

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 921 207**

51 Int. Cl.:

**A01H 1/00** (2006.01)

**A01H 5/00** (2008.01)

**C12N 15/82** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.03.2014 E 19151652 (5)**

97 Fecha y número de publicación de la concesión europea: **01.06.2022 EP 3527068**

54 Título: **Procedimientos y composiciones para aumentar la eficiencia de la modificación genética direccionada utilizando la reparación genética mediada por oligonucleótidos**

30 Prioridad:

**15.03.2013 US 201361801333 P**

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

**19.08.2022**

73 Titular/es:

**CIBUS US LLC (50.0%)  
6455 Nancy Ridge Drive  
San Diego, California 92121, US y  
CIBUS EUROPE B.V. (50.0%)**

72 Inventor/es:

**BEETHAM, PETER R.;  
GOCAL, GREGORY F.W.;  
SCHOPKE, CHRISTIAN;  
SAUER, NOEL JOY;  
PEARCE, JAMES;  
SEGAMI, ROSA y  
MOZORUK, JERRY**

74 Agente/Representante:

**GONZÁLEZ PECES, Gustavo Adolfo**

Observaciones:

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 921 207 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

Procedimientos y composiciones para aumentar la eficiencia de la modificación genética direccionada utilizando la reparación genética mediada por oligonucleótidos

## CAMPO DE LA INVENCION

5 La presente invención se refiere en general a procedimientos novedosos para mejorar la eficacia del direccionamiento de las modificaciones a lugares específicos en las secuencias genómicas o de otros nucleótidos. Además, esta invención se relaciona con el ADN diana que ha sido modificado, mutado o marcado por los enfoques divulgados en el presente documento. La invención también se refiere a células, tejidos y organismos que han sido modificados por los procedimientos de la invención.

## 10 ANTECEDENTES DE LA INVENCION

La siguiente discusión de los antecedentes de la invención se proporciona meramente para ayudar al lector a entender la invención y no se admite que describa o constituya el arte previo a la presente invención.

15 La modificación del ADN genómico es fundamental para los avances de la biotecnología, en general, y los avances médicos de base biotecnológica, en particular. Los procedimientos eficientes para las modificaciones genómicas dirigidas al sitio son deseables para la investigación y posiblemente para las aplicaciones de terapia génica. Uno de los enfoques utiliza oligonucleótidos formadores de triplex (TFO) que se unen como terceras cadenas al ADN dúplex de forma específica para la secuencia, para mediar en la mutagénesis dirigida. Este tipo de TFO puede actuar suministrando un mutágeno enlazado, tal como el psoraleno o el clorambucil (Havre et al., Proc Nat'l Acad Sci, U.S.A. 90:7879-7883, 1993 Havre et al., J Virol 67:7323-7331, 1993 Wang et al., Mol Cell Biol 15:1759-1768, 1995 Takasugi et al., Proc Nat'l Acad Sci, U.S.A. 88:5602-5606, 1991; Belousov et al., Nucleic Acids Res 25:3440-3444, 1997), o uniéndose con suficiente afinidad para provocar una reparación propensa a errores (Wang et al., Science 271:802-805, 1996).

25 Otra estrategia de modificación genómica implica la inducción de recombinación homóloga entre un fragmento de ADN exógeno y el gen diana. Este enfoque se ha utilizado con éxito para dirigir y alterar genes seleccionados en células de mamíferos y ha permitido la producción de ratones transgénicos portadores de anulaciones de genes específicos (Capeechi et al., Science 244:1288-1292, 1989 U.S. Pat. N° 4.873.191 a Wagner). Este enfoque, sin embargo, se basa en la transferencia de marcadores seleccionables para permitir el aislamiento de los recombinantes deseados. Sin selección, la relación entre la integración homóloga y no homóloga del ADN transfectedo en los experimentos típicos de transferencia de genes es baja, usualmente en el intervalo de 1:1000 o menos (Sedivy et al., Gene Targeting, W. H. Freeman and Co., Nueva York, 1992). Esta baja eficacia de la integración homóloga limita la utilidad de la transferencia de genes para uso experimental o terapia génica. La frecuencia de la recombinación homóloga puede verse aumentada por el daño que la irradiación UV y determinados carcinógenos producen en el sitio diana (Wang et al., Mol Cell Biol 8:196-202, 1988) así como por endonucleasas específicas del sitio (Sedivy et al., Gene Targeting, W. H. Freeman and Co., Nueva York, 1992 Rouet et al., Proc Nat'l Acad Sci, U.S.A. 91:6064-6068, 1994 Segal et al., Proc Nat'l Acad Sci, U.S.A. 92:806-810, 1995). Además, el daño en el ADN inducido por los fotoaductos de psoraleno dirigidos al triplex puede estimular la recombinación dentro y entre vectores extracromosómicos (Segal et al., Proc Nat'l Acad Sci, U.S.A. 92:806-810, 1995 Faruqi et al., Mol Cell Biol 16:6820-6828, 1996 Pat. de EE.UU. N° 5.962.426 a Glazer).

40 Otros trabajos han ayudado a definir los parámetros que influyen en la recombinación en las células de mamíferos. En general, los fragmentos donantes lineales son más recombinogénicos que sus homólogos circulares (Folger et al., Mol Cell Biol 2:1372-1387, 1982). La recombinación también se ve influida por la longitud de la homología ininterrumpida entre los sitios donante y diana, ya que los fragmentos cortos parecen ser sustratos ineficaces para la recombinación (Rubnitz et al., Mol Cell Biol 4:2253-2258, 1984). No obstante, varios esfuerzos recientes se han centrado en el uso de fragmentos cortos de ADN o híbridos de ADN/ARN para la corrección de genes. (Kunzelmann et al., Gene Ther 3:859-867, 1996).

50 Las propiedades de unión específicas de la secuencia del TFO se han utilizado para suministrar una serie de moléculas diferentes a sitios diana en el ADN. Por ejemplo, un procedimiento de diagnóstico para examinar las interacciones triplex utilizó el TFO acoplado al Fe-EDTA, un agente de escisión del ADN (Moser et al., Science 238:645-650, 1987). Otros han vinculado enzimas biológicamente activas, como la nucleasa micrococócica y la nucleasa estreptocócica, al TFO y han demostrado la escisión específica del ADN (Pei et al., Proc Nat'l Acad Sci U.S.A. 87:9858-9862, 1990 Landgraf et al., Biochemistry 33:10607-10615, 1994). Además, el daño al ADN dirigido al sitio y la mutagénesis pueden lograrse utilizando TFO conjugado con psoraleno (Havre et al., Proc Nat'l Acad Sci U.S.A. 90:7879-7883, 1993 Takasugi et al., Proc Nat'l Acad Sci U.S.A. 88:5602-5606, 1991) o agentes alquilantes (Belousov et al., Nucleic Acids Res 25:3440-3444, 1997 Posvic et al., J Am Chem Soc 112:9428-9430, 1990).

55 La solicitud de patente de la OMPI WO/2001/025460 describe procedimientos para mutar una secuencia de ADN diana de una planta que incluyen los pasos de (1) electroporar en una microspora de la planta una oligonucleobase recombinogénica que contiene una primera región homóloga que tiene una secuencia idéntica a la secuencia de al menos 6 pares de bases de un primer fragmento de la secuencia de ADN diana y una segunda región homóloga que

- tiene una secuencia idéntica a la secuencia de al menos 6 pares de bases de un segundo fragmento de la secuencia de ADN diana y una región interviniente que contiene al menos 1 nucleobase heteróloga a la secuencia de ADN diana, cuya región interviniente conecta la primera región homóloga y la segunda región homóloga; (2) cultivar la microspora para producir un embrión; y (3) producir a partir del embrión una planta que tenga una mutación situada entre el primer y el segundo fragmento de la secuencia de ADN diana, por ejemplo, cultivando la microspora para producir un embrión somático y regenerando la planta a partir del embrión. En diversas realizaciones de la invención, la oligonucleobase recombinogénica es una MDON y cada una de las regiones homólogas contiene un segmento de ARN de al menos 6 nucleótidos de tipo ARN; la región interviniente tiene al menos 3 nucleótidos de longitud; el primer y o segundo segmento de ARN contiene al menos 8 ribonucleótidos 2'-sustituidos contiguos.
- Uno de los principales objetivos de la investigación biológica es la modificación direccionada del genoma. Como se ha señalado anteriormente, aunque los procedimientos para suministrar genes en las células de mamíferos están bien desarrollados, la frecuencia de modificación y/o recombinación homóloga es limitada (Hanson et al., Mol Cell Biol 15:45-51 1995). Por ello, la modificación de los genes es un proceso que requiere mucho tiempo. Se han contemplado o intentado numerosos procedimientos para mejorar la modificación y/o recombinación entre el ADN donante y el genómico. Sin embargo, las técnicas actuales a menudo presentan bajas tasas de modificación y/o recombinación, o inconsistencia en la tasa de modificación y/o recombinación, lo que dificulta la investigación y la tecnología de terapia génica.

### SUMARIO DE LA INVENCION

- La presente invención se refiere en general a procedimientos novedosos para mejorar la eficacia del direccionamiento de las modificaciones a lugares específicos en las secuencias genómicas o de otros nucleótidos.

El asunto objeto de la presente invención es un procedimiento como se define en la reivindicación 1 para introducir una mutación mediada por una oligonucleobasa de reparación de genes (GRON) en una secuencia de ácido desoxirribonucleico (ADN) diana en una célula vegetal. Las reivindicaciones dependientes se refieren a realizaciones preferidas de las mismas.

- El procedimiento comprende el suministro de una GRON en la célula vegetal; y el suministro de una endonucleasa específica de sitio en la célula vegetal, en la que la endonucleasa específica de sitio se selecciona entre nucleasas de dedo de zinc, nucleasas efectoras similares al activador de la transcripción (TALEN) y complejos de Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas con Regularidad (complejos CRISPR).

- La GRON comprende una o más alteraciones de los nucleótidos convencionales de ARN y ADN seleccionadas entre uno o más nucleótidos 2'O-metilos en el extremo 5' o 3' del mismo, una tapa en el terminal 5', uno o más tintes fluorescentes unidos covalentemente al mismo, una o más modificaciones de fosfotioato, una base inversa en el extremo 3' del mismo. La endonucleasa específica del sitio induce una rotura de la cadena simple o doble en las proximidades del sitio direccionado para conversión con la GRON.

- Como se describe a continuación, las GRON para su uso en la presente invención pueden comprender además una o más de las siguientes alteraciones de los nucleótidos de ARN y ADN convencionales:

la GRON tiene una longitud superior a 55 bases, la GRON comprende opcionalmente dos o más sitios de mutación para su introducción en el ADN diana;

la GRON comprende uno o más nucleótidos abásicos;

la GRON comprende uno o más nucleótidos 8'oxo dA y/o 8'oxo dG;

- la GRON comprende uno o más nucleótidos 2'O-metilo en el extremo 5' del mismo;

la GRON comprende al menos 2'O-metil nucleótidos de ARN en el extremo 5' del mismo;

la GRON comprende un tinte intercalado;

- la GRON comprende una modificación de la columna vertebral seleccionada del grupo que consiste en una modificación de fosfonato de metilo, una modificación de ácido nucleico bloqueado (LNA), una modificación de O -(2-metoxietil) (MOE), una modificación de di PS y una modificación de ácido nucleico peptídico (PNA);

la GRON comprende uno o más enlaces cruzados intracadena;

la GRON comprende una o más bases que aumentan la energía de hibridación. Esta lista no pretende ser limitativa.

- Como se describe a continuación, en ciertas realizaciones la calidad y la eficiencia de conversión de la GRON pueden mejorarse sintetizando todo o una parte de la GRON utilizando multímeros de nucleótidos, como dímeros, trímeros, tetrámeros, etc. mejorando su pureza.

En ciertas realizaciones, la endonucleasa específica del sitio es un complejo de Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas con Regularidad (complejo CRISPR). En ciertas realizaciones, la endonucleasa de sitio específico es una nucleasa efectora similar al activador de la transcripción (TALEN). En ciertas realizaciones, la endonucleasa de sitio específico es una nucleasa de dedo de zinc.

- 5 En ciertas realizaciones, la secuencia de ácido desoxirribonucleico (ADN) diana está dentro del genoma de la célula vegetal. La célula vegetal puede ser no transgénica o transgénica, y la secuencia de ADN diana puede ser un transgén o un gen endógeno de la célula vegetal.

Los procedimientos y composiciones descritos en el presente documento son aplicables a las plantas en general. Sólo a modo de ejemplo, una especie vegetal puede seleccionarse del grupo que consiste en canola, girasol, maíz, tabaco, remolacha azucarera, algodón, maíz, trigo, cebada, arroz, alfalfa, cebada, sorgo, tomate, mango, melocotón, manzana, pera, fresa, plátano, melón, patata, zanahoria, lechuga, cebolla, soja, soja spp, caña de azúcar, guisante, garbanzo, guisante de campo, haba, lentejas, nabo, colinabo, coles de Bruselas, altramuz, coliflor, col rizada, judías de campo, álamo, pino, eucalipto, uva, cítricos, triticale, alfalfa, centeno, avena, césped y hierbas forrajeras, lino, colza, mostaza, pepino, gloria de la mañana, bálsamo, pimienta, berenjena, caléndula, loto, col, margarita, clavel, tulipán, iris y lirio. También pueden aplicarse, en todo o en parte, a todos los demás sistemas biológicos, incluidos, pero no limitados a las bacterias, los hongos y las células de mamíferos, e incluso sus orgánulos (por ejemplo, las mitocondrias y los cloroplastos).

En ciertas realizaciones, los procedimientos comprenden además la regeneración de una planta que tiene una mutación introducida por la GRON a partir de la célula vegetal, y pueden comprender la recolección de semillas de la planta.

En aspectos relacionados, la presente invención se refiere a células vegetales que comprenden una modificación genómica introducida por una GRON según los procedimientos descritos en el presente documento, una planta que comprende una modificación genómica introducida por una GRON según los procedimientos descritos en el presente documento, o una semilla que comprende una modificación genómica introducida por una GRON según los procedimientos descritos en el presente documento.

Otras realizaciones de la invención serán evidentes a partir de la siguiente descripción detallada, las realizaciones ejemplares y las reivindicaciones.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

30 La Fig. 1 representa la conversión de BFP a GFP mediada por GRONs marcados con fosfotioato (PS) (que tienen 3 fracciones PS en cada extremo de la GRON) y GRONs marcados con 5'Cy3/ 3'idC.

La Fig. 2 representa GRONs que comprenden ARN/ADN, denominados en este documento "GRONs de fragmentos de Okazaki"

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

35 Desarrollada en los últimos años, la modificación genética dirigida mediada por oligonucleótidos ha demostrado ser una técnica valiosa para su uso en la alteración específica de tramos cortos de ADN para crear deleciones, inserciones cortas y mutaciones puntuales. Estos procedimientos implican el emparejamiento/apareamiento del ADN, seguido de un evento de reparación/recombinación del ADN. En primer lugar, el ácido nucleico se aparea con su cadena complementaria en el ADN de doble cadena en un proceso mediado por factores proteicos celulares. Este apareamiento crea un par de bases no coincidentes situado en el centro (en el caso de una mutación puntual), lo que da lugar a una perturbación estructural que muy probablemente estimula la maquinaria proteica endógena para iniciar el segundo paso del proceso de reparación: la modificación específica del sitio de la secuencia cromosómica e incluso de sus orgánulos (por ejemplo, mitocondrias y cloroplastos). Esta no coincidencia recién introducida induce a la maquinaria de reparación del ADN a realizar un segundo evento de reparación, que conduce a la revisión final del sitio diana. Los presentes procedimientos mejoran estos procedimientos proporcionando enfoques novedosos que aumentan la disponibilidad de los componentes de reparación del ADN, incrementando así la eficiencia y la reproducibilidad de las modificaciones mediadas por la reparación de genes en los ácidos nucleicos diana.

#### Definiciones

Para facilitar la comprensión de la invención, a continuación se definen una serie de términos.

50 "Secuencia de ácido nucleico", "secuencia de nucleótidos" y "secuencia de polinucleótidos", tal como se utilizan en el presente documento, se refieren a un oligonucleótido o polinucleótido, y a fragmentos o porciones de los mismos, y a ADN o ARN de origen genómico o sintético que puede ser de una o dos cadenas, y representan la cadena en sentido o antisentido.

Tal y como se utiliza en el presente documento, los términos "oligonucleótidos" y "oligómeros" se refieren a una secuencia de ácido nucleico de al menos aproximadamente 10 nucleótidos y hasta aproximadamente 201 nucleótidos, preferentemente de aproximadamente 15 a 30 nucleótidos, y más preferentemente de aproximadamente 20-25 nucleótidos, que puede utilizarse como sonda o amplímero.

