endonucleasa de FokI. La secuencia espaciadora puede contener de 12 a 30 nucleótidos.

45 **[0051]** La relación entre la secuencia de aminoácidos y el reconocimiento de ADN del dominio de unión de TALEN permite proteínas diseñables. En este caso, la síntesis genética artificial resulta problemática debido al apareamiento inadecuado de la secuencia de repetición que se encuentra en el dominio de unión de TALE. Una solución para esto es utilizar un programa de *software* disponible de manera pública (DNABWorks) para calcular los oligonucleótidos adecuados para su ensamblaje en una PCR de dos etapas; el ensamblaje de oligonucleótidos seguido por una amplificación genética completa. También se ha informado de un número de esquemas de ensamblaje modular para generar constructos de TALE modificados. Ambos métodos ofrecen un enfoque sistemático para los dominios de unión al ADN modificados que son conceptualmente similares al método de ensamblaje modular para generar dominios de reconocimiento del ADN de dedos de zinc.

50 **[0052]** Una vez se hayan ensamblado los genes TALEN, estos se insertan en plásmidos; a continuación, los plásmidos se utilizan para transfectar la célula diana donde se expresan los productos génicos y se introducen en el núcleo para acceder al genoma. Las TALEN se pueden utilizar para editar genomas induciendo roturas de doble cadena (DSB), a las que las células responden con mecanismos de reparación. De este modo, se pueden utilizar para corregir mutaciones en el genoma que, por ejemplo, provoquen enfermedades.

55 Vectores y ácidos nucleicos

[0053] Se pueden introducir varios ácidos nucleicos en células para obtener la expresión de un gen. Según se utiliza en el presente documento, el término ácido nucleico incluye ADN, ARN y análogos de ácido nucleico, y ácidos nucleicos que sean bicatenarios o monocatenarios (es decir, una cadena simple de sentido o antisentido). Los análogos de ácido nucleico se pueden modificar en el grupo básico, grupo azúcar o cadena principal de

fosfato para mejorar, por ejemplo, la estabilidad, hibridación o solubilidad del ácido nucleico. Las modificaciones en el grupo básico incluyen desoxiuridina por desoxitimidina, y 5-metil-2'-desoxicitidina y 5-bromo-2'-doxicitidina por desoxicitidina. Las modificaciones del grupo azúcar incluyen la modificación del hidroxilo 2' del azúcar ribosa para formar azúcares 2'-O-metil o 2'-O-alil. La cadena principal de desoxirribosa fosfato se puede modificar para producir ácidos nucleicos morfolinos, en los que cada grupo básico está conectado a un anillo morfolino de seis miembros, o ácidos nucleicos peptídicos, en los que la cadena principal de desoxifosfato se sustituye por una cadena principal pseudopeptídica y las cuatro bases se conservan. Véase, Summerton y Weller (1997) *Antisense Nucleic Acid Drug Dev.* 7(3):187; y Hyrup et al. (1996) *Bioorgan. Med. Chem.* 4:5. Además, la cadena principal de desoxifosfato se puede sustituir, por ejemplo, por una cadena principal de fosforotioato o fosforoditioato, una fosforamidita, o una cadena principal de alquil fosfotriéster.

[0054] Las secuencias de ácido nucleico pueden estar unidas operativamente a una región reguladora, tal como un promotor. Las regiones reguladoras pueden ser de cualquier especie. Según se utiliza en el presente documento, unido/a(s) operativamente se refiere a una región reguladora relativa a una secuencia de ácido nucleico, de tal manera que se permita o facilite la transcripción del ácido nucleico diana. Cualquier tipo de promotor puede estar unido operativamente a una secuencia de ácido nucleico. Entre los ejemplos de promotores se incluyen, sin carácter limitativo, promotores específicos del tejido, promotores constitutivos y promotores que responden o no responden a un estímulo concreto (p. ej., promotores inducibles).

[0055] Las regiones reguladoras adicionales que pueden resultar útiles en constructos de ácido nucleico incluyen, aunque sin carácter limitativo, secuencias de poliadenilación, secuencias de control de la traducción (p. ej., un segmento de entrada interna del ribosoma, IRES), potenciadores, elementos inducibles o intrones. Dichas regiones reguladoras pueden no ser necesarias, aunque pueden incrementar la expresión afectando a la transcripción, estabilidad del ARNm, eficiencia de traducción, etc. Dichas regiones reguladoras se pueden incluir en un constructo de ácido nucleico, como se prefiera, para obtener una expresión óptima de los ácidos nucleicos en la(s) célula(s). No obstante, en ocasiones se puede obtener una expresión suficiente sin dichos elementos adicionales.

[0056] Se puede utilizar un constructo de ácido nucleico que codifique péptidos señal o marcadores seleccionables. Se pueden utilizar péptidos señal de manera que un polipéptido codificado se dirija a una ubicación celular concreta (p. ej., la superficie celular). Entre los ejemplos no limitativos de marcadores seleccionables se incluyen puromicina, ganciclovir, adenosina desaminasa (ADA), aminoglucósido fosfotransferasa (neo, G418, APH), dihidrofolato reductasa (DHFR), higromicina-B-fosfotransferasa, timidina quinasa (TK) y xantina-guanina fosforribosiltransferasa (XGPRT). Dichos marcadores resultan útiles para seleccionar transformantes estables en cultivo. Otros marcadores seleccionables incluyen polipéptidos fluorescentes, como la proteína verde fluorescente o la proteína amarilla fluorescente.

[0057] Los constructos de ácido nucleico se pueden introducir en células de cualquier tipo utilizando varias técnicas. Entre los ejemplos no limitativos de técnicas, se incluye el uso de sistemas transposón, virus recombinantes que puedan infectar células o liposomas u otros métodos no virales, como la electroporación, la microinyección o la precipitación de fosfato de calcio, que sean capaces de administrar ácidos nucleicos a células.

[0058] Los ácidos nucleicos se pueden incorporar en vectores. Un vector es un término amplio que incluye cualquier segmento de ADN específico que esté diseñado para moverse desde una molécula vehículo hacia un ADN diana. Se puede hacer referencia a un vector como un vector de expresión, o sistema vector, que es un conjunto de componentes necesarios para lograr la inserción de ADN en un genoma u otra secuencia de ADN de interés, como un episoma, plásmido o incluso segmento de ADN virus/fago. La mayoría de las veces, los vectores contienen uno o varios casetes de expresión que comprenden una o varias secuencias de control de expresión, donde una secuencia de control de expresión es una secuencia de ADN que controla y regula la transcripción y/o traducción de otra secuencia de ADN o ARNm, respectivamente.

[0059] Se conocen muchos tipos distintos de vectores. Por ejemplo, se conocen plásmidos y vectores virales, por ejemplo, vectores retrovirales. Los plásmidos de expresión en mamíferos poseen normalmente un origen de replicación, un promotor adecuado y un potenciador opcional, y también cualquier sitio de unión del ribosoma, un sitio de poliadenilación, sitios donadores y aceptores, secuencias de terminación de la transcripción y secuencias no transcritas de flanco 5'. Los ejemplos de vectores incluyen: plásmidos (que también pueden ser un vehículo de otro tipo de vector), adenovirus, virus adenoasociados (AAV), lentivirus (p. ej., VIH-1, VIS o VIF), retrovirus (p. ej., ASV, ALV o MoMLV) y transposones (p. ej., *Sleeping Beauty*, elementos P, Tol-2, *Frog Prince*, piggyBac).

Usos terapéuticos

[0060] La corrección genética a base de TALEN posee muchas aplicaciones clínicas y preclínicas (p. ej., de investigación). Por ejemplo, la corrección genética a base de TALEN se puede utilizar para corregir genes en los que las mutaciones deriven en enfermedades. Por ejemplo, cualquier enfermedad caracterizada por alteraciones de base pequeña que incluyen inserciones y deleciones, tales como, sin carácter limitativo, epidermolisis bullosa, osteogénesis imperfecta, disqueratosis congénita, las mucopolisacaridosis, distrofia muscular, fibrosis quística

(CFTR), anemia de Fanconi, las esfingolipidosis, las lipofuscinosis, adrenoleucodistrofia, inmunodeficiencia combinada grave, anemia de células falciformes, talasemia, etc.

5 **[0061]** En una forma de realización, la enfermedad es epidermólisis bullosa. La epidermólisis bullosa distrófica recesiva (EBDR) se caracteriza por un déficit funcional de la proteína de colágeno tipo VII debido a los defectos genéticos en el gen de colágeno tipo VII (COL7A1). Este gen codifica la cadena alfa de colágeno tipo VII. La fibra de colágeno tipo VII, compuesta por tres cadenas idénticas de colágeno alfa, se limita a la zona más baja por debajo de los epitelios escamosos estratificados. Funciona como una fibra de anclaje entre los epitelios externos y el estroma subyacente. Las mutaciones en este gen están asociadas a todas las formas de epidermólisis bullosa distrófica.

10 **[0062]** El COL7A1 se localiza en el brazo corto del cromosoma humano 3, en la región cromosómica denominada 3p21.31 (N.º de conjunto: ENSG00000114270). El gen tiene un tamaño de aproximadamente 31 000 pares de bases, y su secuencia de codificación está fragmentada en 118 exones, véase la SEQ ID NO: 32.

15 **[0063]** El COL7A1 se transcribe en un ARNm de 9287 pares de bases (los n.ºs de acceso para las proteínas y ARNm humanos son, respectivamente, NM_000094 y NP_000085). En la piel, la proteína de colágeno tipo VII se sintetiza mediante queratinocitos y fibroblastos dérmicos. El símbolo para el gen ortólogo en los ratones es Col7a1 (los n.ºs de acceso para ARNm y proteínas de ratones son, respectivamente, NM_00738 y NP_031764).

20 **[0064]** Las personas con EBDR muestran formación de ampollas cutáneas incurables, a menudo letales, y se enfrentan a un mayor riesgo de sufrir carcinoma agresivo de células escamosas¹. Las terapias de aumento de genes resultan prometedoras, aunque se corre el riesgo de sufrir mutagénesis insercional. Por consiguiente, se describen en el presente documento nucleasas efectoras de tipo activador de la transcripción (TALEN) modificadas para la edición precisa del genoma en células de pacientes con EBDR. En el presente documento, se describe la capacidad de TALEN para inducir roturas de ADN bicatenario específicas de sitio (DSB), que dan como resultado una reparación dirigida por homología (HDR) de una plantilla de donador exógeno. Este proceso dio como resultado la corrección de mutación del gen COL7A1 y el restablecimiento de la expresión normal de genes y proteínas. Este estudio proporciona una demostración conceptual para la medicina genómica personalizada y supone la primera corrección mediada por TALEN *in situ* de un gen humano endógeno en fibroblastos.

30 **[0065]** Las células que se han de modificar mediante la corrección de genes a base de TALEN se pueden obtener a partir del paciente o de un donante. Las células pueden ser de cualquier tipo, tales como células de fibroblastos, queratinocitos, células madre pluripotentes, hematopoyéticas, mesenquimales y embrionarias inducibles, células de progenie hematopoyéticas, como células T, células B, células gliales y neuronas, células madre y progenitoras neurogliales, células musculares, células pulmonares, células pancreáticas y células hepáticas y/o células del sistema reticuloendotelial. Una vez modificadas mediante corrección genética a base de TALEN, las células se pueden expandir y/o administrar a un paciente para tratar la enfermedad.

35 **[0066]** Se pueden utilizar matrices para administrar células de la presente invención a sitios anatómicos específicos, donde los factores de crecimiento concretos pueden incorporarse o no en la matriz, o codificarse en plásmidos incorporados en la matriz para la absorción por parte de las células, y se pueden utilizar para dirigir el crecimiento de la población celular inicial. El ADN plasmídico que codifica citoquinas, factores de crecimiento u hormonas puede quedar atrapado en el interior de un portamatrices polimérico activado por genes. A continuación, el polímero biodegradable se implanta cerca del sitio en el que se desea el tratamiento.

[0067] A los efectos descritos en el presente documento, se pueden administrar células autólogas, alogénicas o xenogénicas de la presente invención a un paciente mediante inyección directa a un sitio preseleccionado, sistemáticamente, en la superficie de una matriz aceptable o en torno a esta, o junto con un vehículo farmacéuticamente aceptable.

45 **[0068]** Asimismo, los constructos de ácido nucleico o proteínas se pueden inyectar de manera local o sistemática en un sujeto, por ejemplo, con un vehículo farmacéuticamente aceptable.

Crecimiento/expansión de células

50 **[0069]** Las células que se han de modificar mediante la corrección de genes a base de TALEN se pueden obtener a partir del paciente o de un donante. Las células pueden ser de cualquier tipo, tales como células de fibroblasto. Una vez modificadas mediante corrección genética a base de TALEN, las células se pueden expandir y/o administrar a un paciente para tratar la enfermedad.

55 **[0070]** Las células se pueden cultivar en medio de cultivo establecido en la técnica y disponible comercialmente a través de la Colección estadounidense de cultivos tipo (*American Type Culture Collection* (ATCC)), Invitrogen y otras empresas. Entre dichos medios se incluyen, aunque sin carácter limitativo, el medio Dulbecco's Modified Eagle's Medium (DMEM), el medio DMEM F12, el medio Eagle's Minimum Essential Medium, el medio F-12K, el medio Iscove's Modified Dulbecco's Medium, Knockout D-MEM o RPMI-1640. Un experto en la materia puede modificar o modular concentraciones de medios y/o complementos de medios según sea necesario para las

células utilizadas. También se podrá observar que muchos medios se encuentran disponibles como formulaciones con baja glucosa, con o sin piruvato de sodio.

[0071] También se contempla una complementación de medio de cultivo celular con sueros de mamíferos. Los sueros suelen contener factores y componentes celulares que son necesarios para la viabilidad y la expansión.

5 Entre los ejemplos de sueros se incluye el suero fetal bovino (FBS), suero bovino (BS), suero de ternera (CS), suero fetal de ternera (FCS), suero de ternera recién nacida (NCS), suero de cabra (GS), suero de caballo (HS), suero humano, suero de pollo, suero porcino, suero de oveja, suero de conejo, suero de rata (RS), sustituciones de suero (incluyendo, sin carácter limitativo, sustitución de suero KnockOut (KSR, Invitrogen)) y fluido embrionario bovino. Se entiende que los sueros pueden inactivarse por calor a 55-65°C si se considera necesario para inactivar componentes de la cascada del complemento. La modulación de concentraciones séricas, o la retirada de suero del medio de cultivo, se puede utilizar también para promover la supervivencia de uno o varios tipos de células deseadas. Las células se pueden cultivar en presencia de FBS y/o suero específico para el tipo de célula de la especie. Por ejemplo, las células se pueden aislar y/o expandir con el suero total (p. ej., FBS) o concentraciones de sustitución de suero de aproximadamente un 0,5 % a aproximadamente un 5 % o más, incluyendo aproximadamente de un 5 % a aproximadamente un 15 % o más, tal como aproximadamente un 20 %, aproximadamente un 25 % o aproximadamente un 30 %. Las concentraciones séricas se pueden determinar empíricamente.

[0072] Se pueden utilizar también suplementos adicionales para suministrar oligoelementos a las células para su óptimo crecimiento y expansión. Dichos suplementos incluyen insulina, transferrina, selenio de sodio y combinaciones de los mismos. Estos componentes se pueden incluir en una solución salina tal como, sin carácter limitativo, la solución salina equilibrada de Hank™ (Hanks' Balanced Salt Solution™ (HBSS)), la solución salina de Earle (Earle's Salt Solution™), suplementos antioxidantes, suplementos MCDB-201™, tampón fosfato salino (PBS), ácido N-2-hidroxietilpiperazina-N'-etanosulfónico (HEPES), nicotinamida, ácido ascórbico y/o ácido ascórbico-2-fosfato, así como aminoácidos adicionales. Muchos medios de cultivo celular ya contienen aminoácidos; sin embargo, algunos precisan complementación antes de cultivar las células. Dichos aminoácidos incluyen, aunque sin carácter limitativo, L-alanina, L-arginina, L-ácido aspártico, L-asparagina, L-cisteína, L-cistina, L-ácido glutámico, L-glutamina, L-glicina, L-histidina, L-inositol, L-isoleucina, L-leucina, L-lisina, L-metionina, L-fenilalanina, L-prolina, L-serina, L-treonina, L-triptófano, L-tirosina y L-valina.

[0073] Los antibióticos se suelen utilizar también en cultivos celulares para mitigar la contaminación bacteriana, por micoplasma y fúngica. Normalmente, los compuestos antibióticos o antimicóticos empleados son mezclas de penicilina/estreptomicina, pero también incluyen, aunque sin carácter limitativo, anfotericina (Fungizone™), ampicilina, gentamicina, bleomicina, higromicina, kanamicina, mitomicina, ácido micofenólico, ácido nalidíxico, neomicina, nistatina, paromomicina, polimixina, puromicina, rifampicina, espectinomicina, tetraciclina, tilosina y zeocina.

[0074] Ventajosamente, las hormonas se pueden utilizar también en cultivos celulares e incluyen, aunque sin carácter limitativo, D-alдостерона, dietilestilbestrol (DES), dexametasona, - estradiol, hidrocortisona, insulina, prolactina, progesterona, somatostatina/hormona del crecimiento humano (HGH), tiotropina, tiroxina y L-tironina. El - mercaptoetanol también se puede complementar en medios de cultivo celular.

[0075] Los lípidos y vehículos lipídicos también se pueden utilizar para complementar medios de cultivo celular, en función del tipo de célula y del futuro de la célula diferenciada. Dichos lípidos y vehículos pueden incluir, aunque sin carácter limitativo, ciclodextrina (α β γ), colesterol, ácido linoleico conjugado con albúmina, ácido linoleico y ácido oleico conjugados con albúmina, ácido linoleico no conjugado, ácido linoleico-oleico-araquidónico conjugado con albúmina, ácido oleico no conjugado y conjugado con albúmina, entre otros. Del mismo modo, la albúmina se puede utilizar en formulaciones libres de ácidos grasos.

[0076] Las células en cultivo se pueden mantener bien en suspensión, bien adheridas a un soporte sólido, como componentes de matriz extracelular y polímeros sintéticos o biopolímeros. A menudo las células requieren factores adicionales que potencian su adhesión a un soporte sólido (p. ej., factores de adhesión), como el colágeno tipo I, tipo II y tipo IV, concanavalina A, sulfato de condroitina, fibronectina, «superfibronectina» y/o polímeros similares a fibronectina, gelatina, laminina, poli-D-lisina y poli-L-lisina, Matrigel™, trombospondina y/o vitronectina.

[0077] Las células se pueden cultivar con distintas densidades, p. ej., las células se pueden sembrar o mantener en la placa de cultivo con distintas densidades. Por ejemplo, con densidades, incluyendo, sin carácter limitativo, densidades de menos de aproximadamente 2000 células/pocillo de una placa de 12 pocillos (por ejemplo, área de crecimiento de fondo plano de 12 pocillos: 3,8 cm² volumen de pocillo: 6,0 ml o ID de pocillo x densidad (mm) 22,1x17,5; capacidad de pocillo (ml) 6,5, área de crecimiento (cm²) 3,8), incluyendo menos de aproximadamente 1500 células/pocillo de una placa de 12 pocillos, menos de aproximadamente 1000 células/pocillo de una placa de 12 pocillos, menos de aproximadamente 500 células/pocillo de una placa de 12 pocillos, o menos de aproximadamente 200 células/pocillo de una placa de 12 pocillos. Las células también se pueden sembrar o mantener con densidades mayores, por ejemplo, más de aproximadamente 2000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 2500 células/pocillo de una placa de 12 pocillos, más de aproximadamente 3000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 3500 células/pocillo de una placa de

- 5 12 pocillos, más de aproximadamente 4000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 4500 células/pocillo de una placa de 12 pocillos, más de aproximadamente 5000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 5500 células/pocillo de una placa de 12 pocillos, más de aproximadamente 6000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 6500 células/pocillo de una placa de 12 pocillos, más de aproximadamente 7000 células/pocillo de una placa de 12 pocillos, más de aproximadamente 7500 células/pocillo de una placa de 12 pocillos o más de aproximadamente 8000 células/pocillo de una placa de 12 pocillos.

Ejemplos

- 10 **[0078]** El siguiente ejemplo se proporciona para demostrar y para representar además algunas formas de realización y aspectos de la presente invención y no se interpreta como limitativo del alcance de estas.

Ejemplo 1

Materiales y métodos.

Sujeto de investigación y derivación de línea celular.

- 15 **[0079]** Tras obtener el consentimiento informado de los padres, se obtuvo una biopsia de punción de la piel de un paciente de sexo masculino con EBDR con una mutación homocigota de codón de terminación prematuro c.1837 C>T. Se obtuvo la aprobación para la investigación en sujetos humanos a través del Comité Institucional de Revisión de la Universidad de Minesota. Se derivó y se mantuvo una línea celular primaria de fibroblastos en condiciones de baja concentración de oxígeno.

Construcción de TALEN y donador.

- 20 **[0080]** La TALEN candidata descrita en la figura 1A se generó mediante el método de ensamblaje Golden Gate y se insertó en una forma homodimérica de una endonucleasa *FokI* conducida por el promotor CAGGs, según lo descrito [1, 2]. El brazo donador izquierdo se amplificó con los cebadores LAF y LAR mostrados en la tabla 1. El brazo derecho se sintetizó en dos fragmentos (interno y externo) utilizando una estrategia de ensamblaje de oligonucleótidos superpuestos conforme se describe en [3, 4]. Todos los conjuntos de cebadores se muestran en la tabla 1; los brazos izquierdo y derecho se clonaron en un casete de puromicina PGK entre dos sitios loxp.
- 25

Tabla 1. (SEQ ID NO: 1-21)

Corrección de TALEN para EBDR		
C06	fragmento externo 1-12	TCACGGGTAGCCAAACGCTATGTCTGATAGCGGTCCGCTTAGGAGAGAAAGCGGAGGAATC
C07	C7GT1	Atcgtccacatccctgtctctt
C08	C7APAF	CAAAGGACCAATGAGGGTA
C09	C7GT2	tctagtgggagagagcaatg
C10	RT1	TCGACTTGGATGACGTTTCAG
C11	RT2	GTTGAGCCACGATGACTG
C12	Surveyor F	ttcagccatatcccagctc
D01	Surveyor R	tgctccagctaataccgaaat
D02	Oligodúplex superior	G*T*CCGTACGGATCCAAGCTTCGTCGACCTAGCC
D03	Oligodúplex inferior	CATGCCCTAGGTTCGAAGCAGCTGGATCGGGG*A*C
D04	Conector F	GGATCCAAGCTTCGTCGACCTAGCC
D05	Donador ssODN (purificado con PAGE)	tcctgctccc tgtccatcac tgcctatgctc ccacatccct gtcctttct gaccctgcccacct
D06		agtaagtgtagccacgtctgtg t ggt gtagcctctctgtctaa ctacagg atccctccagacacctttacgta actctggatgaa gtttccctgtacg gttctctgctc gaaatggagttccc ccgtgacgggagaa a
D07	Cebadores Surveyor inespecíficos	
D08	1q23.3 directo	TCTCAGGCAAGAAAAATTGGA-
D09	1q23.3 inverso	TGTGCATTTATTCTGTGCTTGT
D10	5q33.1 directo	GAGTCCCTTGGCCCTATTC
D11	5q33.1 inverso	GGCTGCAGTGAGCTATGATG
D12	7q21.3 directo	ACTCCAAGTCACAGGGGATG
E01	7q21.3 inverso	CAGCTCTGACTGCTGTTTGC
E02	16p13.3 directo	TTGCTCACAGAAGGACCACA
E03	143p13.3 inverso	ACGTGGGTGTGACGGTTATT

Transferencia genética.

