

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 459 880**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**C12N 15/90** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.04.2009** **E 09734442 (8)**

97 Fecha y número de publicación de la concesión europea: **02.04.2014** **EP 2281050**

54 Título: **Construcciones donantes lineales para integración dirigida**

30 Prioridad:

**14.04.2008 US 124047 P**

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

**12.05.2014**

73 Titular/es:

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

72 Inventor/es:

**DEKELVER, RUSSELL;  
HOLMES, MICHAEL C.;  
URNOV, FYODOR y  
GREGORY, PHILIP D.**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 459 880 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

Construcciones donantes lineales para integración dirigida

### Campo técnico

5 La presente descripción pertenece al campo de ingeniería genómica, particularmente construcciones donantes lineales para integración dirigida en el genoma de una célula.

### Antecedentes

10 Un área principal de interés en biología genómica, especialmente a la luz de la determinación de las secuencias completas de nucleótidos de varios genomas, es la integración dirigida en secuencias genómicas. Se han hecho intentos por alterar secuencias genómicas en células cultivadas aprovechando el fenómeno natural de la recombinación homóloga. Véase, por ejemplo, Capecchi (1989) *Science* 244:1288-1292; patentes de Estados Unidos N° 6.528.313 y 6.528.314.

15 Además, se han descrito diversos métodos y composiciones para la escisión dirigida de ADN genómico. Dichos eventos de escisión dirigida pueden usarse, por ejemplo, para inducir mutagénesis dirigida, inducir deleciones dirigidas de secuencias de ADN celular, y facilitar la recombinación dirigida e integración dirigida en un locus cromosómico predeterminado. Véanse, por ejemplo, las publicaciones de patente de Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; y 20060188987, y la publicación internacional WO 2007/014275. Por ejemplo, se ha demostrado la integración dirigida usando nucleasas de dedos de zinc con ADN circulares (plasmídicos) que tienen largos (~750 pares de bases) brazos de homología. Véase, Moehle et al. (2007) *Proc. Nat'l. Acad. Sci. USA* 104(9):3055-3060.

20 Sin embargo, sigue existiendo la necesidad de composiciones adicionales que comprendan polinucleótidos exógenos lineales más cortos que puedan resistir opcionalmente la degradación por exonucleasa y el uso de estas composiciones en métodos para integración dirigida.

### Compendio

25 La presente descripción proporciona ácidos nucleicos exógenos (donantes) lineales, composiciones que comprendan estos ácidos nucleicos y métodos para preparar y usar estas moléculas donantes lineales. En líneas generales, las moléculas donantes descritas en la presente memoria tienen dos brazos de homología de entre 50 y 100 pares de bases flanqueando una secuencia de interés.

30 Las secuencias donantes pueden integrarse de un modo dirigido en el genoma de una célula, por ejemplo usando nucleasas de dedos de zinc (ZFN) y/o meganucleasas. La integración de las secuencias exógenas de ácido nucleico en el genoma se facilita mediante escisión bicatenaria dirigida del genoma (cromosoma) en la región de interés. La escisión se dirige preferiblemente a la región de interés a través del uso de proteínas de fusión que comprenden un dominio de unión de dedos de zinc, que se diseña para que se una a una secuencia dentro de la región de interés, y un dominio de escisión o un semi-dominio de escisión. Dicha escisión estimula la integración de las secuencias polinucleotídicas exógenas en o cerca del sitio de escisión.

35 En un aspecto, en la presente memoria se describe una molécula de ácido nucleico lineal (molécula donante) que comprende brazos de homología de 50-100 pares de bases flanqueando una secuencia de interés y se proporciona. En ciertas realizaciones, la molécula donante lineal persiste de forma estable en la célula en la que se ha introducido. En otras realizaciones, la molécula donante lineal se modifica para resistir la escisión exonucleolítica, por ejemplo colocando uno o más enlaces fosfodiéster fosforotioato entre uno o más pares de bases en los extremos de la molécula donante.

40 La secuencia de interés de la molécula donante puede comprender una o más secuencias que codifican un polipéptido funcional (por ejemplo, un ADNc), con o sin un promotor. En ciertas realizaciones, la secuencia de ácido nucleico comprende una secuencia sin promotor que codifica un anticuerpo, un antígeno, una enzima, un factor de crecimiento, un receptor (de superficie celular o nuclear), una hormona, una linfoquina, una citoquina, un indicador, fragmentos funcionales de cualquiera de los anteriores y combinaciones de los anteriores. La expresión de la secuencia integrada entonces se asegura por transcripción dirigida por un promotor endógeno u otro elemento de control en la región de interés. En otras realizaciones, se integra un casete en "tándem" en el sitio seleccionado de este modo, comprendiendo el primer componente del casete una secuencia sin promotor como se ha descrito anteriormente, seguida por una secuencia de terminación de la transcripción, y una segunda secuencia, que codifica un casete de expresión autónomo. Pueden incluirse secuencias adicionales (secuencias codificantes o no codificantes) en la molécula donante entre los brazos de homología, incluyendo, pero sin limitarse a, secuencias que codifican un péptido 2A, sitio SA, IRES, etc.

55 Las moléculas donantes de la descripción pueden insertarse en una localización específica en un genoma después de la escisión del genoma, por ejemplo usando una o más moléculas de fusión que comprenden un dominio de unión a ADN dirigido a la localización específica en el genoma y un dominio de escisión (por ejemplo, nucleasa de

dedos de zinc (ZFN) o meganucleasa de origen natural o no natural a un locus particular). Por tanto, en otro aspecto, en la presente memoria se proporciona un método para integrar una secuencia exógena como se describe en la presente memoria en una región de interés en el genoma de una célula., comprendiendo el método: (a) expresar una proteína de fusión en la célula, comprendiendo la proteína de fusión un dominio de unión a ADN (por ejemplo, dominio de unión de dedos de zinc) y un dominio de escisión o semi-dominio de escisión, en que el dominio de unión a ADN (por ejemplo, dominio de unión de dedos de zinc) se ha diseñado para que se una a un sitio diana en la región de interés en el genoma de la célula; y (b) poner en contacto la célula con un polinucleótido donante como se describe en la presente memoria, en que la unión de la proteína de fusión al sitio diana escinde el genoma de la célula en la región de interés, provocando de este modo la integración de la secuencia exógena en el genoma de la célula dentro de la región de interés.

En ciertas realizaciones, los métodos comprenden las etapas de (a) expresar una primera proteína de fusión en la célula, comprendiendo la primera proteína de fusión un primer dominio de unión de dedos de zinc y un primer semi-dominio de escisión, en que el dominio de unión de dedos de zinc se ha diseñado para que se una a un primer sitio diana en la región de interés en el genoma de la célula; (b) expresar una segunda proteína de fusión en la célula, comprendiendo la segunda proteína de fusión un segundo dominio de unión de dedos de zinc y un segundo semi-dominio de escisión, en que el segundo dominio de unión de dedos de zinc se une a un segundo sitio diana en la región de interés en el genoma de la célula, en que el segundo sitio diana es diferente del primer sitio diana; y (c) poner en contacto la célula con una molécula donante exógena como se describe en la presente memoria, en que la unión de la primera proteína de fusión al primer sitio diana, y la unión de la segunda proteína de fusión al segundo sitio diana, posiciona los semi-dominios de escisión de tal modo que el genoma de la célula se escinde en la región de interés, provocando de este modo la integración de la molécula donante exógena en el genoma de la célula dentro de la región de interés.

En cualquiera de los métodos descritos en la presente memoria, el polinucleótido donante comprende una secuencia que codifica un polipéptido funcional, cuya secuencia se inserta en el genoma de la célula.

Además, en cualquiera de los métodos descritos en la presente memoria, el primer y segundo semi-dominios de escisión son de una endonucleasa de restricción Tipo IIS, por ejemplo, *FokI* o *Sfsl*. Además, en cualquiera de los métodos descritos en la presente memoria, al menos una de las proteínas de fusión puede comprender una alteración en la secuencia de aminoácidos de la superficie de contacto de dimerización del semi-dominio de escisión, por ejemplo de tal modo que se formen heterodímeros obligados de los semi-dominios de escisión. Como alternativa, en cualquiera de los métodos descritos en la presente memoria, el dominio de escisión puede ser una meganucleasa de origen natural o no natural.

En cualquiera de los métodos descritos en la presente memoria, la célula puede ser una célula de mamífero, por ejemplo, una célula humana, de rata, de ratón o de conejo, o una célula vegetal. Adicionalmente, la célula puede obtenerse de un sistema de insecto, xenopus o nematodo. Además, la célula puede estar detenida en la fase G2 del ciclo celular.

### Breve descripción de los dibujos

La **Figura 1** es un diagrama esquemático que representa la construcción de un polipéptido donante lineal como se describe en la presente memoria. La "x" indica enlaces fosfodiéster fosforotioato como el primer y el segundo enlaces en los extremos 5' y 3' del polinucleótido.

La **Figura 2** representa la secuencia de un donante lineal ejemplar (SEC ID N° 1) que tiene brazos de homología de 100 pares de bases. La molécula donante lineal comprende un brazo izquierdo de homología desde los nucleótidos 1 a 100 (en minúscula, subrayado); un sitio aceptor de corte y ajuste (SA), de los nucleótidos 107 a 132 (en minúscula, negrita); una secuencia que codifica una secuencia de auto-procesamiento 2A (péptido 2A) derivada del virus de la enfermedad de la fiebre aftosa (FMDV) de los nucleótidos 141 a 212 (en mayúscula, sin subrayar); una secuencia que codifica poli(A) de la proteína fluorescente verde (GFP) de los nucleótidos 219 a 1.215 (en mayúscula, subrayado); y un brazo derecho de homología de los nucleótidos 1235 a 1334 (en minúscula, subrayado).

La **Figura 3** representa la secuencia de un donante lineal ejemplar (SEC ID N° 2) que tiene brazos de homología de 75 pares de bases. La molécula donante lineal comprende un brazo izquierdo de homología de los nucleótidos 1 a 75 (en minúscula, subrayado); un sitio SA de los nucleótidos 82 a 107 (en minúscula, negrita); una secuencia que codifica un péptido 2A de los nucleótidos 116 a 187 (en mayúscula, sin subrayar); una secuencia que codifica poli(A) de GFP de los nucleótidos 194 a 1.190 (en mayúscula, subrayado); y un brazo derecho de homología de los nucleótidos 1210 a 1284 (en minúscula, subrayado).

La **Figura 4** representa la secuencia de un donante lineal ejemplar (SEC ID N° 3) que tiene brazos de homología de 50 pares de bases. La molécula donante lineal comprende un brazo izquierdo de homología de los nucleótidos 1 a 50 (en minúscula, subrayado); un sitio SA de los nucleótidos 57 a 82 (en minúscula, negrita); una secuencia que codifica un péptido 2A de los nucleótidos 91 a 162 (en mayúscula, sin subrayar); una secuencia que codifica poli(A) de GFP de los nucleótidos 169 a 1.165 (en mayúscula, subrayado); y un brazo derecho de homología de los

nucleótidos 1.185 a 1.234 (en minúscula, subrayado).

La **Figura 5** representa la secuencia de otro donante lineal ejemplar (SEC ID N° 4) que tiene brazos de homología de 50 pares de bases. La molécula donante lineal comprende un brazo izquierdo de homología de los nucleótidos 1 a 50 (en minúscula, subrayado); una secuencia promotora hPGK de los nucleótidos 79 a 594 (en minúscula, negrita); una secuencia que codifica poli(A) de GFP de los nucleótidos 615 a 1.611 (en mayúscula, subrayado); y un brazo derecho de homología de los nucleótidos 1.639 a 1.688 (en minúscula, subrayado).

La **Figura 6** representa los resultados de un ensayo de PCR y muestra la modificación del locus PPP1R12C (AAVS1) cuando se introducen diversas moléculas donantes como se describe en la presente memoria en células K562 en ausencia (carriles 2-7) o presencia de ZFN dirigidos a AAVS1 (carriles 8-13).

La **Figura 7** es una transferencia de Southern que muestra la modificación del locus PPP1R12C (AAVS1) cuando se introducen diversas moléculas donantes como se describe en la presente memoria en células K562 en ausencia (carriles 3-7) o presencia de ZFN dirigidos a AAVS1 (carriles 9-13). El porcentaje de cromosomas modificados se enumera bajo los carriles 9-13.

La **Figura 8** representa el porcentaje de células GFP positivas evaluadas por FACS.

## 15 Descripción detallada

La presente descripción se refiere a polinucleótidos exógenos (donantes) útiles para integración dirigida (TI) dependiente de homología en una región de interés en un genoma. En particular, los polinucleótidos donantes descritos en la presente memoria son moléculas lineales que comprenden brazos de homología (HA) de aproximadamente 50-100 pares de bases. Los brazos de homología flanquean una o más secuencias de interés a insertar en el genoma de una célula. Estas moléculas donantes son útiles para escisión dirigida y recombinación en una región específica de interés en un genoma cuando se usan en combinación con proteínas de fusión (nucleasas de dedos de zinc) que comprenden un dominio de escisión (o un semi-dominio de escisión) y un dominio de unión de dedos de zinc (y/o polinucleótidos que codifican estas proteínas). Un dominio de unión de dedos de zinc puede comprender uno o más dedos de zinc (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de zinc), y puede diseñarse para que se una a cualquier secuencia dentro de la región de interés. En presencia de ZFP, los polinucleótidos donantes lineales descritos se integran a elevadas tasas en el sitio de escisión por métodos dependientes de homología.

Las ventajas de las moléculas donantes lineales descritas en la presente memoria incluyen la provisión rápida y eficaz de moléculas donantes para su uso con ZFN. Actualmente, las moléculas donantes usadas en combinación con nucleasas de dedos de zinc (ZFN) para inserción dirigida en un locus específico del genoma son construcciones plasmídicas que contienen largos (~750 pares de bases) brazos de homología que flanquean un transgén de interés. La construcción de dichos donantes plasmídicos consume mucho tiempo, tardando al menos 2 semanas. En contraste, las moléculas donantes lineales descritas en la presente memoria pueden construirse en horas y usarse inmediatamente. Además, el uso de donantes lineales como se describe en la presente memoria reduce o elimina los fenómenos de inserción estable del donante plasmídico en la célula hospedadora.

## General

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

## Definiciones

Los términos "ácido nucleico", "polinucleótido", y "oligonucleótido" se usan de forma intercambiable y hacen referencia a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en forma mono o bicatenaria. Para los propósitos de la presente descripción, estos términos no deben entenderse como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en los restos de base, azúcar y/o fosfato (por ejemplo, estructuras fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de apareamiento de bases; es decir, un análogo de A se apareará con T.

Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable para hacer referencia a un

polímero de restos aminoacídicos. El término también se aplica a polímeros de aminoácidos en que uno o más aminoácidos son análogos químicos o derivados modificados de un aminoácido de origen natural correspondiente.

5 "Unión" se refiere a una interacción no covalente específica de secuencia entre macromoléculas (por ejemplo, entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión tienen que ser específicos de secuencia (por ejemplo, contactos con restos fosfato en una estructura de ADN), siempre que la interacción como un conjunto sea específica de secuencia. Dichas interacciones se caracterizan generalmente por una constante de disociación ( $K_d$ ) de  $10^{-6} \text{ M}^{-1}$  o inferior. "Afinidad" se refiere a la fuerza de la unión: correlacionándose una afinidad de unión aumentada con una inferior  $K_d$ .

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

15 Una "proteína de unión a ADN de dedos de zinc" (o dominio de unión) es una proteína, o un dominio dentro de una proteína más grande, que se une a ADN de un modo específico de secuencia a través de uno o más dedos de zinc, que son regiones de secuencia de aminoácidos dentro del dominio de unión cuya estructura se estabiliza a través de la coordinación de un ión zinc. La expresión proteína de unión a ADN de dedos de zinc a menudo se abrevia como proteína de dedos de zinc o ZFP.

20 Los dominios de unión de dedos de zinc pueden "diseñarse" para que se unan a una secuencia de nucleótidos predeterminada. Ejemplos no limitantes de métodos para diseñar proteínas de dedos de zinc son diseño y selección. Una proteína de dedos de zinc diseñada es una proteína que no existe en la naturaleza cuyo diseño/composición resulta principalmente de criterios racionales. Los criterios racionales para el diseño incluyen aplicación de normas de sustitución y algoritmos computarizados para procesar información en una base de datos que clasifica información de diseños y datos de unión existentes de ZFP. Véanse, por ejemplo, las patentes de Estados Unidos 6.140.081; 6.453.242; y 6.534.261; véanse también los documentos WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

30 Una proteína de dedos de zinc "seleccionada" es una proteína no hallada en la naturaleza cuya producción resulta principalmente de un proceso empírico tal como presentación in fago, retención por interacción o selección de híbridos. Véanse, por ejemplo, los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197 y WO 02/099084.

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

40 Una "secuencia no idéntica homóloga" se refiere a una primera secuencia que comparte un grado de identidad de secuencia con una segunda secuencia, pero cuya secuencia no es idéntica a la de la segunda secuencia. Por ejemplo, un polinucleótido que comprende la secuencia de tipo silvestre de un gen mutante es homólogo y no idéntico a la secuencia del gen mutante. En ciertas realizaciones, el grado de homología entre las dos secuencias es suficiente para permitir la recombinación homóloga entre ellas, utilizando mecanismos celulares normales. Dos secuencias no idénticas homólogas pueden ser de cualquier longitud y su grado de no homología puede ser tan pequeño como de un único nucleótido (por ejemplo, para la corrección de una mutación puntual genómica por recombinación homóloga dirigida) o tan larga como de 10 o más kilobases (por ejemplo, para la inserción de un gen en un sitio ectópico predeterminado en un cromosoma). Dos polinucleótidos que comprenden las secuencias no idénticas homólogas no tienen que ser de la misma longitud. Por ejemplo, puede usarse un polinucleótido exógeno (es decir, polinucleótido donante) de entre 20 y 10.000 nucleótidos o pares de nucleótidos.