5 Los términos "molécula modificadora del ADN" y "reactivo modificador del ADN", tal como se utilizan en el presente documento, se refieren a una molécula que es capaz de reconocer y unirse específicamente a una secuencia de ácido nucleico en el genoma de una célula, y que es capaz de modificar una secuencia de nucleótidos diana dentro del genoma, donde el reconocimiento y la unión específica de la molécula modificadora del ADN a la secuencia de ácido nucleico es independiente de las proteínas. El término "independiente de las proteínas", tal y como se utiliza en el presente documento en relación con una molécula modificadora del ADN, significa que la molécula modificadora del ADN no requiere la presencia y/o la actividad de una proteína y/o enzima para el reconocimiento y/o la unión específica a una secuencia de ácido nucleico. Se ejemplifican las moléculas modificadoras del ADN, pero sin limitarse a los oligonucleótidos formadores de tríplex, los ácidos nucleicos peptídicos, las poliamidas y los oligonucleótidos destinados a promover la conversión de genes. Las moléculas modificadoras del ADN de la invención se distinguen de las secuencias de ácido nucleico de la técnica anterior que se utilizan para la recombinación homóloga [Wong & Capecchi, *Molec. Cell. Biol.* 7:2294-2295, 1987] en el sentido de que las secuencias de ácido nucleico de la técnica anterior que se utilizan para la recombinación homóloga son dependientes de las proteínas. El término "dependiente de proteínas", tal y como se utiliza en el presente documento en relación con una molécula, significa que la molécula requiere la presencia y/o la actividad de una proteína y/o enzima para el reconocimiento de, y/o la unión específica de la molécula a, una secuencia de ácido nucleico. Los procedimientos para determinar si una molécula modificadora del ADN requiere la presencia y/o la actividad de una proteína y/o enzima para el reconocimiento y/o la unión específica a una secuencia de ácido nucleico están dentro de la experiencia en la técnica [véase, por ejemplo, Dennis et al. *Nucl. Acids Res.* 27:4734-4742, 1999]. Por ejemplo, la molécula modificadora del ADN puede incubarse *in vitro* con la secuencia de ácido nucleico en ausencia de proteínas y/o enzimas. La detección de la unión específica entre la molécula modificadora del ADN y la secuencia de ácido nucleico demuestra que la molécula modificadora del ADN es independiente de las proteínas. Por otro lado, la ausencia de unión específica entre la molécula modificadora del ADN y la secuencia de ácido nucleico demuestra que la molécula modificadora del ADN depende de las proteínas y/o requiere factores adicionales.

El "oligonucleótido formador de tríplex" (TFO) se define como una secuencia de ADN o ARN capaz de unirse en el surco mayor de una hélice dúplex de ADN o ARN para formar una triple hélice. Aunque el TFO no está limitado a ninguna longitud particular, una longitud preferida del TFO es de 200 nucleótidos o menos, más preferentemente de 100 nucleótidos o menos, aún más preferentemente de 5 a 50 nucleótidos, aún más preferentemente de 10 a 25 nucleótidos, y más preferentemente de 15 a 25 nucleótidos. Aunque es necesario un grado de especificidad de secuencia entre el TFO y el ADN dúplex para la formación de la triple hélice, no se requiere ningún grado particular de especificidad, siempre que la triple hélice sea capaz de formarse. Asimismo, no se requiere ningún grado específico de aidez o afinidad entre el TFO y la hélice dúplex, siempre que la triple hélice sea capaz de formarse. Aunque no se pretende limitar la longitud de la secuencia de nucleótidos a la que se une específicamente el TFO, en una realización, la secuencia de nucleótidos a la que se une específicamente el TFO es de 1 a 100, más preferentemente de 5 a 50, aún más preferentemente de 10 a 25, y más preferentemente de 15 a 25, nucleótidos. Además, la "triple hélice" se define como un ácido nucleico de doble hélice con un oligonucleótido unido a una secuencia diana dentro del ácido nucleico de doble hélice. El ácido nucleico "doble helicoidal" puede ser cualquier ácido nucleico de doble cadena, incluyendo ADN de doble cadena, ARN de doble cadena y dúplex mixtos de ADN y ARN. El ácido nucleico de doble cadena no está limitado a ninguna longitud en particular. Sin embargo, en las realizaciones preferidas tiene una longitud superior a 500 pb, más preferentemente superior a 1 kb y más preferentemente superior a aproximadamente 5 kb. En muchas aplicaciones, el ácido nucleico doblemente helicoidal es el ácido nucleico celular, genómico. El oligonucleótido formador de tríplex puede unirse a la secuencia diana de forma paralela o antiparalela.

Los "ácidos nucleicos peptídicos", "poliamidas" o "PNA" son ácidos nucleicos en los que la columna vertebral de fosfato se sustituye por una estructura de poliamida basada en N-aminoetilglicina. Los PNAs tienen una mayor afinidad por los ácidos nucleicos complementarios que sus homólogos naturales siguiendo las reglas de emparejamiento de bases Watson-Crick. Los PNAs pueden formar estructuras de triple hélice altamente estables con el ADN de la siguiente estequiometría: (PNA)<sub>2</sub>.ADN. Aunque los ácidos nucleicos peptídicos y las poliamidas no están limitados a ninguna longitud particular, una longitud preferida de los ácidos nucleicos peptídicos y las poliamidas es de 200 nucleótidos o menos, más preferentemente de 100 nucleótidos o menos, y más preferentemente de 5 a 50 nucleótidos de longitud. Aunque no se pretende limitar la longitud de la secuencia de nucleótidos a la que se unen específicamente el ácido nucleico peptídico y la poliamida, en una realización, la secuencia de nucleótidos a la que se unen específicamente el ácido nucleico peptídico y la poliamida es de 1 a 100, más preferentemente de 5 a 50, aún más preferentemente de 5 a 25, y más preferentemente de 5 a 20, nucleótidos.

El término "célula" se refiere a una sola célula. El término "células" se refiere a una población de células. La población puede ser una población pura que comprende un tipo de célula. Asimismo, la población puede comprender más de un tipo de célula. En la presente invención, no hay límite en el número de tipos celulares que puede comprender una población celular.

- 5 El término "sincronizar" o "sincronizado", cuando se refiere a una muestra de células, o "células sincronizadas" o "población celular sincronizada" se refiere a una pluralidad de células que han sido tratadas para hacer que la población de células esté en la misma fase del ciclo celular. No es necesario que todas las células de la muestra estén sincronizadas. Un pequeño porcentaje de células puede no estar sincronizado con la mayoría de las células de la muestra. Un intervalo preferido de células que se sincronizan es entre el 10-100 %. Un intervalo más preferido es entre el 30-100 %. Además, no es necesario que las células sean una población pura de un solo tipo celular. La muestra puede contener más de un tipo de célula. En este sentido, sólo uno de los tipos de células puede estar sincronizado o puede estar en una fase diferente del ciclo celular en comparación con otro tipo de células en la muestra.
- 10 El término "célula sincronizada" cuando se hace referencia a una sola célula significa que la célula ha sido manipulada de tal manera que se encuentra en una fase del ciclo celular que es diferente de la fase del ciclo celular de la célula antes de la manipulación. Alternativamente, una "célula sincronizada" se refiere a una célula que ha sido manipulada para alterar (es decir, aumentar o disminuir) la duración de la fase del ciclo celular en la que se encontraba la célula antes de la manipulación en comparación con una célula de control (por ejemplo, una célula en ausencia de la manipulación).
- 15 El término "ciclo celular" se refiere a la progresión fisiológica y morfológica de los cambios que sufren las células cuando se dividen (es decir, proliferan). En general, se reconoce que el ciclo celular está compuesto por fases denominadas "interfase", "profase", "metafase", "anafase" y "telofase". Además, las partes del ciclo celular pueden denominarse "M (mitosis)", "S (síntesis)", "G0", "G1 (brecha 1)" y "G2 (brecha 2)". Además, el ciclo celular incluye periodos de progresión intermedios a las fases mencionadas anteriormente.
- 20 El término "inhibición del ciclo celular" se refiere al cese de la progresión del ciclo celular en una célula o población de células. La inhibición del ciclo celular suele ser inducida por la exposición de las células a un agente (químico, proteínico o de otro tipo) que interfiere con aspectos de la fisiología celular para impedir la continuación del ciclo celular.
- 25 "Proliferación" o "crecimiento celular" se refiere a la capacidad de una célula madre de dividirse en dos células hijas de forma repetida, lo que resulta en un aumento total de células en la población. La población celular puede estar en un organismo o en un aparato de cultivo.
- 30 El término "capaz de modificar el ADN" o "medios modificadores del ADN" se refiere a los procedimientos, así como a los agentes o reactivos endógenos o exógenos que tienen la capacidad de inducir, o pueden ayudar a inducir, cambios en la secuencia de nucleótidos de un segmento de ADN diana. Dichos cambios pueden realizarse mediante la supresión, adición o sustitución de una o más bases en el segmento de ADN diana. No es necesario que los cambios en la secuencia de ADN confieran cambios funcionales a cualquier gen codificado por la secuencia diana. Además, no es necesario que los cambios en el ADN se realicen en una porción o porcentaje concreto de las células.
- 35 El término "secuencia de nucleótidos de interés" se refiere a cualquier secuencia de nucleótidos, cuya manipulación puede considerarse deseable por cualquier razón, por un experto en la técnica. Dichas secuencias de nucleótidos incluyen, pero no se limitan a, secuencias codificantes de genes estructurales (por ejemplo, genes reporteros, genes marcadores de selección, oncogenes, genes de resistencia a fármacos, factores de crecimiento, etc.), y secuencias reguladoras no codificantes que no codifican un ARNm o un producto proteico (por ejemplo, secuencia promotora, secuencia potenciadora, secuencia de poliadenilación, secuencia de terminación, ARN reguladores como miARN, etc.).
- 40 "Secuencia de aminoácidos", "secuencia polipeptídica", "secuencia peptídica" y "péptido" se utilizan indistintamente en el presente documento para referirse a una secuencia de aminoácidos.
- 45 "Secuencia diana", tal como se utiliza en el presente documento, se refiere a un ácido nucleico doblemente helicoidal que comprende una secuencia preferentemente de más de 8 nucleótidos de longitud pero de menos de 201 nucleótidos en longitud. En algunas realizaciones, la secuencia diana está preferentemente entre 8 y 30 bases. La secuencia diana, en general, está definida por la secuencia de nucleótidos en una de las cadenas del ácido nucleico doble helicoidal.
- 50 Tal como se utiliza en el presente documento, una "secuencia rica en purinas" o "secuencia polipurina" cuando se hace en referencia a una secuencia de nucleótidos en una de las cadenas de una secuencia de ácido nucleico doble helicoidal se define como una secuencia contigua de nucleótidos en la que más del 50 % de los nucleótidos de la secuencia diana contienen una base purina. Sin embargo, se prefiere que la secuencia diana rica en purina contenga más del 60 % de nucleótidos de purina, más preferentemente más del 75 % de nucleótidos de purina, después más preferentemente más del 90 % de nucleótidos de purina y más preferentemente el 100 % de nucleótidos de purina.
- 55 Tal como se utiliza en el presente documento, una "secuencia rica en pirimidina" o "secuencia de polipirimidina" cuando se hace referencia a una secuencia de nucleótidos en una de las cadenas de una secuencia de ácido nucleico doble helicoidal se define como una secuencia contigua de nucleótidos en la que más del 50 % de los nucleótidos de la secuencia diana contienen una base de pirimidina. Sin embargo, se prefiere que la secuencia diana rica en pirimidina contenga más del 60 % de nucleótidos de pirimidina y, más preferentemente, más del 75 % de nucleótidos de

pirimidina. En algunas realizaciones, la secuencia contiene preferentemente más del 90 % de nucleótidos de pirimidina y, en otras realizaciones, es más preferentemente el 100 % de nucleótidos de pirimidina.

Una "variante" de una primera secuencia de nucleótidos se define como una secuencia de nucleótidos que difiere de la primera secuencia de nucleótidos (por ejemplo, por tener una o más supresiones, inserciones o sustituciones que pueden detectarse mediante ensayos de hibridación o utilizando la secuenciación del ADN). Dentro de esta definición se incluye la detección de alteraciones o modificaciones en la secuencia genómica de la primera secuencia de nucleótidos. Por ejemplo, los ensayos de hibridación pueden utilizarse para detectar (1) alteraciones en el patrón de los fragmentos de enzimas de restricción capaces de hibridarse con la primera secuencia de nucleótidos cuando está comprendida en un genoma (es decir, análisis RFLP), (2) la incapacidad de una porción seleccionada de la primera secuencia de nucleótidos para hibridarse con una muestra de ADN genómico que contiene la primera secuencia de nucleótidos (por ejemplo, utilizando sondas de oligonucleótidos específicas para cada alelo), (3) hibridación inadecuada o inesperada, como la hibridación a un locus distinto del locus cromosómico normal para la primera secuencia de nucleótidos (por ejemplo, utilizando hibridación in situ fluorescente (FISH) a cromosomas en metafase extendidos, etc.). Un ejemplo de variante es una secuencia de tipo salvaje mutada.

Los términos "ácido nucleico" y "ácido nucleico no modificado", tal y como se utilizan en el presente documento, se refieren a una cualquiera de las cuatro bases conocidas del ácido desoxirribonucleico (es decir, guanina, adenina, citosina y timina). El término "ácido nucleico modificado" se refiere a un ácido nucleico cuya estructura está alterada en relación con la estructura del ácido nucleico no modificado. Un ejemplo de estas modificaciones serían las modificaciones covalentes de sustitución de las bases, como la alquilación de los amino y los nitrógenos del anillo, así como la saturación de los dobles enlaces.

Tal y como se utilizan en el presente documento, los términos "mutación" y "modificación" y sus equivalentes gramaticales cuando se utilizan en referencia a una secuencia de ácido nucleico se usan indistintamente para referirse a una supresión, inserción, sustitución, rotura de cadena y/o introducción de un aducto. Una "delección" se define como un cambio en una secuencia de ácido nucleico en la que uno o más nucleótidos están ausentes. Una "inserción" o "adición" es aquel cambio en una secuencia de ácido nucleico que ha dado lugar a la adición de uno o más nucleótidos. Una "sustitución" resulta del reemplazo de uno o más nucleótidos por una molécula que es una molécula diferente de los nucleótidos reemplazados. Por ejemplo, un ácido nucleico puede ser sustituido por un ácido nucleico diferente, como por ejemplo la sustitución de una timina por una citosina, adenina, guanina o uridina. Las sustituciones de pirimidina a pirimidina (por ejemplo, de C a T o de T a C) o de purina a purina (por ejemplo, de G a A o de A a G) se denominan transiciones, mientras que las de pirimidina a purina o de purina a pirimidina (por ejemplo, de G a T o de G a C o de A a T o A a C) se denominan transversiones. Alternativamente, un ácido nucleico puede ser sustituido por un ácido nucleico modificado, como por ejemplo la sustitución de una timina por timina glicol. Las mutaciones pueden dar lugar a una no coincidencia. El término "no coincidencia" se refiere a una interacción no covalente entre dos ácidos nucleicos, cada uno de los cuales reside en una secuencia de ácido polinucleico diferente, que no sigue las reglas de apareamiento de bases. Por ejemplo, para las secuencias parcialmente complementarias 5'-AGT-3' y 5'-AAT-3', existe una no coincidencia G-A (una transición). Los términos "introducción de un aducto" o "formación de un aducto" se refieren a la unión covalente o no covalente de una molécula a uno o más nucleótidos de una secuencia de ADN, de tal manera que la unión da lugar a una reducción (preferentemente del 10 % al 100 %, más preferentemente del 50 % al 100 %, y más preferentemente del 75 % al 100 %) del nivel de replicación y/o transcripción del ADN.

El término "rotura de cadena", cuando se hace referencia a una secuencia de ácido nucleico de doble cadena, incluye una rotura de una cadena y/o una rotura de doble cadena. Una rotura de una sola cadena (un nick) se refiere a una interrupción en una de las dos cadenas de la secuencia de ácido nucleico de doble cadena. Esto contrasta con una rotura de doble cadena, que se refiere a una interrupción en ambas cadenas de la secuencia de ácido nucleico de doble cadena. Las roturas de cadena pueden introducirse en una secuencia de ácido nucleico de doble cadena directamente (por ejemplo, mediante radiación ionizante o tratamiento con ciertos productos químicos) o indirectamente (por ejemplo, mediante incisión enzimática en una base del ácido nucleico).

Los términos "célula mutante" y "célula modificada" se refieren a una célula que contiene al menos una modificación en la secuencia genómica de la célula.

El término "porción" cuando se utiliza en referencia a una secuencia de nucleótidos se refiere a fragmentos de esa secuencia de nucleótidos. Los fragmentos pueden tener un tamaño que va desde 5 residuos de nucleótidos hasta la secuencia completa de nucleótidos menos un residuo de ácido nucleico.