- 5 **[0081]** Todos los tratamientos con TALEN consistían en la administración de 2,5 µg de cada TALEN y una cantidad de 10 µg de donador a través del sistema de transfección Neon (Life Sciences) con el siguiente reglaje de los instrumentos: 1500 V, ancho de pulso de 20 ms, y un pulso único. Las células se incubaron a 31 °C durante 48 horas tras la transferencia genética [5].

Cultivo celular.

- 10 **[0082]** Las células se mantuvieron en medios de crecimiento constituidos por DMEM complementado con FBS al 20 %, 100 U/ml de aminoácidos no esenciales y 0,1 mg/ml tanto de penicilina como de estreptomicina, respectivamente (Invitrogen) y se cultivaron con O₂ al 2 %, CO₂ al 5 % y 37 °C.

10 Nucleasa Surveyor.

- 15 **[0083]** El ADN genómico se aisló 48 horas después de la transferencia genética con TALEN y se amplificó durante 30 ciclos con los cebadores Surveyor F y Surveyor R y se sometió a un tratamiento con nucleasa Surveyor conforme a lo descrito [6]. Los productos se resolvieron en un gel PAGE con TBE al 10 % (Invitrogen). En el caso de los amplicones inespecíficos, la reacción de PCR se llevó a cabo durante 35 ciclos y todos los cebadores se listan en la tabla 1.

Análisis de reparación dirigida por homología.

- 20 **[0084]** Para la cuantificación de HDR, se transfectaron TALEN y 5 µl de un donador de oligonucleótidos de cadena simple de 40 µM en células y se cribaron mediante PCR a las 48 horas utilizando tres cebadores: cebadores directos Surveyor F, Surveyor R y conector. Se realizó una densitometría según lo descrito [6]. Para la corrección genética, se introdujeron 10 µg del plásmido donador junto con los 2,5 µg de cada ADN de TALEN y se llevó a cabo una selección conforme a lo descrito posteriormente.

Selección.

- 25 **[0085]** Las células se seleccionaron en bloque en 0,2 µg/ml de puomicina, se segregaron en subgrupos, se cribaron para determinar su HDR, y a continuación se sembraron a baja densidad (250-275 células totales) en una placa de 10 cm². Se colocó un disco de clonación con grasa siliconada (todos de Corning) sobre células únicas en presencia de medios básicos complementados con 10 ng/ml de factor de crecimiento epidérmico y 0,5 ng/ml de factor de crecimiento fibroblástico. Las células se expandieron a recipientes sucesivamente más grandes. Se añadió una cre recombinasa adenoviral con una MOI de 20 para suprimir el casete de puomicina PGK (Vector BioLabs).

30 Cribado molecular de corrección celular.

- 35 **[0086]** Los pares de cebadores C7GT1 y C7GT2 se emplearon para amplificar una unión desde el donador hacia el locus endógeno (cribado de SPMP en dirección 5'). La región SPMP de *Apal* se analizó en ADN genómico tratado con *Apal* antes y después de la amplificación por PCR con C7APAF y C7GT2. El ARN mensajero de clones aislados se convirtió a ADNc y se cribó con RT1 y RT2 y, a continuación, se digirió con *Apal*. Los amplicones resistentes a *Apal* se clonaron y se secuenciaron mediante Sanger.

Análisis de expansión celular.

- 40 **[0087]** Los fibroblastos corregidos por genes se expandieron en matraces T150 y se tripsinizaron para obtener suspensiones de células únicas. A continuación, las células se volvieron a suspender en 100 µl de PBS + BSA al 0,5 % + yoduro de propidio (eBiosciences) y, posteriormente, se añadió un volumen equivalente de microesferas de referencia PKH26 (SIGMA). Se recogieron cinco mil microesferas y se calculó el número absoluto de células viables conforme al protocolo de fabricación (SIGMA).

Generación de iPSC y ensayo de teratoma.

- 45 **[0088]** Los fibroblastos corregidos por genes (o células no corregidas como control) se reprogramaron a iPSC conforme a lo descrito [7, 8] y, a continuación, se sembraron en el matraz de un ratón SCID hasta que se formó una masa visible. La masa se extirpó para su integración y tinción.

Inmunofluorescencia.

- 50 **[0089]** Las células corregidas por genes se sembraron en un portaobjetos de cámara y se fijaron 24 horas después con paraformaldehído al 4 %, se permeabilizaron con Triton X al 0,2 %, se bloquearon con BSA al 1 % y se tiñeron con un anticuerpo policlonal de colágeno anti-tipo VII (1:1500; proporcionado generosamente por los doctores David Woodley y Mei Chen). La tinción del anticuerpo secundario se llevó a cabo con IgG Cy3 de burro anti-conejo (1:500; Jackson Immunoresearch). La tinción de control de isotipo se realizó utilizando una molécula entera de IgG de ratón (Jackson Immunoresearch). Los núcleos se tiñeron con 4',6-diamidino-2-fenilindol (Vector Laboratories). Se tomaron imágenes utilizando un voltaje PMT de 745 en un microscopio confocal Olympus BX61 FV500 (Olympus Optical Co LTD) y se analizaron utilizando el *software* Fluoview versión 4.3. La microscopía óptica se llevó a cabo en un microscopio Leica.

IDLV y LAM-PCR/nrLAM PCR.

[0090] Se produjeron partículas lentivirales deficientes en integrasa (IDLV) en células 293T mediante la cotransfección basada en lípidos (Lipofectamine 2000, Invitrogen) del vector de transferencia CMV-GFP, el plásmido de empaquetado pCMV- Ψ R8.2 que alberga la mutación de integrasa D64V [9, 10], y el plásmido de codificación de envoltura pMD2.VSV-G. El marcado génico se realizó mediante la nucleofección de células HEK 293 con las TALEN y, 24 horas más tarde, una transducción de GFP IDLV a una MOI de 7. Se analizaron 100 ng de ADN genómico por duplicado mediante LAM-PCR [11] utilizando enzimas MseI y Tsp509I y nrLAM-PCR [12] para asegurar la recuperación a nivel del genoma de los sitios de integración IDLV. Se secuenciaron amplicones (nr)LAM-PCR mediante la plataforma de pirosecuenciación Roche/454 y los datos del sitio de integración se analizaron utilizando la fuente de HISAP [13, 14],[15]. La posición genómica que alberga >1 IS a escasa distancia se examinó para determinar sus potenciales sitios de unión a TALEN inespecíficos utilizando el patrón de emparejamiento *scan-for-matches* (rastreo de emparejamientos) [13].

Resultados/Discusión

[0091] La falta de la proteína de colágeno tipo VII en la unión dermoepidérmica (UDE) da como resultado una pérdida de la integridad estructural de la piel. El restablecimiento de la deposición del colágeno tipo VII en la UDE mediante el trasplante de fibroblasto localizado o de célula hematopoyética sistémica alogénica puede aliviar los síntomas [16-18]. Sin embargo, la eficacia subóptima del trasplante de células alogénicas debido a los riesgos de toxicidad, infección y fallo del injerto, da impulso al desarrollo de nuevas terapias basadas en células autólogas. Por lo tanto, en el presente documento se describe una estrategia de edición genómica para la corrección de COL7A1 basada en la tecnología TALEN. Los fibroblastos son un tipo de célula ideal debido a su facilidad de derivación y a su baja susceptibilidad a que se detenga el crecimiento en cultivo, así como a su capacidad para depositar colágeno tipo VII en la UDE [18,19]. Las TALEN son nucleasas modificadas que pueden inducir una rotura de ADN de doble cadena en un locus genómico definido por el usuario, estimulando así la HDR, y son mejores que otras nucleasas en cuanto a su capacidad de precisión y a su facilidad de generación [20,21].

[0092] El *software* TAL Effector-Nucleotide Targeter [22,23] identificó 68 potenciales sitios TALEN para el locus de COL7A1 humano y los datos experimentales recientes de apoyo en una amplia serie de genes humanos [21] enfatizan la alta capacidad de especificidad para las TALEN, se debe considerar la EBD y otras enfermedades que muestran heterogeneidad en la ubicación y el número de secuencias mutadas. Se utilizó la metodología de clonación Golden Gate para generar una nucleasa específica del paciente próxima a un codón de terminación prematuro en el exón 14 del gen COL7A1 (figura 1A). Una TALEN está compuesta por una matriz de repetición TALE modificada fusionada con el dominio de nucleasa FokI (figura 1B); las especificidades de fijación de repeticiones TALE en la matriz las dictan las identidades de dos residuos hipervariables en cada repetición (figura 1C). Los fibroblastos de EBD tratados con TALEN se analizaron en busca de pruebas de reparación mediante las dos principales rutas de reparación del ADN: unión de extremos no homólogos (NHEJ) propensa a error y HDR. El ensayo de nucleasa Surveyor y la secuenciación de Sanger que mostraron 11 alelos mutados del total de 75 analizados eran coherentes con la NHEJ (figuras 2A y 2E). La escisión de TALEN también derivó en la captación de un dúplex oligonucleotídico en el sitio de rotura del ADN (figuras 2B-F)[24]. Estos datos establecieron que la nucleasa está activa en el sitio diana. A continuación, se determinó si las células EBD podían someterse a HDR tras la coadministración de TALEN y un donador de oligonucleótidos (ODN) que contenía una única secuencia de cebadores flanqueada por brazos de donador cortos (figura 1F). Los fibroblastos EBD transfectados con plásmidos TALEN y el ODN se analizaron a continuación con un enfoque PCR con tres cebadores que detectó simultáneamente los alelos modificados y no modificados. Este ensayo mostró que las TALEN en células EBD pueden estimular la HDR para que incorpore una secuencia exógena procedente del donador ODN (figura 1G) y la tasa del 14,6 % de NHEJ y del 2,1 % de HDR muestra la eficacia del uso de TALEN para la modificación a alto nivel de fibroblastos humanos.

[0093] Para determinar si se podría corregir una mutación de COL7A1 que provoca EBD y expandirse posteriormente las células corregidas genéticamente, se generó un plásmido donador exógeno que permitiría la detección selectiva y la expansión de células corregidas por genes. Este donador consistía en brazos de homología que sobrepasaban ~1 kb del locus COL7A1 entre los exones 12 y 16 (figura 3A). En el donador había un casete de puromicina PGK entre dos sitios loxP orientado para que se insertara en el intrón entre los exones 12 y 13. Los sitios de flanco loxP permiten la supresión del marcador seleccionable con Cre recombinasa, dejando una pequeña «huella» de loxP en el intrón (figura 4). En el brazo de donador derecho, se modificaron cinco alteraciones de pares de bases únicas: la base normal en el sitio de la mutación que restablece un genotipo normal y cuatro polimorfismos de mutación silenciosa (SPMP) que permitían la delimitación de alelos modificados con HDR frente a no modificados (figura 3A). Tres de estos SPMP son en dirección 5' con respecto a la base diana y el que está en dirección 3' suprime un sitio de restricción Apal (alteraciones denominadas en adelante SPMP en la dirección 5' o en la dirección 3').

[0094] De los nueve clones analizados, se obtuvieron cuatro que mostraron indicios de HDR. En un clon, la presencia de los SPMP en la dirección 5' era evidente; no obstante, la mutación COL7A1 patógena de EBD se mantuvo y no se halló el SPMP en la dirección 3' (figura 5). Estos datos sugieren que un entrecruzamiento HDR

tuvo lugar en el brazo del donador en la dirección 5' de la región que restablece un genotipo normal (figura 6). En el caso de los otros tres clones, sin embargo, se podía detectar el SPMP insertado en el donador en la dirección 3', indicando que un alelo había experimentado HDR y el otro no, dando como resultado un locus COL7A1 heterocigoto (figuras 3B y 3C).

- 5 **[0095]** La HDR debería revertir la base mutante y restablecer la expresión génica normal. Por consiguiente, esto se evaluó con una estrategia de RT-PCR para la detección de la base normal y los SPMP en la dirección 3' en el mismo transcrito seguida por un empalme fuera del intrón que interviene (figura 7). Resulta interesante que la secuenciación directa del ADNc en un clon mostró una delección de secuencias en el sitio diana de TALEN (figura 8). Estos datos indican que la TALEN estaba activa después de la HDR e inducía una mutación adicional
10 mediada por NHEJ. Estudios anteriores con endonucleasas con dedos de zinc (ZFN) muestran que las mutaciones silenciosas en la secuencia donadora pueden reducir la frecuencia de este evento no deseado¹²; sin embargo, esto no fue posible en este experimento, ya que el sitio de TALEN se encontraba en un límite intrón/exón y se optó por dejar inalterada la secuencia de TALEN donadora para no interrumpir el empalme. Esto afectó de manera negativa a la recuperación de un clon; no obstante, dos clones mostraron los transcritos normales deseados derivados del donador y basados en HDR (figura 9A). A continuación, se determinó si el
15 tratamiento con TALEN restablecía la expresión proteica de colágeno tipo VII en comparación con las células naturales o mutantes de EBDR sin tratar que presentan transcritos normales o anormales, respectivamente (figuras 9C y 9E). La detección de colágeno tipo VII basada en inmunofluorescencia reveló un rescate de la producción de colágeno tipo VII en células tratadas con TALEN y una ausencia total en fibroblastos EBDR de control sin tratar (figuras 9B y 9D). Estos resultados confirman la capacidad de las TALEN para mediar en una
20 modificación genética en un sitio diana específico de la enfermedad con el restablecimiento de la producción normal de ARNm y proteínas.

- [0096]** El riesgo de efectos inespecíficos se debe considerar en el uso clínico de reactivos de edición del genoma. Entre las opciones de cartografía de sitios inespecíficos de nucleasas de edición genética se incluyen:
25 (i) la realización de una evolución sistemática *in vitro* de ligandos mediante enriquecimiento exponencial (SELEX) con proteínas monoméricas de fijación al ADN de cada nucleasa en un par y utilizando a continuación estos datos para predecir los potenciales sitios inespecíficos [25], (ii) la realización de una selección del sitio de escisión *in vitro* utilizando nucleasas dimericas e interrogando posteriormente a los sitios de esta selección que tienen lugar en el genoma de células de interés para mutaciones inducidas por nucleasas, (iii) la utilización de la
30 tendencia de un lentivirus deficiente en integrasa (IDLV) a integrarse en DSB inducidas por nucleasas y a la identificación posterior de los puntos de inserción mediante LAM-PCR[9]. Aunque los métodos (ii) y (iii) parecen ser mejores a la hora de identificar los sitios inespecíficos de nucleasas que el método (i), los anteriores métodos no consiguen identificar sitios inespecíficos previstos por el otro, lo que sugiere que ningún método es comprehensivo en su detección de eventos inespecíficos. El método (iii) se utilizó con un IDLV con un gen de
35 proteína verde fluorescente (GFP) que puede quedar atrapado en una DSB generada por nucleasa (figura 11A)[9, 26]. Se utilizaron células embrionarias de riñón humano (293) debido a su capacidad proliferativa acelerada, que debería potenciar la rápida dilución de IDLV no integrados y minimizar la integración aleatoria. Además, se ha teorizado que, debido a la estructura abierta de la cromatina de 293 células, cualquier efecto inespecífico se manifestará en mayor medida que en las células primarias y permitirá una cartografía más
40 sensible de eventos inespecíficos. La introducción del IDLV de GFP solo dio como resultado una rápida pérdida de expresión GFP en 293 células (figura 12). La cointroducción de IDLV y TALEN dio como resultado una población estable de células GFP (figura 11B), que se utilizaron para cartografiar los sitios de integración con PCR no restrictiva y lineal mediada por amplificación ((nr)LAM-PCR) (figura 11C). Se recuperaron cinco sitios que mostraban una unión entre el IDLV y la secuencia genómica adyacente (figura 11D). Estos eventos eran
45 inesperados, ya que, incluso las nucleasas utilizadas en ensayos clínicos muestran efectos inespecíficos [9] y las regiones no codificantes recubiertas sugieren que esta TALEN posee un perfil de seguridad que no se prevé que afecte de manera negativa a la expresión génica.

- [0097]** En la resolución de la metodología LAM-PCR, la TALEN descrita en el presente documento muestra una elevada tasa de actividad en la diana. Además, estos estudios, al igual que otros, muestran que una potencial
50 diana para las nucleasas modificadas es el propio constructo donador, y subrayan los beneficios de la inclusión de una secuencia marcadora que pueden ayudar en la selección del evento HDR deseado [27].

- [0098]** En resumen, se obtuvieron células cutáneas de un paciente con EBDR y los reactivos del donador y de TALEN (las secuencias se incluyen más adelante) se diseñaron y se construyeron para dirigirse específicamente a esta única mutación. La aplicación de las herramientas de edición genética dio como resultado la corrección de
55 la mutación EBDR en células de fibroblastos humanos diploides que resultan adecuadas para el uso terapéutico tras la expansión directa o la reprogramación a pluripotencia seguida por expansión [7, 8] - y proporcionan la primera demostración de corrección mediada por TALEN de un gen patológico en el genoma humano. Estos estudios proporcionan la prueba de que las TALEN se pueden utilizar en el desarrollo de terapias individualizadas clínicamente relevantes.

Ejemplo 2

[0099] Un ejemplo de una secuencia de plásmido donador se expone en la SEQ ID NO: 22. Un ejemplo del brazo izquierdo de la secuencia donadora se expone en la SEQ ID NO:31. Un ejemplo del sitio Loxp del donador se expone en la SEQ ID NO:23. Un ejemplo del promotor PGK del donador se expone en la SEQ ID NO:24. Un ejemplo del gen de puromicina de la secuencia donadora se expone en la SEQ ID NO:25. Un ejemplo de la señal de poliadenilación de la hormona del crecimiento bovino del donador se expone en la SEQ ID NO:26. Un ejemplo del sitio Loxp del donador se expone en la SEQ ID NO:27. Un ejemplo del brazo derecho del donador se expone en la SEQ ID NO:28. Un ejemplo de TALEN izquierda (pTAL 286) se expone en la SEQ ID NO:29. Un ejemplo de TALEN derecha (pTAL 287) se expone en la SEQ ID NO:30.

Bibliografía

[0100]

1. Carlson, D. F., *et al.* Efficient TALEN-mediated gene knockout in livestock. *Proceedings of the National Academy of Sciences of the United States of America* 109: 17382-17387.
2. Cermak, T., *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research* 39: e82.
3. Osborn, M. J., Defeo, A. P., Blazar, B. R., y Tolar, J. Synthetic Zinc Finger Nuclease Design and Rapid Assembly. *Human gene therapy*.
4. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3.º, y Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* 6: 343-345.
5. Doyon, Y., Choi, V. M., Xia, D. F., Vo, T. D., Gregory, P. D., y Holmes, M. C. Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. *Nature methods* 7: 459-460.
6. Guschin, D. Y., Waite, A. J., Katibah, G. E., Miller, J. C., Holmes, M. C., y Rebar, E. J. A rapid and general assay for monitoring endogenous gene modification. *Methods in molecular biology* (Clifton, N.J.649: 247-256.
7. Tolar, J., *et al.* Keratinocytes from Induced Pluripotent Stem Cells in Junctional Epidermolysis Bullosa. *The Journal of investigative dermatology*.
8. Tolar, J., *et al.* Induced pluripotent stem cells from individuals with recessive dystrophic epidermolysis bullosa. *The Journal of investigative dermatology* 131: 848-856.
9. Gabriel, R., *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nature biotechnology* 29: 816-823.
10. Vargas, J., Jr., Gusella, G. L., Najfeld, V., Klotman, M. E., y Cara, A. (2004). Novel integrase-defective lentiviral episomal vectors for gene transfer. *Human gene therapy* 15: 361-372.
11. Schmidt, M., *et al.* (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat Methods* 4: 1051-1057.
12. Paruzynski, A., *et al.* (2010). Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat Protoc* 5: 1379-1395.
13. Dsouza, M., Larsen, N., y Overbeek, R. (1997). Searching for patterns in genomic data. *Trends Genet* 13: 497-498.
14. Arens, A., *et al.* Bioinformatic clonality analysis of next-generation sequencing-derived viral vector integration sites. *Human gene therapy methods* 23: 111-118.
15. Arens, A., *et al.* (2012). Bioinformatic clonality analysis of next-generation sequencing-derived viral vector integration sites. *Hum Gene Ther Methods* 23: 111-118.
16. Wagner, J. E., *et al.* Bone marrow transplantation for recessive dystrophic epidermolysis bullosa. *The New England journal of medicine* 363: 629-639.
17. Tolar, J., *et al.* (2009). Amelioration of epidermolysis bullosa by transfer of wild-type bone marrow cells. *Blood* 113: 1167-1174.
18. Wong, T., *et al.* (2008). Potential of fibroblast cell therapy for recessive dystrophic epidermolysis bullosa. *The Journal of investigative dermatology* 128: 2179-2189.
19. Goto, M., *et al.* (2006). Fibroblasts show more potential as target cells than keratinocytes in COL7A1 gene therapy of dystrophic epidermolysis bullosa. *The Journal of investigative dermatology* 126: 766-772.
20. Cermak, T., *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*.

21. Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D., y Joung, J. K. FLASH assembly of TALENs for high-throughput genome editing. *Nature biotechnology* 30: 460-465.
22. Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F., y Dobbs, D. (2007). Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic acids research* 35: W599-605.
- 5 23. Doyle, E. L., *et al.* TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic acids research* 40: W117-122.
24. Orlando, S. J., *et al.* Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic acids research* 38: e152.
- 10 25. Pattanayak, V., Ramirez, C. L., Joung, J. K., y Liu, D. R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nature methods* 8: 765-770.
26. Paruzynski, A., *et al.* Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nature protocols* 5: 1379-1395.
27. Zou, J., *et al.* (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell stem cell* 5: 97-110.