45 Las técnicas para determinar la identidad de secuencia de ácidos nucleicos y aminoácidos son conocidas en la técnica. Típicamente, dichas técnicas incluyen determinar la secuencia de nucleótidos del ARNm de un gen y/o determinar la secuencia de aminoácidos codificada por la misma, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también pueden determinarse y compararse de este modo. En general, la identidad se refiere a una correspondencia exacta nucleótido a nucleótido o aminoácido a aminoácido de dos polinucleótidos o secuencias polipeptídicas, respectivamente. Pueden compararse dos o más secuencias (polinucleótidos o aminoácidos) determinando su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácido nucleico o aminoácidos, es la cantidad de coincidencias exactas entre dos secuencias alineadas dividida por la longitud de la secuencia más corta y multiplicada por 100. Una alineación

aproximada para secuencias de ácido nucleico se proporciona por el algoritmo de homología local de Smith y Waterman, Advances in Applied Mathematics 2:482-489 (1981). Este algoritmo puede aplicarse a secuencias de aminoácidos usando la matriz de valoración desarrollada por Dayhoff, Atlas of Protein Sequences and Structure. M.O. Dayhoff ed., 5 supl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., EE.UU., y normalizada por Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). Una implementación ejemplar de este algoritmo para determinar el porcentaje de identidad de una secuencia se proporciona por el Genetics Computer Group (Madison, WI) en la aplicación de utilidad "BestFit". Los parámetros por defecto para este método se describen en el Wisconsin Sequence Analysis Package Program Manual, Versión 8 (1995) (disponible en Genetics Computer Group, Madison, WI). Un método preferido para establecer el porcentaje de identidad en el contexto de la presente descripción es usar el paquete MPSRCH de programas con copyright de la University of Edinburgh, desarrollado por John F. Collins y Shane S. Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). Para este paquete de contenidos puede emplearse el algoritmo de Smith-Waterman donde se usan parámetros por defecto para la tabla de resultados (por ejemplo, penalización por apertura de hueco de 12, penalización por extensión de hueco de uno, y un hueco de seis). A partir de los datos generados el valor "Coincidencia" refleja la identidad de secuencia. En la técnica se conocen en líneas generales otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias, por ejemplo, otro programa de alineación es BLAST, usado con parámetros por defecto. Por ejemplo, pueden usarse BLASTN y BLASTP usando los siguientes parámetros por defecto: código genético = convencional; filtro = ninguno; hebra = ambas; punto de corte = 60; expectativa = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; clasificado por = HIGH SCORE; Bases de datos = no redundantes, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Pueden hallarse detalles de estos programas en internet. Con respecto a secuencias descritas en la presente memoria, el intervalo de grados deseados de identidad de secuencia es de aproximadamente el 80% al 100% y cualquier valor entero entre los mismos. Típicamente, los porcentajes de identidad entre secuencias son de al menos el 70-75%, preferiblemente el 80-82%, más preferiblemente el 85-90%, incluso más preferiblemente el 92%, aún más preferiblemente el 95%, y mucho más preferiblemente el 98% de identidad de secuencia.

Como alternativa, el grado de similitud de secuencia entre polinucleótidos puede determinarse por hibridación de polinucleótidos en condiciones que permitan la formación de dúplex estables entre regiones homólogas, seguida por digestión con una o más nucleasas específicas de una hebra, y la determinación del tamaño de los fragmentos digeridos. Dos ácidos nucleicos, o dos secuencias polipeptídicas son sustancialmente homólogas entre sí cuando las secuencias muestran al menos aproximadamente el 70%-75%, preferiblemente 80%-82%, más preferiblemente 85%-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y mucho más preferiblemente 98% de identidad de secuencia sobre una longitud definida de las moléculas, determinado usando los métodos anteriores. Como se usa en la presente memoria, sustancialmente homólogo también se refiere a secuencias que muestran identidad completa con un ADN o secuencia polipeptídica específica. Las secuencias de ADN que son sustancialmente homólogas pueden identificarse en un experimento de hibridación de Southern en, por ejemplo, condiciones rigurosas, definidas para ese sistema particular. La definición de las condiciones apropiadas de hibridación pertenece a las habilidades de la técnica. Véase, por ejemplo, Sambrook et al., supra; Nucleic Acid Hybridization: A Practical Approach, editores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

La hibridación selectiva de dos fragmentos de ácido nucleico puede determinarse del siguiente modo. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta a la eficacia y fuerza de los eventos de hibridación entre dichas moléculas. Una secuencia de ácido nucleico parcialmente idéntica inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica a una molécula diana. La inhibición de la hibridación de la secuencia completamente idéntica puede evaluarse usando ensayos de hibridación que son bien conocidos en la técnica (por ejemplo, transferencia de Southern (DNA), transferencia de Northern (RNA), hibridación en solución, o similares, véase Sambrook, et al., Molecular Cloning: A Laboratory Manual, Segunda Edición, (1989) Cold Spring Harbor, N.Y.). Dichos ensayos pueden realizarse usando grados variables de selectividad, por ejemplo, usando condiciones que varían de baja a alta rigurosidad. Si se emplean condiciones de baja rigurosidad, puede evaluarse la ausencia de unión no específica usando una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda que tiene menos de aproximadamente el 30% de identidad de secuencia con la molécula diana), de modo que, en ausencia de eventos de unión no específica, la sonda secundaria no hibridará con la diana.

Cuando se utiliza un sistema de detección basado en hibridación, se elige una sonda de ácido nucleico que es complementaria a una secuencia de ácido nucleico referencia, y después por selección de condiciones apropiadas la sonda y la secuencia de referencia hibridarán de forma selectiva, o se unirán, entre sí para formar una molécula dúplex. Una molécula de ácido nucleico que es capaz de hibridar de forma selectiva con una secuencia de referencia en condiciones de hibridación moderadamente rigurosas típicamente hibridará en condiciones que permitan la detección de una secuencia de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tiene al menos aproximadamente el 70% de identidad de secuencia con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación rigurosas típicamente permiten la detección de secuencias de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencia de más de aproximadamente el 90-95% con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de la sonda/secuencia de referencia, donde la sonda y la secuencia de referencia tienen un grado específico de identidad de secuencia, pueden determinarse de

formas conocidas en la técnica (véase, por ejemplo, Nucleic Acid Hybridization: A Practical Approach, editores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; ERL Press).

Las condiciones de hibridación son bien conocidas para los expertos en la técnica. Rigurosidad de hibridación se refiere al grado al cual las condiciones de hibridación desfavorecen la formación de híbridos que contienen nucleótidos desapareados, correlacionándose una mayor rigurosidad con una inferior tolerancia por híbridos desapareados. Los factores que afectan a la rigurosidad de hibridación son bien conocidos para los expertos en la técnica e incluyen, pero sin limitarse a, la temperatura, el pH, la fuerza iónica, y la concentración de disolventes orgánicos tales como, por ejemplo, formamida y dimetilsulfóxido. Como saben los expertos en la técnica, la rigurosidad de hibridación se aumenta a temperaturas superiores, inferior fuerza iónica e inferiores concentraciones de disolvente.

Con respecto a las condiciones de rigurosidad para la hibridación, es bien sabido en la técnica que pueden emplearse numerosas condiciones equivalentes para establecer una rigurosidad particular variando, por ejemplo, los siguientes factores: la longitud y naturaleza de las secuencias, la composición de bases de las diversas secuencias, las concentraciones de sales y otros componentes de la solución de hibridación, la presencia o ausencia de agentes bloqueantes en las soluciones de hibridación (por ejemplo, sulfato de dextrano, y polietilenglicol), la temperatura de reacción de hibridación y los parámetros de tiempo, así como, variando las condiciones lavado. La selección de una serie particular de condiciones de hibridación se hace siguiendo métodos convencionales en la técnica (véase, por ejemplo, Sambrook, et al., Molecular Cloning: A Laboratory Manual Segunda Edición, (1989) Cold Spring Harbor, N.Y.).

"Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, "recombinación homóloga (HR)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble cadena en células. Este proceso requiere homología de la secuencia de nucleótidos, usa una molécula "donante" para reparar el molde de una molécula "diana" (es decir, la que ha experimentado la rotura de doble cadena), y se conoce de forma variada como "conversión génica no cruzada" o "conversión génica a corto plazo", porque conduce a la transferencia de información genética del donante a la diana. Sin el deseo de limitarse por ninguna teoría particular, dicha transferencia puede implicar la corrección de desapareamientos de ADN heterodúplex que se forma entre la diana rota y el donante, y/o "hibridación de hebra dependiente de síntesis", en que el donante se usa para resintetizar información genética que llegará a formar parte de la diana, y/o procesos relacionados. Dicha HR especializada a menudo provoca una alteración de la secuencia de la molécula diana de modo que parte de o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

"Escisión" se refiere a la rotura de la estructura covalente de una molécula de ADN. La escisión puede iniciarse por una diversidad de métodos incluyendo, pero sin limitarse a, hidrólisis enzimática o química de un enlace fosfodiéster. Es posible tanto una escisión monocatenaria como una escisión bicatenaria, y la escisión bicatenaria puede suceder como resultado de dos eventos de escisión monocatenaria distintos. La escisión del ADN puede provocar la producción de extremos romos o extremos escalonados. En ciertas realizaciones, se usan polipéptidos de fusión para escisión dirigida de ADN bicatenario.

Un "dominio de unión" comprende una o más secuencias polipeptídicas que tienen actividad catalítica para la escisión del ADN. Un dominio de escisión puede estar contenido en una única cadena polipeptídica o la actividad de escisión puede resultar de la asociación de dos (o más) polipéptidos.

Un "semi-dominio de escisión" es una secuencia polipeptídica que, junto con un segundo polipéptido (idéntico o diferente) forma un complejo que tiene actividad de escisión (preferiblemente actividad de escisión bicatenaria).

"Cromatina" es la estructura nucleoproteica que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteínas, incluyendo histonas y proteínas cromosómicas no histonas. La mayor parte de la cromatina celular eucariota existe en forma de nucleosomas, en donde un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociadas con un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; un ADN enlazador (de longitud variable dependiendo del organismo) se extiende entre núcleos de nucleosoma. Una molécula de histona H1 generalmente se asocia con el ADN enlazador. Para los propósitos de la presente descripción, el término "cromatina" se entiende como abarcando todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye cromatina tanto cromosómica como episómica.

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

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

Una "región accesible" es un sitio en la cromatina celular en que un sitio diana presente en el ácido nucleico puede

unirse por una molécula exógena que reconoce el sitio diana. Sin el deseo de limitarse por ninguna teoría particular, se cree que una región accesible es una que no está empaquetada en una estructura nucleosómica. La estructura distinta de una región accesible a menudo puede detectarse por su sensibilidad a sondas químicas y enzimáticas, por ejemplo, nucleasas.

- 5 Un "sitio diana" o "secuencia diana" es una secuencia de ácido nucleico que define una parte de un ácido nucleico a la cual se unirá una molécula de unión, con la condición de que existan suficientes condiciones para la unión. Por ejemplo, la secuencia 5'-GAATTC-3' es un sitio diana para la endonucleasa de restricción Eco RI.

10 Una molécula "exógena" es una molécula que no está normalmente presente en una célula, pero que puede introducirse en una célula por uno o más métodos genéticos, bioquímicos u otros métodos. La "presencia normal en la célula" se determina con respecto a la fase de desarrollo y condiciones ambientales particulares de la célula. Por tanto, por ejemplo, una molécula que está presente solamente durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular adulta. Así mismo, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula sin choque térmico. Una molécula exógena puede comprender, por ejemplo, una secuencia codificante de cualquier polipéptido o fragmento del mismo, una versión funcional de una molécula endógena que no funciona bien o una versión que no funciona bien de una molécula endógena que funciona normalmente. Una molécula exógena también puede ser el mismo tipo de molécula que una molécula endógena pero obtenida de una especie diferente que la especie de la que se obtiene la molécula endógena. Por ejemplo, una secuencia humana de ácido nucleico puede introducirse en una línea celular originaria de un hámster o ratón.

20 Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como la generada por un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, carbohidrato, lípido, glucoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprenda una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser monocatenarios o bicatenarios; pueden ser lineales, ramificados o circulares; y pueden ser de cualquier longitud. Los ácidos nucleicos incluyen aquellos capaces de formar dúplex, así como ácidos nucleicos que forman tríplex. Véanse, por ejemplo, las patentes de Estados Unidos N° 5.176.996 y 5.422.251. Las moléculas de ácido nucleico exógenas que pueden dirigirse para su inserción en un genoma también se mencionan como polinucleótidos "donantes". Las proteínas incluyen, pero sin limitarse a, proteínas de unión a ADN, factores de transcripción, factores de remodelado de la cromatina, proteínas de unión a ADN metilado, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, quinasas, fosfatasa, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicinas.

30 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína o ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma viral de infección, un plásmido o episoma introducido en una célula, o un cromosoma que no está normalmente presente en la célula. Los métodos para la introducción de moléculas exógenas en células son conocidos para los expertos en la técnica e incluyen, pero sin limitarse a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo con partículas, co-precipitación con fosfato cálcico, transferencia mediada por DEAE-dextrano y transferencia mediada por un vector viral.

40 En contraste, una molécula "endógena" es una que está normalmente presente en una célula particular en una fase particular del desarrollo en condiciones ambientales particulares. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, cloroplasto u otro orgánulo, o un ácido nucleico episómico de origen natural. Moléculas endógenas adicionales pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

45 Una molécula "de fusión" es una molécula en que están unidas dos o más moléculas de subunidad, preferiblemente de forma covalente. Las moléculas de subunidad pueden ser el mismo tipo químico de molécula, o pueden ser tipos químicos diferentes de moléculas. Ejemplos del primer tipo de molécula de fusión incluyen, pero no se limitan a, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión a ADN ZFP y un dominio de escisión) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita *supra*). Ejemplos del segundo tipo de molécula de fusión incluyen, pero no se limitan a, una fusión entre un ácido nucleico que forma tríplex y un polipéptido, y una fusión entre un agente de unión al surco menor y un ácido nucleico.

50 La expresión de una proteína de fusión en una célula puede resultar del suministro de la proteína de fusión a la célula o del suministro de un polinucleótido que codifica la proteína de fusión a una célula, en donde el polinucleótido se transcribe, y el transcrito se traduce, para generar la proteína de fusión. El corte y ajuste en trans, la escisión de polipéptidos y el ligamiento de polipéptidos también puede estar implicado en la expresión de una proteína en una célula. Se presentan métodos para el suministro de polinucleótidos y polipéptidos a células en otra parte de esta descripción.

55 Un "gen", a los efectos de la presente descripción, incluye una región de ADN que codifica un producto génico (véase *infra*), así como todas las regiones de ADN que regulan la producción del producto génico, estén o no dichas secuencias reguladoras adyacentes a secuencias codificantes y/o transcritas. Por consiguiente, un gen incluye, pero



no se limita necesariamente a, secuencias promotoras, terminadores, secuencias reguladoras de la transcripción tales como sitios de unión al ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aisladores, elementos limitadores, orígenes de replicación, sitios de unión a la matriz y regiones de control de locus.

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

10 "Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitarse a, activación génica y represión génica.

Las células "eucariotas" incluyen, pero no se limitan a, células fúngicas (tales como levaduras), células vegetales, células animales, células de mamífero y células humanas.

15 Las células "vegetales" incluyen, pero sin limitarse a, células de plantas monocotiledóneas (monocot) o dicotiledóneas (dicot). Ejemplos no limitantes de monocot incluyen plantas de cereal tales como maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, cebolla, banana, y coco. Ejemplos no limitantes de dicot incluyen tabaco, tomate, girasol, algodón, remolacha azucarera, patata, lechuga, melón, soja, canola (semilla de colza), y alfalfa. Las células vegetales pueden ser de cualquier parte de la planta y/o de cualquier fase de desarrollo de la planta.

20 Una "región de interés" es cualquier región de cromatina celular tal como, por ejemplo, un gen o una secuencia no codificante dentro de o adyacente a un gen, en que es deseable la unión a una molécula exógena. La unión puede ser con fines de escisión dirigida de ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma organular (por ejemplo, de mitocondria, cloroplasto), o un genoma viral infeccioso, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias líder, secuencias de avance o intrones, o dentro de regiones no transcritas, cadena arriba o cadena abajo de la región codificante. Una región de interés puede ser tan pequeña como un único par de nucleótidos o de hasta 2.000 pares de nucleótidos de longitud, o cualquier valor entero de pares de nucleótidos.

30 Las expresiones "unión funcional" y "unido de forma funcional" (o "unido funcionalmente") se usan de forma intercambiable con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en que los componentes están dispuestos de tal modo que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unido de forma funcional a una secuencia codificante si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción está generalmente unida de forma funcional en *cis* con una secuencia codificante, pero no necesita estar directamente adyacente a la misma. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está unida de forma funcional a una secuencia codificante, aunque no esté contigua.

40 Con respecto a polipéptidos de fusión, la expresión "unido de forma funcional" puede hacer referencia al hecho de que cada uno de los componentes realiza la misma función en unión al otro componente que si no estuvieran así unidos. Por ejemplo, con respecto a un polipéptido de fusión en que un dominio de unión a ADN ZFP está fusionado a un dominio de escisión, el dominio de unión a ADN ZFP y dominio de escisión están en unión funcional si, en el polipéptido de fusión, la parte del dominio de unión a ADN ZFP es capaz de unirse a su sitio diana y/o su sitio de unión, mientras que el dominio de escisión es capaz de escindir el ADN en las cercanías del sitio diana.