Se dice que las moléculas de ADN tienen "extremos 5'" y "extremos 3'" porque los mononucleótidos reaccionan para formar oligonucleótidos de manera que el fosfato 5' de un anillo de pentosa mononucleótido se une al oxígeno 3' de su vecino en una dirección mediante un enlace fosfodiéster. Por lo tanto, un extremo de un oligonucleótido se denomina "extremo 5'" si su fosfato 5' no está unido al oxígeno 3' de un anillo pentoso mononucleótido. Un extremo de un oligonucleótido se denomina "extremo 3'" si su oxígeno 3' no está unido a un fosfato 5' de otro anillo pentoso mononucleótido. Tal y como se utiliza en el presente documento, una secuencia de ácido nucleico, incluso si es interna a un oligonucleótido más grande, también puede decirse que tiene extremos 5' y 3'. En una molécula de ADN lineal o circular, los elementos discretos se denominan "corriente arriba" o 5' de los elementos "corriente abajo" o 3'. Esta terminología refleja que la transcripción procede en dirección 5' a 3' a lo largo de la cadena de ADN. Los elementos

promotores y potenciadores que dirigen la transcripción de un gen ligado suelen estar situados a 5' o corriente arriba de la región codificante. Sin embargo, los elementos potenciadores pueden ejercer su efecto incluso cuando están situados a 3' del elemento promotor y de la región codificante. Las señales de terminación de la transcripción y de poliadenilación se sitúan a 3' o corriente abajo de la región codificante.

- 5 El término "molécula de ADN recombinante", tal y como se utiliza en este documento, se refiere a una molécula de ADN que está compuesta por segmentos de ADN unidos mediante técnicas de biología molecular.

El término "proteína recombinante" o "polipéptido recombinante", tal y como se utiliza en el presente documento, se refiere a una molécula de proteína que se expresa utilizando una molécula de ADN recombinante.

- 10 Tal y como se utiliza en el presente documento, los términos "vector" y "vehículo" se utilizan indistintamente en referencia a las moléculas de ácido nucleico que transfieren segmento(s) de ADN de una célula a otra.

- 15 Los términos "en combinación operable", "en orden operable" y "operablemente enlazado", tal y como se utilizan en el presente documento, se refieren a la unión de secuencias de ácido nucleico de tal manera que se produce una molécula de ácido nucleico capaz de dirigir la transcripción de un gen determinado y/o la síntesis de una molécula de proteína deseada. Los términos también se refieren a la unión de secuencias de aminoácidos de tal manera que se produce una proteína funcional.

- 20 El término "transfección", tal como se utiliza en el presente documento, se refiere a la introducción de ADN extraño en las células. La transfección puede llevarse a cabo por diversos medios conocidos en la técnica, como la coprecipitación de fosfato de calcio-ADN, la transfección mediada por DEAE-dextrano, la transfección mediada por polibreno, la electroporación, la microinyección, la fusión de liposomas, la lipofectina, la fusión de protoplastos, la infección retroviral, la biolística (es decir, el bombardeo de partículas) y otros similares.

- 25 Tal y como se utilizan en el presente documento, los términos "complementario" o "complementariedad" se emplean en referencia a "polinucleótidos" y "oligonucleótidos" (que son términos intercambiables que se refieren a una secuencia de nucleótidos) relacionados por las reglas de apareamiento de bases. Por ejemplo, la secuencia "5'-CAGT-3'", es complementaria a la secuencia "5'-ACTG-3'". La complementariedad puede ser "parcial" o "total". La complementariedad "parcial" se da cuando una o más bases de ácido nucleico no coinciden según las reglas de emparejamiento de bases. La complementariedad "total" o "completa" entre ácidos nucleicos es aquella en la que todas y cada una de las bases del ácido nucleico es coincidente con otra base según las reglas de emparejamiento de bases. El grado de complementariedad entre las cadenas de ácido nucleico puede tener efectos significativos en la eficiencia y la fuerza de la hibridación entre las cadenas de ácido nucleico. Esto puede ser especialmente importante en las reacciones de amplificación, así como en los procedimientos de detección que dependen de la unión entre ácidos nucleicos. Por comodidad, los términos "polinucleótidos" y "oligonucleótidos" incluyen moléculas que incluyen nucleósidos.

- 35 Los términos "homología" y "homólogo", tal como se utilizan en el presente documento en referencia a las secuencias de nucleótidos, se refieren a un grado de complementariedad con otras secuencias de nucleótidos. Puede haber homología parcial o completa (es decir, identidad). Cuando se utiliza en referencia a una secuencia de ácido nucleico de doble cadena, como un ADNc o un clon genómico, el término "sustancialmente homólogo" se refiere a cualquier secuencia de ácido nucleico (por ejemplo, una sonda) que pueda hibridarse con una o ambas cadenas de la secuencia de ácido nucleico de doble cadena en condiciones de baja rigurosidad como las descritas anteriormente. Una secuencia de nucleótidos que es parcialmente complementaria, es decir, "sustancialmente homóloga", a una secuencia de ácido nucleico es aquella que inhibe, al menos parcialmente, la hibridación de una secuencia completamente complementaria con una secuencia de ácido nucleico objetivo. La inhibición de la hibridación de la secuencia completamente complementaria a la secuencia diana puede examinarse mediante un ensayo de hibridación (transferencia Southern o Northern, hibridación en solución y similares) en condiciones de baja rigurosidad. Una secuencia o sonda sustancialmente homóloga competirá e inhibirá la unión (es decir, la hibridación) de una secuencia completamente homóloga a una secuencia diana en condiciones de baja rigurosidad. Esto no quiere decir que las condiciones de baja rigurosidad sean tales que permitan la unión no específica; las condiciones de baja rigurosidad requieren que la unión de dos secuencias entre sí sea una interacción específica (es decir, selectiva). La ausencia de unión inespecífica puede comprobarse mediante el uso de una segunda secuencia diana que carezca incluso de un grado parcial de complementariedad (por ejemplo, menos de un 30 % de identidad); en ausencia de unión inespecífica, la sonda no hibridará con la segunda diana no complementaria.

- 55 Las condiciones de baja rigurosidad comprenden condiciones equivalentes a la unión o hibridación a 68° C. en una solución que consiste en 5×SSPE (43,8 g/l de NaCl, 6,9 g/l de NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O y 1,85 g/l de EDTA, pH ajustado a 7,4 con NaOH), 0,1 % de SDS, 5× de reactivo de Denhardt (50× de Denhardt contiene por 500 ml: 5 g de Ficoll (Tipo 400, Pharmacia), 5 g de BSA (Fracción V; Sigma)) y 100 µg/ml de ADN de esperma de salmón desnaturalizado, seguido de un lavado en una solución que comprende 2,0×SSPE, 0,1 % de SDS a temperatura ambiente cuando se emplea una sonda de aproximadamente 100 a aproximadamente 1000 nucleótidos de longitud.

Además, las condiciones que promueven la hibridación en condiciones de alta rigurosidad (por ejemplo, el aumento de la temperatura de los pasos de hibridación y/o lavado, el uso de formamida en la solución de hibridación, etc.) son

bien conocidas en la técnica. Las condiciones de alta rigurosidad, cuando se utilizan en referencia a la hibridación de ácidos nucleicos, comprenden condiciones equivalentes a la unión o hibridación a 68°C. en una solución que consiste en 5×SSPE, 1 % SDS, 5×Reactivo de Denhardt y 100 µg/ml de ADN de esperma de salmón desnaturalizado, seguido de un lavado en una solución que comprende 0,1×SSPE y 0,1 % SDS a 68°C. cuando se emplea una sonda de aproximadamente 100 a aproximadamente 1000 nucleótidos de longitud.

Es bien conocido en la técnica que pueden emplearse numerosas condiciones equivalentes para comprender las condiciones de baja rigurosidad; factores como la longitud y la naturaleza (ADN, ARN, composición de bases) de la sonda y la naturaleza de la diana (ADN, ARN, composición de bases, presente en solución o inmovilizada, etc.) y la concentración de las sales y otros componentes (por ejemplo, la presencia o ausencia de formamida, sulfato de dextrano, polietilenglicol), así como los componentes de la solución de hibridación pueden variarse para generar condiciones de hibridación de baja rigurosidad diferentes, pero equivalentes, a las condiciones enumeradas anteriormente.

El término "equivalente", cuando se hace referencia a una condición de hibridación en relación con una condición de hibridación de interés, significa que la condición de hibridación y la condición de hibridación de interés dan lugar a la hibridación de secuencias de ácido nucleico que tienen el mismo intervalo de porcentaje ( %) de homología. Por ejemplo, si una condición de hibridación de interés resulta en la hibridación de una primera secuencia de ácido nucleico con otras secuencias de ácido nucleico que tienen de 50 % a 70 % de homología con la primera secuencia de ácido nucleico, entonces se dice que otra condición de hibridación es equivalente a la condición de hibridación de interés si esta otra condición de hibridación también resulta en la hibridación de la primera secuencia de ácido nucleico con las otras secuencias de ácido nucleico que tienen de 50 % a 70 % de homología con la primera secuencia de ácido nucleico.

Tal y como se utiliza en el presente documento, el término "hibridación" se usa en referencia al emparejamiento de ácidos nucleicos complementarios utilizando cualquier proceso por el que una cadena de ácido nucleico se une con una cadena complementaria a través del emparejamiento de bases para formar un complejo de hibridación. La hibridación y la fuerza de la hibridación (es decir, la fuerza de la asociación entre los ácidos nucleicos) se ven afectadas por factores como el grado de complementariedad entre los ácidos nucleicos, la rigurosidad de las condiciones implicadas, la Tm del híbrido formado y la relación G:C dentro de los ácidos nucleicos.

Tal y como se utiliza en el presente documento, el término "complejo de hibridación" se refiere a un complejo formado entre dos secuencias de ácido nucleico en virtud de la formación de enlaces de hidrógeno entre las bases G y C complementarias y entre las bases A y T complementarias; estos enlaces de hidrógeno pueden ser estabilizados además por interacciones de apilamiento de bases. Las dos secuencias complementarias de ácido nucleico se unen por hidrógeno en una configuración antiparalela. Un complejo de hibridación puede formarse en solución (por ejemplo, en el análisis de Cot o Rot) o entre una secuencia de ácido nucleico presente en solución y otra secuencia de ácido nucleico inmovilizada a un soporte sólido (por ejemplo, una membrana de nailon o un filtro de nitrocelulosa como los empleados en transferencia Southern y Northern transferencia, transferencia de mancha o un portaobjetos de vidrio como los empleados en hibridación *in situ*, incluyendo FISH (hibridación *in situ* fluorescente)).

Tal y como se utiliza en el presente documento, el término "Tm" se usa en referencia a la "temperatura de fusión" La temperatura de fusión es la temperatura a la que una población de moléculas de ácido nucleico de doble cadena se disocia a medias en cadenas simples. La ecuación para calcular la Tm de los ácidos nucleicos es bien conocida en la técnica. Como se indica en las referencias estándar, una estimación simple del valor de Tm puede calcularse mediante la ecuación:  $T_m = 81,5 + 0,41(\% G+C)$ , cuando un ácido nucleico está en solución acuosa a 1 M de NaCl (véase, por ejemplo Anderson y Young, Quantitative Filter Hybridization, en Nucleic Acid Hybridization, 1985). Otras referencias incluyen cálculos más sofisticados que tienen en cuenta las características estructurales y de secuencia para el cálculo de la Tm.

Tal y como se utiliza en el presente documento, el término "rigurosidad" se emplea en referencia a las condiciones de temperatura, fuerza iónica y presencia de otros compuestos, como los disolventes orgánicos, bajo las que se realizan las hibridaciones de ácidos nucleicos. La "rigurosidad" se produce típicamente en un intervalo de aproximadamente Tm°C. a aproximadamente 20°C. a 25°C. por debajo de Tm. Como comprenderán los expertos en la técnica, una hibridación rigurosa puede utilizarse para identificar o detectar secuencias polinucleotídicas idénticas o para identificar o detectar secuencias polinucleotídicas similares o relacionadas.

Los términos "unión específica", "especificidad de unión" y sus equivalentes gramaticales, cuando se hacen en referencia a la unión de una primera secuencia de nucleótidos a una segunda secuencia de nucleótidos, se refieren a la interacción preferente entre la primera secuencia de nucleótidos con la segunda secuencia de nucleótidos en comparación con la interacción entre la segunda secuencia de nucleótidos con una tercera secuencia de nucleótidos. La unión específica es un término relativo que no requiere una especificidad absoluta de la unión; en otras palabras, el término "unión específica" no requiere que la segunda secuencia de nucleótidos interactúe con la primera secuencia de nucleótidos en ausencia de una interacción entre la segunda secuencia de nucleótidos y la tercera secuencia de nucleótidos. Más bien, es suficiente que el nivel de interacción entre la primera secuencia de nucleótidos y la segunda secuencia de nucleótidos sea mayor que el nivel de interacción entre la segunda secuencia de nucleótidos con la tercera secuencia de nucleótidos. la "unión específica" de una primera secuencia de nucleótidos con una segunda

secuencia de nucleótidos también significa que la interacción entre la primera secuencia de nucleótidos y la segunda secuencia de nucleótidos depende de la presencia de una estructura particular en la primera secuencia de nucleótidos o dentro de ella; en otras palabras, la segunda secuencia de nucleótidos reconoce y se une a una estructura específica en la primera secuencia de nucleótidos o dentro de ella, en lugar de a ácidos nucleicos o a secuencias de nucleótidos en general. Por ejemplo, si una segunda secuencia de nucleótidos es específica para la estructura "A" que está sobre o dentro de una primera secuencia de nucleótidos, la presencia de una tercera secuencia de ácido nucleico que contenga la estructura A reducirá la cantidad de la segunda secuencia de nucleótidos que se une a la primera secuencia de nucleótidos.

Tal y como se utiliza en el presente documento, el término "ácido nucleico amplificable" se utiliza en referencia a los ácidos nucleicos que pueden ser amplificados por cualquier procedimiento de amplificación. Se contempla que el "ácido nucleico amplificable" comprenderá normalmente la "plantilla de la muestra"

Los términos "secuencia de ácido nucleico heteróloga" o "ADN heterólogo" se utilizan indistintamente para referirse a una secuencia de nucleótidos que está ligada a una secuencia de ácido nucleico a la que no está ligada en la naturaleza, o a la que está ligada en un lugar diferente en la naturaleza. El ADN heterólogo no es endógeno a la célula en la que se introduce, sino que se ha obtenido de otra célula. Generalmente, aunque no necesariamente, este ADN heterólogo codifica ARN y proteínas que no son producidos normalmente por la célula en la que se expresa. Algunos ejemplos de ADN heterólogo son los genes informadores, las secuencias reguladoras de la transcripción y la traducción, las proteínas marcadoras seleccionables (por ejemplo, las proteínas que confieren resistencia a los fármacos), etc.

La "amplificación" se define como la producción de copias adicionales de una secuencia de ácido nucleico y generalmente se lleva a cabo utilizando tecnologías de reacción en cadena de la polimerasa bien conocidas en la técnica (Dieffenbach C W y G S Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y.). Tal y como se utiliza en el presente documento, el término "reacción en cadena de la polimerasa" ("PCR") se refiere al procedimiento de K. B. Mullis U.S. Pat. Números 4.683.195, y 4.683.202 que describen un procedimiento para aumentar la concentración de un segmento de una secuencia diana en una mezcla de ADN genómico sin clonación o purificación. La longitud del segmento amplificado de la secuencia diana deseada viene determinada por las posiciones relativas de dos cebadores oligonucleótidos entre sí y, por tanto, esta longitud es un parámetro controlable. En virtud del aspecto de repetición del proceso, el procedimiento se denomina "reacción en cadena de la polimerasa" (en adelante, "PCR"). Dado que los segmentos amplificados deseados de la secuencia diana se convierten en las secuencias predominantes (en términos de concentración) en la mezcla, se dice que están "amplificados por PCR"

Con la PCR, es posible amplificar una sola copia de una secuencia diana específica en el ADN genómico a un nivel detectable por varias metodologías diferentes (por ejemplo, hibridación con una sonda marcada; incorporación de cebadores biotinilados seguida de detección con conjugado de avidina-enzima; incorporación de desoxinucleótidos trifosfatos marcados con  $^{32}\text{P}$ , como dCTP o dATP, en el segmento amplificado). Además del ADN genómico, se puede amplificar cualquier secuencia de oligonucleótidos con el conjunto adecuado de moléculas de cebadores. En particular, los segmentos amplificados creados por el propio proceso de PCR son, en sí mismos, plantillas eficaces para posteriores amplificaciones por PCR.