15 LISTADO DE SECUENCIAS

[0101]

<110> REGENTS OF THE UNIVERSITY OF MINNESOTA Osborn, Mark Tolar, Jakub Blazar, Bruce Voytas, Daniel

20 <120> CORRECCIÓN DE GEN A BASE DE TALEN

<130> J1100/20003

<150> 61771735

25 < 151> 2013-03-01

<160> 75

<170> PatentIn versión 3.5

30

<210> 1

< 211> 61

< 212> ADN

< 213> Secuencia artificial

35

<220>

< 223> Fragmento externo 1-12

<400> 1

40

tcacgggtag ccaacgctat ggtcctgata gcgggtccgct taggagagaa gcggaggaat 60

c 61

<210> 2

< 211> 22
 < 212> ADN
 < 213> Secuencia artificial

5 <220>
 < 223> C7GT1

<400> 2
 atcgtccaca tccctgtctc tt 22

10 <210> 3
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

15 <220>
 < 223> C7APAF

<400> 3
 20 caaagggacc aatgagggta 20

<210> 4
 < 211> 21
 < 212> ADN

25 < 213> Secuencia artificial

<220>
 < 223> C7GT2

30 <400> 4
 tctagtgggg agaggccaat g 21

<210> 5
 < 211> 20
 < 212> ADN

35 < 213> Secuencia artificial

<220>
 < 223> RT1

40 <400> 5

ES 2 750 550 T3

	tcgacttgga tgacgttcag	20
	<210> 6	
	< 211> 19	
5	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
	< 223> RT2	
10		
	<400> 6	
	gttcgagcca cgatgactg	19
	<210> 7	
15	< 211> 21	
	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
20	< 223> Surveyor F	
	<400> 7	
	ttttcagcca tatcccagct c	21
25	<210> 8	
	< 211> 20	
	< 212> ADN	
	< 213> Secuencia artificial	
30	<220>	
	< 223> Surveyor R	
	<400> 8	
	tgctccagct aatccgaaat	20
35		
	<210> 9	
	< 211> 33	
	< 212> ADN	
	< 213> Secuencia artificial	
40		
	<220>	

ES 2 750 550 T3

	< 223> Oligodúplex superior	
	<400> 9	
5	gtccgtacgg atccaagctt cgtcgaccta gcc	33
	<210> 10	
	< 211> 33	
	< 212> ADN	
	< 213> Secuencia artificial	
10	<220>	
	< 223> Oligodúplex inferior	
	<400> 10	
15	catgcctagg ttcgaagcag ctggatcggg gac	33
	<210> 11	
	< 211> 25	
	< 212> ADN	
20	< 213> Secuencia artificial	
	<220>	
	< 223> Conector F	
25	<400> 11	
	ggatccaagc ttcgtcgacc tagcc	25
	<210> 12	
	< 211> 64	
30	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
	< 223> Donador ssODN (purificado con PAGE)	
35	<400> 12	
	tctgcgtccc tgtcatcact gccatcgtcc cacatccctg tctctttctg acccctgccc	60
	acct	64
40	<210> 13	

ES 2 750 550 T3

< 211> 133
 < 212> ADN
 < 213> Secuencia artificial

5 <220>
 < 223> Cebadores

<400> 13

agtagtggtg	gccagctgtt	gtgtgactct	ggtaactaga	gatccctcag	acccttttag	60
tcacttggat	gacgttcagg	ctgggcttag	ctacactgtg	cggtgtctg	ctcgagtgg	120
ccccgtgagg	gca					133

<210> 14
 < 211> 20
 < 212> ADN
 15 < 213> Secuencia artificial

<220>
 < 223> Cebadores

20 <400> 14
 tctcaggcaa gaaaattgga 20

<210> 15
 < 211> 24
 25 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> 1Q23.3 inverso

30 <400> 15
 tgtgcattta ttctgtgtct tggt 24

<210> 16
 35 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

<220>

ES 2 750 550 T3

< 223> 5q33.1 directo

 <400> 16
 gagttccctt gggcctattc 20
 5
 <210> 17
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial
 10
 <220>
 < 223> 5q33.1 inverso

 <400> 17
 15 ggctgcagtg agctatgatg 20

 <210> 18
 < 211> 20
 < 212> ADN
 20 < 213> Secuencia artificial

 <220>
 < 223> 7q21.3 directo

 25 <400> 18
 actccaagtc acaggggatg 20

 <210> 19
 < 211> 20
 30 < 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> 7q21.3 inverso
 35
 <400> 19
 cagctctgac tgctgtttgc 20

 <210> 20
 40 < 211> 20
 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> 16p13.3 directo

5

<400> 20

ttgctcacag aaggaccaca 20

<210> 21

10 < 211> 20

< 212> ADN

< 213> Secuencia artificial

<220>

15 < 223> 16p13.3 inverso

<400> 21

acgtgggtgt gacggttatt 20

20 <210> 22

< 211> 2950

< 212> ADN

< 213> Secuencia artificial

25 <220>

< 223> Secuencia de plásmido donador

<400> 22

ES 2 750 550 T3

gcctgtgagc cctgtaacag acctgcaagc caccgagctg cccgggcagc ggggtgcgagt	60
gtcctggagc ccagtccttg gtgccacca gtaccgcatc attgtgcgca gcacccaggg	120
tgagggtggac gcagccagca cccccaccac acacactgaa gttccagcct gaggggtctg	180
agtgtacctc cagccctctc cttccacacc tgggtcccta tagtctcctc cctgtctcgt	240
tatccatctg cttcccaaga tggcactgag ccctgaactg tcaacagggt gtgtctgccc	300
ctaccctacc ccaaccaatc tcctccctgc cccacacccc atccacggct ttcacctctg	360
cacagccaca agctcaatca cctgtctttc tgttaccta gtgggtacca ctgtacacca	420
acctctcatt gcatgtcccc atccagcact gacctctgcc atcccacatc agtgtctgtc	480
aatgggtcat cagttctcac tgcagtccac tgactcctgt catcctacat gctaagatcc	540
cacagactcc tgtcctccag tgtccctgac catcactggc ccctttcatc ctatgtcctt	600
gtcttgact cacttttttc agccatatcc cagctctctg gtccccacca cccacatcc	660
ccacccccat ggcctttctc actctgcgtc cctgtccatc actgccataa cttcgtataa	720
tgtatgctat acgaagttat caaggcagtc tggagcatgc gcttttagcag ccccgctggg	780
cacttggcgc tacacaagtg gcctctggcc tcgcacacat tccacatcca ccggtaggcg	840

ES 2 750 550 T3

ccaacccggct ccgttctttg gtggcccctt cgcgccacct tctactcctc ccctagttag	900
gaagttcccc cccgccccgc agctcgcgtc gtgcaggacg tgacaaatgg aagtagcacg	960
tctcactagt ctctgtcaga tggacagcac cgctgagcaa tggagcggg taggcctttg	1020
gggcagcggc caatagcagc ttgtctcctt cgctttctgg gctcagaggc tgggaagggg	1080
tgggtccggg ggcgggctca ggggggggct cagggggggg gcggggcgcc gaaggtcctc	1140
cggaggcccg gcattctgca cgtttcaaaa gcgcacgtct gcgcgctgt tctcctcttc	1200
ctcatctccg ggcctttcga cctgcagccc aagcttacca tgaccgagta caagcccacg	1260
gtgcgcctcg ccaccgcga cgcgtcccc agggccgtac gcaccctcgc cgcgcgttc	1320
gccgactacc ccgccacggc ccacaccgtc gatccggacc gccacatcga gcgggtcacc	1380
gagctgcaag aactcttct cagcgcgtc gggctcgaca tcggcaaggt gtgggtcgcg	1440
gacgacggcg ccgcggtggc ggtctggacc acgccggaga gcgtcgaagc gggggcgggtg	1500
ttcgccgaga tcggcccgcg catggccgag ttgagcgggt cccggctggc cgcgcagcaa	1560
cagatggaag gcctcctggc gccgcaccgg cccaaggagc ccgcgtgggt cctggccacc	1620
gtcggcgtct cgcgcgacca ccagggaag ggtctgggca gcgcgctcgt gctccccgga	1680
gtggaggcgg ccgagcgcgc cggggtgccc gccttctctg agacctccgc gccccgaac	1740
ctcccccttct acgagcggct cggtctcacc gtcaccgccg acgtcgaggt gcccgaagga	1800
ccgcgcacct ggtgcatgac ccgcaagccc ggtgcccatc atcaccatca ccattgagtt	1860
taaaaccgct gatcagcctc gactgtgcct tctagtgtcc agccatctgt tgtttgcccc	1920
tccccctgac cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat	1980
gaggaaattg catcgattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg	2040
caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc	2100
tctatggctt ctgagcgga aagaaccagc tataacttcg tataatgtat gctatacgaa	2160
gttatctcga gatcgtccca catccctgtc tctttctgac ccctgccac ctaccctgac	2220
ttctctctta ggggttgagc ggaccctggt gcttctctgg agtcagacag cattcgactt	2280
ggatgacgtt caggctgggc ttagctacac tgtgcgggtg tctgctagag tgggtccccg	2340
tgagggcagc gctagcgtgc tgaccgtccg ccggggtgag tactgcagga ggcttgtgga	2400
ggacagctgc ctgcctcact ctggtcctgg ttctgacttc tgacttctgt ctgtaactcc	2460
tagagccgga aactccactt gctgttcag ggctgcgggt tgtggtgtca gatgcaacgc	2520
gagtgagggt ggcctgggga ccggtccctg gagctagtgg atttcggatt agctggagca	2580
caggcagtgg tcagtgtggg gtgtgtgggg gactgccaaa gggaccaatg agggatatggg	2640
tgccagaggg gacaggcagg agccatgcc gcatctccct ctgacctcag gtccggagtc	2700
cagccagaca ctgccccag actctactgc cacagacatc acagggtgc agcctggaac	2760
cacctaccag gtggctgtgt cgggtactgc aggcagagag gagggacctg ctgcagtcac	2820
cgtggctcga acgggtcagg ccctgcccc gtcccttggc tctctgcctc cattgctctt	2880
tcagaccccc atgccttccc ttgcagaccc ttgcttctcc ccagaactcc tgcctcccc	2940
ttcagaatcc	2950

ES 2 750 550 T3

	<210> 23	
	< 211> 34	
	< 212> ADN	
	< 213> Secuencia artificial	
5	<220>	
	< 223> Sitio Loxp del donador	
	<400> 23	
10	ataacttcgt ataattgtat ctatacgaag ttat	34
	<210> 24	
	< 211> 484	
	< 212> ADN	
15	< 213> Secuencia artificial	
	<220>	
	< 223> Promotor PGK del donador	
20	<400> 24	
	caaggcagtc tggagcatgc gcttttagcag ccccgctggg cacttggcgc tacacaagtg	60
	gcctctggcc tcgcacacat tccacatcca ccggtaggcg ccaaccggct ccgttctttg	120
	gtggccctt cgcgccacct tctactcctc ccctagtcag gaagttcccc cccgccccgc	180
	agctcgcgtc gtgcaggacg tgacaaatgg aagtagcacg tctcactagt ctcgtagcaga	240
	tggacagcac cgctgagcaa tgggaagcggg taggcctttg gggcagcggc caatagcagc	300
	tttgcctctt cgctttcttg gctcagagggc tgggaagggg tgggtccggg ggcgggctca	360
	ggggcgggct caggggcggg gcgggcgccc gaaggctctc cggaggcccg gcattctgca	420
	cgcttcaaaa gcgcacgtct gccgcgctgt tctcctcttc ctcatctccg ggcctttcga	480
	cctg	484
	<210> 25	
25	< 211> 612	
	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
30	< 223> Gen de la puromicina del donador	
	<400> 25	

ES 2 750 550 T3

```

cagcccaagc ttacatgac cgagtacaag cccacgggtgc gcctcgccac ccgcgacgac      60
gtccccaggg ccgtacgcac cctcgccgcc gcgttcgccg actacccgc cagcgccac      120

accgtcgatc cggaccgcca catcgagcgg gtcaccgagc tgcaagaact cttcctcacg      180
cgcgtcgggc tcgacatcgg caaggtgtgg gtcgcgagc acggcgccgc ggtggcggtc      240
tggaaccacgc cggagagcgt cgaagcgggg gcggtgttcg ccgagatcgg cccgcgcatg      300
gccgagttga gcggttcccg gctggccgcg cagcaacaga tggaaggcct cctggcgccg      360
caccggccca aggagcccgc gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag      420
ggcaagggtc tgggcagcgc cgtcgtgctc cccggagtgg aggcggccga gcgcgccggg      480
gtgcccgcct tcctggagac ctccgcgcc cgcaacctcc cttctacga gcggctcggc      540
ttcacgtca ccgccgacgt cgaggtgccc gaaggaccgc gcacctggtg catgaccgc      600
aagcccgtg cc      612

```

5 <210> 26
 < 211> 295
 < 212> ADN
 < 213> Secuencia artificial

10 <220>
 < 223> Señal de poliadenilación de la hormona del crecimiento bovino del donador
 <400> 26

```

catcatcacc atcaccattg agtttaaacc cgctgatcag cctcgactgt gccttctagt      60
tgccagccat ctgttggttg cccctcccc gtgccttcct tgaccctgga aggtgccact      120
cccatgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat      180
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc      240
aggcatgctg gggatgcggt gggctctatg gcttctgagg cggaaagaac cagct      295

```

15 <210> 27
 < 211> 34
 < 212> ADN
 20 < 213> Secuencia artificial

<220>
 < 223> Sitio Loxp del donador

25 <400> 27
 ataactcgt ataattgat ctatacgaag ttat 34

<210> 28

< 211> 785

< 212> ADN

< 213> Secuencia artificial

5

<220>

< 223> Brazo derecho del donador

<400> 28

10

ctcgagatcg tcccacatcc ctgtctcttt ctgaccctcg cccacctacc ctgacttctc	60
tcttaggggt tgagcggacc ctggtgcttc ctgggagtca gacagcattc gacttgatg	120
acgttcaggc tgggcttagc tacactgtgc ggggtgtctgc tagagtgggt ccccgtagg	180
gcagcgctag cgtgctgacc gtccgccggg gtgagtactg caggaggctt gtggaggaca	240
gctgcctgcc tcaactctgg cctggttctg acttctgact tctgtctgta actcctagag	300
ccggaaactc cacttgctgt tccagggctg cgggttgtgg tgtcagatgc aacgcgagtg	360
aggggtggcct ggggacctgt ccctggagct agtggatttc ggattagctg gagcacaggc	420
agtggtcagt gtggggtgtg tgggggactg ccaaaggac caatgagggt atgggtgcca	480
gaggggacag gcaggagcca tgccagcatt tccctctgac ctgaggtccg gagtccagcc	540
agacactgcc cccagactct actgccacag acatcacagg gctgcagcct ggaaccacct	600
accagggtggc tgtgtcggta ctgcgaggca gagaggagg acctgctgca gtcacgtgg	660
ctcgaacggg tcaggccctg ccccgctccc ttggctctct gcctccattg ctctttcaga	720
cccccatgcc ttcccttgca gacccttgc tctccccaga actcctgcct ccccttcag	780
aatcc	785

<210> 29

15

< 211> 3656

< 212> ADN

< 213> Secuencia artificial

<220>

20

< 223> TALEN izquierda

<400> 29

ES 2 750 550 T3

gatcccatc gtccgcgcag gccaaagtcc gcccgcgagc ttctgcccg accccaaccg	60
gataggggtc agccgactgc agatcgtggg gtgtctgcgc ctgctggcag ccctctggat	120
ggcttgcccg ctggcgagc ggtgtcccg acccggtgc catctcccc tgcgcctca	180
cctgcgttct cggcgggcag cttcagcgat ctgctcgcgc cgttcgatcc gtcgcttctt	240
gatacatcgc ttcttgattc gatgcctgcc gtcggcacgc cgcatacagc ggctgcccc	300
gcagagtggg atgaggcgca atcggtctg cgtgcagccg atgaccgcc acccacgtg	360
cgtgtcgtg tcaactgccgc gggcgcccg cgcgccaagc cggccccgc acggcgtgct	420
gcgcaaccct ccgacgctc gccggccgc caggtggatc tacgcacgct cggctacagt	480
cagcagcagc aagagaagat caaacgaag gtgcgttcga cagtggcgca gcaccacgag	540
gcactgggtg gccatgggtt tacacacgc cacatcgtt cgctcagcca acaccggca	600
gcgttaggga ccgtcgtgt cacgtatcag cacataatca cggcgttgcc agaggcgaca	660
cacgaagaca tcgttggcgt cggcaaacag tggccggcg cagcgccct ggaggcctt	720
ctcacggatg cgggggagtt gagaggtcc cgttacagt tggacacag ccaacttgt	780
aagattgcaa aacgtggcg cgtgaccgca atggaggcag tgcattgcat gcgcaatgca	840

ES 2 750 550 T3

ctgacgggtg ccccccctgaa cctgacccccg gaccaagtgg tggctatcgc cagcaacaag 900
 ggcggcaagc aagcgctcga aacggtgcag cggctgttgc cgggtgctgtg ccaggaccat 960
 ggcctgactc cggaccaagt ggtggctatc gccagccacg atggcggcaa gcaagcgctc 1020
 gaaacggtgc agcggctgtt gccggtgctg tgccaggacc atggcctgac tccggaccaa 1080
 gtggtggcta tcgccagcca cgatggcggc aagcaagcgc tcgaaacggt gcagcggctg 1140
 ttgccggtgc tgtgccagga ccatggcctg accccggacc aagtgggtggc tatcgccagc 1200
 aacattggcg gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag 1260
 gaccatggcc tgaccccga ccaagtgggt gctatcgcca gcaacaaggg cggcaagcaa 1320
 gcgctcgaaa cgggtgcagcg gctgttgccg gtgctgtgcc aggaccatgg cctgacccccg 1380
 gaccaagtgg tggctatcgc cagcaacggt ggcggcaagc aagcgctcga aacggtgcag 1440
 cggctgttgc cgggtgctgtg ccaggaccat ggcctgaccc cggaccaagt ggtggctatc 1500
 gccagcaaca agggcggcaa gcaagcgctc gaaacggtgc agcggctgtt gccggtgctg 1560
 tgccaggacc atggcctgac cccggaccaa gtggtggcta tcgccagcaa cggtgggcgc 1620
 aagcaagcgc tcgaaacggt gcagcggctg ttgccggtgc tgtgccagga ccatggcctg 1680
 actccggacc aagtgggtggc tatcgccagc cacgatggcg gcaagcaagc gctcgaaacg 1740
 gtgcagcggc tgttgccggt gctgtgccag gaccatggcc tgactccgga ccaagtgggtg 1800
 gctatcgcca gccacgatgg cggcaagcaa gcgctcgaaa cgggtgcagcg gctgttgccg 1860
 gtgctgtgcc aggaccatgg cctgacccccg gaccaagtgg tggctatcgc cagcaacggt 1920
 ggcggcaagc aagcgctcga aacggtgcag cggctgttgc cgggtgctgtg ccaggaccat 1980
 ggcctgactc cggaccaagt ggtggctatc gccagccacg atggcggcaa gcaagcgctc 2040
 gaaacggtgc agcggctgtt gccggtgctg tgccaggacc atggcctgac cccggaccaa 2100
 gtggtggcta tcgccagcaa cattggcggc aagcaagcgc tcgaaacggt gcagcggctg 2160
 ttgccggtgc tgtgccagga ccatggcctg actccggacc aagtgggtggc tatcgccagc 2220
 cacgatggcg gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag 2280
 gaccatggcc tgaccccga ccaagtgggt gctatcgcca gcaacggtgg cggcaagcaa 2340
 gcgctcgaaa gcattgtggc ccagctgagc cggcctgac cggcgttggc cgcgttgacc 2400
 aacgaccacc tcgtcgctt ggctgcctc ggcggacgct ctgccatgga tgcagtgaaa 2460
 aagggttgc cgcacgcgc ggaattgatc agaagagtca atcgccgtat tggcgaacgc 2520
 acgtcccatc gcgttgccga ctacgcgcaa gtggttcgcg tgctggagtt tttccagtgc 2580
 cactccacc cagcgtacgc atttgatgag gccatgacgc agttcgggat gagcaggaac 2640
 gggttggtac agctctttcg cagagtgggc gtcaccgaac tcgaagccccg cggtggaacg 2700

ES 2 750 550 T3

```

ctccccccag cctcgccagcg ttgggaccgt atcctccagg catcagggat gaaaagggcc 2760
aaaccgtccc ctacttcagc tcaaaccaccg gatcaggcgt ctttgcatgc attcgccgat 2820
tcgctggagc gtgaccttga tgcgcccagc ccaatgcacg agggagatca gacgcgggca 2880
agcagccgta aacgggtcccg atcggatcgt gctgtcaccg gccctccgc acagcaggct 2940
gtcgagggtgc gcgttcccga acagcgcgat gcgctgcatt tgcccctcag ctggagggtga 3000
aaacgcccgc gtaccaggat ctggggcggc ctcccgatc cgatatctag atcccagcta 3060
gtgaaatctg aattggaaga gaagaaatct gaacttagac ataaattgaa atatgtgcc 3120
catgaatata ttgaattgat tgaaatcgca agaaattcaa ctcaggatag aatccttgaa 3180
atgaagggtga tggagttctt tatgaagggt tatgggtatc gtggtaaaca tttgggtgga 3240
tcaaggaaac cagacggagc aatttatact gtcggatctc ctattgatta cgggtgtgatc 3300
gttgatacta aggcataatc aggaggttat aatcttcaa ttggtcaagc agatgaaatg 3360
caaagatatg tcgaagagaa tcaaacaaga aacaagcata tcaaccctaa tgaatggtgg 3420
aaagtctatc catcttcagt aacagaattt aagttcttgt ttgtgagtgg tcatttcaaa 3480
ggaaactaca aagctcagct tacaagattg aatcatatca ctaattgtaa tggagctggt 3540
cttagtgtag aagagctttt gattggtgga gaaatgatta aagctggtac attgacactt 3600
gaggaagtga gaaggaaatt taataacggt gagataaact tttaatagga gctcga 3656

```

<210> 30

< 211> 3645

5 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> TALEN derecha

10

<400> 30

ES 2 750 550 T3

gatcccattc gtccgcgcag gccaaagtcc gcccgcgagc ttctgcccgg accccaaccg	60
gatagggttc agccgactgc agatcgtggg gtgtctgcgc ctgctggcag ccctctggat	120
ggcttgcccg ctccggcgac ggtgtcccgg acccggtgc catctcccc tgcgcctca	180
cctgcgttct cggcgggcag cttcagcgat ctgctccgc cgttcgatcc gtcgcttctt	240
gatacatcgc ttcttgattc gatgcctgcc gtcggcacgc cgcatacagc ggctgcccc	300
gcagagtggg atgaggcgca atcggtctg cgtgcagccg atgaccgcc acccaccgtg	360
cgtgtcgctg tcaactgccgc gcggccgcg cgcgccaagc cggccccgcg acggcgtgct	420
gcgcaaccct ccgacgcttc gccggccgcg caggtggatc tacgcacgct cggctacagt	480
cagcagcagc aagagaagat caaacgaag gtgcgttcga cagtggcgca gcaccacgag	540
gcactggtgg gccatgggtt tacacacgcg cacatcgttg cgctcagcca acaccggca	600
gcgttaggga ccgtcgctgt cacgtatcag cacataatca cggcgttgcc agaggcgaca	660

ES 2 750 550 T3

cacgaagaca tcgttggcgt cggcaaacag tggccggcg cacgcgccct ggaggccttg	720
ctcacggatg cgggggagtt gagaggtccg ccgttacagt tggacacagg ccaacttggtg	780
aagattgcaa aacgtggcgg cgtgaccgca atggaggcag tgcattgcatc gcgcaatgca	840
ctgacgggtg cccccctgaa cctgaccccg gaccaagtgg tggctatcgc cagcaacaag	900
ggcggcaagc aagcgctcga aacggtgcag cggctgttgc cgggtgctgtg ccaggacccat	960
ggcctgaccc cggaccaagt ggtggctatc gccagcaacg gtggcgga gcaagcgctc	1020
gaaacgggtg agcggctgtt gccggtgctg tgccaggacc atggcctgac tccggaccaa	1080
gtggtggcta tcgccagcca cgatggcggc aagcaagcgc tcgaaacggt gcagcggctg	1140
ttgccggtg tgtgccagga ccatggcctg actccggacc aagtgggtggc tatcgccagc	1200
cacgatggcg gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag	1260
gaccatggcc tgaccccgga ccaagtgggt gctatcgcca gcaacggtgg cggcaagcaa	1320
gcgctcgaaa cgggtgcagc gctgttgccg gtgctgtgcc aggaccatgg cctgactccg	1380
gaccaagtgg tggctatcgc cagccacgat ggcggcaagc aagcgctcga aacggtgcag	1440
cggctgttgc cgggtgctgtg ccaggacccat ggcctgactc cggaccaagt ggtggctatc	1500
gccagccacg atggcgga gcaagcgctc gaaacgggtg agcggctgtt gccggtgctg	1560
tgccaggacc atggcctgac cccggaccaa gtggtggcta tcgccagcaa cattggcggc	1620
aagcaagcgc tcgaaacggt gcagcggctg ttgccggtg tgtgccagga ccatggcctg	1680
actccggacc aagtgggtggc tatcgccagc cacgatggcg gcaagcaagc gctcgaaacg	1740
gtgcagcggc tgttgccggt gctgtgccag gaccatggcc tgaccccgga ccaagtgggtg	1800
gctatcgcca gcaacattgg cggcaagcaa gcgctcgaaa cgggtgcagc gctgttgccg	1860
gtgctgtgcc aggaccatgg cctgaccccg gaccaagtgg tggctatcgc cagcaacatt	1920
ggcggcaagc aagcgctcga aacggtgcag cggctgttgc cgggtgctgtg ccaggacccat	1980
ggcctgaccc cggaccaagt ggtggctatc gccagcaaca agggcgga gcaagcgctc	2040
gaaacgggtg agcggctgtt gccggtgctg tgccaggacc atggcctgac tccggaccaa	2100
gtggtggcta tcgccagcca cgatggcggc aagcaagcgc tcgaaacggt gcagcggctg	2160
ttgccggtg tgtgccagga ccatggcctg actccggacc aagtgggtggc tatcgccagc	2220
cacgatggcg gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag	2280
gaccatggcc tgaccccgga ccaagtgggt gctatcgcca gcaacggtgg cggcaagcaa	2340
gcgctcgaaa gcattgtggc ccagctgagc cggcctgac cggcgttggc cgcgttgacc	2400
aacgaccacc tcgtcgctt ggcctgcctc ggcggacgctc ctgccatgga tgcagtgaaa	2460
aagggtattg cgcacgcgc ggaattgatc agaagagtca atcgccgtat tggcgaacgc	2520