45 Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero retiene la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede poseer más, menos, o la misma cantidad de restos que la molécula nativa correspondiente, y/o puede contener una o más sustituciones de aminoácido o nucleótido. Los métodos para determinar la función de un ácido nucleico (por ejemplo, función codificante, capacidad de hibridar con otro ácido nucleico) son bien conocidos en la técnica. Así mismo, los métodos para determinar la función proteica son bien conocidos. Por ejemplo, la función de unión a ADN de un polipéptido puede determinarse, por ejemplo, por unión a filtro, desplazamiento de movilidad electroforética, o ensayos de inmunoprecipitación. La escisión del ADN puede ensayarse por electroforesis en gel. Véase Ausubel *et al.*, *supra*. La capacidad de una proteína de interactuar con otra proteína puede determinarse, por ejemplo, por co-inmunoprecipitación, ensayos de doble híbrido o complementación, tanto genéticos como bioquímicos. Véanse, por ejemplo, Fields *et al.* (1989) *Nature* 340:245-246; la patente de Estados Unidos N° 5.585.245 y el documento PCT WO 98/44350.

**Polinucleótidos exógenos (donantes)**

En la presente memoria se describen polinucleótidos para su inserción en el genoma, también conocidos como polinucleótidos "exógenos" o polinucleótidos "donantes". Se ha demostrado que los donantes plasmídicos que portan brazos de homología de 750 pb que flanquean un transgén de interés, en combinación con nucleasas de dedos de zinc (ZFN) diseñadas pueden usarse para la alteración génica dirigida. Véanse, por ejemplo, Moehle et al. (2007) *Proc. Nat'l. Acad. Sci. USA* 104(9):3055-3060 y la publicación de patente de Estados Unidos N° 20050064474. La construcción de dichos polinucleótidos donantes plasmídicos con largos brazos de homología es un procedimiento que lleva mucho tiempo, implicando: diseño de cebadores de PCR que amplifiquen un fragmento de ~1,5 kb del locus de interés; identificación (por amplificación, clonación y secuenciación) de un clon individual que porte el fragmento deseado y no tenga mutaciones inducidas por PCR; introducción (típicamente por mutagénesis dirigida al sitio) de un único RFLP en el centro de ese fragmento; clonación de la ORF de interés en ese fragmento; identificación (típicamente por digestión de restricción) de un clon que porte la ORF en la orientación deseada; y amplificación del plásmido en cantidades suficientes para su uso en alteración genómica dirigida. En las mejores circunstancias, este proceso dura aproximadamente dos semanas y produce un polinucleótido donante circular (plasmídico).

Sorprendentemente, en la presente memoria demostramos que las secuencias donantes lineales de la descripción que comprenden cortos brazos de homología de aproximadamente 50-100 pares de bases pueden integrarse de forma eficaz en el genoma de la célula. Las secuencias donantes lineales descritas en la presente memoria lleva sólo horas construirlas.

En ciertas realizaciones, las secuencias donantes lineales descritas en la presente memoria son de 25 a 50 pares de bases de longitud (o cualquier valor entre ellos, incluyendo 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40,41, 42, 43,44, 45,46, 47, 48,49 o 50 nucleótidos). En otras realizaciones, las secuencias de entre 50 y 75 nucleótidos de longitud (incluyendo 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74 o 75 nucleótidos de longitud). En otras realizaciones más, las secuencias son de entre 75 y 100 nucleótidos de longitud (incluyendo 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, o 100 nucleótidos de longitud). En otras realizaciones más, los polinucleótidos donantes son de entre 100 y 150 nucleótidos de longitud (o cualquier entre ellos). En otras realizaciones, los polinucleótidos donantes son de entre 50 y 750 nucleótidos de longitud (por ejemplo, 50 y 100, 50 y 150, 50 y 200, 50 y 250, 50 y 300, 50 y 350, 50 y 400, 50 y 450, 50 y 500, 50 y 550, 50 y 600, 50 y 650, 50 y 700).

Las secuencias donantes descritas en la presente memoria pueden aislarse de los plásmidos, células u otras fuentes usando técnicas convencionales conocidas en la técnica tales como PCR. Como alternativa, pueden sintetizarse de forma química usando técnicas convencionales de síntesis de oligonucleótidos. Típicamente, los polinucleótidos donantes se preparan por PCR usando un cebador con una parte 5' de 50-100 pb homóloga a la diana genómica, y una parte de 15-18 pb idéntica a la ORF de interés (FIG. 1).

Los polinucleótidos donantes lineales descritos en la presente memoria pueden incluir uno o más enlaces fosfodiéster fosforotioato entre pares de bases terminales para proteger al polinucleótido donante lineal de la degradación exonucleolítica. Estos enlaces pueden estar en dos o más posiciones en los extremos 5' y/o 3' de la molécula y pueden añadirse durante el aislamiento o síntesis usando metodología convencional. Véanse, por ejemplo, Ciafre et al. (1995) *Nucleic Acids Res.* 23(20):4134-42; Johansson et al. (2002) *Vaccine* 20(27-28):3379-88. En realizaciones en que el polinucleótido donante se aísla por PCR usando cebadores (FIG. 1), los extremos 5' del cebador (y polinucleótido donante) son típicamente enlaces fosfodiéster fosforotioato. Como alternativa, los polinucleótidos donantes lineales pueden incluir uno o más desoxinucleótidos 5', biotina y/o uno o más grupos amina, todos los cuales han demostrado reducir la degradación exonucleolítica.

El polinucleótido exógeno (donante) puede comprender cualquier secuencia de interés (secuencia exógena). Las secuencias exógenas ejemplares incluyen, pero no se limitan a, cualquier secuencia que codifique un polipéptido (por ejemplo, ADNc), secuencias promotoras, secuencias potenciadoras, marcas epitópicas, genes marcadores, sitios de reconocimiento por enzimas de restricción, marcas epitópicas y diversos tipos de construcciones de expresión. Los genes marcadores incluyen, pero no se limitan a, secuencias que codifican proteínas que median la resistencia a antibióticos (por ejemplo, resistencia a ampicilina, resistencia a neomicina, resistencia a G418, resistencia a puromicina), secuencias que codifican proteínas coloreadas o fluorescentes o luminiscentes (por ejemplo, proteína fluorescente verde, proteína fluorescente verde potenciada, proteína fluorescente roja, luciferasa), y proteínas que median el crecimiento celular y/o amplificación génica potenciada (por ejemplo, dihidrofolato reductasa). Las marcas epitópicas incluyen, por ejemplo, una o más copias de FLAG, His, myc, Tap, HA o cualquier secuencia detectable de aminoácidos.

El polinucleótido exógeno (donante) también puede comprender secuencias que no codifican polipéptidos si no cualquier tipo de secuencia no codificante, así como uno o más elementos de control (por ejemplo, promotores). Además, la secuencia exógena de ácido nucleico puede producir una o más moléculas de ARN (por ejemplo, ARN pequeños de horquilla (ARNhp), ARN inhibidores (ARNi), microARN (miARN), etc.).

Una molécula donante puede contener varias regiones discontinuas de homología a la cromatina celular. Por

ejemplo, las regiones de homología pueden flanquear dos o más regiones que contienen las alteraciones deseadas. En una realización preferida, la secuencia exógena comprende un polinucleótido que codifica cualquier polipéptido del cual se desea su expresión en la célula, incluyendo, pero sin limitarse a, anticuerpos, antígenos, enzimas, receptores (de superficie celular o nucleares), hormonas, linfoquinas, citoquinas, polipéptidos indicadores, factores de crecimiento, y fragmentos funcionales de cualquiera de los anteriores. Las secuencias codificantes pueden ser, por ejemplo, ADNc.

Una molécula donante puede ser una molécula lineal tras su linealización, como resultado de escisión dirigida por ZFN, de un plásmido captado por una célula. En otra realización, la molécula donante lineal puede residir en el genoma de la célula en donde la molécula donante llega a estar disponible para su integración dirigida realizada por homología tras la escisión dirigida por ZFN y liberación del donante del genoma.

Por ejemplo, la secuencia exógena puede comprender una secuencia que codifica un polipéptido que carece o no es funcional en el sujeto que tiene una enfermedad genética, incluyendo, pero sin limitarse a, cualquiera de las siguientes enfermedades genéticas: acondroplasia, acromatopsia, deficiencia de maltasa ácida, deficiencia de adenosina desaminasa (OMIM N° 102700), adrenoleucodistrofia, síndrome de aicardi, deficiencia de alfa-1 antitripsina, alfa-talasemia, síndrome de insensibilidad a andrógenos, síndrome de apert, enfermedad arritmogénica del ventrículo derecho, displasia, ataxia telangiectasia, síndrome de Barth, beta-talasemia, síndrome del nevus gomoso azul, enfermedad de Canavan, enfermedades granulomatosas crónicas (CGD), síndrome de cri du chat, fibrosis quística, enfermedad de dercum, displasia ectodérmica, anemia de Fanconi, fibrodisplasia osificante progresiva, síndrome del X frágil, galactosemia, enfermedad de Gaucher, gangliosidosis generalizada (por ejemplo, GM1), hemocromatosis, la mutación de hemoglobina C en el 6° codón de beta-globina (HbC), hemofilia, enfermedad de Huntington, síndrome de Hurler, hipofosfatasa, síndrome de Klinefelter, enfermedad de Krabbes, síndrome de Langer-Giedion, deficiencia de adhesión de leucocitos (LAD, OMEM N° 116920), leucodistrofia, síndrome de QT largo, síndrome de Marfan, síndrome de Moebius, mucopolisacaridosis (MPS), síndrome uña-rótula, diabetes insípida nefrogénica, neurofibromatosis, enfermedad de Neimann-Pick, osteogénesis imperfecta, porfiria, síndrome de Prader-Willi, progeria, síndrome de Proteus, retinoblastoma, síndrome de Rett, síndrome de Rubinstein-Taybi, síndrome de Sanfilippo, inmunodeficiencia combinada severa (SCID), síndrome de Shwachman, enfermedad falciforme (anemia falciforme), síndrome de Smith-Magenis, síndrome de Stickler, enfermedad de Tay-Sachs, síndrome de trombocitopenia con radios ausentes (TAR), síndrome de Treacher Collins, trisomía, esclerosis tuberosa, síndrome de Turner, trastorno del ciclo de la urea, enfermedad de von Hippel-Landau, síndrome de Waardenburg, síndrome de Williams, enfermedad de Wilson, síndrome de Wiskott-Aldrich, síndrome linfoproliferativo ligado al X (XLP, OMIM N° 308240).

Enfermedades ejemplares adicionales que pueden tratarse por integración dirigida incluyen inmunodeficiencias adquiridas, enfermedades de almacenamiento lisosómico (por ejemplo, enfermedad de Gaucher, GM1, enfermedad de Fabry y enfermedad de Tay-Sachs), mucopolisacaridosis (por ejemplo, enfermedad de Hunter, enfermedad de Hurler), hemoglobinopatías (por ejemplo, enfermedades falciformes, HbC,  $\alpha$ -talasemia,  $\beta$ -talasemia) y hemofilias.

En ciertas realizaciones, las secuencias exógenas pueden comprender un gen marcador (descrito anteriormente), que permite la selección de células que han experimentado integración dirigida, y una secuencia ligada que codifica una función adicional. Ejemplos no limitantes de genes marcadores incluyen GFP, marcador o marcadores de selección por fármacos y similares.

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

### Sitios diana

Los métodos descritos y composiciones incluyen proteínas de fusión que comprenden un dominio de escisión (o un semi-dominio de escisión) y un dominio de dedos de zinc, en que el dominio de dedos de zinc, mediante la unión a una a una secuencia en la región de interés en el genoma de una célula dirige la actividad del dominio de escisión (o semi-dominio de escisión) a las cercanías de la secuencia y, por tanto, induce escisión (por ejemplo, una rotura bicatenaria) en la región de interés. Como se expone en otra parte en esta descripción, puede diseñarse un dominio de dedos de zinc para que se una a casi cualquier secuencia deseada. Por consiguiente, puede diseñarse uno o más dominios de unión de dedos de zinc para que se unan a una o más secuencias en la región de interés. La expresión de una proteína de fusión que comprende un dominio de unión de dedos de zinc y un dominio de escisión (o de dos proteínas de fusión, comprendiendo cada una un dominio de unión de dedos de zinc y un semi-dominio de escisión), en una célula, logra la escisión en la región de interés.

La selección de una secuencia en una región de interés para la unión por un dominio de dedos de zinc (por ejemplo, un sitio diana) puede conseguirse, por ejemplo, de acuerdo con los métodos descritos en la patente de Estados Unidos N. 6.453.242 (17 de septiembre de 2002) del mismo propietario que la presente, que también describe métodos para diseñar ZFP que se unen a una secuencia seleccionada. Quedará claro para los expertos en la técnica que la simple inspección visual de una secuencia de nucleótidos también puede usarse para la selección de un sitio diana. Por consiguiente, puede usarse cualquier medio para la selección de un sitio diana en los métodos

descritos en la presente memoria.

Los sitios diana generalmente están compuestos por una pluralidad de subsitios diana adyacentes. Un subsitio diana se refiere a la secuencia (habitualmente un triplete de nucleótidos, o un cuadruplete de nucleótidos que puede solaparse en un nucleótido con un cuadruplete adyacente) unida por un dedo de zinc individual. Véase, por ejemplo, el documento WO 02/077227. Si la hebra con la cual una proteína de dedos de zinc hace la mayoría de los contactos, se señala como la hebra diana "hebra primaria de reconocimiento" o "hebra primaria de contacto", algunas proteínas de dedos de zinc se unirán a un triplete de tres bases en la hebra diana y una cuarta base en la hebra no diana. Un sitio diana generalmente tiene una longitud de al menos 9 nucleótidos y, por consiguiente, se une mediante un dominio de unión de dedos de zinc que comprende al menos tres dedos de zinc. Sin embargo, la unión de, por ejemplo, un dominio de unión 4 dedos a un sitio diana de 12 nucleótidos, un dominio de unión de 5 dedos a un sitio diana de 15 nucleótidos o un dominio de unión de 6 dedos a un sitio diana de 18 nucleótidos, también es posible. Como será evidente, la unión de dominios de unión más grandes (por ejemplo, de 7, 8, 9 dedos y más) a sitios diana más largos también es posible.

No es necesario que un sitio diana sea un múltiplo de tres nucleótidos. Por ejemplo, en casos en que suceden interacciones cruzadas entre hebras (véase, por ejemplo, la patente de Estados Unidos 6.453.242 y el documento WO 02/077227), puede unirse uno o más de los dedos de zinc individuales de un dominio de unión de múltiples dedos a subsitios de cuadruplete solapantes. Como resultado, una proteína de tres dedos puede unirse a una secuencia de 10 nucleótidos, en donde el décimo nucleótido es parte de un cuadruplete unido por un dedo terminal, una proteína de cuatro dedos puede unirse a una secuencia de 13 nucleótidos, en donde el decimotercer nucleótido es parte de un cuadruplete unido por un dedo terminal, etc.

La longitud y naturaleza de las secuencias enlazadoras de aminoácidos entre dedos de zinc individuales en un dominio de unión de múltiples dedos también afecta a la unión a una secuencia diana. Por ejemplo, la presencia de un llamado "enlazador no canónico", "enlazador largo" o "enlazador estructurado" entre dedos de zinc adyacentes en un dominio de unión de múltiples dedos puede permitir que esos dedos se unan a subsitios que no están inmediatamente adyacentes. Ejemplos no limitantes de dichos enlazadores se describen, por ejemplo, en la patente de Estados Unidos N° 6.479.626 y el documento WO 01/53480. Por consiguiente, uno o más subsitios, en un sitio diana para un dominio de unión de dedos de zinc, pueden estar separados entre sí por 1, 2, 3, 4, 5 o más nucleótidos. Para proporcionar solamente un ejemplo, un dominio de unión de cuatro dedos puede unirse a un sitio diana de 13 nucleótidos que comprende, en secuencia, dos subsitios contiguos de 3 nucleótidos, un nucleótido intermedio, y dos subsitios contiguos de triplete.

La distancia entre secuencias (por ejemplo, sitios diana) se refiere a la cantidad de nucleótidos o pares de nucleótidos intermedios entre dos secuencias, medida desde los bordes de las secuencias más cercanas entre sí.

En ciertas realizaciones en que la escisión depende de la unión de dos moléculas de fusión de dominio de dedos de zinc/semi-dominio de escisión para separar los sitios diana, los dos sitios diana pueden estar en hebras opuestas de ADN (Ejemplo 1). En otras realizaciones, ambos sitios diana están en la misma hebra de ADN.

### **Dominios de unión a ADN**

Puede usarse cualquier dominio de unión a ADN en los métodos descritos en la presente memoria. En ciertas realizaciones, el dominio de unión a ADN comprende una proteína de dedos de zinc. Un dominio de unión de dedos de zinc comprende uno o más dedos de zinc. Miller *et al.* (1985) *EMBO J.* 4:1609-1614; Rhodes (1993) *Scientific American* Feb.:56-65; patente de Estados Unidos N° 6.453.242. Los dominios de unión de dedos de zinc descritos en la presente memoria generalmente incluyen 2, 3, 4, 5, 6 o incluso más dedos de zinc.

Típicamente, un dominio de dedos de zinc individual es de aproximadamente 30 aminoácidos de longitud. Estudios estructurales han demostrado que cada dominio (motivo) de dedos de zinc contiene dos láminas beta (sostenidas en un giro beta que contiene los dos restos de cisteína invariables) y una hélice alfa (que contiene los dos restos de histidina invariables), que se mantienen en una conformación particular a través de la coordinación de un átomo de zinc por las dos cisteínas y las dos histidinas.