Uno de estos procedimientos preferidos, especialmente para aplicaciones comerciales, se basa en la tecnología de PCR en tiempo real TaqMan®, ampliamente utilizada, y combina la PCR específica para el alelo con un reactivo de bloqueo (ASB-PCR) para suprimir la amplificación del alelo salvaje. La ASB-PCR puede utilizarse para la detección de mutaciones somáticas o de la línea germinal en el ADN o el ARN extraído de cualquier tipo de tejido, incluidas las muestras tumorales embebidas en parafina fijadas en formalina. Se ha desarrollado un conjunto de reglas de diseño de reactivos que permiten la detección sensible y selectiva de sustituciones, inserciones o deleciones puntuales sobre un fondo de alelos de tipo salvaje en exceso de mil veces o más. (Morlan J, Baker J, Sinicropi D Mutation Detection by Real-Time PCR: A Simple, Robust and Highly Selective Method. PLoS ONE 4(2): e4584, 2009)

Los términos "reacción en cadena de la polimerasa de transcripción inversa" y "RT-PCR" se refieren a un procedimiento de transcripción inversa de una secuencia de ARN para generar una mezcla de secuencias de ADNc, seguido de un aumento de la concentración de un segmento deseado de las secuencias de ADNc transcritas en la mezcla sin clonación ni purificación. Normalmente, el ARN se transcribe de forma inversa utilizando un único cebador (por ejemplo, un cebador oligo-dT) antes de la amplificación por PCR del segmento deseado del ADN transcrito utilizando dos cebadores.

Tal y como se utiliza en el presente documento, el término "cebador" se refiere a un oligonucleótido, ya sea que se produzca de forma natural como en un digerido de restricción purificado o que se produzca sintéticamente, que es capaz de actuar como punto de iniciación de la síntesis cuando se coloca bajo condiciones en las que se induce la síntesis de un producto de extensión del cebador que es complementario a una cadena de ácido nucleico, (es decir, en presencia de nucleótidos y de un agente inductor como la ADN polimerasa y a una temperatura y pH adecuados). El cebador es preferentemente monocatenario para una máxima eficiencia en la amplificación, pero puede ser alternativamente bicatenario. Si es de doble cadena, el cebador se trata primero para separar sus cadenas antes de utilizarlo para preparar los productos de extensión. Preferentemente, el cebador es un oligodesoxirribonucleótido. El

cebador debe ser lo suficientemente largo para cebar la síntesis de los productos de extensión en presencia del agente inductor. Las longitudes exactas de los cebadores dependerán de muchos factores, como la temperatura, la fuente del cebador y el uso del procedimiento.

5 Tal como se utiliza en el presente documento, el término "sonda" se refiere a un oligonucleótido (es decir, una secuencia de nucleótidos), ya sea que se produzca naturalmente como en un digerido de restricción purificado o producido sintéticamente, recombinantemente o por amplificación de PCR, que es capaz de hibridarse con otro oligonucleótido de interés. Una sonda puede ser monocatenaria o bicatenaria. Las sondas son útiles para la detección, identificación y aislamiento de determinadas secuencias genéticas. Se contempla que cualquier sonda utilizada en la presente invención estará marcada con cualquier "molécula informadora", de modo que sea detectable en cualquier sistema de detección, incluyendo, pero sin limitarse a los sistemas enzimáticos (por ejemplo, ELISA, así como los ensayos histoquímicos basados en enzimas), fluorescentes, radiactivos y luminiscentes. No se pretende que la presente invención se limite a ningún sistema de detección o marcado en particular.

15 Tal y como se utiliza en el presente documento, los términos "endonucleasas de restricción" y "enzimas de restricción" se refieren a enzimas bacterianas, cada una de las cuales corta o mella el ADN de doble o simple cadena en o cerca de una secuencia de nucleótidos específica, por ejemplo, se puede utilizar un dominio de endonucleasa de una endonucleasa de restricción de tipo IIS (por ejemplo, FokI), tal y como enseñan Kim et al., 1996, Proc. Nacional. Acad. Sci. USA 6:1 (60)).

20 Tal como se utiliza en el presente documento, el término "un oligonucleótido que tiene una secuencia de nucleótidos que codifica un gen" significa una secuencia de ácido nucleico que comprende la región codificante de un gen, es decir, la secuencia de ácido nucleico que codifica un producto génico. La región codificante puede estar presente en forma de ADNc, ADN genómico o ARN. Cuando se presenta en forma de ADN, el oligonucleótido puede ser de una sola cadena (es decir, la cadena sensitiva) o de dos cadenas. Además, "un oligonucleótido que tiene una secuencia de nucleótidos que codifica un gen" puede incluir elementos de control adecuados, como potenciadores, promotores, uniones de empalme, señales de poliadenilación, etc., si son necesarios para permitir el inicio adecuado de la transcripción y/o el procesamiento correcto del transcrito de ARN primario. Además, la región codificante de la presente invención puede contener potenciadores endógenos, uniones de empalme, secuencias intervinientes, señales de poliadenilación, etc.

30 Las señales de control transcripcional en eucariotas comprenden elementos "potenciadores". Los potenciadores consisten en conjuntos cortos de secuencias de ADN que interactúan específicamente con las proteínas celulares que participan en la transcripción (Maniatis, T. et al., Science 236:1237, 1987). Se han aislado elementos potenciadores de diversas fuentes eucariotas, como genes de plantas, levaduras, insectos, células de mamíferos y virus. La selección de un potenciador concreto depende del tipo de célula que se vaya a utilizar para expresar la proteína de interés.

35 La presencia de "señales de empalme" en un vector de expresión a menudo da lugar a mayores niveles de expresión del transcrito recombinante. Las señales de empalme median la eliminación de los intrones del transcrito de ARN primario y consisten en un sitio donante y aceptor de empalme (Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, pp. 16.7-16.8, 1989). Un sitio donante y aceptor de empalme comúnmente utilizado es la unión de empalme del ARN 16S del SV40.

40 La expresión eficiente de secuencias de ADN recombinante en células eucariotas requiere la expresión de señales que dirijan la terminación y poliadenilación eficiente del transcrito resultante. Las señales de terminación de la transcripción se encuentran generalmente corriente abajo de la señal de poliadenilación y tienen una longitud de unos cientos de nucleótidos. El término "sitio poli A" o "secuencia poli A", tal como se utiliza en el presente documento, denota una secuencia de ADN que dirige tanto la terminación como la poliadenilación del transcrito de ARN naciente. Es deseable una poliadenilación eficiente del transcrito recombinante, ya que los transcritos que carecen de cola poli A son inestables y se degradan rápidamente. La señal de poli A utilizada en un vector de expresión puede ser "heteróloga" o "endógena". Una señal de poli A endógena es la que se encuentra de forma natural en el extremo 3' de la región codificante de un determinado gen en el genoma. Una señal poli A heteróloga es aquella que se aísla de un gen y se coloca a 3' de otro gen.

50 El término "promotor", "elemento promotor" o "secuencia promotora", tal y como se utiliza en el presente documento, se refiere a una secuencia de ADN que cuando se coloca en el extremo 5' de (es decir, precede) una secuencia de oligonucleótidos es capaz de controlar la transcripción de la secuencia de oligonucleótidos en ARNm. Un promotor suele estar situado a 5' (es decir, corriente arriba) de una secuencia de oligonucleótidos cuya transcripción en ARNm controla, y proporciona un sitio para la unión específica de la ARN polimerasa y para el inicio de la transcripción.

55 El término "actividad promotora" cuando se hace en referencia a una secuencia de ácido nucleico se refiere a la capacidad de la secuencia de ácido nucleico para iniciar la transcripción de una secuencia de oligonucleótidos en ARNm.

El término "específico de tejido", tal como se aplica a un promotor, se refiere a un promotor que es capaz de dirigir la expresión selectiva de una secuencia de oligonucleótidos a un tipo específico de tejido en la ausencia relativa de expresión del mismo oligonucleótido en un tipo diferente de tejido. La especificidad tisular de un promotor puede

5 evaluarse, por ejemplo, vinculando de forma operativa un gen informador a la secuencia promotora para generar un constructo informador, introduciendo el constructo informador en el genoma de una planta o un animal de forma que el constructo informador se integre en todos los tejidos del animal transgénico resultante, y detectando la expresión del gen informador (por ejemplo, detectando ARNm, proteínas o la actividad de una proteína codificada por el gen informador) en diferentes tejidos de la planta o el animal transgénicos. La selectividad no tiene por qué ser absoluta. La detección de un mayor nivel de expresión del gen informador en uno o más tejidos en relación con el nivel de expresión del gen informador en otros tejidos muestra que el promotor es específico para los tejidos en los que se detectan mayores niveles de expresión.

10 El término "específico del tipo de célula" aplicado a un promotor se refiere a un promotor que es capaz de dirigir la expresión selectiva de una secuencia de oligonucleótidos en un tipo específico de célula en ausencia relativa de expresión de la misma secuencia de oligonucleótidos en un tipo diferente de célula dentro del mismo tejido. El término "específico del tipo de célula" cuando se aplica a un promotor también significa un promotor capaz de promover la expresión selectiva de un oligonucleótido en una región dentro de un solo tejido. De nuevo, la selectividad no tiene por qué ser absoluta. La especificidad del tipo de célula de un promotor puede evaluarse utilizando procedimientos bien conocidos en la técnica, por ejemplo, la tinción inmunohistoquímica como se describe en el presente documento. En resumen, las secciones de tejido se embeben en parafina, y las secciones de parafina se hacen reaccionar con un anticuerpo primario que es específico para el producto polipeptídico codificado por la secuencia de oligonucleótidos cuya expresión está controlada por el promotor. Como alternativa al seccionamiento en parafina, las muestras pueden ser crioseccionadas. Por ejemplo, las secciones pueden congelarse antes y durante el seccionamiento, evitando así la posible interferencia de la parafina residual. Se permite que un anticuerpo secundario marcado (por ejemplo, conjugado con peroxidasa) que es específico para el anticuerpo primario se una al tejido seccionado y se detecte la unión específica (por ejemplo, con avidina/biotina) mediante microscopía.

15 Los términos "expresión selectiva", "expresar selectivamente" y sus equivalentes gramaticales se refieren a una comparación de los niveles relativos de expresión en dos o más regiones de interés. Por ejemplo, la "expresión selectiva" cuando se utiliza en relación con los tejidos se refiere a un nivel sustancialmente mayor de expresión de un gen de interés en un tejido particular, o a un número sustancialmente mayor de células que expresan el gen dentro de ese tejido, en comparación, respectivamente, con el nivel de expresión y el número de células que expresan el mismo gen en otro tejido (es decir, la selectividad no tiene que ser absoluta). La expresión selectiva no requiere, aunque puede incluir, la expresión de un gen de interés en un tejido concreto y la ausencia total de expresión del mismo gen en otro tejido. Del mismo modo, la "expresión selectiva", tal como se utiliza en el presente documento en referencia a los tipos de células, se refiere a un nivel sustancialmente mayor de expresión de, o un número sustancialmente mayor de células que expresan, un gen de interés en un tipo de célula particular, cuando se compara, respectivamente, con los niveles de expresión del gen y con el número de células que expresan el gen en otro tipo de célula.

25 El término "contiguo" cuando se utiliza en referencia a dos o más secuencias de nucleótidos significa que las secuencias de nucleótidos se ligan en tándem, ya sea en ausencia de secuencias intervinientes, o en presencia de secuencias intervinientes que no comprenden uno o más elementos de control.

30 Tal y como se utilizan en el presente documento, los términos "codificación de moléculas de ácido nucleico", "codificación de nucleótidos", "codificación de secuencias de ADN" y "codificación de ADN" se refieren al orden o secuencia de desoxirribonucleótidos a lo largo de una cadena de ácido desoxirribonucleico. El orden de estos desoxirribonucleótidos determina el orden de los aminoácidos a lo largo de la cadena polipeptídica (proteína). Así, la secuencia de ADN codifica la secuencia de aminoácidos.

35 El término "aislado", cuando se utiliza en relación con un ácido nucleico, como en "un oligonucleótido aislado", se refiere a una secuencia de ácido nucleico que está separada de al menos un ácido nucleico contaminante con el que está normalmente asociado en su fuente natural. El ácido nucleico aislado es un ácido nucleico presente en una forma o entorno diferente al que se encuentra en la naturaleza. Por el contrario, los ácidos nucleicos no aislados son ácidos nucleicos como el ADN y el ARN que se encuentran en el estado en que existen en la naturaleza. Por ejemplo, una determinada secuencia de ADN (por ejemplo, un gen) se encuentra en el cromosoma de la célula huésped en proximidad a los genes vecinos; las secuencias de ARN, como una secuencia específica de ARNm que codifica una proteína específica, se encuentran en la célula como una mezcla con otros numerosos ARNm que codifican una multitud de proteínas. Sin embargo, el ácido nucleico aislado que codifica un polipéptido de interés incluye, a modo de ejemplo, dicho ácido nucleico en células que expresan ordinariamente el polipéptido de interés cuando el ácido nucleico se encuentra en una ubicación cromosómica o extracromosómica diferente a la de las células naturales, o está flanqueado por una secuencia de ácido nucleico diferente a la que se encuentra en la naturaleza. El ácido nucleico u oligonucleótido aislado puede estar presente en forma monocatenaria o bicatenaria. El ácido nucleico aislado puede identificarse fácilmente (si se desea) mediante diversas técnicas (por ejemplo, hibridación, precipitación de mancha, etc.). Cuando un ácido nucleico u oligonucleótido aislado se va a utilizar para expresar una proteína, el oligonucleótido contendrá como mínimo la cadena en sentido o codificante (es decir, el oligonucleótido puede ser monocatenario). Alternativamente, puede contener tanto la cadena en sentido como la antisentido (es decir, el oligonucleótido puede ser de doble cadena).

60 Tal como se utiliza en el presente documento, el término "purificado" o "purificar" se refiere a la eliminación de uno o más componentes (no deseados) de una muestra. Por ejemplo, cuando los polipéptidos recombinantes se expresan

en células huésped bacterianas, los polipéptidos se purifican mediante la eliminación de las proteínas de la célula huésped, aumentando así el porcentaje de polipéptidos recombinantes en la muestra.

5 Tal y como se utiliza en este documento, el término "sustancialmente purificado" se refiere a moléculas, ya sean secuencias nucleicas o de aminoácidos, que se retiran de su entorno natural, se aíslan o se separan, y están libres al menos en un 60 %, preferentemente en un 75 % y más preferentemente en un 90 %, de otros componentes con los que están asociados de forma natural. Un "polinucleótido aislado" es, por tanto, un polinucleótido sustancialmente purificado.

10 Tal y como se utiliza en el presente documento, el término "región codificante" cuando se utiliza en referencia a un gen estructural se refiere a las secuencias de nucleótidos que codifican los aminoácidos que se encuentran en el polipéptido naciente como resultado de la traducción de una molécula de ARNm. En los eucariotas, la región codificante está delimitada en el lado 5' generalmente por el triplete de nucleótidos "ATG" que codifica la metionina iniciadora y en el lado 3' por uno de los tres tripletes que especifican los codones de parada (es decir, TAA, TAG, TGA).

15 Por "secuencia codificante" se entiende una secuencia de un ácido nucleico o su complemento, o una parte del mismo, que puede transcribirse y/o traducirse para producir el ARNm para y/o el polipéptido o un fragmento del mismo. Las secuencias codificantes incluyen los exones de un ADN genómico o los transcritos de ARN primario inmaduro, que son unidos por la maquinaria bioquímica de la célula para proporcionar un ARNm maduro. La cadena antisentido es el complemento de dicho ácido nucleico, y la secuencia codificante puede deducirse de ella.

20 Por "secuencia no codificante" se entiende una secuencia de un ácido nucleico o su complemento, o una parte del mismo que no se transcribe en aminoácido in vivo, o donde el ARNt no interactúa para colocar o intentar colocar un aminoácido. Las secuencias no codificantes incluyen tanto las secuencias de intrones en el ADN genómico o en los transcritos primarios de ARN inmaduros, como las secuencias asociadas a los genes, como promotores, potenciadores, silenciadores, etc.

25 Tal como se utiliza en el presente documento, el término "gen estructural" o "secuencia de nucleótidos estructural" se refiere a una secuencia de ADN que codifica un ARN o una proteína que no controla la expresión de otros genes. En cambio, un "gen regulador" o "secuencia reguladora" es un gen estructural que codifica productos (por ejemplo, factores de transcripción) que controlan la expresión de otros genes.

30 Tal como se utiliza en el presente documento, el término "elemento regulador" se refiere a un elemento genético que controla algún aspecto de la expresión de las secuencias de ácido nucleico. Por ejemplo, un promotor es un elemento regulador que facilita el inicio de la transcripción de una región codificante operable. Otros elementos reguladores son las señales de empalme, las señales de poliadenilación, las señales de terminación, etc.

35 Tal como se utiliza en el presente documento, el término "sitio de unión a factores de transcripción peptídicos" o "sitio de unión a factores de transcripción" se refiere a una secuencia de nucleótidos que se une a factores de transcripción proteicos y, por lo tanto, controla algún aspecto de la expresión de secuencias de ácidos nucleicos. Por ejemplo, los sitios de unión de Sp-1 y API (proteína activadora 1) son ejemplos de sitios de unión de factores de transcripción peptídicos.