ES 2 750 550 T3

```

acgtcccatc gcgttgccga ctacgcgcaa gtggttcgcg tgctggagtt tttccagtgc 2580
cactcccacc cagcgtacgc atttgatgag gccatgacgc agttcgggat gagcaggaac 2640
gggttggtac agctctttcg cagagtgggc gtcaccgaac tcgaagcccg cgggtggaacg 2700
ctccccccag cctcgcagcg ttgggaccgt atcctccagg catcagggat gaaaagggcc 2760
aaaccgtccc ctacttcagc tcaaacaccg gatcaggcgt ctttgcattgc attcgcgat 2820
tcgctggagc gtgaccttga tgcgcccagc ccaatgcacg agggagatca gacgcgggca 2880
agcagccgta aacgggtccc atcggatcgt gctgtcaccg gccctccgc acagcaggct 2940
gtcagagtg cgcgttcccga acagcgcgat gcgctgcatt tgcccctcag ctggagggtgta 3000
aaacgcccgc gtaccaggat ctggggcggc ctcccgatc cgatatctag atcccagcta 3060
gtgaaatctg aattggaaga gaagaaatct gaacttagac ataaattgaa atatgtgcca 3120
catgaatata ttgaattgat tgaaatcgca agaaattcaa ctcaggatag aatccttgaa 3180
atgaaggtga tggagttctt tatgaaggtt tatggttattc gtggtaaaca tttgggtgga 3240
tcaaggaaac cagacggagc aatttatact gtcggatctc ctattgatta cgggtgtgatc 3300
gttgatacta aggcattatc aggaggttat aatcttccaa ttggtcaagc agatgaaatg 3360
caaagatatg tcgaagagaa tcaaacaaga aacaagcata tcaaccctaa tgaatggtgg 3420
aaagtctatc catcttcagt aacagaattt aagttcttgt ttgtgagtgg tcatttcaaa 3480
ggaaactaca aagctcagct tacaagattg aatcatatca ctaattgtaa tggagctggt 3540
cttagtgtag aagagctttt gattggtgga gaaatgatta aagctggtac attgacactt 3600
gaggaagtga gaaggaaatt taataacggt gagataaact tttaa 3645

```

<210> 31

< 211> 706

5 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> Brazo izquierdo del donador

10

<400> 31

ES 2 750 550 T3

```

gcctgtgagc cctgtaacag acctgcaagc caccgagctg cccgggcagc ggggtgcgagt      60
gtcctggagc ccagtccctg gtgccacca gtaccgcatc attgtgcgca gcacccaggg      120
tgaggtggac gcagccagca cccccaccac acacactgaa gttccagcct gaggggtctg      180
agtgtacctc cagccctctc cttccacacc tgggtcccta tagtctcctc cctgtctcgt      240
tatccatctg cttcccaaga tggcactgag ccctgaactg tcaacagggt gtgtctgccc      300
ctaccctacc ccaaccaatc tcctccctgc cccacacccc atccacggct ttcacctctg      360
cacagccaca agctcaatca cctgtctttc tgttaccta gtggtgacca ctgtacacca      420
acctctcatt gcatgtcccc atccagcact gacctctgcc atccacatc agtgtctgtc      480
aatgggtcat cagttctcac tgcagtccac tgactcctgt catcctacat gctaagatcc      540
cacagactcc tgtcctccag tgtccctgac catcactggc ccctttcatc ctatgtcctt      600
gtcttgcaact cacttttttc agccatatcc cagctctctg gtccccacca cccacatcc      660
ccacccccat ggccttctc actctgcgtc cctgtccatc actgcc                        706

```

- 5
- <210> 32
 - < 211> 32395
 - < 212> ADN
 - < 213> Homo sapiens

<400> 32

ES 2 750 550 T3

agagggggtg agttacttgg atccaggcca aggggacctt ggtttcccta agaccggccc	60
agagtcactc atttgccagg gcttcttgcc tgtcaaggag atccgggtgg ggcccaggag	120
gccaccaga cagatggctg aatcacagga gtggccggcg ggaccatgg cctgagggct	180
tgtctgggca cccctactgg attgggggtg agtcatcccc aactgcagcc ccacccccca	240
cggcgctgct gccttgtggc tctgcaggaa cctgtccact cctcagcctg gtcactgtga	300
ttgacctaaa gcagccaaga cctgtgacct tagatggagt taggggtact ccctcagcat	360
ctgcccctgc agaaccttct gggaaattcc cagaagccac ggggggtcgg ggggtttata	420
gttaagtgcg tcatatcggt tgtctggggg aggggtgggg ggggcggcga cctctcaggg	480
atatgggtga gggcgggtgc ctgggttccc gcctgccgct ccgcccccg agatcaggga	540
cttttctctg ctctgccga gagactgcag cggcggcggc gggagcggc ggacgcgcag	600
gcaagaccag gactcgggct ggaggggcgc tgggctcgga cctgccaaag ccacggggga	660
gcaagggaca gaggcggggg tcctagctga cggcttttac tgcctaggat gacgctgcgg	720
cttctgggtg ccgcgctctg cggcgggac ctggcagagg cggcccgagt gcgagcccag	780
cacagggaga gaggtgggca ccgcaaggga ggaccggcc cggggcactg caccgtgcca	840
ccctccgctc cactcggcct tcatcccaa cccccccgc ccacaacca gccaaactcca	900
cgacgccccg cggattcctc ctaattctgg gactccccga accctccgcg atcgggtgctg	960
ggttccctaa accgcctcct atcctgttcc taccctaaatt ctggctccct ggatgcgtgc	1020
ggggccccct gccttatgcc aatccacgcc ggcccctaga ggctgacctc agtcccaagt	1080
ccacaccctg gctggcgact ctccctcccc cgcacacccc caaccctccg cgggagtctg	1140
ctctgggacc caccgagctc aactccctcc aaggccctca ggctcaggcc gattcggccc	1200
tggtttgcat aagatgaggc ctactctccc caccatccca agtcccagtg aggccaccgt	1260
gcgggtaggc ggggtggttg cctcccagaa gaggccctga tctctcctgc ccgccccagt	1320
gacctgcacg cgcctttacg ccgctgacat tgtgttctta ctggatggct cctcatccat	1380
tggccgcagc aatttccgcg aggtccgcag ctttctcgaa gggctggtgc tgcctttctc	1440

ES 2 750 550 T3

tggagcagcc agtgcacagg gtgtgcgctt tgccacagtg cagtacagcg atgacccacg	1500
gtgagtaggg gccctggggg ctggggacca cgagatcccc accaagactc cttccccaga	1560
gagggaggag accctggaaa ccccaggcca ggtctttgca gaaacaagtt ttggggaacc	1620
tagagacccc ctaccagga cctccttttg agacagaggt aaggattccc taggtagaac	1680
cttcgtgccc tgaggcacct ccaggctact gcaagacagg aagggtgcaag acagtgaggc	1740
tgtgttttct gatctggaga tggggagccc aggagggggc cccacgcct gttggattct	1800
ccacgttcta gccctctaac ccttgacgc gttcaccaga ggtgctccat ctggccagca	1860
tccactttcc tacatctgga cccctcctcc tgcaattatt taattaagca gcatcttggt	1920
ccccacagac tgtgagcccc taagggtccc ttcaggctctg cccaggccca cctgcagatg	1980
gtgcacaaag gagccccctg ggagtttgtt aacctggggg ttccagaggc aagtgggtgt	2040
ggagaatgac agaacgaagg gacctctcaa gagagcctga taccgtaac cctcaccccta	2100
gaggcctcct caaggccagg gccagaagag atcctgagtc ctaccctgtt gccaccccta	2160
ccctccaaca ggacagagtt cgccctggat gcacttggtc ctgggggtga tgtgatccgc	2220
gccatccgtg agcttagcta caaggggggc aacctcgca caggggctgc aattctccat	2280
gtggctgacc atgtcttctc gcccagctg gcccagctg gtgtcccaa ggtgatccct	2340
accctacca tgcctccaa gatgaccca aatgaagtgt ccagggaac cgtgatttga	2400
ccctgcacc tgtcccagg ctgcatctg atcacagcg ggaagtcca ggacctggtg	2460
gacacagctg cccaaaggct gaaggggcag ggggtcaagc tatttgctgt gggtaaggac	2520
cgagcaggag tgacaggtca gctgggggt gggggcagtc agagagcatg tgggtgactg	2580
agtcctgatg ggtcgtcact tcagggatca agaattgctga ccctgaggag ctgaagcgag	2640
ttgcctcaca gccaccagt gacttcttct tcttcgtcaa tgacttcagc atcttgagga	2700
cactactgcc cctcgtttcc cgagagagtg gcacgactgc tgggtggcgtg cctgtgaccc	2760
gacctcgtga gttcctgccc acacgggtga ccctgaccta gaccccgac cccaatcccc	2820
acttggcagt gctgattcca tcctatgtgc tcctgacccc gaccccaact ctgcatcata	2880
catgcccacc ccaatcctcc tgcctgatcc cctgatcccg cattcccagc ggatgactcg	2940
acctctgctc cacgagacct ggtgctgtct gagccaagca gccaatcctt gagagtacag	3000
tggacagcgg ccagtggccc tgtgactggc tacaaggctc agtacactcc tctgacgggg	3060
ctgggacagc cactgccgag tgagcggcag gaggtaggat gtcaggagtg ataggtggtg	3120
gctggggact tggctgggca agataaagtg acctcttgcc ctgggcagggt gaacgtccca	3180
gctggtgaga ccagtgtgcg gctgcggggt ctccggccac tgaccgagta ccaagtgact	3240
gtgattgccc tctacgcca cagcatcggg gaggtgtga gcgggacagc tcggaccagt	3300

ES 2 750 550 T3

gagcaattct gccagcctct gaccccatc acctaaccac cgaccccagt accccctccc	3360
acttctgact ccatgaatcc ctggtgggac tctccccag ctgccctaga agggccggaa	3420
ctgaccatcc agaataccac agcccacagc ctccctggtg cctggcggag tgtgccaggt	3480
gccactggct accgtgtgac atggcgggtc ctgagtggtg agtgagagat gtgggctgag	3540
gggagtcccc gcgcctcaga caaggctgta gagtccctgag tctgcaaggc cactggccc	3600
cttgggtgtcc cccatgcagc tgggcccaca cagcagcagg agctgggccc tgggcagggt	3660
tcagtgttgc tgcgtgactt ggagcctggc acggactatg aggtgaccgt gagcacccta	3720
tttgcccgca gtgtggggcc cggcacttcc ctgatggctc gcaactggtga gaaggctggg	3780
cactttcttc aggtctgggac gggcaggcag ggcaaggccc ggaggccact gacatcccat	3840
gtgccctggg cagacgcttc tgttgagcag accctgcgcc cggtcacccct gggccccaca	3900
tccatcctcc tttcctggaa cttggtgcct gaggcccggtg gctaccggtt ggaatggcgg	3960
cgtgagactg gtcagtgcgg gggaagggat ggacaggcaa gggtcagggc atggcctctg	4020
ctggtctgac cctgttatct tgcaggcttg gagccaccgc agaaggtggt actgccctct	4080
gatgtgacct gctaccagtt ggatgggctg cagccgggca ctgagtaccg cctcacactc	4140
tacactctgc tggagggcca cgaggtggcc acccctgcaa ccgtggttcc cactggtgag	4200
ggctctgtgg gctggggcag tgagtggggg aggtgtcgga gcccaggct gcctctatgc	4260
tgtgctctcc aacaggacca gagctgcctg tgagccctgt aacagacctg caagccaccg	4320
agctgcccg ggcaggggtg cgagtgtcct ggagcccagt ccctggtgcc acccagtacc	4380
gcatcattgt gcgcagcacc cagggtgagg tggacgcagc cagcaccccc accacacaca	4440
ctgaagttcc agcctgaggg gtctgagtgat acctccagcc ctctccttcc acacctgggt	4500
ccctatagtc tcctccctgt ctggttatcc atctgcttcc caagatggca ctgagccctg	4560
aactgtcaac aggtgtgtgc tgcccctacc ctaccccaac caatctcctc cctgccccac	4620
accccatcca cggttttcac ctctgcacag ccacaagctc aatcacctgt ctttctgtta	4680
ccctagtggg gaccactgta caccaacctc tcattgcatg tcccatcca gcaactgacct	4740
ctgccatccc acatcagtgat ctgtcaatgg ctcatcagtt ctactgcag tccactgact	4800
cctgtcatcc tacatgctaa gatcccacag actcctgtcc tccagtgtcc ctgaccatca	4860
ctggcccctt tcacccatg tccttgtctt gcaactcactt ttttcagcca tatcccagct	4920
ctctggtccc caccacccca catccccacc cccatggccc ttctcactct gcgtccctgt	4980
ccatcactgc catcgtccca catccctgtc tctttctgac ccctgcccac ctaccctgac	5040
ttctctctta ggggttgagc ggaccctggg gcttcctggg agtcagacag cattcgactt	5100
ggatgacgtt caggctgggc ttagctacac tgtgcgggtg tctgctcgag tgggtccccg	5160
tgagggcagt gccagtgtcc tcaactgtccg ccggggtgag tactgcagga ggcttgtgga	5220

ES 2 750 550 T3

ggacagctgc ctgcctcaact ctggtccttg ttctgacttc tgacttctgt ctgtaactcc	5280
tagagccgga aactccactt gctgttccag ggctgcgggt tgtggtgtca gatgcaacgc	5340
gagtgagggg ggcctgggga cccgtccctg gagccagtgg atttcggatt agctggagca	5400
caggcagtgg tcagtgtggg gtgtgtgggg ggactgccaa agggaccaat gagggtatgg	5460
gtgccagagg ggacaggcag gagccatgcc agcatttccc tctgacctca ggtccggagt	5520
ccagccagac actgccccca gactctactg ccacagacat cacagggtg cagcctggaa	5580
ccacctacca ggtggctgtg tcggtactgc gaggcagaga ggagggccct gctgcagtca	5640
tcgtggctcg aacgggtcag gccctgcccc cgtcccttgg ctctctgcct ccattgctct	5700
ttcagacccc catgccttcc cttgcagacc cttgcttctc cccagaactc ctgcctcccc	5760
cttcagaatc ccactccctc ctccctcaga gacgtgtgtt cctcttcagt ttcagctgtc	5820
tccccttaag ccctccccgt attcagattc ctccgttctt ctctaagac tcccacagcc	5880
tttgccctca aatatccacc acttcaaccc tcagaccccc atcccttttt tgctggaccc	5940
aaatctctc tgagatttac actttctcct ccactcagacc tgctcactgc ttcctaaac	6000
ctccacctcc ccactcctcc actggaccta ttcaccatct cctctatcaa attcccgct	6060
cctgttagac tcccatcatc tccccctatc agatagcccc actccttcct gcctgttcca	6120
gcagactccc cattgcctct cccactaga cccactgggc ccagtgagga cggcccatgt	6180
gactcaggcc agcagctcat ctgtcaccat tacctggacc agggttcctg gcgccacagg	6240
atacaggggt tcctggcact cagcccacgg tggggactgg ggtttgggag ggggcaggtc	6300
agggtggaat gggggctggg ggtttattgg ggggtccagg ggggctctgg ggcacagagt	6360
ttgctagccc tggagctgcc tccatccctg tcccaggcc cagagaaatc ccagttgggt	6420
tctggggagg ccacggtggc tgagctggat ggactggagc cagatactga gtatacgggtg	6480
catgtgaggg cccatgtggc tggcgtggat gggccccctg cctctgtggt tgtgaggact	6540
ggtgagtgga ccctggccag ctactagcca caccgcatta gctaccctgc cctgctgtgt	6600
gttcctgatt gccagctgc ctccaccccc ctgctcagtg actgtcctgt ccacactgac	6660
ccaccccaca ctgattaagt gtccaccac ggtgacctcc tgggtggccc aactgccc	6720
ccccaggca ctgacctct cctcctcagg gctgcctaaa gtgacctgtc cacactaacc	6780
tcacactgac tttcagatca cccctgcgtc aaccattcct gcactgcctc tctgttctct	6840
accaggaccc cttaccttgc ctctgtcccc agccccag cctgtgggtc gtgtgtcgag	6900
gctgcagatc ctcaatgctt ccagcgacgt tctacggatc acctgggtag gggtcactgg	6960
agccacagct tacagactgg cctggggccg gagtgaaggt atggctccct gacgccaccc	7020
ctgtccttcc tggctgggac tgetcacccc taaccattgc tgtatgccca cctggccagg	7080

ES 2 750 550 T3

cgccccatg aggcaccaga tactcccagg aaacacagac tctgcagaga tccggggtct	7140
cgaaggtgga gtcagctact cagtgcgagt gactgcactt gtcggggacc gcgagggcac	7200
acctgtctcc attgttgtca ctacgcgtag gcagagcatg cgctggagag cttcggatgg	7260
gtggtgtgga tggtttgggg atccgggctt gtgctctgga ttggagaaag gaccaggatt	7320
ggtagtgagc ctttgggggc agggctctgag aggagggaga ggggtctgag aggctgggcg	7380
gggtgcgtgt gccaggggtg gcctgggatt ggtgcagggg ccatgggggc agagcctccc	7440
tgattcctga gctttctctc cagcgctga ggctccgcca gccctgggga cgcttcacgt	7500
ggtgcagcgc ggggagcact cgctgaggct gcgctgggag ccggtgccca gagcgagg	7560
cttccttctg cactggcaac ctgaggggtga gaggtgtccc caggaggaag ttagggacca	7620
ttgggggcag ggctgtggca gactccgcaa ggtaggcgag gacagtgatg gtgggcgggg	7680
tctgtcactg ggggtctgcg ggatccgtga cagtaggtaa gatcaatcaa tgaagtgggt	7740
gggatcaatg agttcatgga gggtcagtct ggcagggctc gtaaagtggg cagggtcagt	7800
gaggatgagg aagatcagtc aggaggggtga acccagttaa cagagccagt gaagtgggca	7860
ggccctttaa tgcccccag tgtcactatc ccacatgctc ttggcccca ccctcacgcc	7920
tgccccaggt ggccaggaac agtcccggt cctggggccc gagctcagca gctatcacct	7980
ggacgggctg gagccagcga cacagtaccg cgtgaggctg agtgtcctag ggccagctgg	8040
agaagggccc tctgcagag tgactgcgcg cactggtgaa cctgcctcac cttggcgtgc	8100
tcctcccctg ctgatgacct accctgactt cctgcacccc gactctgagt gactcctcct	8160
gtacccctac ccctcacct ctgacctgg gtgaccccag catgacctcc catgatgctt	8220
caataagaca aaccggacct aggatctcag atctctccct tctgggttct caggactcag	8280
cctctgatcc tcgatcttct attcctcctt cagagtcacc tcgtgttcca agcattgaac	8340
tacgtgtggt ggacacctcg atcgactcgg tgactttggc ctggactcca gtgtccagg	8400
catccagcta catcctatcc tggcgggcac tcagaggccc tggccagggt gagggggagg	8460
ccaggatttg ggtgggctgg cagttggggc tctgtggaga cagcttttta atcaagttct	8520
gtctcctgca gaagtgcctg ggtccccgca gacacttcca gggatctcaa gctcccagcg	8580
ggtgacagg ctagagcctg gcgtctctta catcttctcc ctgacgcctg tcctggatgg	8640
tgtcggggt cctgaggcat ctgtcacaca gacgccagg atagtgggcg tagtgggaag	8700
ggcagaaaag tgtgtctggg tgggtgccc gcttcagtaa cttgttcccc ttctacagt	8760
gtgccccgt ggcctggcgg atgtggtgtt cctaccacat gccactcaag acaatgctca	8820
ccgtgcggag gctacgagga gggctcctgga gcgtctggtg ttggcacttg ggcctcttg	8880
gccacaggca gttcaggttt ggcctgggg cagccagatt ctacctctcc ctgaggcccc	8940
ccaccaggc tccataccag ttaccatctc ctgaccaccc aataggttca ggcattctcc	9000

ES 2 750 550 T3

cagctctctc acaggtcca gacctcacct actgcctggc tccctatcct gccatcactg	9060
ctgactccct aatagatagg gtttggtcac acacaccctg atgtgtttct ccaggctctg	9120
tgcaccccca tccctgctct gtcattcctcc tcaggctctg tgtttccttc caccgccagt	9180
tggcctgctg tcttacagtc atcgccctc cccactgttc ccactgaatg gctcccatga	9240
ccttggcatt atcttgcaaa ggatccgtga catgccctac atggacccaa gtgggaacaa	9300
cctgggtgag gactgcagca ggcatggact cctggggcta tccactggga gcttgggtgcc	9360
ccagggttct gacatgtcct tccttcagg cacagccgtg gtcacagctc acagatacat	9420
gttggcacca gatgtcctg ggcgccgcca gcacgtacca ggggtgatgg ttctgctagt	9480
ggatgaacct ttgagaggtg acatattcag ccccatccgt gagggccagg cttctggtaa	9540
ggagtaggct gatggggaag gggctctggga ataggggtgg ccctgaaaag gctatcatgc	9600
agccactgga cccacacctta gggcttaatg tgggtgatgtt gggaatggct ggagcggacc	9660
cagagcagct gcgtcgcttg gcgccgggta tggactctgt ccagaccttc ttcgccgtgg	9720
atgatgggcc aagcctggac caggcagtcg gtggtctggc cacagccctg tgtcaggcat	9780
ccttcactac tcaggtttgg aagaggcctc tgggggactg ggtggtagaa tataaggggt	9840
ctgggggttc ttggtagaaa tgttcaacct caaagagacc ctatgtccgc agccccggcc	9900
agagccctgc ccagtgtatt gtccaaaggt aagagtcctt catcgagagg tgaacagagg	9960
ctgcacctgg ggctagccac tctgaccac atctgacatg tctttcccca gggccagaag	10020
ggggaacctg gagagatggt gagtggccct gcggggggg aatagagaat gggttgagtg	10080
tcagtggggc aggagaggtt ggcagcacag gtctcactga ttaccttcc tagggcctga	10140
gaggacaagt tgggcctcct ggcgaccctg gcctcccgtt gagtgtcccc cacatctgtc	10200
cctgcaccct gactggttta tccccacctt actgtatct tctttcaggg caggaccggt	10260
gctcccgcc cccaggggcc ccctggaagt gccactgcca agggcgagag ggtgagtgtg	10320
gagaccatcg aggtttctct gaggcactcc tcaaagctgt ctgggggttt ttacaagaga	10380
gaaggaagaa gcagagttag aatcatcggg aattcctaga agcccagctg gaggttataa	10440
accaccacag agactttggg ggacactctg aaatagaagg gccagtggg cctctagccc	10500
cgctgcagag cctgagcagc agctccaagt cccagaggca gggaagccct ctgccttgtc	10560
gctgcccttg tgcctggaat tgggctctgt gaagctctga ggggccattt ctctgcctca	10620
ctgttcacc cccaataaca cttgggggtc agtggggtag gcgaccctt gttgatgggg	10680
tgaggcccct ccagcaaacc aggggccttg ggggagggtc ccttcctgtc ctgacctctt	10740
cacctcctca gggcttccct ggagcagatg ggcgtccagg cagccctggc cgcgccggga	10800
atcctgggac ccctggagcc cctggcctaa aggtgagcaa gccttgtcct gcaggtcagg	10860