Los dedos de zinc incluyen tanto dedos de zinc  $C_2H_2$  canónicos (es decir, aquellos en que el ión zinc está coordinado por dos restos de cisteína y dos de histidina) y dedos de zinc no canónicos tales como, por ejemplo, dedos de zinc  $C_3H$  (aquellos en que el ión zinc está coordinado por tres restos de cisteína y un resto de histidina) y dedos de zinc  $C_4$  (aquellos en que el ión zinc está coordinado por cuatro restos de cisteína). Véase también el documento WO 02/057293.

Los dominios de unión de dedos de zinc pueden diseñarse para que se unan a un sitio diana (véase anteriormente) usando técnicas convencionales. Véase el Ejemplo 1; las patentes de Estados Unidos 6.453.242 y 6.534.261 del mismo propietario que la presente, incluyendo las referencias citadas en las mismas. Un dominio de unión de dedos de zinc diseñado por ingeniería puede tener una nueva especificidad de unión, en comparación con una proteína de dedos de zinc de origen natural. Los métodos de diseño por ingeniería incluyen, pero no se limitan a, el diseño racional y diversos tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos de tripletes (o cuadrupletes) y secuencias de aminoácidos de dedos de zinc

individuales, en que cada secuencia de nucleótidos de tripletes o cuadrupletes está asociada con una o más secuencias de aminoácidos de dedos de zinc que se unen a la secuencia particular de tripletes o cuadrupletes.

Se describen métodos de selección ejemplares, incluyendo presentación en fagos y sistemas de doble híbrido, en las patentes de Estados Unidos 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como los documentos WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.

Se ha descrito la potenciación de la especificidad de unión para dominios de unión de dedos de zinc, por ejemplo, en el documento WO 02/077227 del mismo propietario que la presente.

Como un dedo de zinc individual se une a una secuencia de tres nucleótidos (es decir, triplete) (o una secuencia de cuatro nucleótidos que puede solapar, en un nucleótido, con el sitio de unión de cuatro nucleótidos de un dedo de zinc adyacente), la longitud de una secuencia para la cual se diseña un dominio de unión de dedos de zinc para que se una a la misma (por ejemplo, una secuencia diana) determinará la cantidad de dedos de zinc en un dominio de unión de dedos de zinc diseñado por ingeniería. Por ejemplo, para ZFP en que los motivos de dedo no se unen a subsitios solapantes, una secuencia diana de seis nucleótidos se une mediante un dominio de unión de dos dedos; una secuencia diana de nueve nucleótidos se une mediante un dominio de unión de tres dedos, etc. Como se indica en la presente memoria, los sitios de unión para dedos de zinc individuales (es decir, subsitios) en un sitio diana no tienen que ser contiguos, sino que pueden estar separados por uno o varios nucleótidos, dependiendo de la longitud y naturaleza de las secuencias de aminoácidos entre los dedos de zinc (es decir, los enlazadores inter-dedos) en un dominio de unión de múltiples dedos.

En un dominio de unión de dedos de zinc de múltiples dedos, los dedos de zinc adyacentes pueden estar separados por secuencias enlazadoras de aminoácidos de aproximadamente 5 aminoácidos (los llamados enlazadores inter-dedos "canónicos") o, como alternativa, por uno o más enlazadores no canónicos. Véanse, por ejemplo, las patentes de Estados Unidos N° 6.453.242 y 6.534.261 del mismo propietario que la presente. Para dominios de unión de dedos de zinc diseñados por ingeniería que comprenden más de tres dedos, puede preferirse la inserción de enlazadores inter-dedos más largos ("no canónicos") entre ciertos dedos de zinc ya que esto puede aumentar la afinidad y/o especificidad de la unión por el dominio de unión. Véase, por ejemplo, la patente de Estados Unidos N° 6.479.626 y el documento WO 01/53480. Por consiguiente, los dominios de unión de dedos de zinc de múltiples dedos también pueden caracterizarse con respecto a la presencia y localización de enlazadores inter-dedos no canónicos. Por ejemplo, un dominio de unión de dedos de zinc de seis dedos que comprende tres dedos (unidos por dos enlazadores inter-dedos canónicos), un enlazador largo y tres dedos adicionales (unidos por dos enlazadores inter-dedos canónicos) se señala como una configuración 2x3. Asimismo, un dominio de unión que comprende dos dedos (con un enlazador canónico entre los mismos), un enlazador largo y dos dedos adicionales (unidos por un enlazador canónico) se señala como una proteína 2x2. Una proteína que comprende tres unidades de dos dedos (en cada una de las cuales los dos dedos están unidos por un enlazador canónico), y en que cada unidad de dos dedos está unida a la unidad de dos dedos adyacentes por un enlazador largo, se menciona como una proteína 3x2.

La presencia de un enlazador inter-dedos largo o no canónico entre dos dedos de zinc adyacentes en un dominio de unión de múltiples dedos a menudo permite que los dos dedos se unan a subsitios que no están inmediatamente contiguos en la secuencia diana. Por consiguiente, puede haber huecos de uno o más nucleótidos entre los subsitios en un sitio diana; es decir, un sitio diana puede contener uno o más nucleótidos que no están en contacto con un dedo de zinc. Por ejemplo, un dominio de unión de dedos de zinc 2x2 puede unirse a dos secuencias de seis nucleótidos separadas por un nucleótido, es decir, se une a un sitio diana de 13 nucleótidos. Véase también Moore *et al.* (2001a) Proc. Natl. Acad. Sci. USA **98**:1432-1436; Moore *et al.* (2001b) Proc. Natl. Acad. Sci. USA **98**:1437-1441 y el documento WO 01/53480.

Como se ha mencionado previamente, un subsitio diana es una secuencia de tres o cuatro nucleótidos que se une por un único dedo de zinc. Para ciertos propósitos, una unidad de dos dedos se indica como un módulo de unión. Un módulo de unión puede obtenerse, por ejemplo, seleccionando dos dedos adyacentes en el contexto de una proteína de múltiples dedos (generalmente tres dedos) que se une a una secuencia diana particular de seis nucleótidos. Como alternativa, pueden construirse módulos mediante el ensamblaje de dedos de zinc individuales. Véanse también los documentos WO 98/53057 y WO 01/53480.

Como alternativa, el dominio de unión a ADN puede obtenerse de una nucleasa. Por ejemplo, las secuencias de reconocimiento de endonucleasas de asentamiento y meganucleasas tales como I-SceI, I-CeuI, P1-Pspl, VI-Sce, 1-SceIV, I-Csml, I-ParI, I-SceU, I-PpoI, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII son conocidas. Véase también la patente de Estados Unidos N° 5.420.032; la patente de Estados Unidos N° 6.833.252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25**:3379-3388; Dujon *et al.* (1989) *Gene* **82**:115-118; Perler *et al.* (1994) *Nucleic Acids Res.* **22**, 1125-1127; Jasin (1996) *Trends Genet.* **12**:224-228; Gimble *et al.* (1996) *J. Mol. Biol.* **263**:163-180; Argast *et al.* (1998) *J. Mol. Biol.* **280**:345-353 y el catálogo de New England Biolabs. Además, la especificidad de unión a ADN de las endonucleasas de asentamiento y meganucleasas puede diseñarse para que se unan a sitios diana no naturales. Véase, por ejemplo, Chevalier *et al.* (2002) *Molec. Cell* **10**:895-905; Epinat *et al.* (2003) *Nucleic Acids Res.* **31**:2952-2962; Ashworth *et al.* (2006) *Nature* **441**:656-659; Paques *et al.* (2007) *Current Gene Therapy* **7**:49-66; la publicación de patente de Estados Unidos N° 20070117128.

## Dominios de escisión

La parte de dominio de escisión de las proteínas de fusión descritas en la presente memoria puede obtenerse de cualquier endonucleasa o exonucleasa. Endonucleasas ejemplares a partir de las cuales puede obtenerse un dominio de escisión incluyen, pero no se limitan a, endonucleasas de restricción y endonucleasas de asentamiento. Véase, por ejemplo, el catálogo 2002-2003, New England Biolabs, Beverly, MA; y Belfort *et al.* (1997) *Nucleic Acids Res.* **25**:3379-3388. Se conocen enzimas adicionales que escinden el ADN (por ejemplo, nucleasa S1; nucleasa de frijol mungo; DNasa I pancreática; nucleasa de micrococos; endonucleasa HO de levaduras; véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Ejemplos no limitantes de endonucleasas de asentamiento y meganucleasas incluyen I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-Scell, I-Ppol, I-ScellI, I-Crel, I-TeVl, I-TeVlI y I-TeVlIII que son conocidas. Véase también la patente de Estados Unidos N° 5.420.032; la patente de Estados Unidos N° 6.833.252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25**:3379-3388; Dujon *et al.* (1989) *Gene* **82**:115-118; Perler *et al.* (1994) *Nucleic Acids Res.* **22**, 1125-1127; Jasin (1996) *Trends Genet.* **12**:224-228; Gimble *et al.* (1996) *Mol. Biol.* **263**:163-180; Argast *et al.* (1998) *J. Mol. Biol.* **280**:345-353 y el catálogo de New England Biolabs. Puede usarse una o más de estas enzimas (o fragmentos funcionales de las mismas) como fuente de dominios de escisión y semi-dominios de escisión.

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

Una enzima de restricción ejemplar de Tipo IIS, cuyo dominio de escisión se puede separar del dominio de unión, es *Fok I*. Esta enzima particular es activa como dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95**: 10.570-10.575. Por consiguiente, para los propósitos de la presente descripción, la parte de la enzima *Fok I* usada en las proteínas de fusión descritas se considera un semi-dominio de escisión. Por tanto, para escisión bicatenaria dirigida y/o remplazo dirigido de secuencias celulares usando fusiones dedo de zinc-*Fok I*, pueden usarse dos proteínas de fusión, comprendiendo cada una un semi-dominio de escisión de *FokI*, para reconstituir un dominio de escisión catalíticamente activo. Como alternativa, también puede usarse una única molécula polipeptídica que contenga un dominio de unión de dedos de zinc y dos semi-dominios de escisión de *Fok I*. Los parámetros para la escisión dirigida y la alteración dirigida de la secuencia usando fusiones de dedo de zinc-*Fok I* se proporcionan en otra parte en esta descripción.

Un dominio de escisión o semi-dominio de escisión puede ser cualquier parte de una proteína que retenga la actividad de escisión, o que retenga la capacidad de multimerizar (por ejemplo, dimerizar) para formar un dominio de escisión funcional.

Se describen enzimas de restricción ejemplares de Tipo IIS en la publicación internacional WO 2007/014275 del mismo propietario que la presente, incorporada por referencia en la presente memoria en su totalidad.

Para potenciar la especificidad de escisión, también pueden modificarse los dominios de escisión. En ciertas realizaciones, se emplean variantes del semi-dominio de escisión, que son variantes que minimizan o evitan la homodimerización de los semi-dominios de escisión. Se describen en detalle ejemplos no limitantes de dichos semi-dominios de escisión modificados en el documento WO 2007/014275, incorporado por referencia en su totalidad en la presente memoria. Véanse también los Ejemplos. En ciertas realizaciones, el dominio de escisión comprende un semi-dominio de escisión diseñado por ingeniería (también mencionado como dominio de dimerización mutante) que minimiza o evita la homodimerización, lo que es conocido para los expertos en la técnica y se describe, por ejemplo, en las publicaciones de patente de Estados Unidos N° 20050064474 y 20060188987, incorporadas por referencia en sus totalidades en la presente memoria. Los restos aminoacídicos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, y 538 de *Fok I* son todos dianas para influenciar en la dimerización de los semi-dominios de escisión de *Fok I*.

También pueden usarse semi-dominios de escisión adicionales de *Fok I* diseñados por ingeniería que forman heterodímeros obligados en las ZFN descritas en la presente memoria. El primer semi-dominio de escisión incluye mutaciones en los restos aminoacídicos en las posiciones 490 y 538 de *Fok I* y el segundo semi-dominio de escisión incluye mutaciones en los restos aminoacídicos 486 y 499.

En ciertas realizaciones, el dominio de escisión comprende dos semi-dominios de escisión, ambos cuales son parte de un único polipéptido que comprende un dominio de unión, un primer semi-dominio de escisión y un segundo semi-dominio de escisión. Los semi-dominios de escisión pueden tener la misma secuencia de aminoácidos o

diferentes secuencias de aminoácidos, siempre que funcionen escindiendo el ADN.

En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semi-dominios de escisión. Como alternativa, puede usarse una única proteína que comprenda dos semi-dominios de escisión. Los dos semi-dominios de escisión pueden obtenerse de la misma endonucleasa (o fragmentos funcionales de la misma), o cada semi-dominio de escisión puede obtenerse de una endonucleasa diferente (o fragmentos funcionales de las mismas). En otra realización más, se usan dos semi-dominios de escisión en donde uno de los semi-dominios es enzimáticamente inactivo, de modo que se introduce una mella monocatenaria en el sitio diana (véase, por ejemplo, la solicitud provisional de Estados Unidos 61/189.800 del mismo propietario que la presente). Además, los sitios diana para las dos proteínas de fusión están dispuestos preferiblemente, uno con respecto al otro, de tal modo que la unión de las dos proteínas de fusión a sus respectivos sitios diana sitúe los semi-dominios de escisión en una orientación espacial entre sí que permita que los semi-dominios de escisión formen un dominio de escisión funcional, por ejemplo, por dimerización. Por tanto, en ciertas realizaciones, los bordes cercanos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, no puede interponerse ningún número integral de nucleótidos o pares de nucleótidos entre dos sitios diana (por ejemplo, de 2 a 50 nucleótidos o más). En general, el punto de escisión recae entre los sitios diana.

### Fusiones de dominio de unión a ADN-dominio de escisión

Los métodos para el diseño y construcción de proteínas de fusión (y polinucleótidos que codifican los mismos) son conocidos para los expertos en la técnica. Por ejemplo, se describen métodos para el diseño y construcción de proteínas de fusión que comprenden proteínas de dedos de zinc (y polinucleótidos que codifican las mismas) en las patentes de Estados Unidos 6.453.242 y 6.534.261 del mismo propietario que la presente; y la publicación internacional WO 2007/014275. En ciertas realizaciones, se construyen los polinucleótidos que codifican dichas proteínas de fusión. Estos polinucleótidos pueden insertarse en un vector y el vector puede introducirse en una célula (véase a continuación para una descripción adicional respecto a vectores y métodos para introducir polinucleótidos en células).

En ciertas realizaciones de los métodos descritos en la presente memoria, una proteína de fusión comprende un dominio de unión de dedos de zinc y un semi-dominio de escisión de la enzima de restricción *Fok I*, y se expresan dos de estas proteínas de fusión en una célula. La expresión de dos proteínas de fusión en una célula puede resultar del suministro de las dos proteínas a la célula; el suministro de una proteína y un ácido nucleico que codifica una de las proteínas a la célula; el suministro de dos ácidos nucleicos, conteniendo cada uno una de las proteínas, a la célula; o el suministro de un único ácido nucleico, que codifica ambas proteínas, a la célula. En realizaciones adicionales, una proteína de fusión comprende una única cadena polipeptídica que comprende dos semi-dominios de escisión y un dominio de unión de dedos de zinc. En este caso, se expresa una única proteína de fusión en una célula y, sin el deseo de limitarse por ninguna teoría, se cree que escinde el ADN como resultado de la formación de un dímero intramolecular de los semi-dominios de escisión.

Pueden expresarse dos proteínas de fusión, comprendiendo cada una un dominio de unión de dedos de zinc y un semi-dominio de escisión, en una célula, y unirse a sitios diana que están yuxtapuestos de tal modo que se reconstituya un dominio de escisión funcional y se escinda el ADN en las cercanías de los sitios diana. En una realización, la escisión sucede entre los sitios diana de los dos dominios de unión de dedos de zinc. Pueden diseñarse por ingeniería los dominios de unión de dedos de zinc y/o los dominios de escisión.

Los componentes de las proteínas de fusión (por ejemplo, fusiones ZFP-*Fok I*) pueden disponerse de tal modo que el dominio de dedos de zinc esté lo más cerca posible del extremo amino terminal de la proteína de fusión, y el semi-dominio de escisión esté lo más cerca posible del extremo carboxi-terminal. La dimerización de los semi-dominios de escisión para formar una nucleasa funcional se consigue uniendo las proteínas de fusión a sitios en hebras opuestas del ADN, estando los extremos 5' de los sitios de unión proximales entre sí.

Como alternativa, los componentes de las proteínas de fusión (por ejemplo, fusiones ZFP-*Fok I*) pueden disponerse de tal modo que el semi-dominio de escisión esté lo más cerca posible del extremo amino terminal de la proteína de fusión, y el dominio de dedos de zinc esté lo más cerca posible del extremo carboxi-terminal. En estas realizaciones, la dimerización de los semi-dominios de escisión para formar una nucleasa funcional se consigue uniendo las proteínas de fusión a sitios en hebras opuestas del ADN, estando los extremos 3' de los sitios de unión proximales entre sí.

En realizaciones adicionales más, una primera proteína de fusión contiene el semi-dominio de escisión lo más cerca posible del extremo amino terminal de la proteína de fusión, y el dominio de dedos de zinc lo más cerca posible del extremo carboxi-terminal, y una segunda proteína de fusión está dispuesta de tal modo que el dominio de dedos de zinc esté lo más cerca posible del extremo amino terminal de la proteína de fusión, y el semi-dominio de escisión esté lo más cerca posible del extremo carboxi-terminal. En estas realizaciones, ambas proteínas de fusión se unen a la misma hebra de ADN, conteniendo el sitio de unión de la primera proteína de fusión el dominio de dedos de zinc lo más cerca posible del extremo carboxi terminal localizado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio de dedos de zinc lo más cerca posible del extremo amino terminal.