40 Tal como se utiliza en el presente documento, el término "gen" significa las secuencias desoxirribonucleotídicas que comprenden la región codificante de un gen estructural. Un "gen" también puede incluir secuencias no traducidas situadas adyacentes a la región codificante tanto en el extremo 5' como en el 3', de manera que el gen corresponde a la longitud del ARNm completo. Las secuencias que se encuentran a 5' de la región codificante y que están presentes en el ARNm se denominan secuencias no traducidas 5'. Las secuencias que se encuentran a 3' o corriente abajo de la región codificante y que están presentes en el ARNm se denominan secuencias no traducidas 3'. El término "gen" abarca tanto el ADNc como las formas genómicas de un gen. Una forma genómica o clon de un gen contiene la región codificante interrumpida con secuencias no codificantes denominadas "intrones" o "regiones intervinientes" o "secuencias intervinientes". Los intrones son segmentos de un gen que se transcriben en ARN nuclear heterogéneo (ARNhn); los intrones pueden contener elementos reguladores como los potenciadores. Los intrones se eliminan o "empalman" del transcrito nuclear o primario; por tanto, los intrones están ausentes en el transcrito del ARN mensajero (ARNm). El ARNm funciona durante la traducción para especificar la secuencia o el orden de los aminoácidos en un polipéptido naciente.

50 Además de contener intrones, las formas genómicas de un gen también pueden incluir secuencias situadas en los extremos 5' y 3' de las secuencias que están presentes en el transcrito de ARN. Estas secuencias se denominan secuencias o regiones "flanqueantes" (estas secuencias flanqueantes se sitúan a 5' o 3' de las secuencias no traducidas presentes en el transcrito del ARNm). La región flanqueante 5' puede contener secuencias reguladoras como promotores y potenciadores que controlan o influyen en la transcripción del gen. La región flanqueante 3' puede contener secuencias que dirigen la terminación de la transcripción, la escisión postranscripcional y la poliadenilación.

55 Un "animal no humano" se refiere a cualquier animal que no sea humano e incluye vertebrados como roedores, primates no humanos, ovinos, bovinos, rumiantes, lagomorfos, porcinos, caprinos, equinos, caninos, felinos, aves, etc.

Los animales no humanos preferidos se seleccionan del orden Rodentia. "Animal no humano" se refiere además a anfibios (por ejemplo, *Xenopus*), reptiles, insectos (por ejemplo, *Drosophila*) y otras especies animales no mamíferas.

5 Tal como se utiliza en el presente documento, el término "transgénico" se refiere a un organismo o célula que tiene ADN derivado de otro organismo insertado en el que se integra en el genoma, ya sea de las células somáticas y/o de la línea germinal de la planta o el animal. Un "transgén" es una secuencia de ADN que es parcial o totalmente heteróloga (es decir, que no está presente en la naturaleza) para la planta o el animal en el que se encuentra, o que es homóloga a una secuencia endógena (es decir, una secuencia que se encuentra en el animal en la naturaleza) y se inserta en el genoma de la planta o el animal en un lugar que difiere del de la secuencia natural. Las plantas o animales transgénicos que incluyen uno o más transgenes están dentro del ámbito de esta invención. Además, el término "transgénico", tal y como se utiliza en este documento, se refiere a un animal al que se le han modificado y/o "anulado" uno o más genes (haciéndolos no funcionales o haciéndolos funcionar a un nivel reducido, es decir, una mutación "anulada") mediante los procedimientos de la invención, por recombinación homóloga, mutación TFO o por procesos similares. Por ejemplo, en algunas realizaciones, un organismo o célula transgénica incluye ADN insertado que incluye un promotor y/o una región codificante ajena.

15 Una "célula transformada" es una célula o línea celular que ha adquirido la capacidad de crecer en cultivo celular durante múltiples generaciones, la capacidad de crecer en agar blando, y/o la capacidad de no tener un crecimiento celular inhibido por el contacto célula a célula. En este sentido, la transformación se refiere a la introducción de material genético extraño en una célula u organismo. La transformación puede llevarse a cabo mediante cualquier procedimiento conocido que permita la introducción con éxito de ácidos nucleicos en las células y que dé lugar a la expresión del ácido nucleico introducido. "Transformación" incluye, pero no se limita a procedimientos tales como la transfección, la microinyección, la electroporación, la nucleofección y la lipofección (transferencia de genes mediada por liposomas). La transformación puede realizarse mediante el uso de cualquier vector de expresión. Por ejemplo, se contempla el uso de baculovirus para introducir ácido nucleico extraño en las células de los insectos. El término "transformación" también incluye procedimientos como la transformación de la línea germinal de insectos enteros mediada por elementos P. Además, la transformación se refiere a las células que se han transformado de forma natural, normalmente a través de una mutación genética.

Tal como se utiliza en el presente documento, "exógeno" significa que el gen que codifica la proteína no se expresa normalmente en la célula. Además, "exógeno" se refiere a un gen transfectado en una célula para aumentar el nivel normal (es decir, natural) de expresión de ese gen.

30 Una secuencia peptídica y una secuencia nucleotídica pueden ser "endógenas" o "heterólogas" (es decir, "extrañas"). El término "endógeno" se refiere a una secuencia que se encuentra de forma natural en la célula en la que se introduce, siempre que no contenga alguna modificación relativa a la secuencia natural. El término "heterólogo" se refiere a una secuencia que no es endógena a la célula en la que se introduce. Por ejemplo, el ADN heterólogo incluye una secuencia de nucleótidos que se liga a, o se manipula para que se ligue a, una secuencia de ácido nucleico a la que no está ligada en la naturaleza, o a la que está ligada en un lugar diferente en la naturaleza. El ADN heterólogo también incluye una secuencia de nucleótidos que se encuentra de forma natural en la célula en la que se introduce y que contiene alguna modificación con respecto a la secuencia natural. Generalmente, aunque no necesariamente, el ADN heterólogo codifica ARN heterólogo y proteínas heterólogas que no son producidas normalmente por la célula en la que se introduce. Algunos ejemplos de ADN heterólogo son los genes informadores, las secuencias reguladoras de la transcripción y la traducción, las secuencias de ADN que codifican proteínas marcadoras seleccionables (por ejemplo, proteínas que confieren resistencia a los fármacos), etc.

### Constructos

45 Las moléculas de ácido nucleico divulgadas en el presente documento (por ejemplo, nucleasas de sitio específico o ARN guía para CRISPR) pueden utilizarse en la producción de constructos de ácido nucleico recombinante. En una realización, las moléculas de ácido nucleico de la presente divulgación pueden utilizarse en la preparación de constructos de ácido nucleico, por ejemplo, casetes de expresión para la expresión en la planta de interés. Esta expresión puede ser transitoria, por ejemplo, cuando el constructo no se integra en el genoma del huésped o se mantiene bajo el control que ofrece el promotor y la posición del constructo dentro del genoma del huésped si se integra.

50 Los casetes de expresión pueden incluir secuencias reguladoras unidas de forma operable a las secuencias de nucleasas específicas del sitio o a las secuencias de ARN guía divulgadas en el presente documento. El casete puede contener, además, al menos un gen adicional para ser co-transformado en el organismo. Alternativamente, el/los gen(es) adicional(es) puede(n) proporcionarse en múltiples casetes de expresión.

55 Los constructos de ácido nucleico pueden estar provistos de una pluralidad de sitios de restricción para la inserción de la secuencia codificante de la nucleasa específica del sitio que estará bajo la regulación transcripcional de las regiones reguladoras. Los constructos de ácido nucleico pueden contener además moléculas de ácido nucleico que codifican para genes marcadores seleccionables.

Se puede utilizar cualquier promotor en la producción de los constructos de ácido nucleico. El promotor puede ser nativo o análogo, o ajeno o heterólogo, a las secuencias de ácido nucleico del hospedador de la planta divulgadas en el presente documento. Además, el promotor puede ser la secuencia natural o alternativamente una secuencia sintética. Cuando el promotor es "extraño" o "heterólogo" para la planta huésped, se pretende que el promotor no se encuentre en la planta nativa en la que se introduce el promotor. Tal y como se utiliza en el presente documento, un gen quimérico comprende una secuencia codificante unida operativamente a una región de iniciación de la transcripción que es heteróloga a la secuencia codificante.

Las secuencias de nucleasas dirigidas al sitio divulgadas en el presente documento pueden expresarse utilizando promotores heterólogos.

Se puede utilizar cualquier promotor en la preparación de constructos para controlar la expresión de las secuencias de nucleasas dirigidas al sitio, tales como promotores que proporcionen una expresión constitutiva, preferida por el tejido, inducible u otros promotores para la expresión en plantas. Los promotores constitutivos incluyen, por ejemplo, el promotor central del promotor Rsyn7 y otros promotores constitutivos divulgados en WO 99/43 838 y Patente de EE.UU. nº 6.072.050, el promotor central del CaMV 35S (Odell et al. Nature 313:810-812; 1985); la actina del arroz (McElroy et al., Plant Cell 2:163-171, 1990); ubiquitina (Christensen et al., Plant Mol. Biol. 12:619-632, 1989 Datta et al., Plant Sci. Biol. 18:675-689, 1992); pEMU (Last et al., Theor. Appl. Genet. 81:581-588, 1991); MAS (Velten et al., EMBO J. 3:2723-2730, 1984); promotor de ALS (Patente de EE.UU. nº 5.659.026), y otros similares. Otros promotores constitutivos incluyen, por ejemplo, la patente estadounidense nº 5.608.149, 5.608.144. 5.604,121, 5.569.597. 5.466.785. 5.399.680;. 5.268.463;. 5.608.142y. 6.177.611.

Los promotores preferidos por el tejido pueden utilizarse para dirigir la expresión de la nucleasa dirigida al sitio dentro de un tejido vegetal particular. Dichos promotores preferidos de tejidos incluyen, pero no se limitan a, promotores preferidos de hojas, promotores preferidos de raíces, promotores preferidos de semillas y promotores preferidos de tallos. Los promotores preferidos por los tejidos son Yamamoto et al., Plant J. 12(2):255-265, 1997; Kawamata et al., Plant Cell Physiol. 7:792-803, 1997 Datta et al., Plant Sci. Gen Genet. 254:337-343, 1997 Datta et al., Plant Sci. 6:157-168, 1997 Datta et al., Plant Sci. 1:1331-1341, 1996 Datta et al., Plant Sci. 1:525-535, 1996 Datta et al., Plant Sci. 112-2 513-524, 1996; Yamamoto et al., Plant Cell Physiol. 35(5):773-778, 1994 Lam, Resultados Probl. Cell Differ. 20:181-196, 1994 Datta et al., Plant Sci. Plant Mol Biol. 23(6):1129-1138, 1993 Matsuoka et al., Proc Nat'l. Acad. Sci. USA 90(20):9586- 9590, 1993 y Guevara-Garcia et al., Plant J. 4(3):495-505, 1993.

Los constructos de ácido nucleico también pueden incluir regiones de terminación de la transcripción. Cuando se utilizan regiones de terminación de la transcripción, puede utilizarse cualquier región de terminación en la preparación de los constructos de ácido nucleico. Por ejemplo, la región de terminación puede proceder de otra fuente (es decir, ajena o heteróloga al promotor). Ejemplos de regiones de terminación que están disponibles para su uso en los constructos de la presente divulgación incluyen las del plásmido Ti de *A. tumefaciens*, tal como las regiones de terminación de la octopina sintasa y la nopalina sintasa. Véase también Guerinneau et al., Mol. Gen. Genet. 262:141-144, 1991 Proudfoot, Cell 64:671-674, 1991 Sanfacon et al., Genes Dev. 5:141-149, 1991 Mogen et al., Plant Cell 2:1261-1272, 1990 Munroe et al., Gene 91:151-158, 1990 Ballas et al., Nucleic Acids Res. 17:7891-7903, 1989 y Joshi et al., Nucleic Acid Res. 15:9627-9639, 1987).

En conjunción con cualquiera de los aspectos, realizaciones, procedimientos y/o composiciones divulgados en el presente documento, los ácidos nucleicos pueden optimizarse para aumentar la expresión en la planta transformada. Es decir, los ácidos nucleicos que codifican las proteínas nucleasas dirigidas al sitio pueden sintetizarse utilizando codones preferidos por las plantas para mejorar la expresión. Véase, por ejemplo, Campbell y Gowri, (Plant Physiol. 92:1-11, 1990) para una discusión sobre el uso de codones preferidos por el huésped. Existen procedimientos en la técnica para sintetizar los genes preferidos por las plantas. Véase, por ejemplo, Patente de EE.UU. nº 5.380.831 y 5,436,391 y Murray et al., Nucleic Acids Res. 17:477-498, 1989).

Además, se pueden realizar otras modificaciones de la secuencia a las secuencias de ácido nucleico divulgadas en el presente documento. Por ejemplo, se sabe que las modificaciones adicionales de la secuencia mejoran la expresión de los genes en un huésped celular. Esto incluye la eliminación de secuencias que codifican señales de poliadenilación espurias, señales de sitios de empalme exón/intrón, repeticiones tipo transposón y otras secuencias bien caracterizadas que pueden ser perjudiciales para la expresión del gen. El contenido de G-C de la secuencia también puede ajustarse a los niveles medios de un huésped celular objetivo, calculados por referencia a genes conocidos expresados en la célula huésped. Además, la secuencia puede modificarse para evitar las estructuras secundarias en horquilla de ARNm predichas.

También pueden utilizarse otras secuencias de ácido nucleico en la preparación de los constructos de la presente divulgación, por ejemplo para mejorar la expresión de la secuencia codificadora de la nucleasa dirigida al sitio. Tales secuencias de ácido nucleico incluyen los intrones del gen Adh1, intrón del maíz (Callis et al., Genes and Development 1:1183-1200, 1987), y las secuencias líderes, (secuencia W) del virus del mosaico del tabaco (TMV), del virus del moteado clorótico del maíz y del virus del mosaico de la alfalfa (Gallie et al., Nucleic Acid Res. 15:8693-8711, 1987 Datta et al., Plant Sci. Biol. 15:65-79, 1990). Se ha demostrado que el primer intrón del locus shrunken-1 del maíz aumenta la expresión de los genes en los constructos de genes quiméricos. Las Pat. de EE.UU. Números 5.424.412 y 5.593.874 divulgan el uso de intrones específicos en constructos de expresión genética, y Gallie et al. Plant Physiol.

106:929-939, 1994) también han demostrado que los intrones son útiles para regular la expresión de los genes sobre una base específica de tejido. Para mejorar u optimizar aún más la expresión del gen de la nucleasa dirigida al sitio, los vectores de expresión de plantas divulgados en el presente documento pueden contener también secuencias de ADN que contengan regiones de unión a la matriz (MAR). Las células vegetales transformadas con dichos sistemas de expresión modificados, por tanto, pueden mostrar una sobreexpresión o una expresión constitutiva de una secuencia nucleotídica de la divulgación.

Los constructos de expresión divulgados en el presente documento también pueden incluir secuencias de ácido nucleico capaces de dirigir la expresión de la secuencia de nucleasa dirigida al sitio al cloroplasto. Tales secuencias de ácido nucleico incluyen secuencias de direccionamiento al cloroplasto que codifican un péptido de tránsito del cloroplasto para dirigir el producto génico de interés a los cloroplastos de las células vegetales. Estos péptidos de tránsito son conocidos en la técnica. Con respecto a las secuencias direccionadas al cloroplasto, "unidas operativamente" significa que la secuencia de ácido nucleico que codifica un péptido de tránsito (es decir, la secuencia direccionada al cloroplasto) está unida a las moléculas de ácido nucleico de la nucleasa dirigida al sitio divulgadas en el presente documento, de manera que las dos secuencias son contiguas y están en el mismo marco de lectura. Véase, por ejemplo, Von Heijne et al., *Plant Mol. Biol. Rep.* 9:104-126, 1991 Datta et al., *Plant Sci. Chem.* 264:17544-17550, 1989 Datta et al., *Plant Sci.* 84:965-968, 1987 Datta et al., *Plant Sci. Biophys. Res. Commun.* 196:1414-1421, 1993y Shah et al., *Science* 233:478-481, 1986.