ES 2 750 550 T3

gtgggcgctg	cctgagtggg	tggggtggct	ccgactgttc	tgcctctggc	ctccatttgc	10920
agggctctcc	agggttgcct	ggccctcgtg	gggacccggg	aagggtgcctt	cccttctttg	10980
ctctctaagt	gtcttcccag	ggttcttcca	cagggtgagg	gcctggggtg	gtggtgcagt	11040
gcccacgttg	acattcgcct	gagcccaagc	accaccctct	gctctgtttc	gtcctcaggg	11100
agagcgagga	cctcgaggcc	caaaggggga	gccggtagg	gaagggggaa	gggaggcggc	11160
cgggatgtcc	cagggaggag	caggactgcc	ccacaccaga	ccctgtgcag	ggcctaaggc	11220
gcgaatagga	atagctggac	atgtctaggg	gcttcttcca	gtcgaaggcc	cccatagcct	11280
gaatcctgcc	cactgctctc	tgtccttaca	gggggctccc	ggacaagtca	tgggaggtga	11340
aggacctggg	cttctggggc	ggaaagggga	ccctggacca	tgggtaagt	caggggtatgt	11400
ggaggcaagt	gatgtgtagt	ggggggacca	acacgagggg	ggcgagagt	aggtctgtgg	11460
ggttgccact	tatactttgt	ctcctccatc	agggccccc	tggacctcgt	ggaccactgg	11520
gggacccagg	accccggtgc	ccccagggc	ttcctggaac	agccatgaag	gtgacagcct	11580
catgagtgcc	atgtgatgca	gagacctgg	gaccccat	gaaccacat	aaccctgcc	11640
agttactctg	gcccttgtga	ccctttgatt	accccatcc	tcaccatgac	gcctcagttc	11700
tcccaaatc	cttgaaatcc	aattggaccc	catgacctc	atcactcctg	gtatctttgg	11760
gagtgaggtg	tggcccagg	tcatggggc	gtcatctgtt	ttctagggg	acaaaggcga	11820
tcgtggggag	cgggtaagt	agggacagg	tgtgctagg	gtggcttgg	gtctgattcc	11880
cctgttcatt	ccctgacctg	ctgttctctc	ccagggtccc	cctggaccag	gtgaagggtg	11940
cattgctcct	ggggagcctg	ggctgccgg	gaggggcctt	gaggctctgc	tgggggccct	12000
gctcaggggt	gtgggtctct	cctggggcag	tgggtgggtg	ctgggcttca	tagttcttgg	12060
ctcatatttt	tactcacttc	ttcctagggt	cttcccgga	gccctggacc	ccaaggcccc	12120
gttgcccccc	ctggaaagaa	aggagaaaa	gtaggaaggc	tgacttgatg	atgtcccagt	12180
tctgggggtg	gaggtgcgt	gctgggggca	gggcctccct	tgggtcttcc	caccctgtg	12240
tttctccttc	aggtgactc	tgaggatgga	gctccaggcc	tcccaggaca	acctgggtct	12300
ccgggtgagc	aggtgagtgg	aggggccagg	gattctgaat	atgggtgggca	cagctccagc	12360
ccctacctca	atcatcaacc	actgctccat	cctcatgccc	aaacccaaat	ctctgaaccc	12420
ccaaattcat	ccctccagg	gcccacgggg	acctcctgga	gctattggcc	ccaaagtga	12480
taccagttgg	gggattcagg	tgtgaggggt	gctactctgg	gctcccatg	gtgttagggg	12540
aggtggaag	ataaggagat	aagagtcccc	tccaggtcag	aggtcgtgg	tttgaggggg	12600
gtggttgag	tttgggaccc	cttgtctggg	gtttgacgtt	caagccccgc	caccaaccct	12660
ctctctctct	gtctttctct	caccctctct	cttcagggtg	accggggcct	tccagggcc	12720
ctgggtgagg	ctggagagaa	ggtaagtgc	acctgggggg	tgccaagggc	cctggaggat	12780

ES 2 750 550 T3

ctgggccc	aa	ctcagctctg	acctcttctt	ttccatcagg	gcgaacgtgg	acccccaggc	12840
ccagcgggat	cccgggtaaa	cccactggct	gcaatgctca	taccagctga	cctggctgtg		12900
ccctttctgg	ttctgacttc	ttgcccttga	cccctgctac	ccctgctcct	cacccctcct		12960
caatgaccac	ttatccctgc	tgatacaggc	tctaaccctc	agccccaggg	acctggcttt		13020
gaacctctga	ccctgtgaa	ctgaccttga	ttttactga	cctggctctct	gttctcctgc		13080
caagtcttac	ccctgccaac	ctaaatccca	atcttccttg	acccctctcc	agccccacc		13140
ccagcctcta	gccctgtctg	tccatatccc	ccgtcccccac	ccacctgcac	agctcttccc		13200
ttctctcct	ccaggggctg	ccaggggttg	ctggacgtcc	tggagccaag	ggctcctgaag		13260
tgagtctgtg	actgtgggtg	gaccaggagt	gggacttttg	tgtgtccctc	ccctttccct		13320
ttccctcctg	ggctcacact	ttctctacat	tcaggggcca	ccaggacca	ctggccgcca		13380
aggagagaag	gtgggtcctc	ggctgggggt	ggcactgtct	ggtactaggg	atgtggcaga		13440
tgggacactg	ggattttggg	ctcctaggtg	actccctgac	ctgtccctgc	tcctatcctc		13500
tctccacagg	gggagcctgg	tcgccctggg	gacctgcag	tgggtggtgag	tgacgggagg		13560
atggcgctct	gagcacagca	cagcccttga	gcagtgacct	tcctatagaa	cactatctgg		13620
gctgtgattc	cacagtgtctg	ggcccgtgag	caggctggga	gctctgcggc	tctccttctg		13680
ctagaacctg	ccccagact	cttggctatg	atcctgtgac	cccaagaccg	ccatgcaggt		13740
catgagctct	ttgtgtcagt	ccattttgta	taacccttc	cctgctgtca	gcggtgactc		13800
tgtgacttct	gggcggggac	tgagctgtat	gacttccaat	tccatgtgac	ctccattcca		13860
atgaagactt	tgatcataca	acccaaggc	agggccaagc	tgtatctgtc	ctgtttgttt		13920
tcagggacct	gctgttgctg	gacccaaagg	agaaaaggta	agcctggtat	ggggcaaggg		13980
gaggtttcta	cagggttgag	gtctaggtca	tagggcctat	ctatgggact	tggggggtca		14040
caggacttgc	tgggtcaggg	ggttaactgg	agcctgggac	tagcactgat	ggtctttgtc		14100
acctccaggg	agatgtgggg	cccgttgggc	ccagaggagc	taccggagtc	caaggggaac		14160
gggtaagtga	agcgaagtgt	ttagggggca	gttggtgaag	gttgtcttcc	tgacttctta		14220
tccttccatc	cacagggcc	accgggcttg	gttcttctctg	gagaccctgg	ccccaagggga		14280
gaccttgag	accgggtgaa	tcaatgtggg	aatggggagt	gtgacagagg	gagatgaggt		14340
ggtgggacct	tgactaagtc	ctgccccct	tctgtcccct	tcagggtccc	attggcctta		14400
ctggcagagc	aggaccccca	gtgagtacct	gttaccctgg	gcaacctcaa	ggcttctggg		14460
gtccccctccc	cttgagaact	gcttgcttcg	agcgtcctgc	atcacctccc	tcttgccctcc		14520
tccacagggt	gactcagggc	ctcctggaga	gaaggagagc	cctgggcggc	ctggcccccc		14580
aggacctgtt	ggcccccgag	gacgagatgt	aagaggctgg	agtcggggga	gtcatggcgg		14640

ES 2 750 550 T3

gtaagggagt	agggctgttg	ccagcatcat	gggggttctt	ggaaccaggg	ctgactctca	14700
tgtttcacag	ggtgaagttg	gagagaaagg	tgacgaggg	cctccggtga	gactccttcc	14760
cactgtggtt	tctgatcctt	tacccttgaa	ctaggatccc	agtaggctgg	tgctccacca	14820
gttcatccat	ccactccctg	cctcctgtcc	agctgctgct	cagacccttc	tctgtccctt	14880
ctccctgagt	gagttagatc	ctgactgccc	tgtgcagtat	gacttttctc	tctatcacca	14940
gggtgaccgg	ggtttgcctg	gaaaagcagg	cgagcgtggc	cttcgggtga	gtcttggcag	15000
agagaagtaa	caggggtgat	gggaggtggg	catgaaggtg	ataggaaagg	ctgagggggg	15060
taaggggtga	tgggagtcct	tgcagggagg	catgggggtg	tgggaacctc	tgatgtggat	15120
tttggagtaa	tggggaacct	ggggcagtg	agcgggtcat	aggggcacct	gcaaaatatt	15180
gggaggggtc	gtctcaggcc	agctgctctt	cttagggggc	acctggagtt	cgggggcctg	15240
tgggtgaaaa	gggagaccag	ggagatcctg	gagaggatgg	acgaaatgtg	agtcccagcc	15300
tatgactcct	cacccaaccc	ttaacctccc	aaccagccaa	ttcccaatac	ctgatcctac	15360
ccccaacccc	tgtaatgcta	aacctcccca	atcttcactt	cctttgacct	ctgcacacac	15420
gcatctgaag	gctaccaaca	ttcccatgag	tcctcatgg	gcttccaaag	ctctcccaaa	15480
ctgctgccta	ccttgagtgg	ccctgacccc	tgcctttctg	cagggcagcc	ctggatcatc	15540
tggaccaag	ggtgaccgtg	gggagccgg	gagtagggct	ggtgtcctgg	gctcagaagg	15600
gatggaggg	cccctggcgc	tgctgccagt	gtgccttcac	acaaggggtc	tggaggtcca	15660
ggtgggaagc	atggatggcc	acccatgcct	gctgggggct	ggcattgggg	tgcttcgggg	15720
aacttggggg	cagagttgag	cctggggcaa	accagagcca	tggcctggga	cctgagggct	15780
cctattgact	ataatcattt	cctttcccag	ggtccccccag	gacccccggg	acggctggta	15840
agggctgcgc	tgggtctggt	cctctgtcat	tgttctcact	tgcccccttg	gctccacatc	15900
atgtcccctc	attccttcca	ctaactccca	cttccccccac	tgtccctctg	aaacctccgc	15960
ggatccctgg	agccccctac	agaccctgta	tcccctggcc	aactcctctt	ccccttggtg	16020
agccattctt	ctcaactggc	attcccccca	caggtagaca	caggacctgg	agccagagag	16080
aaggtattag	ggtctgtggg	tggagggtaa	gaaagacccc	agtgcccttc	ccagcaggtt	16140
ctaccttggg	catggcttgg	cttcaaggct	gttcctcagc	aagcttatct	ctgccacagg	16200
gagagcctgg	ggaccgcgga	caagagggtc	ctcgagggcc	caaggggtgat	cctggcctcc	16260
ctggagcccc	tggggaaagg	gtgagtgtga	ttggtcctcg	ggggtgcagg	catgggaggg	16320
cctgctctga	tttcttcctc	ccctatcctc	agggcattga	agggtttcgg	ggacccccag	16380
gcccacaggt	gagtgatgca	ctttgccccg	tctgccaagt	ccccatctta	ccctcacctc	16440
ttatctgacc	ctgttccctc	cagggggacc	caggtgtccg	agggccagca	ggagaaaagg	16500
tgagaggggtg	tgggggtttc	tcaggacatg	agcctggggt	atcagatcca	cctcagcctt	16560

ES 2 750 550 T3

ggtggcctct taccactcta ttttccacag ggtgaccggg gtccccctgg gctggatggc 16620
 cggagcggac tggatgggaa accaggagcc gctgggccct ctgggccgaa tgtgagtctt 16680
 ggtagtcctg cctgggtgtc cccttccccct accccttcac atgaggaccc tagacccag 16740
 cctcattggt tggctctggc cttgacgagc tctggggcga gacctgttgt ctgtgtcttc 16800
 tcccagagcc ctgcctgggt ggggcatgtt cttgccatgc aggccttagg ctcacaggga 16860
 ctagagcccc tgaccccatg accatgatga actgactttg agtctctcct cagggtgctg 16920
 caggcaaagc tggggaccca gggagagacg taagtgaggg gagatgctgg gacagagggg 16980
 gctcggggct gcgtaagctc caaccagatc atcatagtca cagcatccga gagagttggt 17040
 ggggggtcgg ttcagccttg cctttgggag ggtctcagtc cctggcacac aggttttagt 17100
 agaatcccc tttctgacct tctttgcctt gggggtcttc ataggccctc caagtcctgg 17160
 gggttcttcc cttgggagtc tcagcagggg ccccatcca tggggctctgt gcctgaggga 17220
 tctcaatggg acctctcagt cctgggggtc tctaccctgg agggagatct ttttggggtc 17280
 ctcttagtcc atggggtgtg cccaaggga tatctcagag gccctataaa tcttaacaaa 17340
 gactttttcc aggggcttcc aggcctccgt ggagaacagg gcctccctgg cccctctggt 17400
 cccctggat taccggtgag accagacttt catgttacct ccttttcccc cttactacct 17460
 tcacccgatc cccgacatcc aaccagtgat ctgttccac agggaaagcc aggcgaggat 17520
 ggcaaacctg gcctgaatgg aaaaaacgtg agtgtgtcca gggcagctgc ggcgaaacct 17580
 gccaaagaaac gccagcacac ctacacagcc aagatctgta gcaacacatg aggcacacgt 17640
 gtagacacat gcatcctggc ccagacatac agatgcatcc agaaacatag tttttttttt 17700
 tttgagatgg agtttcgctc ttgttgccca ggttgagtg cagtggcaca atctcagctc 17760
 actgcaacct tcgactctcg ggttcaagt attctcctgc ctcaacctcc cgagtaactg 17820
 ggattacagg catgcgccac catgcctggc taattttgta tttttagtag agatgggggt 17880
 tctccatgtt agtcaggctg atcacaaact gacctcaggt gatccgcccg cctcggcctc 17940
 ctaaagtgtc gggattacag gcgtgagcca ttgcgccagg cctggaaaca tagttctata 18000
 tgacacacaca ggcagcctca cagacacata gggagttagg tgaatgggcc aatttatgcc 18060
 aggatgggct gacacatgcc cagataacct acagccatag gcacaggtga agagccacag 18120
 acaggtcgac aggcacaaac aactgggtgt gtagccacat ttggggcact tagagacatg 18180
 ttgataccca gacacttcca cccagactca tgtgtctaaga catggacca agaccgatg 18240
 tggacctgtg ttgagacatg gtgagacaca catggatggc tctgacctgt gacatgtgct 18300
 aggacatgct ggcccatgac tttatatggt caccacacag cttatggtca tacacattga 18360
 gctacacaca agagttgcag acagatgcca agacacatgg ctctcaaggc atgtataagc 18420

ES 2 750 550 T3

ccatgctctt	tggagacatg	tataagccca	tgcccatatg	cagacatggg	ccaagatgtg	18480
caaacagacc	aggacacaca	gacccatggg	tacatgtact	aagacacaca	tggatatgca	18540
cacacccagg	acatgtggac	gatcaccaac	ccagagatgc	ataaatgtgg	agctacctct	18600
agagagccag	agtcaaagga	catgtgtgca	gttgctgtt	ggggtgagca	ggcctatgca	18660
ggcacaccct	tagacatgta	gcagctcact	cagacggcca	caggccatgc	tccaagacac	18720
acacaaccat	ggagctacag	gcacagaaat	acagtcacag	tcatctgaaa	tctagcaagg	18780
tagtgtcttg	cagccagaca	ccagtgaatg	tttgggtga	tgtgagtcct	ctgcccacag	18840
ggagaacctg	gggaccttg	agaagacggg	aggaaggtaa	agtccccac	ctggggtctc	18900
tctggtctct	gcttgaggcc	atgcctgaag	catcctgtc	ttccttaggg	agagaaagga	18960
gattcaggcg	cctctgggag	agaagtgagt	attggagttt	tctgcaacct	ctgacctctg	19020
acctgaccc	tgggggaata	tgactccact	ctttctgag	gtctcagggt	gctgatgctg	19080
gctgcacct	tcacagggtc	gtgatggccc	caagggtgag	cgtggagctc	ctggtatcct	19140
tggacccag	gggcctccag	gcctccagg	gccagtgggc	cctcctggcc	aggtgagtgt	19200
cctgggtcat	tctgggactt	cagagcatta	gaagccatga	tgtttcaatg	ggcaaccctc	19260
ctgggaagcc	atgggccttt	gtgacccctt	tgtgttctgc	cgtatctcag	ggttttctctg	19320
gtgtcccagg	aggcacgggc	cccaagggtga	gtgcggatgt	tgggtagggg	gcgtggtgag	19380
ggggctgacc	aggctggggg	ccatttcccc	acctggtcat	tcttgttttc	agggtgaccg	19440
tggggagact	ggatccaaag	gggagcagg	gaggccccc	ccttttccac	atgccagggt	19500
agccacagca	cccacatgtg	cacatgcaca	cccataggct	ggttcctagc	agtttgtctt	19560
catctctcca	gggcctccct	ggagagcgtg	gcctgcgagg	agagcctgga	agtgtgccgg	19620
tgagccaggc	ttgggatgtc	cccttgactc	tgttttgcat	gccattggg	gccaccttgc	19680
catccccttc	cccttgccat	gaggtccat	aggttctgtg	ctgtgtgttc	agtgccctgc	19740
cccattgggg	ttcttgtagc	ccacactcaa	gggaacttgg	gcagtgggga	cacaccagaa	19800
aggggctccc	aggagtctcc	agccatgcct	caaccaagtg	ctaaagggtg	ctctgggcaa	19860
gaggcctagg	gaaaagggtg	tgaaggtgct	ttcctgaggc	cgtgcggggc	agaaggcagg	19920
agcttctctg	tcatgggcag	cccttcaccc	agactttgtc	cccagaatgt	ggatcggttg	19980
ctggaaactg	ctggcatcaa	ggtgggttgt	ttaggggctg	gggtaggga	caagtggggg	20040
cccttggggc	tagtgggtgc	cacaggcata	ggggctgcgg	cgacgcaccc	cgctcctctg	20100
acctcttgct	gtccctcag	catctgccct	gcgggagatc	gtggagacct	gggatgagag	20160
ctctggtagc	ttcctgcctg	tgcccgaacg	gcgtcgaggc	cccaaggggg	actcaggcga	20220
acagggcccc	ccaggcaagg	aggtgagcag	aagtggctca	gtgggttgtg	ccccgtggag	20280
tgggggtgtag	ctgtacagcc	accagcattc	tctcttcac	tcctgcaggg	cccatcggc	20340

ES 2 750 550 T3

tttcctggag aacgcgggct gaagggcgac cgtggagacc ctggccctca ggggccacct	20400
ggtctggccc ttggggagag gggccccccc gggccttcg gccttgccgg ggagcctgga	20460
aagcctggta ttcccgggct ccagggcagg gctgggggtg tgggagaggc aggaaggcca	20520
ggagagaggg tgaggctggg ggctggccag gagagtgagg gaagaggggt tgggaggggt	20580
gggaccccc atgggcttgg ccctcaccgg ctatttgcac ttcagggaga acggggagag	20640
aaaggagaac gtggagaaca ggtgggctgc gatgggcttc gtggggcagg ctgtctggag	20700
gctgtgctgg ggctgccacc ccattttctt gtttcctgca gggcagagat ggccctcctg	20760
gactccctgg aaccctggg ccccccggac cccctggccc caaggtgatc accccatccc	20820
tgccttagtc ctgtgactag tgaccaggaa gccaccctta gcttggtccc cagaaatatg	20880
gtagtgctgt ccataaccct ggaatttctg accctataac cctctgtgat cctgagatct	20940
gtgatgactc ccccatgcct ctatgacaga gacatctctc ccctgtgacc ttgtgtttgt	21000
aggtgtctgt ggatgagcca ggtcctggac tctctggaga acagggaccc cctggactca	21060
agggtgctaa ggtcagtgtg tggaaacagc tcggggccac cctctgccat ggcactaggg	21120
actgacttga catctcatcc ccacaggggg agccgggcag caatggtgac caaggtccca	21180
aaggagacag ggtgaggcct ctctccaccc ttccatagag tccccctcct ttctgggggc	21240
acactagagg tgggtgtgat atgcacctgg gcacgtggca gggactgggg gctcagggca	21300
cgacactctg ccttcagggg gtgccaggca tcaaaggaga ccggggagag cctggaccga	21360
ggggtcagga cggcaaccgg gtaagtcctt gcccaacagc cacacatgtg caagaagggtg	21420
gctctcacat gtattgtcct gtgtgcaggg ctggggctctg tactgcctgg gactgtctgg	21480
gtcctgactc tgtctagggg gatggtgggt ggaggggaag ttggaactgg gaaactaagg	21540
ccttaaaccct attctctgca ggtctacca ggagagcgtg gtatggctgg gcctgaaggg	21600
aagccggtga gtgggtggctg aagcacctgg ccccaggctc cggaccctcc gctaggtgct	21660
gctgctgctg tgtgtgtgca tgtctgtgtg tgtgtgtttg tgtgtacctg tgtctgtgtg	21720
cctctctgtg tgtgcctgct tgtgtgtgcc tgtgtctgtg tgtgcctgtg tgtgcctgtg	21780
tctgtgtgtg gttgtatgtg gatgtgtgtg tgcaggcccg tgtgtgctat ctatgagagg	21840
cagtccatgg gtaggtatta tctgtgactg gaaaggggtg ggtatggaaa ttgaccccca	21900
aggaaaaagc cccagaggtg tgggaacagg cccaagttag gccagattg aggctcatca	21960
gtgccctctc tatgtagggt ctgcagggtc caagaggccc ccctggccca gtggtgagta	22020
cccaagaacc ttcacctgtc ttgcccccat cctgtgccct gcccagtgga ccagtactgc	22080
ctcagtttcc ttgggtgggg gcggctaact cccctcatc agactctttt tcgccacaac	22140
aggggtggtca tggagaccct ggaccacctg gtgccccggg gagtgaccag ggaacactgc	22200

ES 2 750 550 T3

ctggtgaggg tctggaaggg ctgggatagg cattggccac agctgatgag ccaggccttc	22260
tctgtgttaa tccctgagcc ctgttccctg cccttgaccc ttttctctag ggtcttgctg	22320
gccctgcagg accccaagga ccttctggcc tgaaggtagg tctaggtgtg tggataggag	22380
gaggaggctc cttcaagctg tgtccatgcc tggggtagtg tgcgccaacc tcctgggctg	22440
tcatctcctg caatgaggat gagctccagg agccctggcc acgtgggctc tgctcatgca	22500
gtctctgggt tgtttgcagg gggagcctgg agagacagga cctccaggac gggtagtggt	22560
cctagctcac aggttagggg catagggaga tgggtggggg tggcactgcc ctgaactttc	22620
tcttctcca gggcctgact ggacctactg gagctgtggg acttctgga cccccggcc	22680
cttcaggcct tgtggtgagt gagtccctgt ggccctgta gggacaccgt gttttcactc	22740
cttggggccc atgttctctc atgtcgtcct gtgtccattg tcacctgac atccgacttg	22800
ttctccgtca gggccacag gggctccag gtttgccctg acaagtggg agttctgggg	22860
gtcaaggggt gggctccagg ggtcaagggg cgacaggcag ccctgacaga gctcttccct	22920
ctcaggggga gacagggaa cggggagccc caggctcaga tggtgccagt ggaaaagatg	22980
gagacagagg gagccctggt gtgccagtat gtgttctggg ggcagctcgc taggggtgtg	23040
tgccagctg tgggcctgaa atatgaggag tggggcagca ggggtggtgg tggagaggca	23100
ctgagttcct cacgctgctc tgccataggg gtcaccagggt ctgcctggcc ctgtcggacc	23160
taaaggagaa cctggcccca cgggggcccc tggacagggt atctttgacc ctgacttcca	23220
ccccctgcag caactcctct gcctcaccca caagcctggt tccaaatgcc atgggggtgg	23280
cagggtgggg gcgggggagg aggaggaaac tctactagcat tccccacagg ctgtggtcgg	23340
gctccctgga gcaaaggag agaaggtagg tgtgtgtggg gctgccagtg agggggggctc	23400
aactggtggg ggccaaggaa tccactgac ctctccccct taccaggag cccctggagg	23460
ccttgctgga gacctggtg gtgagccggt aagtagggaa cttctgacag cagatgttct	23520
gggggtcctg tctctccagt ggcctaactt ctgaccttcg acccatagtt taccaccct	23580
catgaccctc agctttcata gtagaccgca tatttaagct ctggcccat gcctccctcc	23640
ggagtctcat ctttctgtga ctgatgcctg tgttgctcc tgacctctgt caacaggag	23700
ccaaagggtga ccgaggactg ccagggccgc gaggcgagaa ggtgaggtgg gttggccctg	23760
gggcctgact actgagcaga gaaggctcag tccagacacc cctcacctgc cattctgtgc	23820
agggtagaagc tggccgtgca ggggagccc gagaccctgg ggaagatgtg agtccggggc	23880
ctaggcaagg gcgagcctg cctgaggagt gtgatggcg gcataagggg cactcttg	23940
gccggggcac atgtctgagc ccctgagtcg ggctgcattg cccaacatcc aggagcctgc	24000
tctctggatc acagggagcc acacactgtc ggcttccaca ctctgagctc tgggggccc	24060
gtcttggtt ctgtgtgcc atccttggtc tgtttctcac cacatctggg cacacgtgtg	24120