Las dos proteínas de fusión puede unirse en la región de interés en la misma polaridad u opuesta, y sus sitios de unión (es decir, sitios diana) pueden estar separados por cualquier cantidad de nucleótidos, por ejemplo, de 0 a 200 nucleótidos o cualquier valor integrante entre los mismos. En ciertas realizaciones, los sitios de unión para dos proteínas de fusión, comprendiendo cada una un dominio de unión de dedos de zinc y un semi-dominio de escisión, pueden estar localizados entre 5 y 18 nucleótidos separados, por ejemplo, 5-8 nucleótidos separados, o 15-18 nucleótidos separados, o 6 nucleótidos separados, o 16 nucleótidos separados, medido desde el borde de cada sitio de unión lo más cerca posible del otro sitio de unión, y la escisión sucede entre los sitios de unión.

El sitio en el que se escinde el ADN generalmente recae entre los sitios de unión para las dos proteínas de fusión. La rotura bicatenaria del ADN a menudo resulta de dos roturas monocatenarias, o "mellas", desplazadas en 1, 2, 3, 4, 5, 6 o más nucleótidos (por ejemplo, la escisión de ADN bicatenario por *Fok I* nativa resulta de roturas monocatenarias desplazadas en 4 nucleótidos). Por tanto, la escisión no necesariamente sucede en sitios exactamente opuestos en cada hebra de ADN. Además, la estructura de las proteínas de fusión y la distancia entre los sitios diana puede influir en si la escisión sucede adyacente a un único par de nucleótidos, o si la escisión sucede en varios sitios. Sin embargo, para la integración dirigida, la escisión dentro de un intervalo de nucleótidos generalmente es suficiente, y no se requiere la escisión entre pares de bases particulares.

En las proteínas de fusión descritas, la secuencia de aminoácidos entre el dominio de dedos de zinc y el dominio de escisión (o semi-dominio de escisión) se indica como el "enlazador ZC". El enlazador ZC tiene que distinguirse de los enlazadores inter-dedos analizados anteriormente. Los enlazadores ZC se describen en detalle, por ejemplo, en el documento WO 2007/014275.

Como se analiza en detalle a continuación, la proteína de fusión (ZFN), o un polinucleótido que codifica la misma, se introduce en una célula. Una vez introducida, o expresada, en la célula, la proteína de fusión se une a la secuencia diana en PPP1R12C y escinde dentro de este locus génico.

### Integración dirigida

Los métodos y composiciones descritos pueden usarse para escindir ADN en la cromatina celular, lo que facilita la integración dirigida de una secuencia exógena (polinucleótido donante) como se describe en la presente memoria. Por "integración" se entiende tanto inserción física (por ejemplo, en el genoma de una célula hospedadora) como, además, integración por copiado de la secuencia donante en el genoma de la célula hospedadora mediante los procesos de replicación de ácido nucleico.

Para la integración dirigida, se diseñan uno o más dominios de unión de dedos de zinc para que se unan a un sitio diana en o cerca del sitio de escisión predeterminado, y una proteína de fusión que comprende el dominio de unión de dedos de zinc diseñado por ingeniería y se expresa un dominio de escisión en una célula. Tras la unión de la parte de dedos de zinc de la proteína de fusión al sitio diana, se escinde el ADN, preferiblemente mediante una rotura bicatenaria, cerca del sitio diana por el dominio de escisión. La presencia de una rotura bicatenaria facilita la integración de secuencias exógenas como se describe en la presente memoria mediante recombinación homóloga.

La integración dirigida de secuencias exógenas, como se describe en la presente memoria, puede usarse para generar células y líneas celulares para la expresión de proteínas. Véase, por ejemplo, la publicación de solicitud de patente de Estados Unidos N° 2006/0063231 (cuya descripción se incorpora por la presente por referencia en la presente memoria, en su totalidad, a todos los efectos) del mismo propietario que la presente. Para una expresión óptima de una o más proteínas codificadas por secuencias exógenas integradas en un genoma, el sitio cromosómico de integración debe ser compatible con la transcripción de alto nivel de las secuencias integradas, preferiblemente en un amplio intervalo de tipos celulares y fases del desarrollo. Sin embargo, se ha observado que la transcripción de secuencias integradas varía dependiendo del sitio de integración debido a, entre otras cosas, la estructura de la cromatina del genoma en el sitio de integración. Por consiguiente, son deseables sitios diana genómicos que soporte una transcripción de alto nivel de las secuencias integradas. En ciertas realizaciones, también será deseable que la integración de las secuencias exógenas no provoque la activación ectópica de uno o más genes celulares (por ejemplo, oncogenes). Por otro lado, en el caso de integración de secuencias promotoras y/o potenciadoras, puede desearse la expresión ectópica.

La secuencia exógena (donante) puede introducirse en la célula antes de, concurrentemente con, o después de, la expresión de la proteína o proteínas de fusión.

También se proporcionan métodos y composiciones que pueden potenciar los niveles de recombinación dirigida incluyendo, pero sin limitarse a, el uso de fusiones adicionales de ZFP-dominio funcional. Véase el documento WO 2007/014275.

Aumentos adicionales en la eficacia de la recombinación dirigida, en células que comprenden una molécula de fusión de dedo de zinc/nucleasa y una molécula de ADN donante, se consiguen bloqueando las células en fase G<sub>2</sub> del ciclo celular, cuando los procesos de reparación dirigidos por homología están en su máxima actividad. Dicha detención puede conseguirse de varios modos. Por ejemplo, las células pueden tratarse con, por ejemplo, fármacos, compuestos y/o moléculas pequeñas que influyen en la progresión del ciclo celular deteniendo las células en fase G<sub>2</sub>. Moléculas ejemplares de este tipo incluyen, pero no se limitan a, compuestos que afectan a la polimerización de



los microtúbulos (por ejemplo, vinblastina, nocodazol, Taxol), compuestos que interactúan con el ADN (por ejemplo, dicloruro de *cis*-platino(II) diamina, cisplatino, doxorubicina) y/o compuestos que afectan a la síntesis del ADN (por ejemplo, timidina, hidroxurea, L-mimosina, etopósido, 5-fluorouracilo). Aumentos adicionales en la eficacia de recombinación se consiguen mediante el uso de inhibidores de histona desacetilasa (HDAC) (por ejemplo, butirato de sodio, tricostatina A) que alteran la estructura de la cromatina para hacer al ADN genómico más accesible a la maquinaria celular de recombinación.

Métodos adicionales para la detención del ciclo celular incluyen la sobreexpresión de proteínas que inhiben la actividad de las quinasas del ciclo celular CDK, por ejemplo, introduciendo un ADNc que codifica la proteína en la célula o introduciendo en la célula una ZFP diseñada por ingeniería que activa la expresión del gen que codifica la proteína. La detención del ciclo celular también se consigue inhibiendo la actividad de las ciclinas y CDK, por ejemplo, usando métodos de ARNi (por ejemplo, patente de Estados Unidos N° 6.506.559) o introduciendo en la célula una ZFP diseñada por ingeniería que reprime la expresión de uno o más genes implicados en la progresión del ciclo celular tales como, por ejemplo, genes de ciclina y/o CDK. Véase, por ejemplo, la patente de Estados Unidos N° 6.534.261 del mismo propietario que la presente, para métodos para la síntesis de proteínas de dedos de zinc diseñadas por ingeniería para la regulación de la expresión génica.

Como alternativa, en ciertos casos, la escisión dirigida se realiza en ausencia de un polinucleótido donante (preferiblemente en fase S o G<sub>2</sub>), y la recombinación sucede entre cromosomas homólogos.

### Suministro

Los ácidos nucleicos descritos en la presente memoria (por ejemplo, un polinucleótido que codifica ZFN y/o secuencia donante) pueden introducirse en una célula usando cualquier método adecuado.

Para células vegetales, las construcciones de ADN pueden introducirse en (por ejemplo, en el genoma de) un hospedador vegetal deseado mediante una diversidad de técnicas convencionales. Para revisiones de dichas técnicas véase, por ejemplo, Weissbach y Weissbach *Methods for Plant Molecular Biology* (1988, Academic Press, N.Y.) Sección VIII, pág. 421-463; y Grierson y Corey, *Plant Molecular Biology* (1988, 2ª Ed.), Blackie, Londres, Cap. 7-9.

Por ejemplo, la construcción de ADN puede introducirse directamente en el ADN genómico de la célula vegetal usando técnicas tales como electroporación y microinyección en protoplastos de células vegetales, o las construcciones de ADN pueden introducirse directamente en el tejido vegetal usando métodos biolísticos, tales como bombardeo con partículas de ADN (véase, por ejemplo, Klein et al. (1987) *Nature* 327:70-73). Como alternativa, las construcciones de ADN pueden combinarse con regiones flanqueantes de ADN-T adecuadas e introducirse en un vector hospedador convencional de *Agrobacterium tumefaciens*. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, incluyendo el desarme y uso de vectores binarios, están bien descritas en la bibliografía científica. Véase, por ejemplo Horsch et al (1984) *Science* 233:496-498, y Fraley et al (1983) *Proc. Natl Acad. Sci. USA* 80:4803,

Además, la transferencia génica puede conseguirse usando bacterias no *Agrobacterium* o virus tales como *Rhizobium sp.* NGR234, *Sinorhizobium meliloti*, *Mesorhizobium loti*, virus X de la patata, virus del mosaico de la coliflor y virus del mosaico de las nervaduras de la mandioca y/o virus del mosaico del tabaco. Véase, por ejemplo, Chung et al. (2006) *Trends Plant Sci.* 11(1): 1-4.

Las funciones de virulencia del hospedador *Agrobacterium tumefaciens* dirigirán la inserción de la construcción y el marcador adyacente en el ADN de la célula vegetal cuando la célula se infecte por las bacterias usando un vector binario de ADN T (Bevan (1984) *Nuc. Acid Res.* 12:8711-8721) o el procedimiento de co-cultivo (Horsch et al (1985) *Science* 227:1229-1231). Generalmente, el sistema de transformación de *Agrobacterium* se usa para diseñar por ingeniería plantas dicotiledóneas (Bevan et al (1982) *Ann. Rev. Genet* 16:357-384; Rogers et al (1986) *Methods Enzymol.* 118:627-641). El sistema de transformación de *Agrobacterium* también puede usarse para transformar, así como transferir, ADN a plantas y células vegetales monocotiledóneas. Véase la patente de Estados Unidos N° 5.591.616; Hernalsteen et al (1984) *EMBO J* 3:3039-3041; Hooykass-Van Slogteren et al. (1984) *Nature* 311:763-764; Grimsley et al. (1987) *Nature* 325:1677-179; Boulton et al. (1989) *Plant Mol. Biol.* 12:31-40; y Gould et al. (1991) *Plant Physiol.* 95:426-434.

Métodos alternativos de transferencia y transformación génica incluyen, pero no se limitan a, transformación de protoplastos a través de captación mediada por calcio, polietilenglicol (PEG) o electroporación de ADN desnudo (véase Paszkowski et al. (1984) *EMBO J* 3:2717-2722, Potrykus et al. (1985) *Molec. Gen. Genet.* 199:169-177; Fromm et al. (1985) *Proc. Nat. Acad. Sci. USA* 82:5824-5828; y Shimamoto (1989) *Nature* 338:274-276) y electroporación de tejidos vegetales (D'Halluin et al. (1992) *Plant Cell* 4:1495-1505). Métodos adicionales para la transformación de células vegetales incluyen microinyección, captación de ADN mediada por carburo de silicio (Kaeppeler et al. (1990) *Plant Cell Reporter* 9:415-418), y bombardeo con microproyectiles (véase Klein et al. (1988) *Proc. Nat. Acad. Sci. USA* 85:4305-4309; y Gordon-Kamm et al. (1990) *Plant Cell* 2:603-618).

Asimismo, la proteína o proteínas de fusión (ZFN) pueden introducirse en forma de polipéptidos y/o polinucleótidos. Por ejemplo, dos polinucleótidos, comprendiendo cada uno secuencias que codifican uno de los

polipéptidos mencionados anteriormente, pueden introducirse en una célula, y cuando se expresen los polipéptidos y cada uno se una a su secuencia diana, sucederá la escisión en o cerca de la secuencia diana. Como alternativa, se introduce un único polinucleótido que comprende secuencias que codifican ambos polipéptidos de fusión en una célula. Los polinucleótidos pueden ser ADN, ARN o cualquier forma modificada o análogo de ADN y/o ARN.

5 En ciertas realizaciones, puede clonarse una o más ZFP o proteínas de fusión ZFP en un vector para su introducción por transformación en células procariotas o eucariotas para su replicación y/o expresión. Los vectores pueden ser vectores procariotas, por ejemplo, plásmidos, o vectores lanzadera, vectores de insecto, o vectores eucariotas. También puede clonarse un ácido nucleico que codifica secuencias descritas en la presente memoria (ZFN) en un vector de expresión, para su administración a una célula vegetal, célula animal, preferiblemente una célula de mamífero o una célula humana, célula fúngica, célula bacteriana, o célula protozoaria usando técnicas convencionales descritas, por ejemplo, en Sambrook et al., *supra* y las publicaciones de patente de Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; y 20060188987, y la publicación internacional WO 2007/014275.

15 En ciertas realizaciones, las ZFN y secuencias donantes se suministran *in vivo* o *ex vivo* para usos de terapia génica. Los sistemas de suministro de vectores no virales para suministrar polinucleótidos a células incluyen plásmidos de ADN, ácido nucleico desnudo, y ácido nucleico en complejo con un vehículo de suministro tal como un liposoma o poloxámero. Los sistemas de suministro de vectores virales para el suministro de las ZFN incluyen virus ADN y ARN, que tienen genomas episódicos o integrados después del suministro a la célula. Para una revisión de los procedimientos de terapia génica, véase Anderson, *Science* 256:808-813 (1992); Nabel y Feigner, *TIBTECH* 11:211-217 (1993); Mitani y Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer y Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al., en *Current Topics in Microbiology and Immunology* Doerfler y Böhm (eds.) (1995); y Yu et al., *Gene Therapy* 1:13-26 (1994).

25 Los métodos de suministro no viral de ácidos nucleicos *in vivo* o *ex vivo* incluyen electroporación, lipofección (véase, la patente de Estados Unidos N° 5.049.386; 4.946.787 y reactivos disponibles en el mercado tales como Transfectam™ y Lipofectin™), microinyección, biolística, virosomas, liposomas (véase, por ejemplo, Crystal, *Science* 270:404-410 (1995); Blaese et al., *Cancer Gene Ther.* 2:291-297 (1995); Behr et al., *Bioconjugate Chem.* 5:382-389 (1994); Remy et al., *Bioconjugate Chem.* 5:647-654 (1994); Gao et al., *Gene Therapy* 2:710-722 (1995); Ahmad et al., *Cancer Res.* 52:4817-4820 (1992); patentes de Estados Unidos N° 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028, y 4.946.787), inmunoliposomas, conjugados polication o lípido:ácido nucleico, ADN desnudo, viriones artificiales, sistemas de vector viral (por ejemplo, vectores retrovirales, lentivirales, adenovirales, adeno-asociados, del virus vaccinia y de herpes simple como se describe en el documento WO 2007/014275 para el suministro de proteínas que comprenden ZFP) y captación potenciada por agente de ADN. También puede usarse sonoporación usando, por ejemplo, el sistema Sonitron 2000 (Rich-Mar) para el suministro de ácido nucleicos.

Sistemas ejemplares adicionales de suministro de ácido nucleico incluyen aquellos proporcionados por Amaxa Biosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Mariland) y BTX Molecular Delivery Systems (Holliston, MA).

40 En ciertas realizaciones, por ejemplo, en que se prefiere la expresión transitoria de una proteína de fusión ZFP, pueden usarse sistemas basados en adenovirus. Los vectores basados en adenovirus tienen capacidad de eficacia de transducción muy elevada en muchos tipos celulares y no requieren división celular. Con dichos vectores, se han obtenido elevados títulos y elevados niveles de expresión. Este vector puede producirse en grandes cantidades en un sistema relativamente simple. También se usan vectores virales adeno-asociados ("AAV") para translucir células con ácidos nucleicos diana, por ejemplo, en la producción *in vitro* de ácidos nucleicos y péptidos, y para procedimientos de terapia génica *in vivo* y *ex vivo* (véase, por ejemplo, West et al., *Virology* 160:38-47 (1987); patente de Estados Unidos N° 4.797.368; documento WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). La construcción de vectores AAV recombinantes se describe en varias publicaciones, incluyendo la patente de Estados Unidos N° 5.173.414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat y Muzyczka, *PNAS* 81:6466-6470 (1984); y Samulski et al., *J. Virol* 63:03822-3828 (1989).

Están disponibles al menos seis enfoques de vector viral para la transferencia génica en ensayos clínicos, que utilizan enfoques que implican la complementación de vectores defectuosos por genes insertados en líneas celulares auxiliares para generar el agente de transducción.

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

Los vectores virales adeno-asociados recombinantes (rAAV) son sistemas alternativos prometedores de suministro génico basados en el virus adeno-asociado tipo 2 parvovirus defectuoso y no patogénico. Todos los vectores se obtienen de un plásmido que retiene solamente las repeticiones terminales invertidas de 145 pb de AAV que flanquean el casete de expresión del transgén. La transferencia génica eficaz y el suministro estable del transgén debidos a la integración en los genomas de la célula transducida son características clave de este sistema de vector. (Wagner *et al.*, *Lancet* 351:9117 1702-3 (1998), Kearns *et al.*, *Gene Ther.* 9:748-55(1996)).