Las secuencias direccionadas al cloroplasto son conocidas en la técnica e incluyen la subunidad pequeña del cloroplasto de la ribulosa-1,5-bisfosfato carboxilasa (Rubisco) (de Castro Silva Filho et al., *Plant Mol. Biol.* 30:769-780, 1996 Datta et al., *Plant Sci. Chem.* 266(5):3335-3342, 1991); 5-(enolpiruvil)shikimato-3-fosfato sintasa (EPSPS) (Archer et al., *J. Bioenerg. Biomemb.* 22(6):789-810, 1990); triptófano sintasa (Zhao et al., *J. Biol. Chem.* 270(11):6081-6087, 1995); plastocianina (Lawrence et al., *J. Biol. Chem.* 272(33):20357-20363, 1997); la corimato sintasa (Schmidt et al., *J. Biol. Chem.* 268(36):27447-27457, 1993); y la proteína de unión a la clorofila a/b que recoge la luz (LHBP) (Lamppa et al., *J. Biol. Chem.* 263:14996-14999, 1988). Véase también Von Heijne et al., *Plant Mol. Biol. Rep.* 9:104-126, 1991 Datta et al., *Plant Sci. Chem.* 264:17544-17550, 1989 Datta et al., *Plant Sci.* 84:965-968, 1987 Datta et al., *Plant Sci. Biophys. Res. Commun.* 196:1414-1421, 1993 y Shah et al., *Science* 233 :478-481, 1986.

En conjunción con cualquiera de los aspectos, realizaciones, procedimientos y/o composiciones divulgados en el presente documento, los constructos de ácido nucleico pueden prepararse para dirigir la expresión de la secuencia codificadora de la nucleasa dirigida al sitio mutante del cloroplasto de la célula vegetal. Los procedimientos de transformación de cloroplastos son conocidos en la técnica. Véase, por ejemplo, Svab et al., *Proc. Nacional. Acad. Sci. USA* 87:8526-8530, 1990 Svab y Maliga, *Proc. Nacional. Acad. Sci. USA* 90:913-917, 1993 Svab y Maliga, *EMBO J.* 12:601-606, 1993. El procedimiento se basa en el suministro con pistola de partículas de ADN que contienen un marcador seleccionable y el direccionamiento del ADN al genoma de los plástidos a través de la recombinación homóloga. Además, la transformación de plástidos puede llevarse a cabo mediante la transactivación de un transgén silencioso transportado por plástidos mediante la expresión preferente en el tejido de una ARN polimerasa codificada en el núcleo y dirigida por los plástidos. Un sistema de este tipo ha sido informado en McBride et al. *Proc. Nacional. Acad. Sci. USA* 91:7301-7305 (1994).

Los ácidos nucleicos de interés para ser direccionados al cloroplasto pueden ser optimizados para su expresión en el cloroplasto para tener en cuenta las diferencias en el uso de codones entre el núcleo de la planta y este orgánulo. De este modo, los ácidos nucleicos de interés pueden sintetizarse utilizando los codones preferidos por los cloroplastos. Véase, por ejemplo, Patente de EE.UU. n° 5.380.831,

Los constructos de ácido nucleico pueden utilizarse para transformar células vegetales y regenerar plantas transgénicas que comprendan las secuencias codificantes de nucleasas dirigidas al sitio. Existen numerosos vectores y procedimientos de transformación de plantas. Véase, por ejemplo, Patente de EE.UU. n° 6.753.458, An, G. et al., *Plant Physiol.* 81:301-305, 1986; Fry, J. et al., *Plant Cell Rep.* 6:321-325, 1987 Block, M., *Theor. Appl Genet.* 76:767-774, 1988 Hinchey et al., *Stadler. Genet. Symp.* 203212.203-212, 1990 Cousins et al., *Aust. J. Plant Physiol.* 18:481-494, 1991 Chee, P. P. y Slightom, J. L., *Gene.* 118:255-260, 1992 Christou et al., *Trends. Biotechnol.* 10:239-246, 1992 D'Halluin et al., *Bio/Technol.* 10:309-3, 1992 Dhir et al., *Plant Physiol.* 99:81-88, 1992 Casas et al., *Proc. Nat'l. Acad. Sci. USA* 90:11212-11216, 1993 Christou, P., *In Vitro Cell. Dev. Biol.-Plant* 29P:1 19-124, 1993 Davies, et al., *Plant Cell Rep.* 12:180-183, 1993 Dong, J. A. y Mc Hughen, A., *Plant Sci.* 91:139-148, 1993 Franklin, C. I. y Trieu, T. N., *Plant. Physiol.* 102:618-167, 1993 Golovkin et al., *Plant Sci.* 90:41-52, 1993 Guo Chin Sci. Bull. 38:2072-2078 Asano, et al., *Plant Cell Rep.* 13, 1994; Ayeres N. M. y Park, W. D., *Crit. Rev. Plant. Sci.* 13:219-239, 1994 Barcelo et al., *Plant.* 5:583-592, 1994 Becker, et al., *Plant.* 5:299-307, 1994 Borkowska et al., *Acta.. Physiol Plant.* 16:225-230, 1994 Christou, P., *Agro. Food. Ind. Hi Tech.* 5:17-27, 1994 Eapen et al., *Plant Cell Rep.* 13:582-586, 1994 Hartman et al., *Bio-Technology* 12:919923, 1994 Ritala et al., *Plant. Mol. Biol.* 24:317-325, 1994y Wan, Y. C. y Lemaux, P. G., *Plant Physiol.* 104, 3748, 1994]. Los constructos también pueden transformarse en células vegetales mediante recombinación homóloga.

El término "tipo salvaje" cuando se hace referencia a una secuencia peptídica y a una secuencia nucleotídica se refiere a una secuencia peptídica y a una secuencia nucleotídica (locus/género/alelo), respectivamente, que tiene las características de esa secuencia peptídica y de esa secuencia nucleotídica cuando se aísla de una fuente de origen natural. Una secuencia peptídica y una secuencia nucleotídica de tipo salvaje es la que se observa con mayor

frecuencia en una población y, por tanto, se designa arbitrariamente como la forma "normal" o "de tipo salvaje" de la secuencia peptídica y la secuencia nucleotídica, respectivamente. "Tipo salvaje" también puede referirse a la secuencia en una posición o posiciones específicas de nucleótidos, o la secuencia en una posición o posiciones particulares de codones, o la secuencia en una posición o posiciones particulares de aminoácidos.

- 5 La "secuencia de consenso" se define como una secuencia de aminoácidos o nucleótidos que contienen aminoácidos o nucleótidos idénticos o funcionalmente equivalentes en al menos el 25 % de la secuencia. No es necesario que los aminoácidos o nucleótidos idénticos o funcionalmente equivalentes sean contiguos.

10 El término "Brassica", tal como se utiliza en el presente documento, se refiere a las plantas del género Brassica. Las especies ejemplares de Brassica incluyen, pero no se limitan a, B. carinata, B. elongata, B. fruticulosa, B. juncea, B. napus, B. narinosa, B. nigra, B. oleracea, B. perviridis, B. rapa (syn B. campestris), B. rupestris, B. septiceps y B. tournefortii.

15 Una nucleobase es una base, que en ciertas realizaciones preferidas es una purina, pirimidina o un derivado o análogo de la misma. Los nucleósidos son nucleobases que contienen una fracción de pentosefuranosilo, por ejemplo, un ribósido opcionalmente sustituido o un 2'-desoxirribósido. Los nucleósidos pueden estar unidos por una de varias fracciones de enlace, que pueden contener o no un fósforo. Los nucleósidos unidos por enlaces fosfodiésteres no sustituidos se denominan nucleótidos. El término "nucleobase" como se usa en el presente documento incluye las nucleobases peptídicas, las subunidades de los ácidos nucleicos peptídicos y las nucleobases de morfolina, así como los nucleósidos y los nucleótidos.

20 Una oligonucleobase es un polímero que comprende nucleobases, preferentemente al menos una porción de la cual puede hibridarse por emparejamiento de bases Watson-Crick a un ADN que tenga la secuencia complementaria. [Una cadena de oligonucleobase puede tener un único terminal 5' y 3', que son las últimas nucleobases del polímero. Una cadena de oligonucleobase concreta puede contener nucleobases de todo tipo. Un compuesto de oligonucleobase es un compuesto que comprende una o más cadenas de oligonucleobase que pueden ser complementarias y se hibridan mediante el emparejamiento de bases Watson-Crick. Las nucleobases de tipo ribo son nucleobases que contienen pentosefuranosilo y en las que el carbono 2' es un metileno sustituido por un hidroxilo, un alquilo o un halógeno. Las nucleobases de tipo desoxirribo son nucleobases distintas de las de tipo ribo e incluyen todas las nucleobases que no contienen una fracción de pentosefuranosilo.

25 En ciertas realizaciones, una cadena de oligonucleobase puede incluir cadenas de oligonucleobase y segmentos o regiones de cadenas de oligonucleobase. Una cadena de oligonucleobase puede tener un extremo 3' y un extremo 5', y cuando una cadena de oligonucleobase es coextensiva con una cadena, los extremos 3' y 5' de la cadena también son los terminales 3' y 5' de la cadena.

El término "oligonucleobase de reparación de genes", tal como se utiliza en el presente documento, denota oligonucleobases, incluyendo oligonucleótidos dúplex mixtos, moléculas que no contienen nucleótidos, oligodeoxinucleótidos monocatenarios y otras moléculas de reparación de genes.

35 Tal y como se utiliza en el presente documento, el término "codón" se refiere a una secuencia de tres nucleótidos adyacentes (ya sea de ARN o ADN) que constituye el código genético que determina la inserción de un aminoácido específico en una cadena polipeptídica durante la síntesis de proteínas o la señal para detener la síntesis de proteínas. El término "codón" también se utiliza para referirse a las secuencias correspondientes (y complementarias) de tres nucleótidos en el ARN mensajero en el que se transcribe el ADN original.

40 Tal y como se utiliza en el presente documento, el término "homología" se refiere a la similitud de secuencia entre las proteínas y el ADN. El término "homología" u "homólogo" se refiere a un grado de identidad. Puede haber homología parcial o completa. Una secuencia parcialmente homóloga es aquella que tiene menos del 100 % de identidad de secuencia cuando se compara con otra secuencia.

45 "Heterocigoto" se refiere a tener diferentes alelos en uno o más loci genéticos en segmentos cromosómicos homólogos. Tal y como se utiliza en el presente documento, "heterocigoto" también puede referirse a una muestra, una célula, una población de células o un organismo en el que se pueden detectar diferentes alelos en uno o más loci genéticos. Las muestras heterocigotas también pueden determinarse mediante procedimientos conocidos en la técnica como, por ejemplo, la secuenciación de ácidos nucleicos. Por ejemplo, si un electroferograma de secuenciación muestra dos picos en un mismo locus y ambos picos son aproximadamente del mismo tamaño, la muestra puede caracterizarse como heterocigota. O bien, si un pico es más pequeño que otro, pero tiene al menos un 25 % del tamaño del pico más grande, la muestra puede caracterizarse como heterocigota. En algunas realizaciones, el pico más pequeño es al menos un 15 % del pico más grande. En otras realizaciones, el pico más pequeño es al menos un 10 % del pico más grande. En otras realizaciones, el pico más pequeño es al menos un 5 % del pico más grande. En otras realizaciones, se detecta una cantidad mínima del pico más pequeño.

55 Tal y como se utiliza en el presente documento, "homocigoto" se refiere a tener alelos idénticos en uno o más loci genéticos en segmentos cromosómicos homólogos. "Homocigoto" también puede referirse a una muestra, una célula, una población de células o un organismo en el que se pueden detectar los mismos alelos en uno o más loci genéticos. Las muestras homocigóticas pueden determinarse mediante procedimientos conocidos en la técnica, como, por

ejemplo, la secuenciación de ácidos nucleicos. Por ejemplo, si un electroferograma de secuenciación muestra un único pico en un locus concreto, la muestra puede denominarse "homocigótica" con respecto a ese locus.

El término "hemizigoto" se refiere a un gen o segmento de gen que está presente sólo una vez en el genotipo de una célula u organismo porque el segundo alelo está eliminado. Tal y como se utiliza en el presente documento, "hemizigoto" también puede referirse a una muestra, una célula, una población celular o un organismo en el que un alelo en uno o más loci genéticos puede detectarse sólo una vez en el genotipo.

El término "estado de cigosidad", tal y como se utiliza en este documento, se refiere a una muestra, una población celular o un organismo que aparece como heterocigoto, homocigoto o hemizigoto, tal y como se determina mediante procedimientos de prueba conocidos en la técnica y descritos en este documento. El término "estado de cigosidad de un ácido nucleico" significa determinar si la fuente de ácido nucleico parece heterocigota, homocigota o hemizigota. El "estado de cigosidad" puede referirse a diferencias en un solo nucleótido de una secuencia. En algunos procedimientos, el estado de cigosidad de una muestra con respecto a una única mutación puede clasificarse como homocigoto de tipo salvaje, heterocigoto (es decir, un alelo de tipo salvaje y un alelo mutante), homocigoto mutante o hemizigoto (es decir, una única copia del alelo de tipo salvaje o mutante).

Tal y como se utiliza en el presente documento, el término "**RTDS**" se refiere a The Rapid Trait Development System™ (RTDS) desarrollado por Cibus. El **RTDS** es un sistema de modificación de genes en sitios específicos que es eficaz para realizar cambios precisos en una secuencia genética sin la incorporación de genes extraños o secuencias de control.

El término "aproximadamente", tal como se utiliza en el presente documento, significa en términos cuantitativos más o menos el 10 %. Por ejemplo, "aproximadamente del 3 %" abarcaría el 2,7-3,3 % y "aproximadamente el 10 %" abarcaría el 9-11 %.

### Oligonucleótidos de reparación

La presente invención se refiere en general a procedimientos novedosos para mejorar la eficacia del direccionamiento de las modificaciones a lugares específicos en las secuencias genómicas o de otros nucleótidos. Además, esta invención se relaciona con el ADN diana que ha sido modificado, mutado o marcado por los enfoques divulgados en el presente documento. La invención también se refiere a células, tejidos y organismos que han sido modificados por los procedimientos de la invención. La presente invención se basa en el desarrollo de composiciones y procedimientos relacionados en parte con el exitoso sistema de conversión, el Rapid Trait Development System (**RTDS**™, Cibus US LLC).

**El RTDS** se basa en la alteración de un gen direccionado utilizando el propio sistema de reparación de genes de la célula para modificar específicamente la secuencia del gen in situ y no insertar ADN extraño y secuencias de control de la expresión génica. Este procedimiento produce un cambio preciso en la secuencia genética, mientras que el resto del genoma permanece inalterado. A diferencia de los GMO transgénicos convencionales, no hay integración de material genético extraño, ni queda material genético extraño en la planta. Los cambios en la secuencia genética introducidos por **el RTDS** no se insertan al azar. Dado que los genes afectados permanecen en su ubicación nativa, no se produce un patrón de expresión aleatorio, incontrolado o adverso.

El **RTDS** que efectúa este cambio es un oligonucleótido sintetizado químicamente que puede estar compuesto tanto de ADN como de bases de ARN modificadas, así como de otras fracciones químicas, y está diseñado para hibridarse en la ubicación del gen diana para crear un par o pares de bases no coincidentes. Este par de bases no coincidentes actúa como señal para atraer al sistema natural de reparación de genes de la propia célula hacia ese lugar y corregir (sustituir, insertar o eliminar) el/los nucleótidos designados dentro del gen. Una vez completado el proceso de corrección, la molécula **RTDS** se degrada y el gen ahora modificado o reparado se expresa bajo los mecanismos normales de control endógeno de ese gen.

Los procedimientos y composiciones divulgados en el presente documento pueden practicarse o hacerse con "oligonucleobases de reparación de genes" (GRON) que tengan las conformaciones y químicas descritas en detalle a continuación. Las "oligonucleobases reparadoras de genes" contempladas en el presente documento también se han descrito en la literatura científica y de patentes publicada con otros nombres, como "oligonucleobases recombinagénicas"; "oligonucleótidos quiméricos de ARN/ADN"; "oligonucleótidos quiméricos"; "oligonucleótidos dúplex mixtos" (MDON); "oligonucleótidos de ADN-ARN (RDOs)"; "oligonucleótidos direccionados a genes"; "genoplastos"; "oligonucleótidos modificados monocatenarios"; "vectores mutacionales oligodeoxinucleótidos monocatenarios" (SSOMVs); "vectores mutacionales dúplex"; y "vectores mutacionales heterodúplex." La oligonucleobase reparadora de genes puede ser introducida en una célula vegetal utilizando cualquier procedimiento comúnmente utilizado en la técnica, incluyendo pero no limitado a, microportadores (suministro biolístico), microfibras, absorción mediada por polietilenglicol (PEG), electroporación, y microinyección.