ES 2 750 550 T3

ctctggtctg atctcccaga tcccaggaca cacctgagtt cttgtgaacc tctccttggg	24180
cctgtgttta caatcctcct tccccctgcc cctggctgcc agcgttcagc ctgtggccat	24240
gcctgctcta cccgggagtg gacgtgttg ggttccccctc tgagcgcccg tgtgtccgca	24300
ctcacgtctg tggagccaga tgtctgcaact cgcgtgtggg ctccctgtgc ctgtgccatg	24360
cctgaactcc cacctgtcta tggtagtata tgagcgtccc tctgtctgtg ctgtcctgag	24420
ataggccatg gtcattgtctg agctcctgtg agccaattct tggtcgcatg tctgagttcc	24480
catgtgttca tggtcacatc agaactcccc tgggaatatt ttcagcccgt gtctgaactc	24540
tgtgtctcatg ttcctacctt ctcaaagtct gtttctgtggg ttttcttagg gtcagaaagg	24600
ggctccagga cccaaagggt tcaaggtagt tgtaccacaga agggtccttg ctggggctct	24660
ggctcgtgaga ctccctgagc ttgatccgat gcctcttttc ctcaaagggt gaccacaggag	24720
tcggggtccc gggctcccc tggcctcctg gccctccagg tgtgaaggta agtcaatgcc	24780
ccatcaccag ttgtaggggc agcaggcccg gccccacag gaggaagagg gagttctgat	24840
gagagtcctg ggagggtgct tgcattgacc attccctccc tttgctgttt ttatttcagg	24900
gagatctggg cctccctggc ctgcccgggt ctccctgggtgt tgttgggttc ccgggtcaga	24960
caggccctcg aggagagatg ggtcagccag gccctagtgg agagcgggtg aggggctggg	25020
aacagcatgt aggggcatgg tgaatccatg gtgggtgagg agggggccct gccttggggg	25080
ttcccagtct tgagggggca gagggtagga gggttcccaa gtcactctct gcctctcttg	25140
ccccattttt ctggtagggg ctggcaggcc cccagggag agaaggaatc ccaggacccc	25200
tggggccacc tggaccaccg gggtcagtgg tgagtagagg tgcctaaag cccacgtat	25260
ttgatttcct gtcctcgtga ggacttagga tggggcgggc cactactgtg ggtgttgggg	25320
gaggcgcttt gagagccaca ggacctcac ctactctga cctacaggga ccacctgggg	25380
cctctggact caaaggagac aaggtaggtg ggacaagtgc tgctgactct ctcttgtgcc	25440
ctggtcaccc ccttcagcct cactgacct ccatgacct tgcgtcccc actgattccc	25500
ctcactgata cccggtgtcc tgttgggtct caggagagacc ctggagtagg gctgcctggg	25560
ccccgaggcg agcgtgggga gccaggcatc cgggtacgta tgtcttactc cacagccgaa	25620
ctcccttcat ccagctctt gcctctgatt tccaacctct gactcaatga acctaattgc	25680
accatccagg gtgaagatgg ccgccccggc caggaggagc cccgaggact cacggtgggt	25740
cccgtgggg aaagtgacag tgctgtgact tcagtcacct gcctgtgccc tttgtacccc	25800
aggactctct gccatcctcc ctagccctct ggccctctc gtctacccct gtcccctttg	25860
tctggtcccc actgtttccc tcctccccct gatcctctc cgtcctccac tgcctgggtt	25920
cctgtagctc acggtcacct tctctcacac aggggcccc tggcagcagg ggagagcgtg	25980

ES 2 750 550 T3

gggagaaggt aggaacgtgg ggaaggctcct ggcagtagtg ggggggtgag tggatcttgt	26040
atgatattag aattcagggt cggggcctcc ccagtttggg cgggactcac actcttttgc	26100
ccaatagggt gatgttggga gtgcaggact aaagggtgac aaggtagtg tgggcatggg	26160
gagggcaggg caggggccag ggtgctccc actcctctga ttcctgcctg cccctcagg	26220
gagactcagc tgtgatcctg gggcctccag gcccacggg tgccaagggg gacatggtga	26280
gtgggcccac gtgtggactg ggtctcccct ggggtgcgaga gggaccccg ggggctggcc	26340
cagggcatct gagagctcaa actcccagag ggtgagggac cttcagcccc ttaccgtgac	26400
ttcttaccba aaccaacca ggggtgaacga gggcctcggg gcttggtatg tgacaaagga	26460
cctcggggag acaatgggga ccctggtgac aaggtagcag gaagagggtg gtgtactctc	26520
tgtggggtcc tgtggctgtg gtggtccac agcatctgtg agggcaggag gggaggacag	26580
gaggaggctt tctggggttt agggggaagg ggttctgttg ctgactatgg agtggccgag	26640
aggaggggct ctagggagg ctcaaggcct tgggtgtagg ttaggagagc tctgtggact	26700
ggccctatga aggggcccta tagggacagg aaaggggctc cttgggtgtc cccatctgaa	26760
ggtgcttcag atgggatgag gaaccacagt ataagagggc ggggcaaggg tctggagtcc	26820
ttggggtgag aaatgagcca gtctgtggac tacagcccc agttccacc gctatggcac	26880
aggctgggag gacgtgctgc tagggccgtt cgggagtgac ttgagctctg cccagctgg	26940
cgggctcgtt gtattctaag ccccagcct gatctggtcc tgactgaatc cctctcctgg	27000
tcactcccac agggcagcaa gggagagcct ggtgacaagg gctcagccgg gttgccagga	27060
ctgcgtggac tcctgggacc ccaggcaagt tctgccccag gccaggccc agggccagg	27120
ccagagccag catcacacac acacatatcc tgggggtcct cttaagtgt gggggttctg	27180
gtcccttgc tgccggcccc atgtgtgttc acccgggac tgtgtgtggc ctgggccctg	27240
gccgctgagc atctgagtgg ccacacatgc agcggatggg ccagtgttg ggacagtccc	27300
actgaagggc tccccatct gtgtttctca cagggtcaac ctggtgcagc agggatccct	27360
ggtgaccgg taagatgccc ctgctccac aggggcccc tgtcccaga gggccccgt	27420
caccatggtg gtcccaccag tgtggtagct tttgacctta tagtgacca gtgctctcat	27480
ggtgtcctca tgtcaccag ggtccccac ctccacagta gccgcctatg atggtgtccc	27540
ctgcctatgg tgacccttgt ccccatggca agccccagt gctctggcct ccagaggtga	27600
ctcccacccc atggtgttcc ctaccacta tccaggcgga ttctctttgg tccctcactg	27660
acctcctctg acctcacatg gacctctcc ctgctagga tcccaggaa aggatggagt	27720
gcctggtatc cgaggagaaa aaggagatgt tggcttcatg ggtccccgg gcctcaagg	27780
aggaaagaaa caagattggt tttttctg tcaaggaagg ctaatgttga agggcgagg	27840
acctggggtc agatattag gcactgggtc agagggttga gtagctgagg cctgttttga	27900

ES 2 750 550 T3

gagttagttt	caggggtgaag	gggggtcacag	agtgaggtgt	ctgggaatca	cttcttcatc	27960
tgcTTTTctc	cctcaggggtg	aacggggaggt	gaaggagagcc	tgtggccttg	atggagagaa	28020
gggagacaag	gtacagaggg	gatgggggct	ggggggctac	gtggcctggg	gcctaaggct	28080
cacccactg	cctgatgcct	gcagggagaa	gctggtcccc	caggccgccc	cgggtctggca	28140
ggacacaaag	gagagatggt	gagtgtgggc	acgctcagaa	tgagggggcc	acgggtggac	28200
gggggccaag	ttcatgtctt	tttttctcca	gggggagcct	ggtgtgccgg	gccagtcggg	28260
ggcccttggc	aaggagggcc	tgatcgggtcc	caaggtacag	ggtctgtgga	cagtggacat	28320
gacagggagt	caggatgggg	cagggaaagct	catggtcggg	tcatgggggt	gtagggacaa	28380
ggaaggctgt	agagggttgg	gactttgggg	gctgaagtgg	ggctggtctc	gaggtgtcaa	28440
ggtgggttgg	ggtcacagtg	gagttaggat	gaaagcatgg	gaggtggagg	gggcatgtgc	28500
atcagggcag	ggggctatag	tggagttgga	gtcatgggga	tcacatctgg	ttggctcaca	28560
ggtgcttggg	ttacagaagg	tttgggcatt	aggccaggga	gctcatggga	gttcagggag	28620
gttccagagc	tgagggaggt	cagggcagaa	ggtctctcat	gcttcttctg	ttcccagggt	28680
gaccgaggct	ttgacgggca	gccaggcccc	aagggtgacc	agggcgagaa	aggggagcgg	28740
gtgagttgaa	gccatggtca	tggtgaattg	aggggagcgt	gcactttggg	agagggtcca	28800
tctctggggc	ctctgattac	tgactctgtc	ctatatctcc	ccagggaacc	ccaggaattg	28860
ggggcttccc	aggccccagt	ggaaatgatg	gctctgctgg	tccccagggy	ccacctggca	28920
gtgttgggtcc	cagaggcccc	gaaggacttc	agggccagaa	ggtaaggggc	cccagctctg	28980
actcctgata	cctgaacctt	ccctaccctt	ccaacctcat	aacctcttc	cagcctgcct	29040
gccccaacct	ctgaagctgt	ggcccccaac	ctgcctgact	ctgatcccca	ctgcctctga	29100
ccctgctcac	ttggtccctg	tgtctgacag	ggtgagcgag	gtccccccgg	agagagagtg	29160
gtgggggctc	ctgggggtccc	tggagctcct	ggcgagagag	gggagcaggt	gagtagggat	29220
tccaaggctt	gggtcagaga	tcgggggtgac	ttctgttgtc	cctgaggtca	gaggtcacag	29280
cctgggtccca	tctgttgcac	atagggggcgg	ccagggcctg	ccggtcctcg	aggcgagaag	29340
ggagaagctg	cactgacggt	gagtgtgggc	ctggatgggc	ctgggagggc	ctgggtgggc	29400
ctgggtgggc	tggggcccta	cctccctcac	ccagaccct	gaccctggg	ccctggctcc	29460
atgcagtctc	accatagtcc	ctgtattatg	tgcctatgt	ccttcctgtg	ggccatgggt	29520
tctttatggt	ccctgtggtc	ttctgctccc	aggaggatga	catccggggc	tttgtgcgcc	29580
aagagatgag	tcagcaactgt	ggtgagtggg	gccagccccg	cagtctccca	ctccacccca	29640
gcacccatag	caagggcagg	caggccccta	gaacttacag	ggcaagcagt	caagaagatg	29700
gggggatgga	tggatacaca	gaaggacaca	tgtgctgcag	gactgacaca	tgacatgtgt	29760

ES 2 750 550 T3

ccccagtgga	gggagacaca	caggcagatg	aggattgccca	tgcagtgtctc	tcagatgtcc	29820
agcttggctg	tgtggggagt	gggatgatgg	tgggagcaga	gctgggtcccc	ttgggcctga	29880
cctggaccgg	gtgggagggg	catcagagtg	aagctgtctc	ttcccgtctc	tgtctgcacc	29940
acctgcctgt	ttctgtgtct	ggcctgcttc	tgtctcttgc	cttttgttgg	cctgttccca	30000
acttccctct	cctctgcctt	cttctctttt	tccatctctc	tatctacctc	ccaccctctc	30060
tcttccctct	tctcctgtta	ccctctcctg	ctatctcttt	gtgtatctct	acccctctgc	30120
ctgtgtgtct	ctgtctgtct	ctccatcttc	ccatccttct	ctctgtcatt	gtctctctat	30180
ccctctctgc	cccctctagc	ctgccagggc	cagttcatcg	catctggatc	acgtgagtag	30240
ttttctactc	ccagaacttt	cttcacccca	ggccctgccc	tgcctatcaa	ctggggctct	30300
ctcagggggg	ttgctgggga	tggctgccc	tgggtgacttc	agggccctga	ggcccctgct	30360
cttggctcca	ggaccctcc	ctagtattgc	tgcagacact	gccggctccc	agctccatgc	30420
tgtgcctgtg	ctccgcgtct	ctcatgcaga	ggaggaaggt	gaggacagct	gaaccctggg	30480
ggcagctatg	ggtggggccg	agacacgcac	atgggtgtcc	atgaatgcag	ggcacacgcc	30540
aagcacgtag	ggtctgcatg	cagggcacac	gcatgggcac	tgtgtgcaca	cagtggaaat	30600
cagtgtgcc	caccttgccc	cggggccagc	agccactgct	cccagcacac	cctgccctac	30660
ctgcagagcg	ggtacccct	gaggatgatg	agtactctga	atactccgag	tattctgtgg	30720
aggagtacca	ggaccctgaa	gctccttggg	atagtgatgg	tgagaatggg	gggctgccag	30780
cggggtctgg	ggaggggcag	gcagggctga	gccctgctga	cctccccctg	acctttcaac	30840
cctctctgat	tcccacaaac	cctgtctgact	tgacccatt	ggcccagacc	cctgttccct	30900
gccactggat	gagggctcct	gcactgccta	cacctgcgc	tgggtaccatc	gggctgtgac	30960
aggcagcaca	gaggcctgtc	acccttttgt	ctatggtggc	tgtggagggga	atgccaaaccg	31020
ttttgggacc	cgtgaggcct	gcgagcgccg	ctgccacccc	cgggtggtcc	agagccaggg	31080
gacaggtatg	ggctgagccc	ccaccgtggg	gaactgggca	ctgagcctgc	ctggatcggg	31140
ttctggggga	ggagtccctg	ggccaggggt	ccaggtcagg	gtcctggagg	agacgtctcc	31200
tcgcagtagg	ggacctgggg	cagacgccc	gaccaaagag	ctgaatatag	agccccagcc	31260
gtggagcccc	cagtagggtc	cccttccatg	ttccctcctt	taaagacct	agtatggacc	31320
cctctgaggt	cagagcccc	acttctgtt	gtagcctccg	ctccctcccc	ttggcggtgc	31380
ctctgcctga	gcgtctccgg	ggaaggtcag	atggctgacg	accgtttcca	acctgtcctc	31440
accaggtact	gcccaggact	gaggcccaga	taatgagctg	agattcagca	tcccctggag	31500
gagtcgggg	ctcagcagaa	ccccactgtc	cctccccttg	gtgctagagg	cttgtgtgca	31560
cgtgagcgtg	cgtgtgcacg	tccgttattt	cagtgaactg	gtcccgtggg	tctagccttc	31620
ccccctgtgg	acaaaccccc	attgtggctc	ctgccaccct	ggcagatgac	tactgtggg	31680

ES 2 750 550 T3

```

ggggtggctg tgggcagtga gcggatgtga ctggcgtctg acccgcccct tgaccaagc 31740
ctgtgatgac atggtgctga ttctgggggg cattaaagct gctgttttaa aaggctcctg 31800
ttgtgactgt ttgggaagat ggggggtttc aagggggaag gttttccttg gggggttgg 31860
attattctgc atgggtacag agtccctctg ccagtcctg gtcactgtct tgtgattctc 31920
agtccccaac ttgtccccgg aaaagagtag ataggggtgg ggctaaggac acccccggga 31980
gggatgagtc ataggtgggg ggctgcctca tgccaggaag catgtaccag ctcccacccc 32040
agggggctga gggagataaa tggggcctga agcggggtag agggtcagac cacaggacag 32100
tagtgcctgg cccagcccc aggcagccac agcaggctgc cttaccccag aagcagctgg 32160
tgccggtagg actgggttgg gtcgggatgg gaagggtctt ggaggttgag tggatgtggg 32220
gtttggcttt atggagggtc tggaccaggg ggactctggg atctctggct gcttttctgc 32280
ctctgagatc cgattcctgc ccttctgttt cctggatcag ctgcaagctc tcctgctgag 32340
aaccgcctgc cctcctgtgg actctgtgtt tctgtctgaa tctttctttc catca 32395

```

<210> 33

< 211> 54

5 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> Cebador

10

<400> 33

gccagtgtcc tcactgtccg ccggggtgag tactgcagga ggctgtgga ggac 54

<210> 34

15 < 211> 54

< 212> ADN

< 213> Secuencia artificial

<220>

20 < 223> Cebador

<400> 34

cggtcacagg agtgacaggc ggccccactc atgacgtcct ccgaacacct cctg 54

25 <210> 35

< 211> 54

< 212> ADN

< 213> Secuencia artificial

ES 2 750 550 T3

	<220>	
	< 223> Cebador	
	<400> 35	
5	gccagtgcc tcactgtccg ccgggggtgag tactgcagga ggcttgtgga ggac	54
	<210> 36	
	< 211> 54	
	< 212> ADN	
10	< 213> Secuencia artificial	
	<220>	
	< 223> Cebador	
15	<400> 36	
	cggtcacagg agtgacaggc ggccccactc atgacgtcct ccgaacacct cctg	54
	<210> 37	
	< 211> 54	
20	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
	< 223> Cebador	
25	<400> 37	
	gccagtgcc tcactgtccg ccgggggtgag tactgcagga ggcttgtgga ggac	54
	<210> 38	
30	< 211> 53	
	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
35	< 223> Cebador	
	<400> 38	
	gccagtgcc tcactgtccg ccgtgagtac tggcaggaag gcttgtggag gac	53
40	<210> 39	
	< 211> 47	

ES 2 750 550 T3

< 212> ADN

< 213> Secuencia artificial

<220>

5 < 223> Cebador

<400> 39

gccagtgccc tcactgtccg ccggactgca ggaggcttgt ggaggac 47

10 <210> 40

< 211> 46

< 212> ADN

< 213> Secuencia artificial

15 <220>

< 223> Cebador

<400> 40

gccagtgccc tcactgtccg ccgggtgcag gaggcttggt gaggac 46

20

<210> 41

< 211> 41

< 212> ADN

< 213> Secuencia artificial

25

<220>

< 223> Cebador

<400> 41

30 gccagtgccc tcactgtccg ccgggtgag tactggagga c 41

<210> 42

< 211> 39

< 212> ADN

35 < 213> Secuencia artificial

<220>

< 223> Cebador

40 <400> 42

gccagtgccc tcactgagtg caggaggctt gtggaggac 39

<210> 43
 < 211> 38
 < 212> ADN
 < 213> Secuencia artificial
 5
 <220>
 < 223> Cebador

 <400> 43
 10 gccagtgctc tcaccgtagc aggaggcttg tggaggac 38

 <210> 44
 < 211> 38
 < 212> ADN
 15 < 213> Secuencia artificial

 <220>
 < 223> Cebador

 20 <400> 44
 gccagtgctc tcactgcagg aggagccttg tggaggac 38

 <210> 45
 < 211> 32
 25 < 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> Cebador
 30
 <400> 45
 gccagtgctc tcactgtcct gacttgagg ac 32

 <210> 46
 35 < 211> 33
 < 212> ADN
 < 213> Secuencia artificial

 <220>
 40 < 223> Cebador

<400> 46
gccagtgccc tcactgtccg ccgtgaggac agc 33

<210> 47
5 < 211> 24
< 212> ADN
< 213> Secuencia artificial

<220>
10 < 223> Cebador

<400> 47
gtacgcagga ggcttggtga ggac 24

<210> 48
15 < 211> 16
< 212> ADN
< 213> Secuencia artificial

<220>
20 < 223> Cebador

<400> 48
gaggcttggtg gaggac 16
25

<210> 49
< 211> 12
< 212> ADN
< 213> Secuencia artificial
30

<220>
< 223> Cebador

<400> 49
35 gggcccggga cc 12

<210> 50
< 211> 10
< 212> ADN
40 < 213> Secuencia artificial

<220>
 < 223> Cebador

 <400> 50
 5 gagggccctg 10

 <210> 51
 < 211> 10
 < 212> ADN
 10 < 213> Secuencia artificial

 <220>
 < 223> Cebador

 15 <400> 51
 gagggacctg 10

 <210> 52
 < 211> 37
 20 < 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> Cebador
 25
 <400> 52
 tggcccttct cactctgcgt ccctgtccat cactgcc 37

 <210> 53
 30 < 211> 12
 < 212> ADN
 < 213> Secuencia artificial

 <220>
 35 < 223> Cebador

 <400> 53
 gggcccggga cc 12

 40 <210> 54
 < 211> 60

ES 2 750 550 T3

	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
5	< 223> Cebador	
	<400> 54	
	gggtggagag aaccctggtg ctctctggga gccagacggc attcgacttg gatgacgttc	60
10	<210> 55	
	< 211> 60	
	< 212> ADN	
	< 213> Secuencia artificial	
15	<220>	
	< 223> Cebador	
	<400> 55	
	gggtggagag aaccctggtg ctctctggga gccagacggc attcgacttg gatgacgttc	60
20	<210> 56	
	< 211> 60	
	< 212> ADN	
	< 213> Secuencia artificial	
25	<220>	
	< 223> Cebador	
	<400> 56	
30	actccacttg ctgttcagg gctgcgggtt gtggtgtcag atgcaacgcg agtgagggtg	60
	<210> 57	
	< 211> 60	
	< 212> ADN	
35	< 213> Secuencia artificial	
	<220>	
	< 223> Cebador	
40	<400> 57	
	actccacttg ctgttcagg gctgcgggtt gtggtgtcag atgcaacgtg agtgagggtg	60

	<210> 58	
	< 211> 39	
	< 212> ADN	
	< 213> Secuencia artificial	
5	<220>	
	< 223> Cebador	
	<400> 58	
10	gtggctgtgt cggctactgcg aggcagagag gagggacct	39
	<210> 59	
	< 211> 39	
	< 212> ADN	
15	< 213> Secuencia artificial	
	<220>	
	< 223> Cebador	
20	<400> 59	
	gtggctgtgt cggctactgcg aggcagagag gagggccct	39
	<210> 60	
	< 211> 54	
25	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
	< 223> Cebador	
30	<400> 60	
	gccagtgtcc tcactgtccg ccgggggtgag tactgcagga ggcttggtga ggac	54
	<210> 61	
35	< 211> 240	
	< 212> ADN	
	< 213> Secuencia artificial	
	<220>	
40	< 223> Cebador	