Los vectores adenovirales (Ad) recombinantes deficientes en replicación pueden producirse a elevado título e infectan fácilmente varios tipos celulares diferentes. La mayoría de los vectores adenovirales se diseñan por ingeniería de modo que un transgén remplace los genes Ad E1a, E1b, y/o E3; posteriormente el vector deficiente en replicación se propaga en células 293 humanas que aportan la función génica eliminada en *trans*. Los vectores Ad pueden introducirse por transducción en múltiples tipos de tejidos *in vivo*, incluyendo células diferenciadas no en división tales como las halladas en hígado, riñón y músculo. Los vectores Ad convencionales tienen una gran capacidad portadora. Un ejemplo del uso de un vector Ad en un ensayo clínico implicó terapia con polinucleótido para inmunización antitumoral con inyección intramuscular (Sterman *et al.*, *Hum. Gene Ther.* 7:1083-9 (1998)). Ejemplos adicionales del uso de vectores adenovirales para transferencia génica en ensayos clínicos incluyen Rosenecker *et al.*, *Infection* 24:1 5-10 (1996); Sterman *et al.*, *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh *et al.*, *Hum. Gene Ther.* 2:205-18 (1995); Alvarez *et al.*, *Hum. Gene Ther.* 5:597-613 (1997); Topf *et al.*, *Gene Ther.* 5:507-513 (1998); Sterman *et al.*, *Hum. Gene Ther.* 7:1083-1089 (1998).

Se usan células empaquetadoras para formar partículas virales que sean capaces de infectar una célula hospedadora. Dichas células incluyen células 293, que empaquetan adenovirus, y células Ψ2 o células PA317, que empaquetan retrovirus. Los vectores virales usados en terapia génica se generan habitualmente por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula viral. Los vectores típicamente contienen las secuencias virales mínimas necesarias para el empaquetamiento y posterior integración en un hospedador (si fuera aplicable), remplazándose las demás secuencias virales por un casete de expresión que codifica la proteína a expresar. Las funciones virales perdidas se proporcionan en *trans* por la línea celular empaquetadora. Por ejemplo, los vectores AAV usados en terapia génica típicamente poseen solamente secuencias de repetidas terminales invertidas (ITR) del genoma AAV que son necesarias para el empaquetado e integración en el genoma hospedador. El ADN viral se empaqueta en una línea celular, que contiene un plásmido auxiliar que codifica los demás genes AAV, concretamente *rep* y *cap*, pero carece de las secuencias ITR. La línea celular también se infecta con adenovirus como auxiliar. El virus auxiliar promueve la replicación del vector AAV y la expresión de genes AAV desde el plásmido auxiliar. El plásmido auxiliar no se empaqueta en cantidades significativas debido a la ausencia de secuencias ITR. La contaminación con adenovirus puede reducirse por, por ejemplo, tratamiento térmico al cual el adenovirus es más sensible que el AAV.

En muchas aplicaciones de terapia génica, es deseable que los polinucleótidos (por ejemplo, secuencia que codifica ZFN) se suministren con un elevado grado de especificidad a un tipo tisular particular. Por consiguiente, puede modificarse un vector viral para que tenga especificidad por un tipo celular dado expresando un ligando como una proteína de fusión con una proteína de envuelta viral sobre la superficie externa del virus. El ligando se elige para que tenga afinidad por un receptor que se sabe que está presente en el tipo celular de interés. Por ejemplo, Han *et al.*, *Proc. Natl. Acad. Sci. USA* 92:9747-9751 (1995), informaron de que el virus de la leucemia murina de Moloney puede modificarse para que exprese la heregulina humana fusionada a gp70, y el virus recombinante infecta ciertas células cancerosas de mama humana que expresan el receptor del factor de crecimiento epidérmico humano. Este principio puede extenderse a otros pares virus-célula diana, en que la célula diana expresa un receptor y el virus expresa una proteína de fusión que comprende un ligando para el receptor de superficie celular. Por ejemplo, puede diseñarse por ingeniería un fago filamentoso para que presente fragmentos de anticuerpo (por ejemplo, FAB o Fv) que tengan afinidad de unión específica por casi cualquier receptor celular elegido. Aunque la descripción anterior se aplica principalmente a los vectores virales, pueden aplicarse los mismos principios a vectores no virales. Dichos vectores pueden diseñarse por ingeniería para que contengan secuencias específicas de captación que favorecen la captación por células diana específicas.

Los vectores de terapia génica pueden suministrarse *in vivo* mediante la administración a un paciente individual, típicamente por administración sistémica (por ejemplo, infusión intravenosa, intraperitoneal, intramuscular, subdérmica, o intracraneal) o aplicación tópica, como se describe a continuación. Como alternativa, los vectores pueden suministrarse a células *ex vivo*, tal como células explantadas de un paciente individual (por ejemplo, linfocitos, aspiraos de médula ósea, biopsia tisular) o células madre hematopoyéticas donantes universales, seguido de reimplante de las células en un paciente, habitualmente después de la selección de las células que hayan incorporado el vector.

La transfección celular *ex vivo* para diagnóstico, investigación, o terapia génica (por ejemplo, mediante re-infusión de las células transfectadas en el organismo hospedador) es bien conocida para los expertos en la técnica. En una realización preferida, se aíslan células del organismo objeto, se transfectan con un ácido nucleico ZFP (gen o ADNc) y secuencia exógena, y se re-infunden de nuevo en el organismo objeto (por ejemplo, paciente). Son bien conocidos diversos tipos celulares adecuados para transfección *ex vivo* para los expertos en la técnica (véase, por ejemplo, Freshney *et al.*, *Culture of Animal Cells, A Manual of Basic Technique* (3ª ed. 1994)) y las referencias citadas en el mismo para un análisis del modo de aislar y cultivar células de pacientes).

- En una realización, se usan células madre en procedimientos *ex vivo* para transfección celular y terapia génica. La ventaja de usar células madre es que pueden diferenciarse en otros tipos celulares *in vitro*, o pueden introducirse en un mamífero (tal como el donante de las células) donde se injertan en la médula ósea. Los métodos para diferenciar células CD34+ *in vitro* en tipos de células inmunes clínicamente importantes usando citoquinas tales como GM-CSF, IFN- $\gamma$  y TNF- $\alpha$  son conocidos (véase Inaba *et al.*, *J. Exp. Med.* 176:1693-1702(1992)).
- Las células madre se aíslan para su transducción y diferenciación usando métodos conocidos. Por ejemplo, las células madre se aíslan de células de médula ósea por lavado en batea de las células de médula ósea con anticuerpos que se unen a células no deseadas, tales como CD4+ y CD8+ (células T), CD45+ (células panB), GR-1 (granulocitos), e lad (células presentadoras de antígeno diferenciadas) (véase Inaba *et al.*, *J. Exp. Med.* 176:1693-1702 (1992)).
- En una realización, la célula a usar es un ovocito.
- En otras realizaciones, pueden usarse células obtenidas de organismo modelo. Éstas pueden incluir células obtenidas de xenopus, células de insecto (por ejemplo, drosophila) y células de nematodo.
- Los vectores (por ejemplo, retrovirus, adenovirus, liposomas, etc.) que comprenden ácidos nucleicos descritos en la presente memoria también pueden administrarse directamente a un organismo para la transducción de células *in vivo*. Como alternativa, puede administrarse ADN desnudo. La administración es mediante cualquiera de las vías normalmente usadas para introducir una molécula en contacto final con células sanguíneas o tisulares incluyendo, pero sin limitarse a, inyección, infusión, aplicación tópica y electroporación. Los métodos adecuados para administrar dichos ácidos nucleicos están disponibles y son bien conocidos para los expertos en la técnica y, aunque puede usarse más de una vía para administrar una composición particular, una vía particular a menudo puede proporcionar una reacción más inmediata y más eficaz que otra vía.
- Los métodos para la introducción de ADN en células madre hematopoyéticas se describen, por ejemplo, en la patente de Estados Unidos N° 5.928.638. Los vectores útiles para la introducción de transgenes en células madre hematopoyéticas, por ejemplo, células CD34<sup>+</sup>, incluyen adenovirus Tipo 35.
- Los vectores adecuados para la introducción de transgenes en células inmunes (por ejemplo, células T) incluyen vectores lentivirales no integrantes. Véase, por ejemplo, Ory *et al.* (1996) *Proc. Natl. Acad. Sci. USA* **93**:11382-11388; Dull *et al.* (1998) *J. Virol.* **72**:8463-8471; Zuffery *et al.* (1998) *J. Virol.* **72**:9873-9880; Follenzi *et al.* (2000) *Nature Genetics* **25**:217-222.
- Los vehículos farmacéuticamente aceptables se determinan en parte por la composición particular que se esté administrando, así como por el método particular usado para administrar la composición. Por consiguiente, existe una amplia diversidad de formulaciones adecuadas de composiciones farmacéuticas disponibles, como se describe a continuación (véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 17<sup>a</sup> ed., 1989).
- Como se ha indicado anteriormente, también puede introducirse una o más proteínas de fusión ZFN en la célula en forma de polipéptidos usando métodos descritos por ejemplo en el documento WO 2007/014275. Ejemplos no limitantes de vehículos de suministro de proteínas incluyen, "polipéptidos de translocación en la membrana", por ejemplo, péptidos que tienen subsecuencias de aminoácidos anfífilos o hidrófobos que tienen la capacidad de actuar como vehículos de translocación en la membrana, moléculas de toxina, liposomas y derivados de liposoma tales como inmunoliposomas (incluyendo liposomas dirigidos).
- Las ZFP y vectores de expresión que codifican ZFP pueden administrarse directamente al paciente para integración por escisión dirigida en el locus PPP1R12C para aplicaciones terapéuticas o profilácticas, por ejemplo, cáncer, isquemia, retinopatía diabética, degeneración macular, artritis reumatoide, psoriasis, infección por VIH, anemia falciforme, enfermedad de Alzheimer, distrofia muscular, enfermedades neurodegenerativas, enfermedad vascular, fibrosis quística, apoplejía, y similares.
- La administración de cantidades terapéuticamente eficaces es mediante cualquiera de las vías normalmente usadas para introducir ZFP en contacto final con el tejido a tratar. Las ZFP se administran de cualquier manera adecuada, preferiblemente con vehículos farmacéuticamente aceptables. Los métodos adecuados para administrar dichos moduladores están disponibles y son bien conocidos para los expertos en la técnica y, aunque puede usarse más de una vía para administrar una composición particular, una vía particular a menudo puede proporcionar una reacción más inmediata y más eficaz que otra vía.
- Los vehículos farmacéuticamente aceptables se determinan en parte por la composición particular que se esté administrando, así como por el método particular usado para administrar la composición. Por consiguiente, existe una amplia diversidad de formulaciones adecuadas de composiciones farmacéuticas que están disponibles (véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 17<sup>a</sup> ed. 1985)).
- Las ZFP, en solitario o en combinación con otros componentes adecuados, pueden prepararse en formulaciones de aerosol (es decir, pueden "nebulizarse") para administrarse mediante inhalación. Las formulaciones en aerosol pueden introducirse en gases inertes aceptables presurizados, tales como diclorodifluorometano, propano,

nitrógeno, y similares.

Las formulaciones adecuadas para administración parenteral tales como, por ejemplo, por vía intravenosa, intramuscular, intradérmica, y subcutánea, incluyen soluciones estériles e isotónicas de inyección, acuosas y no acuosas, que pueden contener antioxidantes, tampones, bacteriostáticos, y solutos que convierten a la formulación en isotónica con la sangre del destinatario pretendido, y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión, solubilizantes, agentes espesantes, estabilizantes, y conservantes. Las composiciones descritas pueden administrarse, por ejemplo, por infusión intravenosa, por vía oral, tópica, intraperitoneal, intravesical o intratecal. Las formulaciones de compuestos pueden presentarse en recipientes precintados de una dosis o múltiples dosis, tales como ampollas y viales. Las soluciones y suspensiones de inyección pueden prepararse a partir de polvos, gránulos, y comprimidos estériles del tipo previamente descrito.

Las células vegetales transformadas que se producen mediante cualquiera de las anteriores técnicas de transformación de células vegetales pueden cultivarse para regenerar una planta completa que posea el genotipo transformado y por tanto el fenotipo deseado. Dichas técnicas de regeneración se basan en la manipulación de ciertas fitohormonas en un medio de crecimiento de cultivo tisular, que típicamente se basa en un biocida y/o herbicida marcador que se ha introducido junto con las secuencias de nucleótidos deseadas. La regeneración de plantas a partir de protoplastos cultivados se describe en Evans, et al., "Protoplasts Isolation and Culture" en *Handbook of Plant Cell Culture*, pág. 124-176, Macmillan Publishing Company, Nueva York, 1983; y Binding, *Regeneration of Plants, Plant Protoplasts*, pág. 21-73, CRC Press, Boca Raton, 1985. La regeneración también puede obtenerse a partir de callos, explantes, órganos, pólenes, embriones de las plantas o partes de los mismos. Dichas técnicas de regeneración se describen en líneas generales en Klee et al. (1987) *Ann. Rev. of Plant Phys.* 38:467-486.

Los ácidos nucleicos introducidos en una célula vegetal pueden usarse para conferir rasgos deseados en esencialmente cualquier planta. Puede diseñarse por ingeniería una amplia diversidad de sistemas de plantas y células vegetales para las características fisiológicas y agronómicas deseadas descritas en la presente memoria usando las construcciones de ácido nucleico de la presente descripción y los diversos métodos de transformación mencionados anteriormente. En realizaciones preferidas, las plantas y células vegetales diana para diseño por ingeniería incluyen, pero no se limitan a, aquellas plantas monocotiledóneas y dicotiledóneas, tales como cultivos que incluyen cultivos de grano (por ejemplo, trigo, maíz, arroz, mijo, cebada), cultivos frutales (por ejemplo, tomate, manzana, pera, fresa, naranja), cultivos forrajeros (por ejemplo, alfalfa), cultivos de tubérculos (por ejemplo, zanahoria, patata, remolacha azucarera, ñame), cultivos de hortalizas de hoja verde (por ejemplo, lechuga, espinaca); plantas de flor (por ejemplo, petunia, rosa, crisantemo), coníferas y pinos (por ejemplo, abeto, píceas); plantas usadas en fitoterapia (por ejemplo, plantas que acumulan metales pesados); cultivos oleaginosos (por ejemplo, girasol, colza) y plantas usadas con fines experimentales (por ejemplo, *Arabidopsis*). Por tanto, los métodos y composiciones descritos tienen uso sobre un amplio intervalo de plantas incluyendo, pero sin limitarse a, especies de los géneros *Asparagus*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucurbita*, *Daucus*, *Erigeron*, *Glycine*, *Gossypium*, *Hordeum*, *Lactuca*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Nicotiana*, *Orychophragmus*, *Oryza*, *Persea*, *Phaseolus*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Solanum*, *Sorghum*, *Triticum*, *Vitis*, *Vigna*, y *Zea*.

Un experto en la técnica reconocerá que tras incorporarse de forma estable el casete de expresión en las plantas transgénicas y confirmarse que es funcional, puede introducirse en otras plantas por cruce sexual. Puede usarse cualquiera de varias técnicas convencionales de reproducción, dependiendo de la especie a cruzar.

Una célula vegetal, callo, tejido o planta transformada puede identificarse y aislarse seleccionando o examinando el material vegetal diseñado por ingeniería para los rasgos codificados por los genes marcadores presentes en el ADN transformante. Por ejemplo, la selección puede realizarse cultivando el material vegetal diseñado por ingeniería en medios que contienen una cantidad inhibitoria del antibiótico o herbicida al cual la construcción génica transformante confiere resistencia. Además, las plantas y células vegetales transformadas también pueden identificarse explorando las actividades de cualquier gen marcador visible (por ejemplo, los genes de  $\beta$ -glucuronidasa, luciferasa, B o C1) que pueden estar presentes en las construcciones recombinantes de ácido nucleico. Dichas metodologías de selección y exploración son bien conocidas para los expertos en la técnica.

También pueden usarse métodos físicos y bioquímicos para identificar plantas o células vegetales transformantes que contengan construcciones génicas insertadas. Estos métodos incluyen, pero no se limitan a: 1) análisis de Southern o amplificación por PCR para detectar y determinar la estructura del inserto de ADN recombinante; 2) transferencia de Northern, protección de RNasa S1, amplificación por PCR con extensión de cebador o transcriptasa inversa para detectar y examinar transcritos de ARN de las construcciones génicas; 3) ensayos enzimáticos para detectar actividad enzimática o ribozimática, donde dichos productos génicos están codificados por la construcción génica; 4) electroforesis en gel de proteínas, técnicas de transferencia de Western, inmunoprecipitación, o inmunoensayos ligados a enzimas, donde los productos de las construcciones génicas son proteínas. También pueden usarse técnicas adicionales, tales como hibridación *in situ*, tinción enzimática, e inmunotinción, para detectar la presencia o expresión de la construcción recombinante en órganos y tejidos vegetales específicos. Los métodos para hacer todos estos ensayos son bien conocidos para los expertos en la técnica.

La presente descripción también abarca semillas de las plantas transgénicas descritas anteriormente en donde la

semilla tiene el transgén o construcción génica. La presente descripción abarca adicionalmente la descendencia, clones, líneas celulares o células de las plantas transgénicas descritas anteriormente en donde dicha descendencia, clon, línea celular o célula tiene el transgén o construcción génica.

**Ejemplos**

**5 Ejemplo 1: Diseño de construcciones donantes lineales**

Se diseñaron y construyeron construcciones donantes lineales con brazos de homología de 50, 75 o 100 pares de bases flanqueando una secuencia codificante de una proteína de interés del siguiente modo. Las construcciones donantes incluían brazos de homología contenidos dentro del locus PPP1R12C (también mencionado como AAVS1 o sitio p84) o dentro del locus endógeno rL2R $\gamma$ . Véase, la solicitud provisional de Estados Unidos N° 60/926.322, presentada el 26 de abril de 2007 para una descripción del locus PPP1R12C, incorporada por referencia en su totalidad en la presente memoria.