En una realización, la oligonucleobase reparadora de genes es un oligonucleótido dúplex mixto en la que los nucleótidos de tipo ARN del oligonucleótido dúplex mixto se hacen resistentes a la RNasa sustituyendo el 2'-hidroxilo por una funcionalidad de fluoro, cloro o bromo o colocando un sustituyente en el 2'-O. Los sustituyentes adecuados incluyen los sustituyentes enseñados por el Kmiec II. Los sustituyentes alternativos incluyen los sustituyentes

enseñados por la Pat. de EE.UU. N° 5.334.71 1 (Sproat) y los sustituyentes enseñados por las publicaciones de patentes EP 629 387 y EP 679 657 (colectivamente, las Solicitudes Martin), Tal como se utiliza en el presente documento, un derivado 2'-fluoro, cloro o bromo de un ribonucleótido o un ribonucleótido que tenga una T- OH sustituida con un sustituyente descrito en las Solicitudes Martin o Sproat se denomina "Ribonucleótido sustituido con T" Tal como se utiliza en el presente documento, el término "nucleótido tipo ARN" indica un nucleótido T'-hidroxilo o 2'-sustituido que está unido a otros nucleótidos de un oligonucleótido dúplex mixto mediante un enlace fosfodiéster no sustituido o cualquiera de los enlaces no naturales enseñados por Kmiec I o Kmiec II. Tal como se utiliza en el presente documento, el término "nucleótido de tipo desoxirribo" significa un nucleótido que tiene un T-H, que puede unirse a otros nucleótidos de una oligonucleobase reparadora de genes mediante un enlace fosfodiéster no sustituido o cualquiera de los enlaces no naturales enseñados por Kmiec I o Kmiec II.

En una realización de la presente invención, la oligonucleobase reparadora de genes es un oligonucleótido dúplex mixto (MDON) que está unido únicamente por enlaces fosfodiéster no sustituidos. En realizaciones alternativas, el enlace es por fosfodiésteres sustituidos, derivados de fosfodiésteres y enlaces no basados en fósforo como enseña Kmiec II. En otra realización, cada nucleótido de tipo ARN en el oligonucleótido dúplex mixto es un nucleótido 2'-sustituido. Las realizaciones particularmente preferidas de los ribonucleótidos 2'-sustituidos son los ribonucleótidos 2'-fluoro, T-metoxi, 2'-propiloxi, 2'-aliloxi, 2'-hidroxiletiloxi, 2'-metoxietiloxi, T-fluoropropiloxi y 2'-trifluoropropiloxi sustituidos. Las formas de realización más preferidas de los ribonucleótidos 2'-sustituidos son los nucleótidos 2'-fluoro, 2'-metoxi, 2'-metoxietiloxi y 2'-aliloxi sustituidos. En otra realización, el oligonucleótido dúplex mixto está unido por enlaces fosfodiéster no sustituidos.

Aunque los oligonucleótidos dúplex mixtos (MDONs) que tienen un solo tipo de nucleótido de tipo ARN 2'-sustituido se sintetizan más convenientemente, los procedimientos de la invención pueden practicarse con oligonucleótidos dúplex mixtos que tienen dos o más tipos de nucleótidos de tipo ARN. La función de un segmento de ARN puede no verse afectada por una interrupción causada por la introducción de un desoxinucleótido entre dos trinucleótidos de tipo ARN, en consecuencia, el término segmento de ARN abarca términos tales como "segmento de ARN interrumpido" Un segmento de ARN ininterrumpido se denomina segmento de ARN contiguo. En una realización alternativa, un segmento de ARN puede contener nucleótidos alternados resistentes a la RNasa y 2'-OH no sustituidos. Los oligonucleótidos dúplex mixtos tienen preferentemente menos de 100 nucleótidos y más preferentemente menos de 85 nucleótidos, pero más de 50 nucleótidos. La primera y la segunda cadena están emparejadas con bases Watson-Crick. En una realización, las cadenas del oligonucleótido dúplex mixto están unidas covalentemente por un enlazador, tal como un hexa, penta o tetranucleótido de una sola cadena, de modo que las cadenas primera y segunda son segmentos de una sola cadena de oligonucleótidos que tiene un solo extremo 3' y un solo extremo 5'. Los extremos 3' y 5' pueden protegerse mediante la adición de una "tapa de horquilla" por el que los nucleótidos terminales 3' y 5' se emparejan mediante Watson-Crick con nucleótidos adyacentes. Una segunda tapa de horquilla puede, además, colocarse en la unión entre las cadenas primera y segunda, distante de los extremos 3' y 5', de modo que se establezca el emparejamiento Watson-Crick entre las cadenas primera y segunda.

Las cadenas primera y segunda contienen dos regiones que son homólogas con dos fragmentos del gen diana, es decir, tienen la misma secuencia que el gen diana. Una región homóloga contiene los nucleótidos de un segmento de ARN y puede contener uno o más nucleótidos de tipo ADN de conexión del segmento de ADN y también puede contener nucleótidos de tipo ADN que no están dentro del segmento de ADN interviniente. Las dos regiones de homología están separadas por, y cada una es adyacente a, una región que tiene una secuencia que difiere de la secuencia del gen objetivo, denominada "región heteróloga" La región heteróloga puede contener uno, dos o tres nucleótidos no coincidentes. Los nucleótidos no coincidentes pueden ser contiguos o, alternativamente, pueden estar separados por uno o dos nucleótidos homólogos con el gen objetivo. Alternativamente, la región heteróloga también puede contener una inserción o uno, dos, tres o de cinco o menos nucleótidos. Alternativamente, la secuencia del oligonucleótido dúplex mixto puede diferir de la secuencia del gen objetivo sólo por la supresión de uno, dos, tres o cinco o menos nucleótidos del oligonucleótido dúplex mixto. En este caso, se considera que la longitud y la posición de la región heteróloga es la longitud de la delección, aunque ningún nucleótido del oligonucleótido dúplex mixto esté dentro de la región heteróloga. La distancia entre los fragmentos del gen diana que son complementarios a las dos regiones homólogas es idéntica a la longitud de la región heteróloga cuando se pretende una sustitución o sustituciones. Cuando la región heteróloga contiene una inserción, las regiones homólogas están así separadas en el oligonucleótido dúplex mixto más que sus fragmentos homólogos complementarios en el gen, y lo contrario es aplicable cuando la región heteróloga codifica una delección.

Los segmentos de ARN de los oligonucleótidos dúplex mixtos son cada uno una parte de una región homóloga, es decir, una región que es idéntica en secuencia a un fragmento del gen diana, cuyos segmentos juntos contienen preferentemente al menos 13 nucleótidos de tipo ARN y preferentemente de 16 a 25 nucleótidos de tipo ARN o aún más preferentemente 18-22 nucleótidos de tipo ARN o más preferentemente 20 nucleótidos. En una realización, los segmentos de ARN de las regiones homológicas están separados y adyacentes, es decir, "conectados por" un segmento de ADN interviniente. En una realización, cada nucleótido de la región heteróloga es un nucleótido del segmento de ADN interviniente. Un segmento de ADN interviniente que contiene la región heteróloga de un oligonucleótido dúplex mixto se denomina "segmento mutador"

En otra realización de la presente invención, la oligonucleobase reparadora de genes (GRON) es un vector mutacional de oligodeoxinucleótidos de una sola cadena o SSOMV, que se divulga en la Solicitud de Patente Internacional

PCT/US00/23457. Números 6.271.360, 6.479.292y 7.060.500. La secuencia del SSOMV se basa en los mismos principios que los vectores mutacionales descritos en las Pat. de EE.UU. Nos. 5.756.325; 5.871.984; 5.760.012; 5.888.983; 5.795.972; 5.780.296; 5.945,339; 6.004.804; y 6.010.907 y en las publicaciones internacionales Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; y WO 99/40789. La secuencia del SSOMV contiene dos regiones homólogas a la secuencia diana separadas por una región que contiene la alteración genética deseada denominada región mutadora. La región mutadora puede tener una secuencia de la misma longitud que la secuencia que separa las regiones homólogas en la secuencia diana, pero que tiene una secuencia diferente. Una región mutadora de este tipo provocará una sustitución. Alternativamente, las regiones homólogas en el SSOMV pueden ser contiguas entre sí, mientras que las regiones en el gen diana que tienen la misma secuencia están separadas por uno, dos o más nucleótidos. Un SSOMV de este tipo provoca una supresión en el gen diana de los nucleótidos que están ausentes en el SSOMV. Por último, la secuencia del gen diana que es idéntica a las regiones homólogas puede ser adyacente en el gen diana pero estar separada por uno, dos o más nucleótidos en la secuencia del SSOMV. Un SSOMV de este tipo provoca una inserción en la secuencia del gen diana.

Los nucleótidos del SSOMV son desoxirribonucleótidos que están unidos por enlaces fosfodiéster no modificados, excepto que el enlace internucleotídico terminal 3' y/o terminal 5' o alternativamente los dos enlaces internucleotídicos terminal 3' y/o terminal 5' pueden ser un fosforotioato o un fosfoamidato. Tal y como se utiliza en el presente documento, un enlace internucleotídico es el enlace entre los nucleótidos del SSOMV y no incluye el enlace entre el nucleótido del extremo 3' o el nucleótido del extremo 5' y un sustituyente de bloqueo. En una realización específica, la longitud del SSOMV está entre 21 y 55 desoxinucleótidos y las longitudes de las regiones de homología son, en consecuencia, una longitud total de al menos 20 desoxinucleótidos y al menos dos regiones de homología deben tener cada una longitudes de al menos 8 desoxinucleótidos.

El SSOMV puede ser diseñado para ser complementario a la cadena codificante o no codificante del gen diana. Cuando la mutación deseada es una sustitución de una sola base, se prefiere que tanto el nucleótido mutador como el nucleótido diana sean una pirimidina. En la medida en que sea coherente con la consecución del resultado funcional deseado, se prefiere que tanto el nucleótido mutador como el nucleótido diana de la cadena complementaria sean pirimidinas. Se prefieren particularmente los SSOMV que codifican mutaciones de transversión, es decir, un nucleótido mutador C o T es no coincidente, respectivamente, con un nucleótido C o T en la cadena complementaria.

### Mejorar la eficiencia

La presente invención describe una serie de enfoques para aumentar la eficacia de la conversión de un gen diana utilizando oligonucleótidos de reparación, y que pueden utilizarse solos o en combinación con otros. Entre ellas se encuentran:

1. Introducir modificaciones en los oligonucleótidos de reparación que atraigan a la maquinaria de reparación del ADN al sitio diana (no coincidente).

A. La introducción de uno o más sitios abásicos en el oligonucleótido (por ejemplo, dentro de las 10 bases, y más preferentemente con 5 bases del sitio no coincidente deseado) genera una lesión que es un intermedio en la reparación por escisión de bases (BER), y que atrae a la maquinaria de BER a la vecindad del sitio direccionado a la conversión por el oligonucleótido de reparación. Los oligonucleótidos modificados con dSpacer (furano abásico) pueden prepararse como se describe en, por ejemplo, Takeshita et al., J. Biol. Chem, 262:10171-79, 1987).

La inclusión de compuestos que inducen roturas de cadena simple o doble, ya sea en el oligonucleótido o junto con el oligonucleótido, genera una lesión que se repara mediante la unión de extremos no homólogos (NHEJ), la unión de extremos mediada por microhomología (MMEJ) y la recombinación homóloga. A modo de ejemplo, la familia de antibióticos de la bleomicina, las nucleasas de dedo de zinc, la FokI (o cualquier clase de enzima de restricción de tipo IIS) y otras nucleasas pueden acoplarse covalentemente al extremo 3' o 5' de los oligonucleótidos de reparación, con el fin de introducir roturas de cadena simple o doble en las proximidades del sitio que se pretende convertir con el oligonucleótido de reparación. Los antibióticos de la familia de la bleomicina son glicopéptidos que escinden el ADN, incluyen la bleomicina, la zeocina, la fleomicina, la tallysomicina y la pepleomicina, entre otros.

C. La introducción de uno o más 8'oxo dA o dG incorporados en el oligonucleótido (por ejemplo, dentro de 10 bases, y más preferentemente con 5 bases del sitio de desajuste deseado) genera una lesión que es similar a las lesiones creadas por las especies reactivas del oxígeno. Estas lesiones inducen el llamado sistema de "reparación por empuje". Véase, por ejemplo Kim et al., J. Biochem. Mol. Biol. 37:657-62, 2004).

2. Aumentar la estabilidad de los oligonucleótidos de reparación:

Introducción de una base inversa (idC) en el extremo 3' del oligonucleótido para crear un extremo 3' bloqueado en el oligonucleótido de reparación.

Introducción de uno o más nucleótidos 2'O-metilos o bases que aumentan la energía de hibridación (véase, por ejemplo, WO2007/073149) en el 5' y/o 3' del oligonucleótido de reparación.

5 Introducción de una pluralidad de nucleótidos de ARN 2'O-metilo en el extremo 5' del oligonucleótido de reparación, que conducen a bases de ADN que proporcionan el sitio no coincidente deseado, creando así una estructura de ácido nucleico similar a la del fragmento Okazaki.

Tintes intercalantes conjugados (5' o 3') como la acridina, el psoraleno, el bromuro de etidio y las tinciones de Syber.

Introducción de una tapa en el extremo 5', como una abrazadera T/A, una fracción de colesterol, SIMA (HEX), riboC y amidita.

10 Modificaciones de la columna vertebral como fosfotioato, 2'-O metil, fosfonatos de metilo, ácido nucleico bloqueado (LNA), MOE (metoxietil), di PS y ácido nucleico peptídico (PNA).

Reticulación del oligonucleótido de reparación, por ejemplo, con agentes reactivos de reticulación intrastrada como el cisplatino y la mitomicina C.

Conjugación con tintes fluorescentes como Cy3, DY547, Cy3.5, Cy3B, Cy5 y DY647.

15 3. Aumentar la energía de hibridación del oligonucleótido de reparación mediante la incorporación de bases que aumentan la energía de hibridación (véase, por ejemplo, WO2007/073149).

4. Aumentar la calidad de la síntesis de oligonucleótidos de reparación utilizando multímeros de nucleótidos (dímeros, trímeros, tetrámeros, etc.) como bloques de construcción para la síntesis. Esto da lugar a menos pasos de acoplamiento y a una separación más fácil de los productos de longitud completa de los bloques de construcción.

20 5. Uso de oligonucleótidos de reparación largos (es decir, de más de 55 nucleótidos de longitud, preferentemente entre 75 y 300 nucleótidos de longitud, más preferentemente de al menos 100 nucleótidos de longitud, aún más preferentemente de al menos 150 nucleótidos de longitud, y más preferentemente de al menos 200 nucleótidos de longitud), preferentemente con dos o más mutaciones dirigidas en el oligonucleótido de reparación.

25 En la siguiente tabla se ofrecen ejemplos de los enfoques anteriores

**Tabla 1.** Productos químicos de GRON que se van a probar

	Tipo de oligo	Modificaciones
5' mods	Abrazadera T/A	Abrazadera T/A
Modificaciones de la columna vertebral	Fosfotioato	PS
Tintes intercalantes	5, 777, 3].	idC Acridina, idC
Fragmentos de Okasaki		ADN/ARN
Sustituciones de Cy3		DY547
Facilitadores	2'OMe oligos diseñados 5' y 3' del oligo convertidor	2'OMe
Abásico	Sitio abásico colocado en varios lugares a 5' y 3' de la base de conversión. 44 mer	Abásico 2
Asistencia	Enfoque de asistencia	Cy3, idC en uno, ninguno en el otro:
	Se solapan: 2 oligos: 1 con Cy3/idC, 1 oligo de reparación sin modificar	
Asistencia	Enfoque de asistencia No hay solapamiento: 2 oligos: 1 con Cy3/idC, 1 oligo de reparación sin modificar	sólo hacen el oligo no modificado
Abásico	Sitio de THF colocado en varios lugares a 5' y 3' de la base de conversión. 44 mer	Tetrahidrofurano (dspacer)

(continuación)

	<b>Tipo de oligo</b>	<b>Modificaciones</b>
Modificaciones de la columna vertebral	9:	2'OMe
Trímeros		Trímero amidita, Cy3. idC
Reparación por empuje		8'oxo dA, 5' Cy3, idC
Reparación por empuje		8'oxo dA, 5' Cy3, idC
Ruptura del doble hilo		Bleomicina
Reticulador		Cisplatino
Reticulador		Mitomicina C
Facilitadores	superbases 5' y 3' del oligo convertidor	2 amino dA y 2- tio T
Super oligos		2'amino d, 5' Cy3, idC
Super oligos		2-thio T, 5' Cy3, idC
Super oligos		7-deaza A, 5' Cy3, idC
Super oligos		7-deaza G,5' Cy3, idC
Super oligos		propanil dC, 5' Cy3, idC
Tintes intercalantes	5' Psoralen/3' idC	Psoralen, idC
Tintes intercalantes	5' Bromuro de etidio	
Tintes intercalantes	5' Manchas de Syber	
5' mods	5' Chol/3' idC	Colesterol
Doble mutación	Oligo largo (100 bases) con 2 mutaciones	Unlmown
5' mods	5' SIMA HEX/3'idC	SIMA HEX, idC
Modificaciones de la columna vertebral	9:	Fosfonatos de metilo
Modificaciones de la columna vertebral		LNA
Modificaciones de la columna vertebral	1.	MOE (metoxietil)
Sustituciones de Cy3		Cy3. 5:
Sustituciones de Cy3		Cy5
Modificaciones de la columna vertebral		di PS
5' mods		riboC para ramificación nm
Modificaciones de la columna vertebral		PNA
Sustituciones de Cy3		DY647
5' mods	ramificación de 5'	ramificación simétrica amidita/idC