ES 2 750 550 T3

<400> 61

tcgacttgga tgacgttcag gctgggctta gctacactgt gcgggtgtct gctcgagtgg	60
gtccccgtga gggcagtgcc agtgtcctca ctgtccgccg ggagccggaa actccacttg	120
ctgttccagg gctgcggggt gtggtgtcag atgcaacgcg agtgaggggtg gcctggggac	180
cctaccaggt ggctgtgtcg gtactgcgag gcagagagga gggacctgct gcagtcacgc	240

5 <210> 62
 < 211> 213
 < 212> ADN
 < 213> Secuencia artificial

10 <220>
 < 223> Cebador

<400> 62

tcgacttgga tgacgttcag gctgggctta gctacactgt gcgggtgtct gctcgagtgg	60
gtccccgtga gggcagtgcc agtgtcctca ctgtccgccg ggggctgcgg gttgtggtgt	120
cagatgcaac gcgagtgagg gtggcctggg gaccctacca ggtggctgtg tcggtactgc	180
gaggcagaga ggagggacct gctgcagtca tcg	213

20 <210> 63
 < 211> 12
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Cebador

25 <400> 63
 gggcccggga cc 12

30 <210> 64
 < 211> 12
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Cebador

ES 2 750 550 T3

	<400> 64		
	caacgcgagt ga	12	
	<210> 65		
5	< 211> 12		
	< 212> ADN		
	< 213> Secuencia artificial		
	<220>		
10	< 223> Cebador		
	<400> 65		
	caacgtgagt ga	12	
15	<210> 66		
	< 211> 12		
	< 212> ADN		
	< 213> Secuencia artificial		
20	<220>		
	< 223> Cebador		
	<400> 66		
	caacgcgagt ga	12	
25	<210> 67		
	< 211> 227		
	< 212> ADN		
	< 213> Secuencia artificial		
30	<220>		
	< 223> Cebador		
	<400> 67		
35	gtgtcagatg caacgcgagt gaggggtggcc tggggacccg tocctggagc cagtggattt	60	
	cggattagct ggagcacagg cagtgggtccg gagtccagcc agacactgcc cccagactct	120	
	actgccacag acatcacagg gctgcagcct ggaaccacct accaggtggc tgtgtcggtta	180	
	ctgcgaggca gagaggaggg acctgctgca gtcacgtgtg ctcgaaac	227	
	<210> 68		

ES 2 750 550 T3

< 211> 227

< 212> ADN

< 213> Secuencia artificial

5

<220>

< 223> Cebador

<400> 68

gtgtcagatg caacgcgagt gaggggtggc tggggaccg tccctggagc cagtggattt	60
cggattagct ggagcacagg cagtgggtccg gagtccagcc agacactgcc cccagactct	120
actgccacag acatcacagg gctgcagcct ggaaccacct accaggtggc tgtgtcggtt	180
ctgcgaggca gagaggagg acctgctgca gtcacgtgg ctcgaac	227

<210> 69

< 211> 228

< 212> ADN

15

< 213> Secuencia artificial

<220>

< 223> Cebador

20

<400> 69

gtgtcagatg caacgcgaag tgaggggtggc ctggggaccg gtccctggag ccagtggatt	60
tccgattagc tggagcacag gcagtgggtcc ggagtccagc cagacactgc cccagactc	120
tactgccaca gacatcacag ggctgcagcc tgggaaccacc taccaggtgg ctgtgtcggt	180
actgcgaggc agagaggagg gacctgctgc agtcacgtg gctcgaac	228

<210> 70

25

< 211> 228

< 212> ADN

< 213> Secuencia artificial

<220>

30

< 223> Cebador

<400> 70

ES 2 750 550 T3

gtgtcagatg caacgcgaag tgagggtggc ctggggaccc gtccctggag ccagtggatt	60
tcggattagc tggagcacag gcagtgggcc ggagtccagc cagacactgc ccccagactc	120
tactgccaca gacatcacag ggctgcagcc tggaaccacc taccaggtgg ctgtgtcggt	180
actgcgagggc agagaggagg gacctgctgc agtcacgtg gctcgaac	228

<210> 71
 < 211> 31
 5 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Cebador

10 <400> 71
 agtactcagg aggcttgg aggacagctg c 31

15 <210> 72
 < 211> 33
 < 212> ADN
 < 213> Secuencia artificial

20 <220>
 < 223> Cebador

<400> 72
 gatgccgccc caccgcgac ctgagccttc tca 33

25 <210> 73
 < 211> 31
 < 212> ADN
 < 213> Secuencia artificial

30 <220>
 < 223> Cebador

35 <400> 73
 gtgcagggtg atgggagggc ttgcaccac g 31

<210> 74
 < 211> 34
 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> Cebador

5

<400> 74

ccagatttca cctttaacc ttggtgctat atag 34

<210> 75

< 211> 33

10

< 212> ADN

< 213> Secuencia artificial

<220>

< 223> Cebador

15

<400> 75

caccttgga gtaggatt taacatgat tta 33

REIVINDICACIONES**1. Composición comprendiendo**

un ácido nucleico que codifica una primera proteína TALEN, donde la primera proteína TALEN es capaz de inducir una rotura de ADN bicatenario específica de sitio en un gen COL7A1 en una célula, presentando el gen COL7A1 una mutación genética capaz de provocar epidermólisis bullosa, y

una secuencia donadora de ácido nucleico,

donde la secuencia donadora es una plantilla para la corrección de la mutación genética en el gen COL7A1.

2. Composición según la reivindicación 1, para su uso en el tratamiento de epidermólisis bullosa.

3. Composición según la reivindicación 1 o composición según la reivindicación 2 para el uso de la reivindicación 2, donde la célula se selecciona de entre el grupo que consiste en un fibroblasto, queratinocito, célula madre pluripotente inducible, célula madre hematopoyética, célula madre mesenquimal, célula madre embrionaria, célula de progenie hematopoyética, célula T, célula B, célula glial, célula neural, célula progenitora neuroglial, célula madre neuroglial, célula muscular, célula pulmonar, célula pancreática, célula hepática y una célula del sistema reticuloendotelial, y

opcionalmente donde la primera proteína TALEN es una TALEN izquierda y el ácido nucleico codifica además una segunda TALEN que es una TALEN derecha que coopera con la TALEN izquierda para realizar la rotura de ADN bicatenario específico del sitio en el gen diana; y

opcionalmente donde el ácido nucleico que codifica la proteína TALEN o la secuencia donadora de ácido nucleico forma parte de un vector o plásmido.

4. Composición según la reivindicación 3, para el uso de la reivindicación 3, donde la primera TALEN o la segunda TALEN comprende una pluralidad de secuencias de repetición de efector TAL y el dominio de endonucleasa y un espaciador entre la pluralidad de secuencias de repetición de efector TAL y el dominio de endonucleasa incluye un espaciador, opcionalmente donde el espaciador contiene de 12 a 30 nucleótidos de longitud.

5. Composición según la reivindicación 1, o composición según la reivindicación 2-4, comprendiendo al menos un ácido nucleico comprendiendo:

(i) un primer ácido nucleico que codifica un monómero de endonucleasa efectora de tipo activador de la transcripción (TAL),

(ii) un segundo ácido nucleico que codifica un segundo monómero de endonucleasa de efector TAL, y

(iii) una secuencia donadora,

donde cada uno del primer y el segundo monómero de endonucleasa de efector TAL comprende una pluralidad de secuencias de repetición de efector TAL y un dominio de endonucleasa FokI, donde cada una de la pluralidad de secuencias de repetición de efector TAL comprende un díresiduo variable de repetición, donde el primer monómero de endonucleasa de efector TAL es capaz de unirse a una primera secuencia de sitio medio de un ADN diana en el interior de la célula, presentando el ADN diana una mutación genética, y el primer monómero de endonucleasa de efector TAL es capaz de escindir el ADN diana cuando el segundo monómero de endonucleasa de efector TAL se une a una segunda secuencia de sitio medio del ADN diana, donde el ADN diana comprende la primera secuencia de sitio medio y la segunda secuencia de sitio medio separada por una secuencia espaciadora, y donde el primer y el segundo sitio medio poseen la misma secuencia de nucleótidos o distintas secuencias de nucleótidos, donde la secuencia donadora comprende una región homóloga al ADN diana, al menos en los extremos 5' y 3' del ADN diana y es una plantilla para la reparación de ADN derivando en una corrección de la mutación genética en el ADN diana, donde además la secuencia donadora es una plantilla para la corrección de la mutación genética en el gen COL7A1.

6. Proteína codificada o expresada por el ácido nucleico que codifica la primera proteína TALEN de la reivindicación 3.

7. Proteína codificada o expresada por el primer ácido nucleico o el segundo ácido nucleico de la reivindicación 5.

8. Vector comprendiendo el ácido nucleico que codifica la primera proteína TALEN de la reivindicación 3.

9. Vector comprendiendo el primer ácido nucleico o el segundo ácido nucleico de la reivindicación 5.

10. Ácido nucleico que codifica una proteína TALEN, donde el ácido nucleico comprende una secuencia seleccionada de entre el grupo que consiste en SEQ ID NO: 29 y SEQ ID NO: 30.

11. Proteína codificada o expresada por el ácido nucleico de la reivindicación 10.

Figura 1.

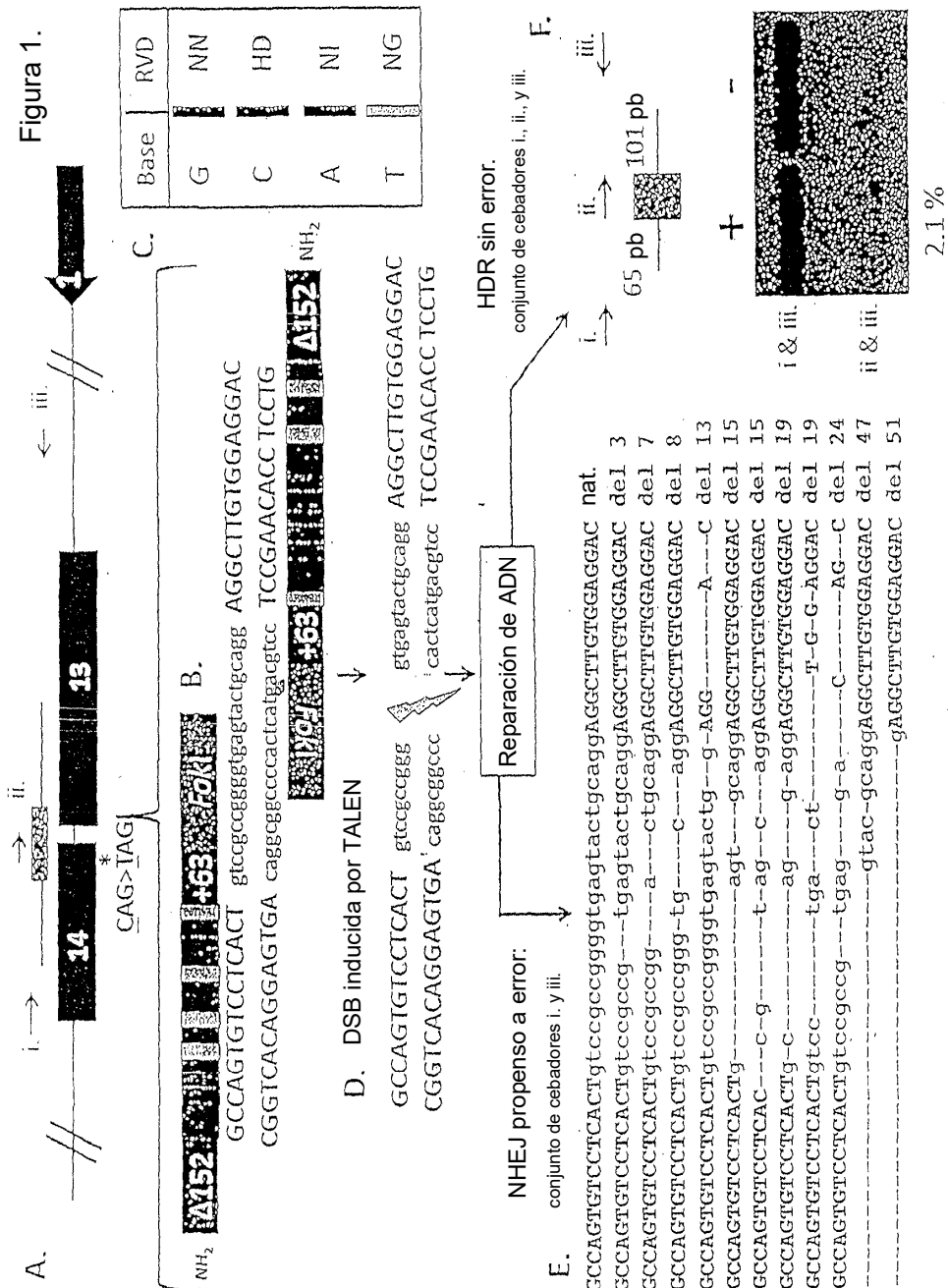


Figura 2.

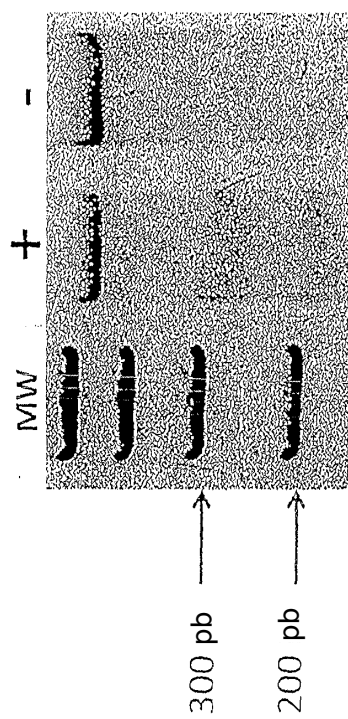


Figura 3.

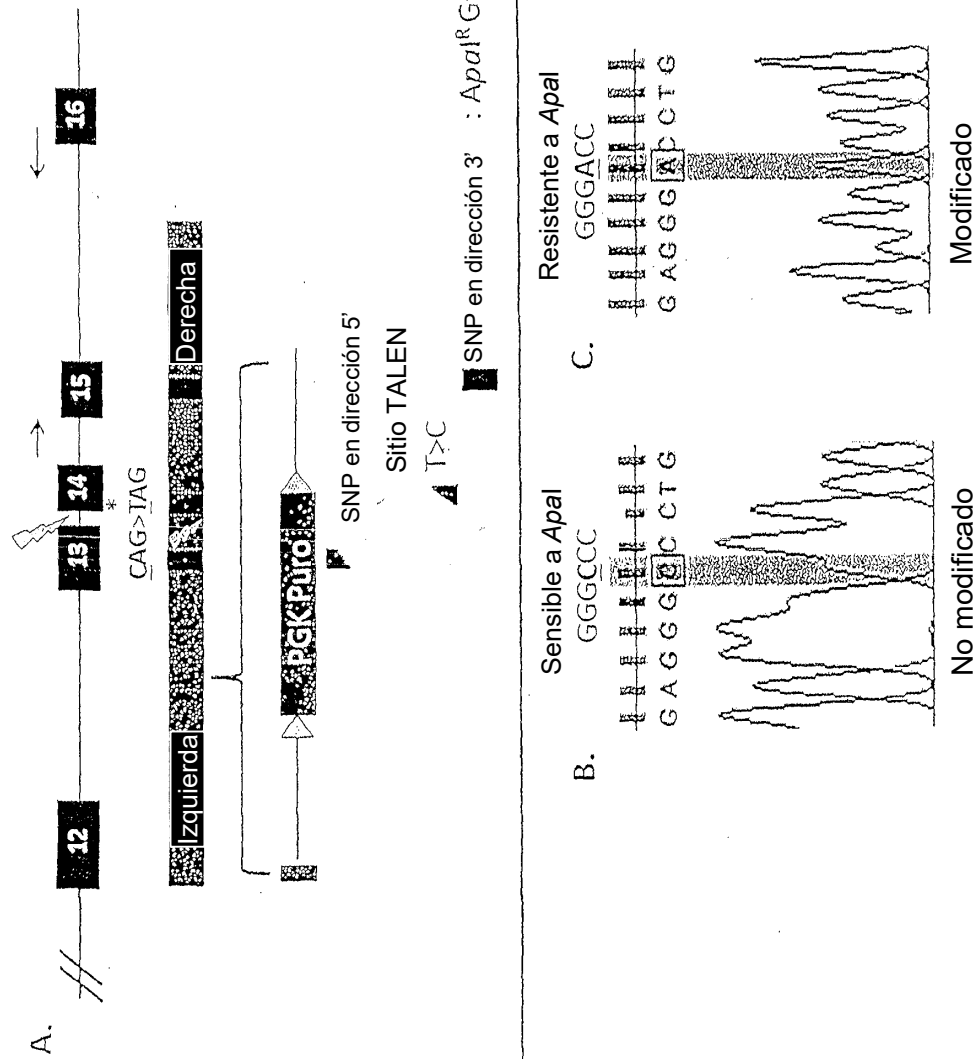


Figura 4

Escisión Cre de puromicina PGK

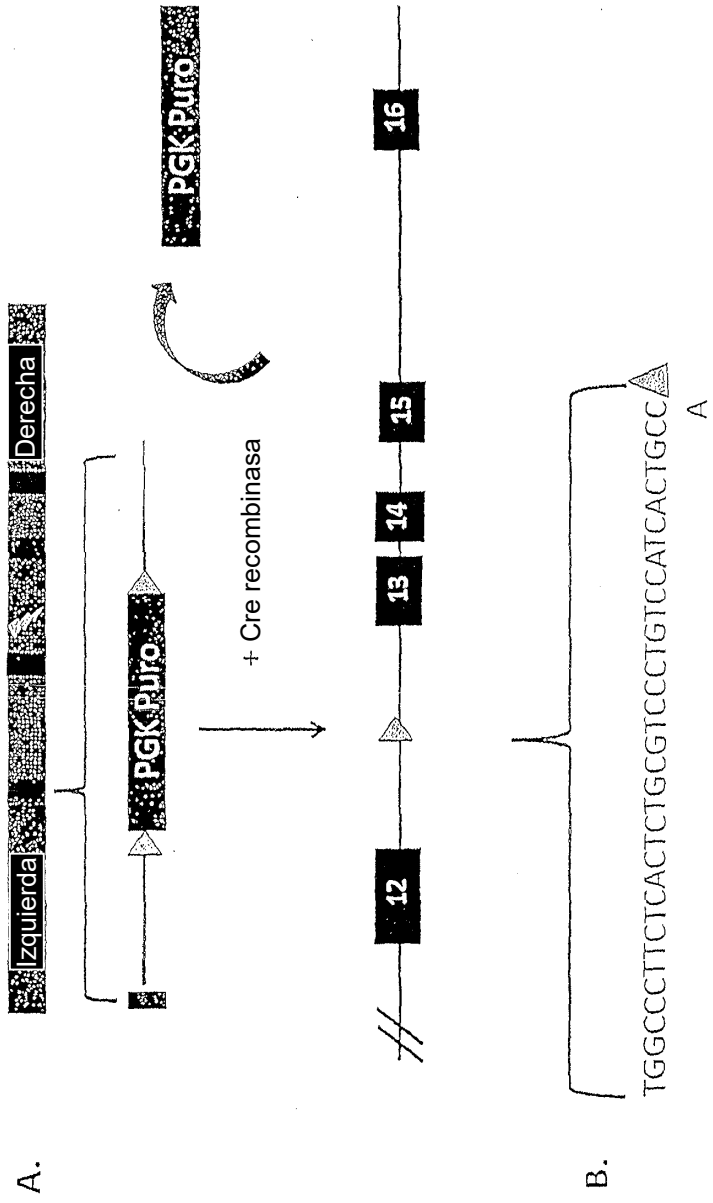
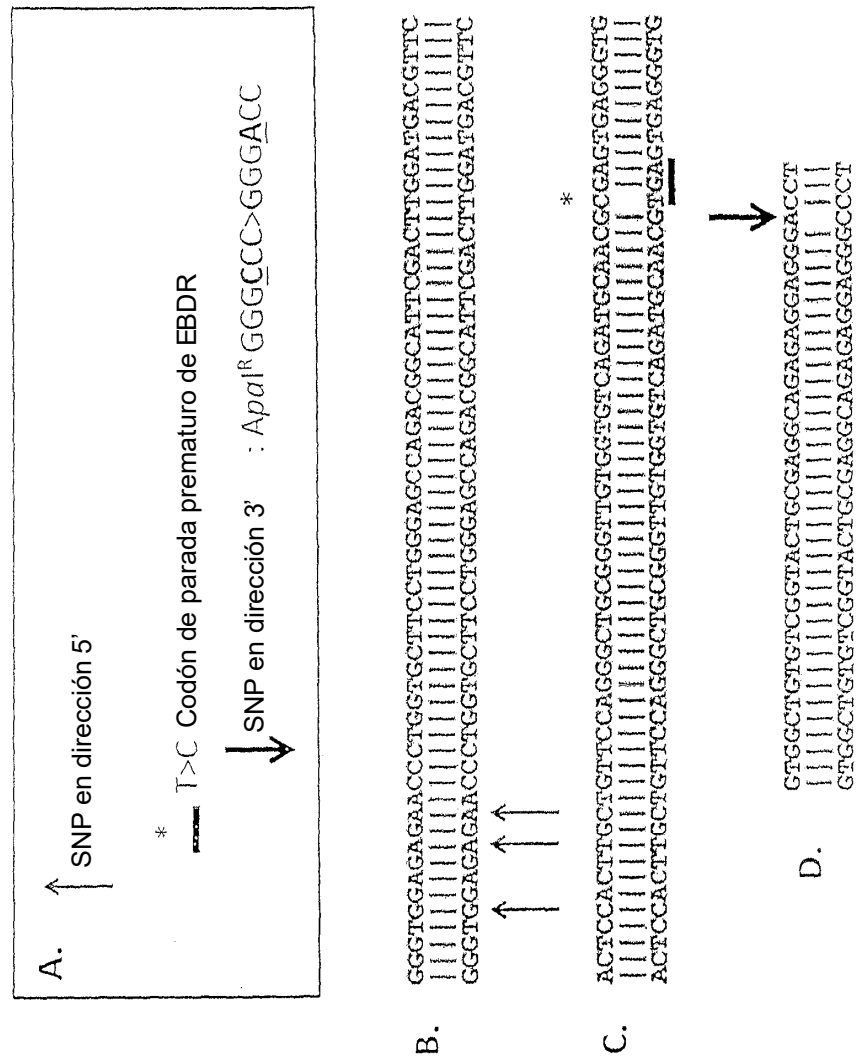