Las construcciones donantes se prepararon por PCR usando cebadores con 50, 75 o 100 pares de bases de homología con la diana genómica (PPP1R12C o rL2R $\gamma$ ). Los moldes usados para estas PCR fueron moléculas plasmídicas que contenían dos largos (aprox. 750 pb) fragmentos homólogos a la diana genómica, flanqueando una construcción GFP (las construcciones GFP se detallan en las secciones 0139 para AAVS1 y 0140 para EL2R $\gamma$ ). Además, los cebadores se construyeron para que incluyeran enlaces fosfodiéster fosforotioato en el primer y segundo enlaces fosfodiéster de los extremos 5' de los cebadores para proteger el donante lineal de la degradación exonucleolítica. Los ensalces fosfodiéster fosforotioato se introdujeron usando técnicas convencionales, por ejemplo como se describe en Ciafre et al. (1995) *Nucleic Acids Res.* 23(20):4134-42 y Johansson et al. (2002) *Vaccine* 20(27-28):3379-88.

Como alternativa, las construcciones donantes pueden prepararse por PCR como se muestra esquemáticamente en la FIG. 1. En resumen, los donantes pueden prepararse por PCR usando un cebador con una parte 5' de 50, 75 o 100 pares de bases homóloga a la diana genómica (PPP1R12C o IL2R $\gamma$ ) y una parte de 15-30 pares de bases idéntica a la fase de lectura abierta (ORF) de interés. Además, los cebadores pueden construirse para que incluyan enlaces fosfodiéster fosforotioato en el primer y segundo enlaces fosfodiéster de los extremos 5' de los cebadores para proteger al donante lineal de la degradación exonucleolítica. Los enlaces fosfodiéster fosforotioato pueden introducirse usando técnicas convencionales, por ejemplo como se describe en Ciafre et al. (1995) *Nucleic Acids Res.* 23(20):4134-42 y Johansson et al. (2002) *Vaccine* 20(27-28):3379-88.

Los cebadores de PCR para construcciones que contienen brazos de homología de 50, 75 y 100 pares de bases para PPP1R12C se muestran en la Tabla 1 y los cebadores de PCR para construcciones que contienen brazos de homología de 50 pares de bases para IL2R $\gamma$  se muestran en la Tabla 2.

Tabla 1

Cebador de PCR	Secuencia	SEC ID N°
AAV-50F	GGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCACAGTGGGGC	5
AAV-50R	AGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCCCAATCCTGTCCCTAGT	6
AAV-75F	TTATATTCCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCACAGTGGGGC	7
AAV-75R	TAGACCCAATATCAGGAGACTAGGAAGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCCCAATCCTGTCCCTAGT	8
AAV-100F	CCTGTGTCCCCGAGCTGGGACCACCTTATATTCCCAGGGCCGGTTAATGTGTGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCACAGTGGGGC	9
AAV-100R	AATCTGCCTAACAGGAGGTGGGGGTTAGACCCAATATCAGGAGACTAGGAA GGAGGAGGCCTAAGGATGGGGCTTTTCTGTCCCAATCCTGTCCCTAGT	10

Tabla 2

Cebador de PCR	Secuencia	SEC ID N°
IL-50F	GTGTGGATGGGCAGAAACGCTACACGTTTCGTGTTCGGAGCCGCTTTAAC	11
IL-50R	TGGATTGGGTGGCTCCATTCACCTCCAATGCTGAGCACTTCCACAGAGTGG	12

Se realizaron múltiples reacciones de PCR para cada construcción donante. Las condiciones para las PCR de ambos donantes AAVS1 e IL2R $\gamma$ : 95°C, 3 min. -> 30x[95°C, 30 s; 72°C, 2 min.] -> 72°C, 5 min. -> mantenimiento

4°C. Las reacciones se combinaron y se purificaron las construcciones usando el kit de purificación de PCR QiaQuick™ (Qiagen) para obtener las construcciones mostradas en las FIG. 2 a 5.

Las FIG. 2, 3 y 4 muestran moléculas donantes dirigidas a PPP1R12C (AAVS1). En particular, la FIG. 2 muestra una molécula donante lineal (SEC ID N° 1) dirigida a AAVS1 y que tiene brazos de homología de 100 pares de bases y mencionada como donante AAVS1 100 pb HA. El brazo izquierdo de homología de AAVS1 100 pb HA se extiende desde los nucleótidos 1 a 100 (en minúscula, subrayado); un sitio SA se extiende desde los nucleótidos 107 a 132 (en minúscula, negrita); una secuencia que codifica un péptido 2A desde los nucleótidos 141 a 212 (mayúscula, sin subrayado); una secuencia que codifica poli(A) de la proteína fluorescente verde (GFP) se extiende desde los nucleótidos 219 a 1.215 (mayúscula, subrayado); y un brazo derecho de homología se extiende desde los nucleótidos 1235 a 1334 (en minúscula, subrayado).

La FIG. 3 muestra una molécula donante lineal (SEC ID N° 2) que tiene brazos de homología de 75 pares de bases y denominada AAVS1 75 pb HA. En esta construcción, el brazo izquierdo de homología se extiende desde los nucleótidos 1 a 75 (en minúscula, subrayado); un sitio SA se extiende desde los nucleótidos 82 a 107 (en minúscula, negrita); una secuencia que codifica un péptido 2A desde los nucleótidos 116 a 187 (mayúscula, sin subrayado); una secuencia que codifica poli(A) de GFP se extiende desde los nucleótidos 194 a 1.190 (mayúscula, subrayado); y un brazo derecho de homología se extiende desde los nucleótidos 1210 a 1284 (en minúscula, subrayado).

La FIG. 4 muestra una molécula donante lineal (SEC ID N° 3) que tiene brazos de homología de 50 pares de bases y denominada AAVS1 50 pb HA. AAVS1 50 pb HA comprende un brazo izquierdo de homología desde los nucleótidos 1 a 50 (en minúscula, subrayado); un sitio SA desde los nucleótidos 57 a 82 (en minúscula, negrita); una secuencia que codifica un péptido 2A desde los nucleótidos 91 a 162 (mayúscula, sin subrayado); una secuencia que codifica poli(A) de GFP desde los nucleótidos 169 a 1.165 (mayúscula, subrayado); y un brazo derecho de homología desde los nucleótidos 1.185 a 1.234 (en minúscula, subrayado).

La secuencia de una molécula donante para rL2R $\gamma$  se muestra en la FIG. 5 (SEC ID N° 4). Esta molécula comprende brazos de homología de 50 pares de bases (brazo izquierdo de homología desde los nucleótidos 1 a 50 (en minúscula, subrayado) y brazo derecho de homología desde los nucleótidos 1.639 a 1.688 (minúscula, subrayado)). La molécula donante IL2R $\gamma$  50 pb HA también comprende una secuencia promotora hPGK desde los nucleótidos 79 a 594 (en minúscula, negrita) y una secuencia que codifica poli(A) de GFP desde los nucleótidos 615 a 1.611 (mayúscula, subrayado).

### Ejemplo 2: Integración dirigida de construcciones donantes lineales

Para evaluar la integración dirigida de construcciones donantes lineales que tienen cortos (50-100 pares de bases) brazos de homología, se introdujeron por transfección donantes lineales y un par de proteínas de fusión que comprendían una nucleasa de proteína de dedos de zinc (ZFN) como se describe en la solicitud provisional de Estados Unidos N° 60/926.322, presentada el 26 de abril de 2007 y mostrado en la Tabla 3 (sitios diana de ADN indicados en letras mayúsculas; los nucleótidos no contactados indicados en minúscula), en células K562 usando el kit Amaxa™ Nucleofection como se muestra en la Tabla 4.

Tabla 3

Nombre ZFN	Sitio diana	F1	F2	F3	F4
2189-11	acTAGGGACAGGATg (SEC ID N° 13)	QSSNLAR (SEC ID N° 14)	RPDFLNQ (SEC ID N° 15)	QSGHLAR (SEC ID N° 16)	RSDNLT (SEC ID N° 17)
r2182-8	ccCCACTGTGGGGTgg (SEC ID N° 18)	QSSHLTR (SEC ID N° 19)	RSDHLTT (SEC ID N° 20)	HNYARD C (SEC ID N° 21)	QKATRIT (SEC ID N° 22)

Tabla 4

N° muestra	Diana ZFN (2,5 $\mu$ g)	Donante	Conc. donante ( $\mu$ g)
1	GFP	ninguno	
2	no ZFN	SA-2A-GFP-pA (circular)	50
3	no ZFN	donante 50 pb HA	5
4	no ZFN	donante 75 pb HA	5
5	no ZFN	donante 100 pb HA	5
6	no ZFN	donante 50 pb HA	6,9
7	no ZFN	donante 100 pb HA	7,5
8	AAVS1	SA-2A-GFP-pA (circular)	50

9	AAVS1	donante 50 pb HA	5
10	AAVS1	donante 75 pb HA	5
11	AAVS1	donante 100 pb HA	5
12	AAVS1	donante 50 pb HA	6,9
13	AAVS1	donante 100 pb HA	7,5

El donante SA-2A-GFP-pA se refiere al fragmento donante circular de 1.647 pb descrito en la solicitud provisional de Estados Unidos N° 60/926.322, que corresponde a las posiciones 60318104-60319750 de PPP1R12C.

5 Cuarenta y ocho horas después de la transfección, se ensayó la tasa de integración dirigida (TI) mediante un ensayo de PCR radiomarcado y transferencia de Southern, como describen Moehle et al. (2007) *Proc. Nat'l Acad. Sci. USA* 104:3055-3060.

Los resultados de PCR y transferencia de Southern se muestran en la FIG. 6 y FIG. 7, respectivamente. La parte superior de cada carril está marcado con el número de muestra (columna izquierda, Tabla 4) y el porcentaje de cromosomas modificados por la misma se enumera debajo de cada carril.

10 Además, una semana después de la transfección, se ensayó el porcentaje de células GFP-positivas por FACS, también como describen Moehle et al. (2007).

Los resultados se muestran en la Tabla 5 y la FIG. 8 y confirman que la ORF de GFP de las secuencias donantes lineales se integraron en el genoma.

Tabla 5

	ZFN (2,5 ug)	Donante	Cantidad	% GFP	MFI Verde	% TI
1	GFP			0,21	17,99	0
2		Donante SA-2A-GFP-pA	50 ug	37,28	46,32	0
3		Donante 50 pb HA	5 ug	2,3	53,75	0
4		Donante 75 pb HA	5 ug	0,8	29,24	0
5		Donante 100 pb HA	5 ug	0,28	20,81	0
6		Donante 50 pb HA	6,9 ug	2,07	12,51	0
7		Donante 100 pb HA	7,5 ug	0,81	12,21	0
8	AAVS1	Donante SA-2A-GFP-pA	50 ug	14,47	18,1	2
9	AAVS1	Donante 50 pb HA	5 ug	13,15	8,71	4
10	AAVS1	Donante 75 pb HA	5 ug	9,31	7,52	3,1
11	AAVS1	Donante 100 pb HA	5 ug	8,71	8,28	3,9
12	AAVS1	Donante 50 pb HA	6,9 ug	17,48	9,26	9,4
13	AAVS1	Donante 100 pb HA	7,5 ug	8,91	8,71	4,6

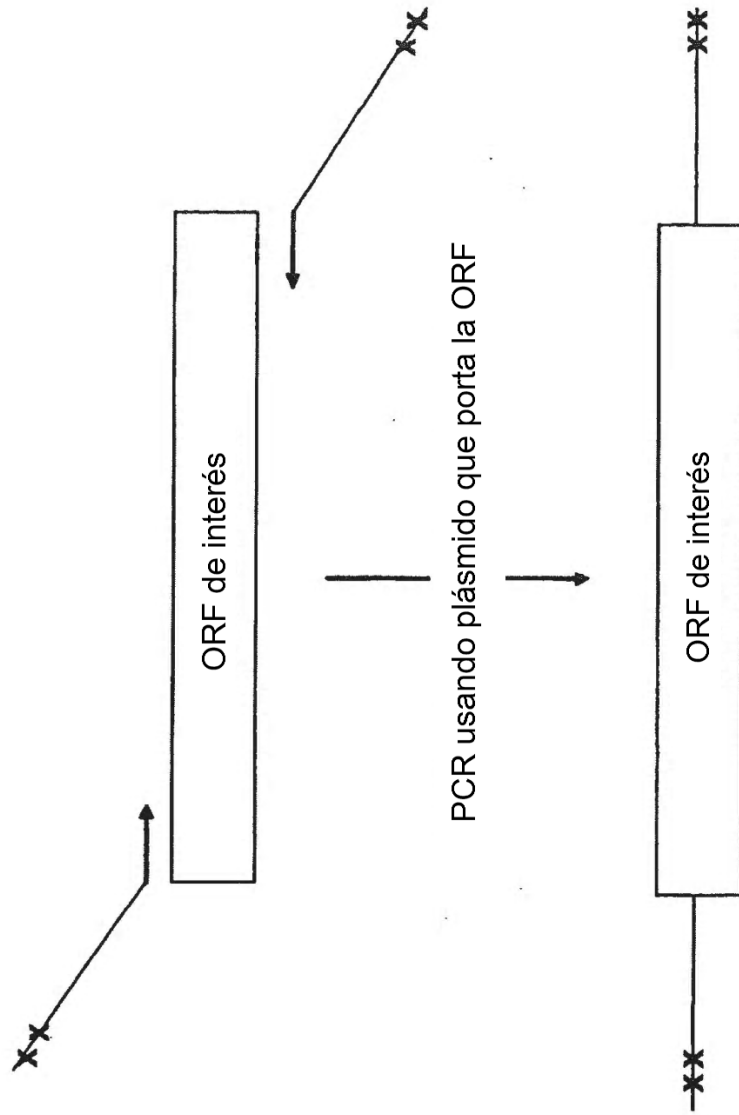
15 Por tanto, estos resultados demuestran que pueden usarse construcciones donantes lineales con cortos brazos de homología (~50-100 pb) para transferir de forma eficaz una secuencia que codifica un polipéptido de interés a una localización genómica específica. Las construcciones donantes lineales descritas en la presente memoria se generan rápidamente por PCR usando un molde plasmídico y pueden protegerse de la degradación exonucleolítica usando modificación de fosforioato.



**REIVINDICACIONES**

- 5 1. Uso de una molécula de ácido nucleico donante lineal para la integración dirigida dependiente de homología de una secuencia de interés en una célula eucariota, comprendiendo la molécula de ácido nucleico donante brazos de homología de entre 50 y 100 pares de bases y la secuencia de interés, en donde los brazos de homología flanquean la secuencia de interés.
2. El uso de la reivindicación 1, en donde uno o más de los pares de bases de los brazos de homología están unidos con un enlace fosfodiéster fosforotioato.
3. El uso de la reivindicación 2, en donde los enlaces fosfodiéster fosforotioato están posicionados en el primero y, opcionalmente, el segundo enlaces de los extremos 5' y 3' del ácido nucleico donante.
- 10 4. El uso de cualquiera de las reivindicaciones 1 a 3, en donde el ácido nucleico donante lineal comprende adicionalmente, entre los brazos de homología, una secuencia seleccionada del grupo que consiste en una secuencia que codifica un péptido 2A, una secuencia que comprende un sitio SA, y una secuencia que comprende una secuencia IRES.
- 15 5. El uso de una cualquiera de las reivindicaciones 1 a 4, en donde la secuencia de interés no codifica un polipéptido.
6. El uso de las reivindicaciones 1 a 4, en donde la secuencia de interés codifica un polipéptido.
7. El uso de la reivindicación 6, en donde el ácido nucleico donante lineal comprende adicionalmente una secuencia promotora unida de forma funcional a la secuencia de interés.
- 20 8. El uso según la reivindicación 6 o 7, en donde el polipéptido se selecciona del grupo que consiste en un anticuerpo, un antígeno, una enzima, un factor de crecimiento, un receptor de superficie celular, un receptor nuclear, una hormona, una linfoquina, una citoquina, un gen receptor, un marcador de selección, un factor secretado, una marca epitópica y fragmentos funcionales de los mismos y combinaciones de los mismos.
9. El uso según la reivindicación 5, en donde el ácido nucleico no codificante se selecciona del grupo que consiste en un ARNm, y SH-ARN, o ARNip.
- 25 10. Un método in-vitro para la integración dirigida dependiente de homología de una secuencia de interés en una región de interés en el genoma de una célula eucariota, comprendiendo el método las etapas de:
  - (a) expresar una proteína de fusión en la célula, comprendiendo la proteína de fusión un dominio de unión a ADN o un semi-dominio de escisión, en donde el dominio de unión a ADN se ha diseñado para que se una a un sitio diana en la región de interés;
  - 30 (b) poner en contacto la célula con un polinucleótido donante de cualquiera de las reivindicaciones 1 a 9, en donde la unión de la proteína de fusión al sitio diana escinde el genoma de la célula en la región de interés, provocando de este modo la integración dirigida dependiente de homología de la secuencia de interés en el genoma de la célula.
11. Un método in-vitro para la integración dirigida dependiente de homología de una secuencia de interés en una célula eucariota, comprendiendo el método:
  - 35 (a) expresar una primera proteína de fusión en la célula, comprendiendo la primera proteína de fusión un primer dominio de unión a ADN y un primer semi-dominio de escisión, en donde el primer dominio de unión a ADN se ha diseñado para unirse a un primer sitio diana en una región de interés en el genoma de la célula;
  - (b) expresar una segunda proteína de fusión en la célula, comprendiendo la segunda proteína de fusión un segundo dominio de ADN y un segundo semi-dominio de escisión, en donde el segundo dominio de unión de dedos de zinc se une a un segundo sitio diana en la región de interés en el genoma de la célula, en donde el segundo sitio diana es diferente del primer sitio diana; y
  - 40 (c) poner en contacto la célula con un polinucleótido que comprende un ácido nucleico donante según cualquiera de las reivindicaciones 1 a 9,
  - 45 en donde la unión de la primera proteína de fusión al primer sitio diana, y la unión de la segunda proteína de fusión al segundo sitio diana, posiciona los semi-dominios de escisión de tal modo que el genoma de la célula se escinde en la región de interés, provocando de este modo la integración dependiente de homología del ácido nucleico donante en el genoma de la célula.
- 50 12. El método in-vitro de la reivindicación 10 o reivindicación 11, en donde al menos un dominio de unión a ADN es un dominio de unión de dedos de zinc o un dominio de unión a ADN de meganucleasa y/o en donde el dominio de escisión es un dominio de escisión de meganucleasa de tipo silvestre o de origen no natural o un dominio de escisión de endonucleasa de restricción Tipo IIS de tipo silvestre o de origen no natural tal como un dominio de escisión *FokI* o *StsI*.
13. El método in-vitro según cualquiera de las reivindicaciones 10 a 11, en donde la célula es una célula de mamífero tal como una célula humana o una célula vegetal.