Figura 5



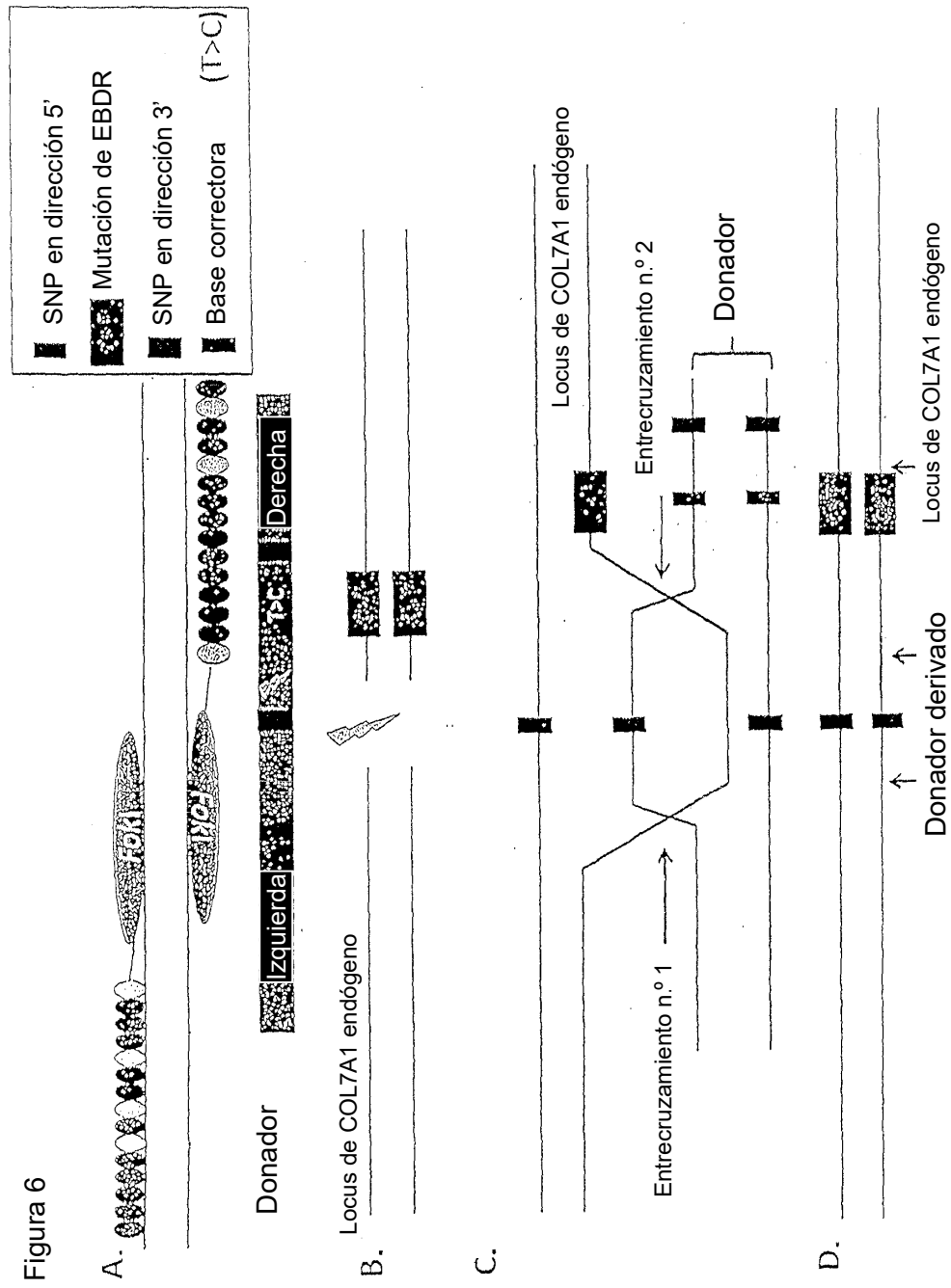


Figura 7

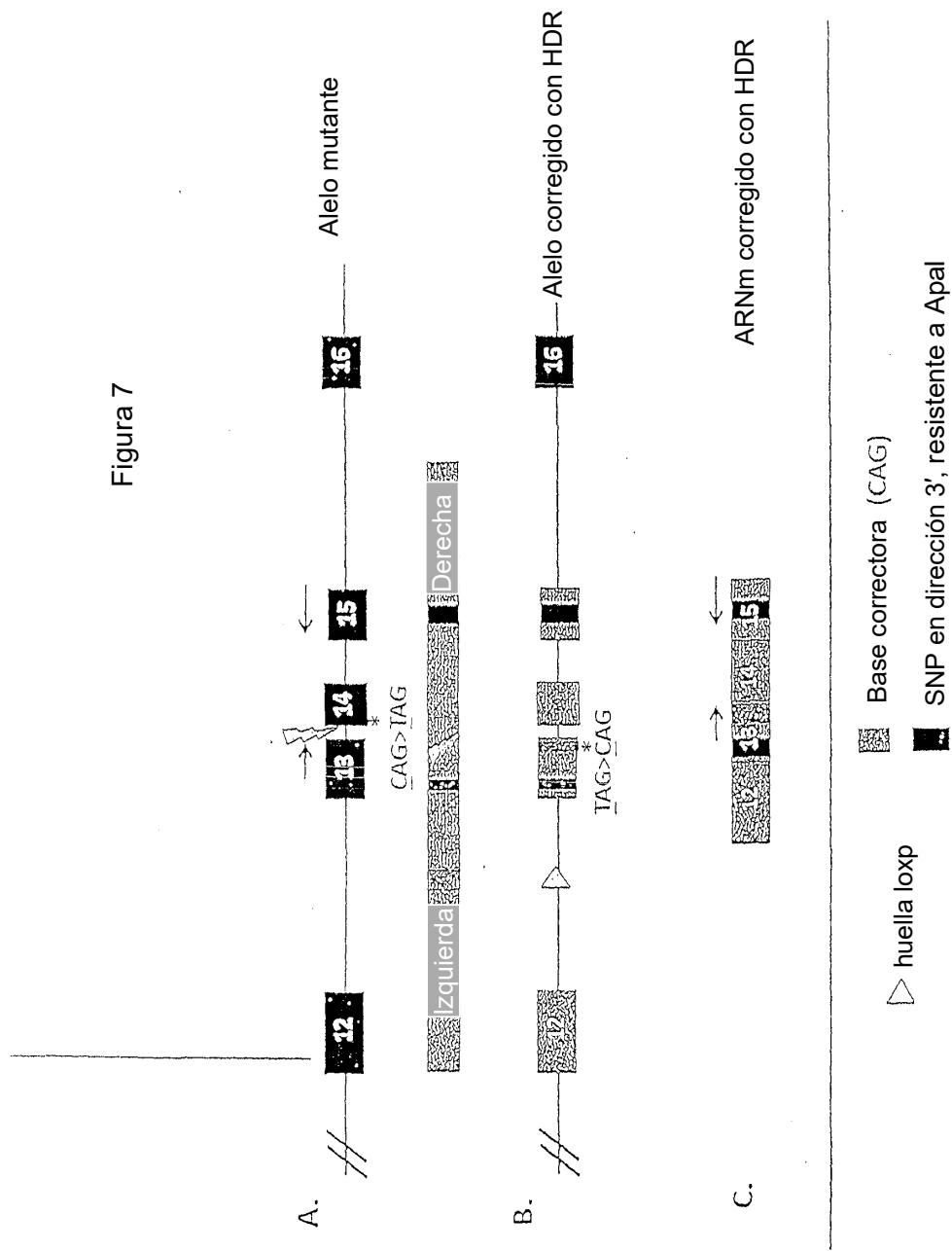


Figura 8

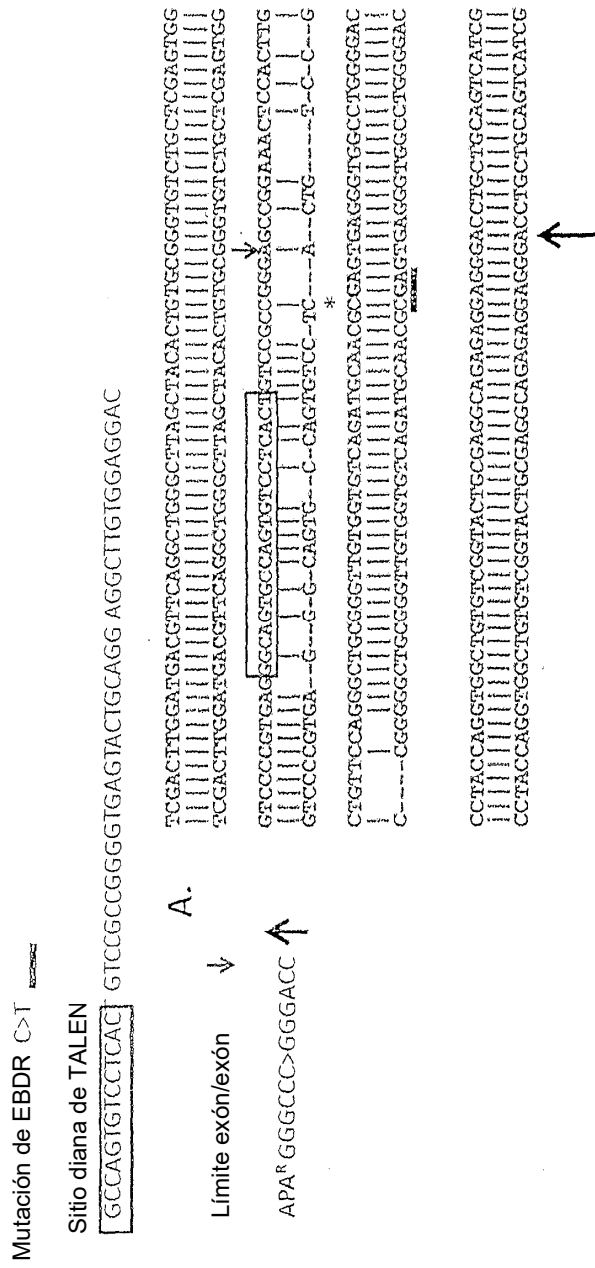


Figura 9.

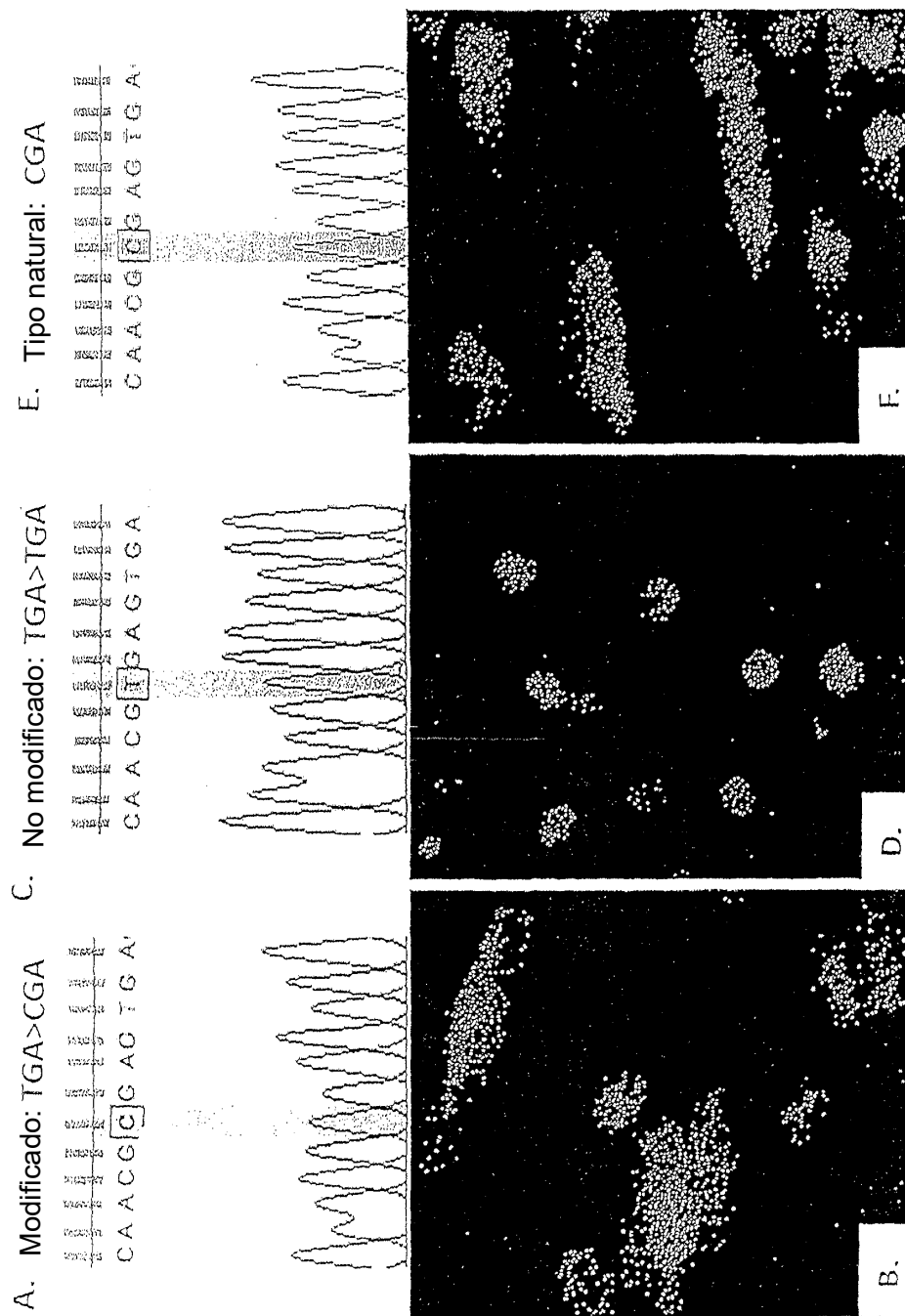
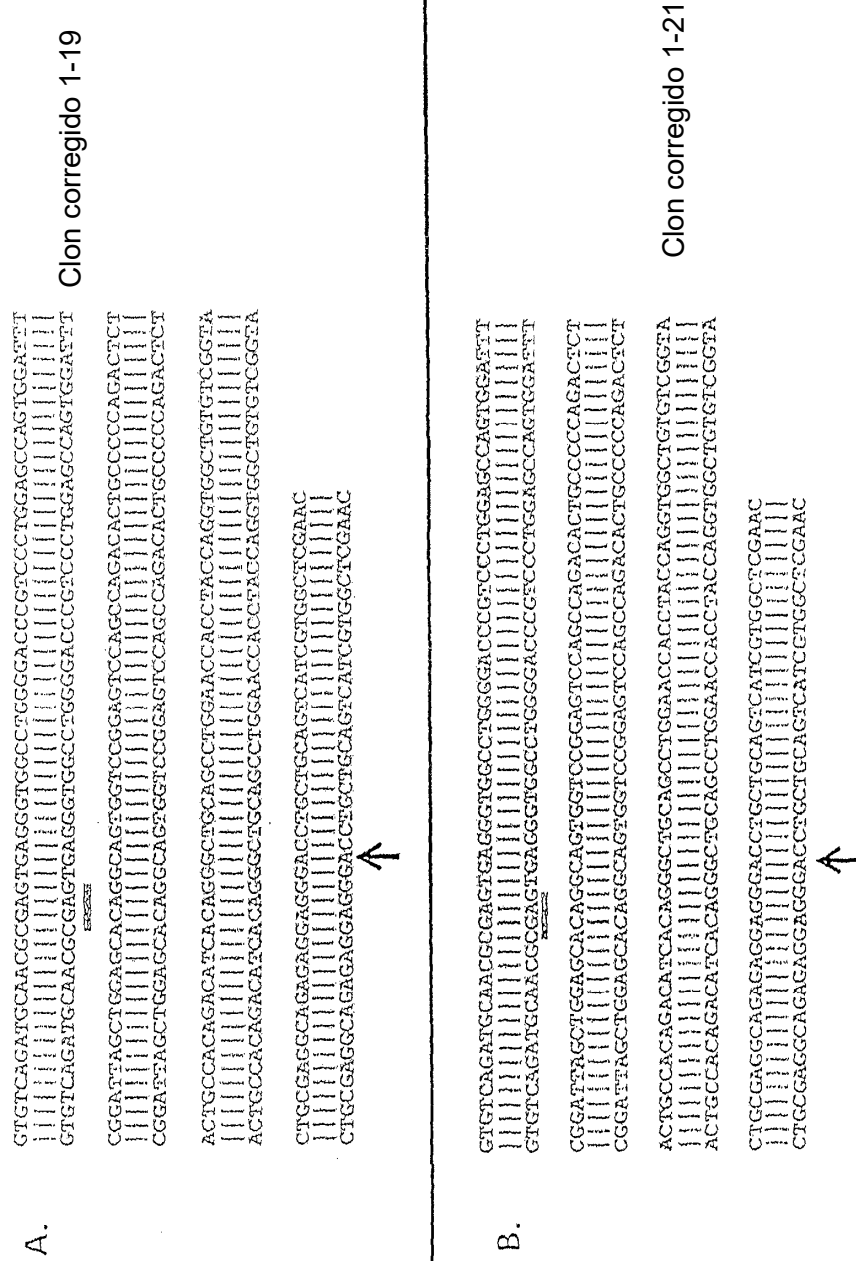


Figura 10



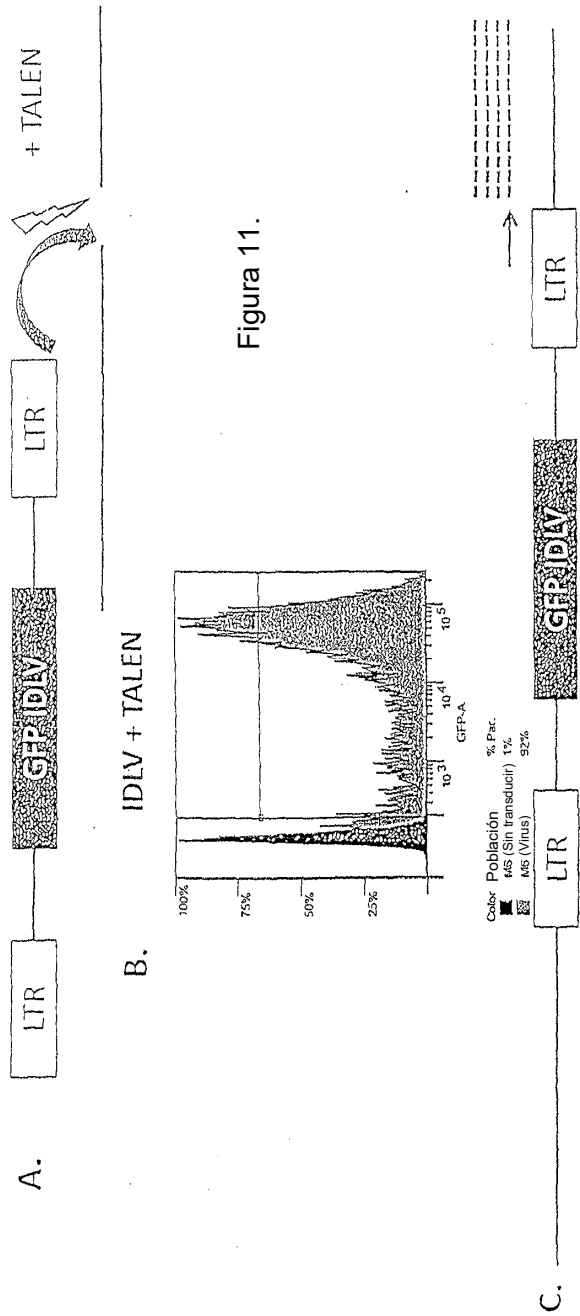


Figura 11.

Sitio	Secuencia	Información de locus
3p21.31	AGTACICAGGAGGGCTTGTGGAGGAGGCTGG	Diana de TALEN
7p22.3	GATGCCGCCACCCGCCACCTCAGCCTTCTCA	No codificante
16p13.3	GTGCAGGGTAATGGGAGGGGCTTGCACGCG	Intrón
1q23.3	CCAGATTTCACCTTTAACTTGGTGCTATATAG	No codificante
5q33.1	CACCTGGGAGCTAGGATTTTACATGGATTTA	No codificante

Figura 12

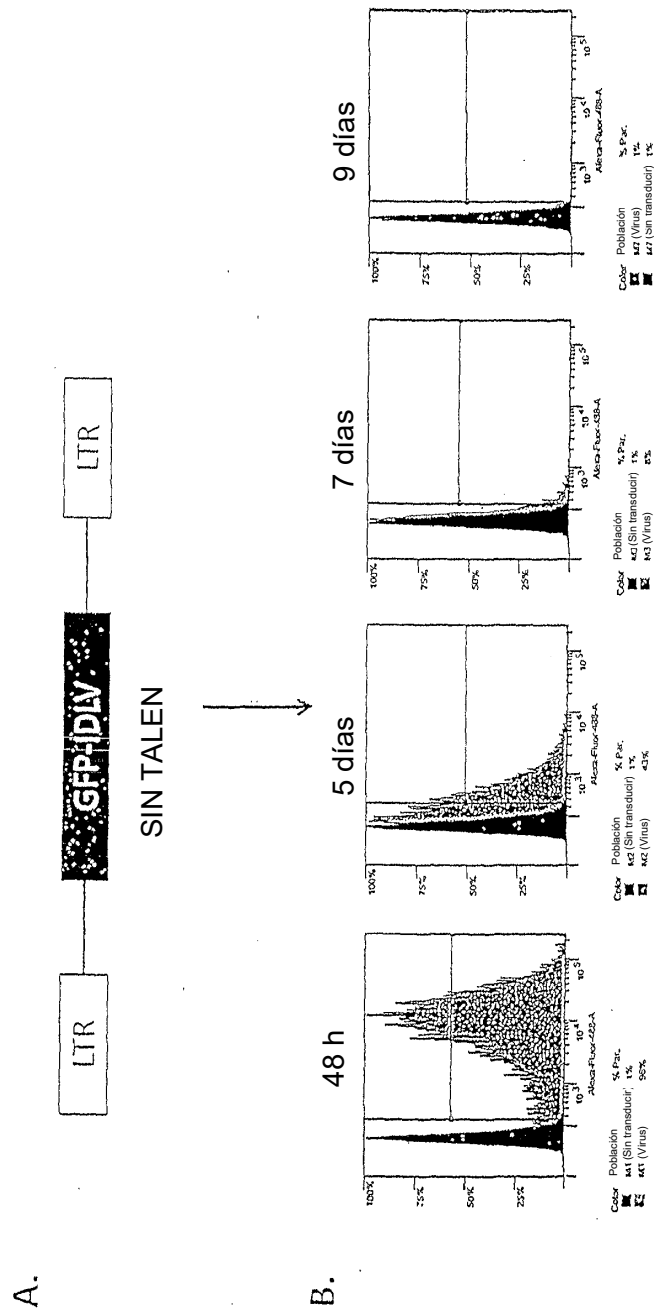


Figure 13

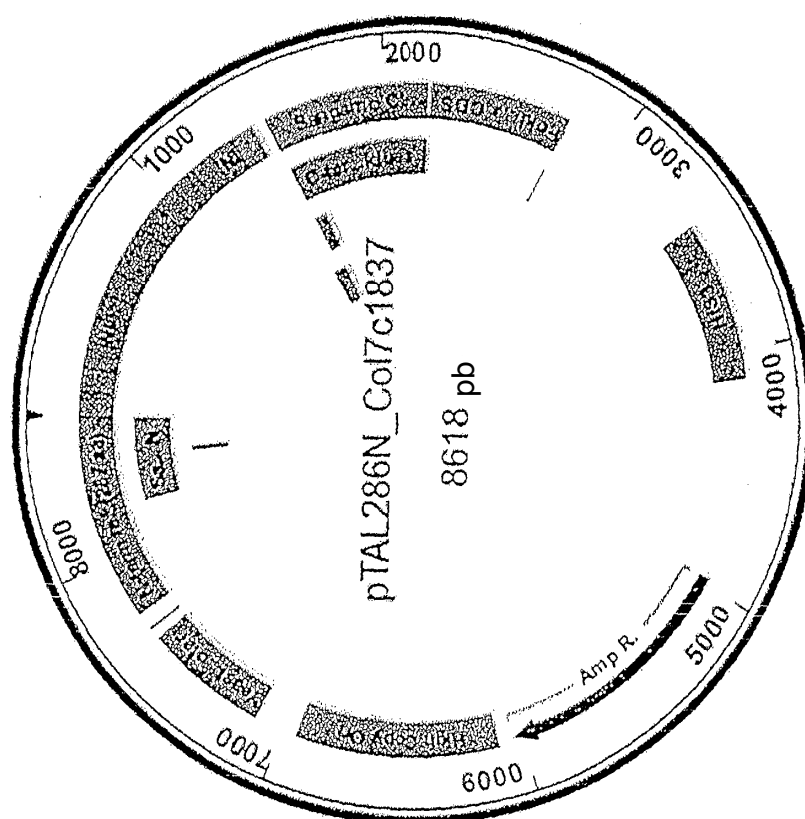


Figura 14