14. El método in-vitro según cualquiera de las reivindicaciones 10 a 13, en donde la célula está detenida en la fase G2 del ciclo celular.



**Fig. 1**

cctgtgtccccgagctgggaaccacctatattcccaggccgggtaagtggctctgggttctgggtacttttatctgtccc  
ctccacccccacagtggggcaagcttctgaccttctctctctcccaaggccctcgagAGATCTGGCAGCGGAGAGGGCA  
GAGGAAGTCTTTAACATGCCGTGACGTGGAGGAGAAATCCCGGCCCTAGGctcgagATGGTGAGCAAGGGCGGAGGAGCTGT  
TCACCGGGTGGTGCCCATCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGGAGGGCGAGG  
GCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGACACCACCGGAAGTGCCTGCCCCCTGGCCCCACCCCTCGTGA  
CCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCAATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC  
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCGAGGTGAAGTTCGAGG  
GCGACACCCCTGGTGAACCCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGT  
ACAACTAACAGCCACAACGCTATATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACA  
ACATCGAGACGGCAGCGTGCAGCTCGCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCCGTGTCTGCCCCG  
ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGGATCACATGGTCCCTGCTGGAGTTCG  
TGACCGCCCGGGGATCACCTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCTGAGTCTAGAGGGCCCCGTTTAAA  
CCCCGTGATCAGCCTCGACTGTGCCCTTCTAGTTGCCAGCCATCTGTGTTTGCCTCCCGTGCCCTTCTTGACCCCTGG  
AAGGTGCCACTCCCCACTGTCCTTTCCCTAATAAATGAGGAAATTGCAATGCTGAGTAGGTGTCAATTCATATTCTGG  
GGGTGGGGTGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG  
gtcgcagactaagctttactagggacaggattggtagacagaaaaagccccatccttaggcctcctccttagtctcct  
gatattgggtctaaccacccacctcctgttaggcagatt

## AAVS1 100 pb HA (SEC ID N° 1)

Fig. 2

ttatattcccagggccgggttaatgtggctctgtggttctggttaacttttatctgtccccctccacccccacagtggggcaagcttctg  
acctcttctctctcccacagggcctcgagAGATCTGGCAGCGGAGAGGGCAGAGGAAGTCTTTAACATGCGGGTGACGTGGA  
GGAGAAATCCCGCCCTAGGctcgagATGGTGAGCAAGGGCGAGGAGCTGTTTACCAGGGGTGGTGCCCATCCTGGTCGAGCTGGA  
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGGATGCCACCTACGGCAAGCTGACCCCTGAAGTTTCAT  
CTGCACCAACCGCAAGCTGCCCGTGCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCAGTGTTCAGCCCGCTACCC  
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCCATCTTCTTCAAGGACGA  
CGGCAACTACAAGACCCGCGGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCCGCATCGAGCTGAAGGGCATCGACTTCAA  
GGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACACTACAACAGCCACACAGTCTATATCATGGCCGACAAGCAGAAAGAA  
CGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCCACTACCAGCAGAACAACCCC  
CATCGGCGACGGCCCCGTGCTGCTGCCCGACAAACCCTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAAGCG  
CGATCACATGGTCTGCTGGAGTTCGTGACCCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCCGCGT  
CGAGTCTAGAGGGCCCCGTTTAAACCCCGCTGATCAGCCTCGACTGTGCCCTTCTAGTTGCCAGCCATCTGTTGTTGCCCCCTCCCC  
CGTGCCCTTCCTTGACCCCTGGAAGGTGCCACTGCCCTTCCCTAATAAAATGAGGAAATTGCAATCGCAATTGTCTGAGTAG  
GTGTCATTCATTCCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCCATGCTGGGGATGC  
GGTGGGCTCTATGGgtcgacagtaagctttactagggacaggattggtgacagaaaagcccccatccttagggcctcctcctt  
cctagtcctctgatattgggtcta

## AAVS1 75 pb HA (SEC ID N° 2)

Fig. 3

ggctctggttctggtgtaacttttatctgtcccctccaccaccacagtggggcaagcttctgacctcttctcttctctcccaaggccctc  
gagAGATCTGGCAGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAAATCCCGGCCCTAGGctcgagATGGTG  
AGCAAGGGCGAGGAGCTGTTCAACCGGGTGGTGCCCATCCTGGTCGAGCTGGACCGCGACGTAACCGGCCACAAGTTCAGCGTGTCC  
GGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCCAGGCAAGCTGCCCGTGCCTGGCCCCACC  
CTCGTGACCACCCCTGACCTACGGCGTGCAAGTGTTCAGCCGCTACCCCGACCATGAAGCAGCACGACTTCTTCAAGTCCGCCCATG  
CCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGCGCACTACAAGACCCCGCCGAGGTGAAGTTCGAGGGCGGAC  
ACCCCTGGTGAACCCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACATAACAAC  
AGCCACAACGTCATATCATGGCCGACAAGCAGAAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCCAGC  
GTGCAGCTCGCCGACCACTACAGCAGAAACCCCATCGGCAGCGGCCCTGTGTGCTGCCCGACAAACCACTACCTGAGCACCCAG  
TCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTCTGTGAGTTCGTGACCCGCCCGGGATCACCTCTCGGCATG  
GACGAGCTGTACAAGTAAAGCGGCCCGGTGAGTCTAGAGGGCCCGTTAAACCCGCTGATCAGCCTCGACTGTGCCCTCTAGTTGC  
CAGCCATCTGTGTTTGCCCTCCCGCTGCCCTTCCCTGACCCCTGAAAGGTGCCACTCCCACTGTCTTCCCTAATAAATGAGGAA  
ATTGCATCGCATTTGCTGAGTAGGTGTCATTTCTATTTCTGGGGGTGGGTGGGCAGGACAGCAAGGGGGAGGATGGGAAGACAAT  
AGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGgtcgacagtactaagcttactagggacaggttggtgacagaaaagccccat  
ccttaggcctcctcct

**AAVS1 50 pb HA (SEC ID N° 3)**

**Fig. 4**

gtgtggatgggcagaaaacgctacaagtttcgtgttcggagccgctttaacctcgatcgagaagcttgatataatccccac  
ggggttggggttggccttttccaaggcagccctgggtttgcgagggacggctgctctgggctggttccgggaaacgc  
agggcgcgacctgggtctgcacatctctcaagtcctgcagcgtcacccggatcttcgcccgtaccttgtgggccc  
ccccggcagccttctgctcccccctaagtgggaaagttccttgcggttcgcggtccggcgtccggaacgaaagc  
cgacgtctactagtaacctcgcagacggacagcgcagggagcaatggcagcgcgcgaccgcgatgggctgtggccaat  
agggctgctcaggggcccagagcagcggccgggaagggcgtgcgggagggcgttggggcgttagtggtggg  
ccctgtctcggcggcgtgttccgca tctgcaagcctccggagcagcagtcggcagtcggctccctcgttgaccgaa  
caccgacctctctcccaggggatccaccggtcgccaccATGGTGAGCAAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCC  
ATCCGGTGGAGCTGGACGGGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA  
AGTGACCCCTGAAGTTCACTGTGACCCCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGT  
GCAGTGCTTACGCCGCTACCCCGACACATGAAGCAGCACGACTTCTTCAAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG  
CGCACCATCTTCAAGGACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCCGA  
TCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCTTGGGGCACAAAGCTGGAGTACAATAACAAGCCACAACGT  
CTATATCATGGCCGACAAGCAGAAGACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGGTGCAG  
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCGTGTGTGCCCGACAACTACTCTGAGCACCCAGT  
CCGCCCTGAGCAAAGACCCCAACGAGAGCGCGATCACATGGTCTGTGGAGTTCGTGACCCCGCCGGGATCACTCTCGG  
CATGGACGAGCTGTACAAGTAAGCGGCCGCTGAGTCTAGAGGGCCCGTTTAAACCCCGTGTATCAGCCCTCGACTGTGCCCT  
TCTAGTTGCCAGCCATCTGTTGTTGCCCTTCCCCCTTCCCTTGACCCCTGGAAGGTGCCACTCCCCACTGCTTCTTCTT  
AATAAATGAGGAAATTCATCGCATTTGTCTGAGTAGGTGTCTATTCGTGGGGGTGGGGTGGGCAGGACAGCAAGGG  
GGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGGTGGGCTCTATGGcttctgagggcggaaagaaaccagtcgagcc  
actctgtggaagtgcagcattggagtgatggagccaccatcca

**IL2R 50 pb HA (SEC ID N° 4)**

**Fig. 5**

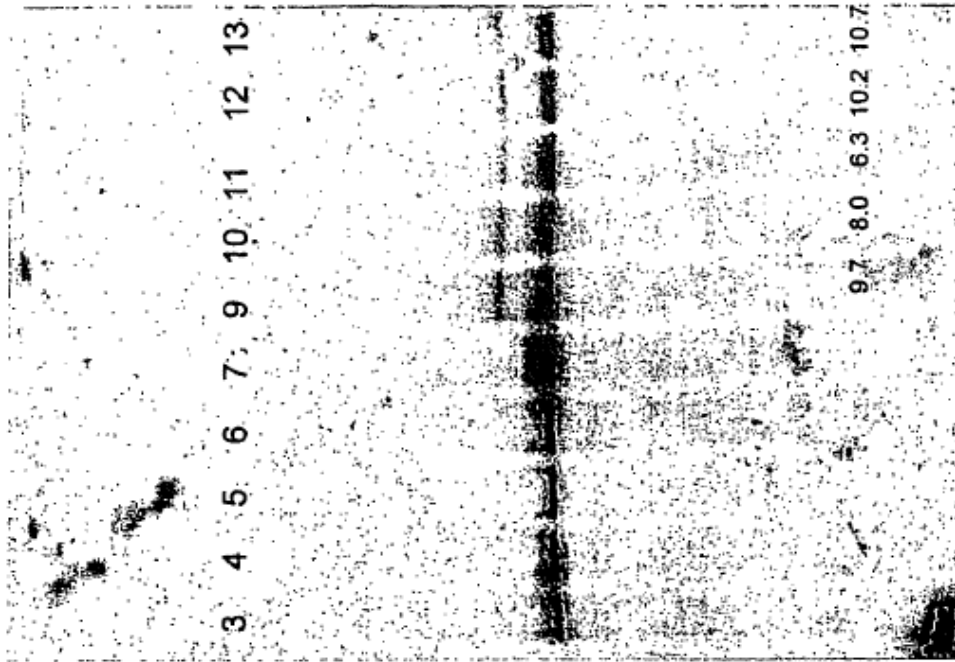


Fig. 7

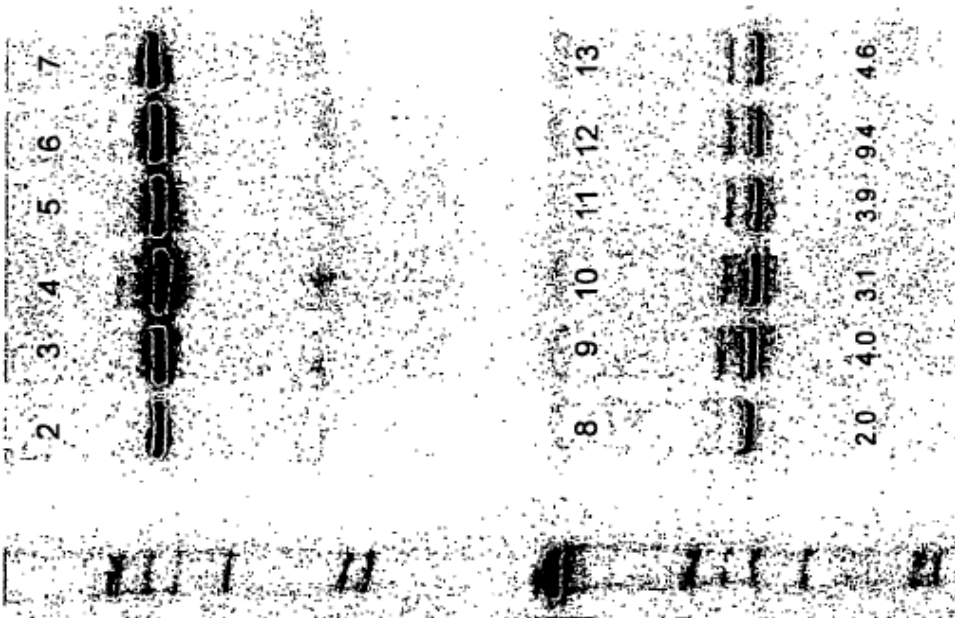
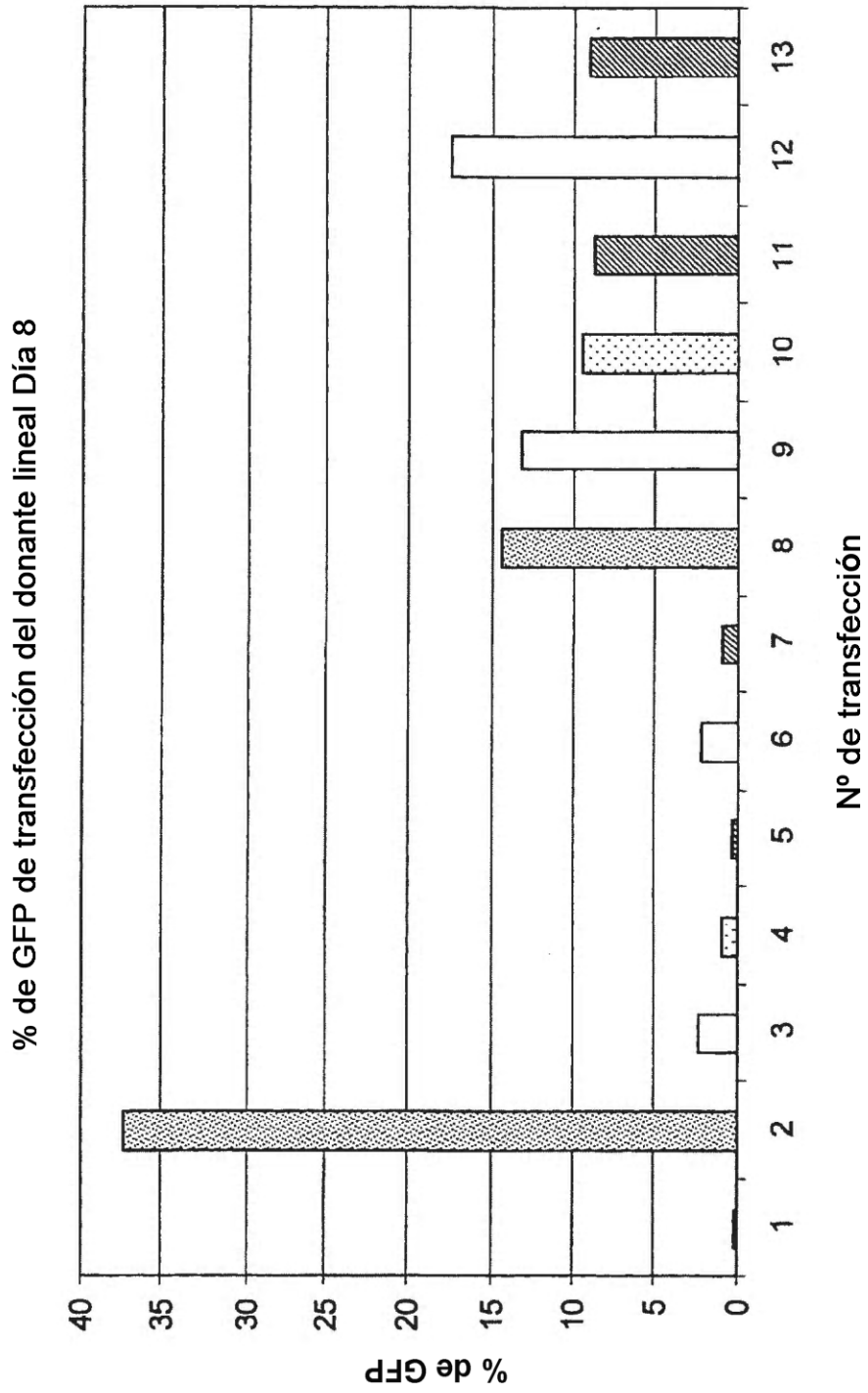


Fig. 6





**Fig. 8**