5 Las modificaciones anteriores también pueden incluir modificaciones conocidas de los nucleótidos, como la metilación, los tintes intercalantes 5', las modificaciones de los extremos 5' y 3', las modificaciones de la columna vertebral, los reticulantes, la ciclización y las "tapas" y la sustitución de uno o más de los nucleótidos naturales por un análogo como la inosina. Las modificaciones de los nucleótidos incluyen la adición de acridina, amina, biotina, azul de cascada, colesterol, Cy3@, Cy5@, Cy5.5@ Daboyl, digoxigenina, dinitrofenilo, Edans, 6-FAM, fluoresceína, 3'- glicerilo, HEX, IRD-700, IRD-800, JOE, fosfato psoraleno, rodamina, ROX, tiol (SH) espaciadores, TAMRA, TET, AMCA-S", SE, BODIPY°, Marina Blue@, Pacific Blue@, Oregon Green@, Rhodamine Green@, Rhodamine Red@, Rhodol Green@ y Texas Red@. Las modificaciones de la columna vertebral de los polinucleótidos incluyen el metilfosfonato, el ARN 2'-OMe-metilfosfonato, el fosfortiorato, el ARN 2'-OMe. Las modificaciones de las bases incluyen 2-amino-dA, 2-aminopurina, 3'- (ddA), 3'dA (cordicepina), 7-deaza-dA, 8-Br-dA, 8- oxo-dA, N6-Me-dA, sitio abásico (dSpacer), biotina

10

dT, 2'-OMe-5Me-C, 2'-OMe-propinil-C, 3'-(5-Me-dC), 3'-(ddC), 5-Br-dC, 5-1-duc, 5-Me-dC, 5-F-dC, carboxi-dT, dA convertible, dC convertible, dG convertible, dT convertible, dU convertible, 7-deaza-dG, 8-Br-dG, 8-oxo-dG, O6-Me-dG, S6-DNP-dG, 4-metilindol, 5-nitroindol, 2'-OMe-inosina, 2'-dl, o6-fenil-dl, 4-metilindol, 2'-desoxinebularina, 5-nitroindol, 2-aminopurina, dP (análogo de la purina), dK (análogo de la pirimidina), 3-nitropirrol, 2-tio-dT, 4-tio-dT, biotina-dT, carboxi-dT, O4-Me-dT, O4-triazol dT, 2'-OMe-propinil-U, 5-Br-dU, 2'-dU, 5-F-dU, 5-I-dU, O4-triazol dU. Dichos términos también abarcan los ácidos nucleicos peptídicos (PNA), un análogo del ADN en el que la columna vertebral es un pseudopéptido formado por unidades de N-(2-aminoetil)-glicina en lugar de un azúcar. Los PNAs imitan el comportamiento del ADN y se unen a cadenas de ácido nucleico complementarias. La columna vertebral neutra del PNA da lugar a una unión más fuerte y a una mayor especificidad de la que se consigue normalmente. Además, las propiedades químicas, físicas y biológicas únicas del PNA se han explotado para producir potentes herramientas biomoleculares, agentes antisentido y antigénicos, sondas moleculares y biosensores.

Las oligonucleobases pueden tener mellas, brechas, nucleótidos modificados, como es el caso de las columnas vertebrales de oligonucleótidos modificados, nucleótidos abásicos u otras fracciones químicas. En otra realización, al menos una cadena de la oligonucleobase incluye al menos un nucleótido modificado adicional, por ejemplo un nucleótido modificado con 2'-O-metilo, como un MOE (metoxietilo), un nucleótido con un grupo 5'-fosforotioato, un nucleótido terminal unido a un derivado del colesterol, un nucleótido modificado con 2'-desoxi-2'-fluro, un nucleótido modificado con 2'-desoxi, un nucleótido bloqueado, un nucleótido abásico (la nucleobase no tiene o tiene un grupo hidroxilo en su lugar (véase, por ejemplo, Glen Research, <http://www.glenresearch.com/GlenReports/GR21-14.html>)), un nucleótido modificado con 2'-amino, un nucleótido modificado con 2'-alquilo, un nucleótido morfolino, un fosforamidito y un nucleótido con base no natural. También se incluyen diversas sales, sales mixtas y formas de ácidos libres.

Las columnas vertebrales de oligonucleótidos modificados preferidos incluyen, por ejemplo, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriésteres, aminoalquilfosfotriésteres, fosfonatos de metilo y otros alquilo incluyendo fosfonatos de 3'-alquilo, fosfonatos de 5'-alquilo y fosfonatos quirales, fosfinatos fosforamidatos, incluidos el 3'-aminofosforamidato y los aminoalquilfosforamidatos, los tionofosforamidatos, los tionoalquilfosfonatos, los tionoalquilfosfotresteres, los selenofosfatos y los boranofosfatos con enlaces normales 3'-5', los análogos con enlaces 2'-5' de los mismos, y los que tienen polaridad invertida en los que uno o más enlaces internucleotídicos es un enlace 3' a 3', 5' a 5' o 2' a 2'. Los oligonucleótidos preferidos que tienen polaridad invertida comprenden un único enlace 3' a 3' en el enlace internucleotídico más 3', es decir, un único residuo de nucleósido invertido que puede ser abásico (la nucleobase falta o tiene un grupo hidroxilo en su lugar). El uso más común de una inversión de enlace es añadir un enlace 3'-3' al final de un oligonucleótido antisentido con un esqueleto de fosforotioato. El enlace 3'-3' estabiliza aún más el oligonucleótido antisentido frente a la degradación por exonucleasas al crear un oligonucleótido con dos extremos 5'-OH y ningún extremo 3'-OH. Las inversiones de enlace pueden introducirse en lugares específicos durante la síntesis de oligonucleótidos mediante el uso de "fosforamiditos invertidos". Estos reactivos tienen los grupos fosforamidita en la posición 5'-OH y el grupo protector dimetoxitritilo (DMT) en la posición 3'-OH. Normalmente, el grupo protector del DMT está en el 5'-OH y el fosforamidito en el 3'-OH.

Ejemplos de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 2'-aminobutiril pireno-uridina, 2'-aminouridina, 2'-desoxiuridina, 2'-fluro-citidina, 2'-fluro-uridina, 2,6-diaminopurina, 4-tio-uridina, 5-bromo-uridina, 5-fluro-citidina, 5-fluorouridina, 5-indo-uridina, 5-metil-citidina, inosina, N3-metil-uridina, 7-deaza-guanina, 8-aminohexil-amino-adenina, 6-tioguanina, 4-tiomina, 2-tiomina, 5-yodo-uridina, 5-yodo-citidina, 8-bromo-guanina, 8-bromo-adenina, 7-deaza-adenina, 7-diaza-guanina, 8-oxo-guanina, 5,6-dihidro-uridina y 5-hidroximetil-uridina. Estas unidades sintéticas están disponibles en el mercado (por ejemplo, compradas a Glen Research Company) y pueden incorporarse al ADN mediante síntesis química.

Ejemplos de modificación de la fracción de azúcar son la 3'-desoxilación, la 2'-fluoración y la arabanosidación, sin embargo, no debe interpretarse como algo limitado. La incorporación de estos en el ADN también es posible mediante síntesis química.

Ejemplos de la modificación del extremo 5' son la 5'-aminación, la 5'-biotinilación, la 5'-fluoresceinilación, la 5'-tetrafluro-fluoreceinilación, la 5'-tionación y la 5'-dabsilación, sin embargo no debe interpretarse como limitada a ello.

Ejemplos de la modificación del extremo 3' son la 3'-aminación, la 3'-biotinilación, la 2,3-dideoxidación, la 3'-tionación, la 3'-dabsilación, la 3'-carboxilación y la 3'-colesterilación, sin embargo, no debe interpretarse como algo limitado a ello.

En una realización preferida, la oligonucleobase puede contener un sustituyente de bloqueo en 5' que está unido a los carbonos terminales en 5' a través de un enlazador. La química del enlazador no es crítica, salvo su longitud, que debe ser preferentemente de al menos 6 átomos y que el enlazador debe ser flexible. Se puede utilizar una variedad de sustituyentes no tóxicos tales como la biotina, el colesterol u otros esteroides o un tinte fluorescente catiónico no intercalante. Reactivos particularmente preferidos para hacer oligonucleobases son los reactivos vendidos como Cy3™ y Cy5™ por Glen Research, Sterling VA (ahora GE Healthcare), que son fosfoamiditas bloqueadas que al incorporarse a un oligonucleótido dan lugar a los tintes 3,3',3'-tetrametil N,N'-isopropil sustituidos indomonocarbocianina e indodicarbocianina, respectivamente. Se prefiere especialmente el Cy3. Cuando la indocarbocianina está sustituida por N-oxialquilo, puede unirse convenientemente al terminal 5' del

oligodeoxinucleótido a través de un fosfodiéster con un fosfato terminal 5'. Cuando se utiliza la fosforamidita Cy3 disponible en el mercado, la modificación 5' resultante consiste en un sustituto de bloqueo y un enlazador juntos que son un N-hidroxiopropilo, N'-fosfatidilpropilo 3,3,3',3'-tetrametil indomonocarbocianina. Otros tintes contemplados son la Rodamina6G, la Tetrametilrhodamina, la Sulforhodamina 101, la Merocianina 540, el Atto565, el Atto550 26, el Cy3.5, el Dy547, el Dy548, el Dy549, el Dy554, el Dy555, el Dy556, el Dy560, el mStrawberry y el mCherry.

En una realización preferida, el tinte de indocarbocianina está tetra sustituido en las posiciones 3 y 3' de los anillos de indol. Sin limitación en cuanto a la teoría, estas sustituciones impiden que el tinte sea un tinte intercalante. La identidad de los sustituyentes en estas posiciones no es crítica.

Los diseños de oligos descritos en el presente documento también podrían utilizarse como plantillas donantes más eficientes en combinación con otras tecnologías de edición o recombinación del ADN, incluyendo, pero sin limitarse a ello, la focalización de genes mediante recombinación homóloga específica de sitio por nucleasas de dedos de zinc, nucleasas efectoras similares a activadores de la transcripción (TALEN) o Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas con Regularidad (CRISPR).

La presente invención se refiere en general a procedimientos para la modificación eficiente del ADN genómico celular y/o la recombinación de ADN en el ADN genómico de las células. Aunque no se limitan a un uso particular, los procedimientos de la presente invención son útiles, por ejemplo, para introducir una modificación en el genoma de una célula con el fin de determinar el efecto de la modificación en la misma. Por ejemplo, se puede introducir una modificación en la secuencia de nucleótidos que codifica una enzima para determinar si la modificación altera la actividad enzimática de la enzima, y/o determinar la ubicación de la región catalítica de la enzima. Alternativamente, la modificación puede introducirse en la secuencia de codificación de una proteína de unión al ADN para determinar si la actividad de unión al ADN de la proteína está alterada, y así delinear la región particular de unión al ADN dentro de la proteína. Otra alternativa es introducir una modificación en una secuencia reguladora no codificante (por ejemplo, un promotor, un potenciador, una secuencia de ARN reguladora (miARN), etc.) para determinar el efecto de la modificación en el nivel de expresión de una segunda secuencia que está vinculada de forma operativa a la secuencia reguladora no codificante. Esto puede ser deseable para, por ejemplo, definir la secuencia particular que posee actividad reguladora.

De acuerdo con la presente invención, la estrategia para producir una disrupción génica dirigida es a través de la generación de roturas de ADN de una sola cadena causadas por nucleasas específicas del sitio. Las endonucleasas se utilizan con mayor frecuencia para la interrupción selectiva de genes direccionados en organismos que tradicionalmente han sido refractarios a los procedimientos de direccionamiento de genes más convencionales, tal como las algas, las plantas y los modelos de animales grandes, incluyendo los humanos. Por ejemplo, actualmente se están realizando ensayos clínicos en humanos que involucran nucleasas de dedos de zinc para el tratamiento y la prevención de la infección por el VIH. Además, la ingeniería de endonucleasas se utiliza actualmente para intentar interrumpir los genes que producen fenotipos indeseables en los cultivos.

Las endonucleasas de anidamiento, también conocidas como meganucleasas, son endonucleasas específicas de secuencia que generan roturas de doble cadena en el ADN genómico con un alto grado de especificidad debido a sus grandes (por ejemplo, >14 pb) sitios de escisión. Si bien la especificidad de las endonucleasas de anidamiento para sus sitios diana permite direccionar con precisión las roturas de ADN inducidas, los sitios de escisión de las endonucleasas de anidamiento son raros y la probabilidad de encontrar un sitio de escisión de origen natural en un gen direccionado es baja.

Una clase de endonucleasas artificiales son las endonucleasas de dedo de zinc. Las endonucleasas de dedo de zinc combinan un dominio de escisión inespecífico, normalmente el de la endonucleasa FokI, con dominios de proteína de dedo de zinc diseñados para unirse a secuencias específicas de ADN. La estructura modular de las endonucleasas de dedos de zinc las convierte en una plataforma versátil para suministrar roturas de doble cadena específicas en el genoma. Una de las limitaciones de las endonucleasas de dedo de zinc es que la baja especificidad para un sitio diana o la presencia de múltiples sitios diana en un genoma puede dar lugar a eventos de escisión fuera del objetivo. Dado que la endonucleasa FokI se escinde como un dímero, una de las estrategias para evitar eventos de escisión fuera del objetivo ha sido diseñar dominios de dedos de zinc que se unen en sitios adyacentes de 9 pares de bases.

Las TALEN son nucleasas direccionables que se utilizan para inducir roturas de una o dos cadenas en sitios específicos del ADN, que luego son reparadas por mecanismos que pueden ser explotados para crear alteraciones de la secuencia en el sitio de escisión.

El bloque de construcción fundamental que se utiliza para diseñar la región de unión al ADN de las TALEN es un dominio de repetición altamente conservado derivado de los TALE de origen natural codificados por las proteobacterias *Xanthomonas* spp. La unión al ADN por parte de un TALEN está mediada por conjuntos de repeticiones de 33-35 aminoácidos altamente conservados que están flanqueados por dominios adicionales derivados de TALE en los extremos amino-terminal y carboxi-terminal de las repeticiones.

Estas repeticiones TALE se unen específicamente a una sola base de ADN, cuya identidad está determinada por dos residuos hipervariables que se encuentran típicamente en las posiciones 12 y 13 de la repetición, con el número de

repeticiones en un conjunto correspondiente a la longitud del ácido nucleico objetivo deseado, la identidad de la repetición seleccionada para que coincida con la secuencia del ácido nucleico objetivo. El ácido nucleico objetivo está preferentemente entre 15 y 20 pares de bases para maximizar la selectividad del sitio diana. La escisión del ácido nucleico diana suele producirse en un plazo de 50 pares de bases desde la unión de TALEN. Se han descrito en la técnica programas informáticos para el diseño del sitio de reconocimiento de TALEN. Véase, por ejemplo Cermak et al., *Nucleic Acids Res.* 2011 Julio; 39(12): e82.

Una vez diseñadas para que coincidan con la secuencia diana deseada, las TALENS pueden ser expresadas recombinantemente e introducidas en protoplastos como proteínas exógenas, o expresadas desde un plásmido dentro del protoplasto.

Otra clase de endonucleasas artificiales son las meganucleasas de ingeniería. Las endonucleasas de anidamiento de ingeniería se generan modificando la especificidad de las endonucleasas de anidamiento existentes. En uno de los enfoques, se introducen variaciones en la secuencia de aminoácidos de las endonucleasas de anidamiento de origen natural y, a continuación, se examinan las endonucleasas de anidamiento diseñadas resultantes para seleccionar las proteínas funcionales que escinden un sitio de unión específico. En otro enfoque, las endonucleasas de anidamiento químicas se diseñan combinando los sitios de reconocimiento de dos endonucleasas de anidamiento diferentes para crear un nuevo sitio de reconocimiento compuesto por un medio sitio de cada endonucleasa de anidamiento.

En la recombinación génica direccionada pueden utilizarse otras moléculas modificadoras del ADN. Por ejemplo, los ácidos nucleicos peptídicos pueden utilizarse para inducir modificaciones en el genoma de la célula o células diana (véase, por ejemplo, U.S. Pat. N° 5.986.053, a Ecker). En resumen, se utilizan nucleótidos sintéticos que comprenden, al menos, una columna vertebral péptida parcial para direccionar una secuencia de nucleótidos genómica homóloga. Al unirse al ADN de doble hélice, o a través de un mutágeno ligado al ácido nucleico peptídico, se induce la modificación de la secuencia de ADN diana y/o la recombinación. La especificidad de direccionamiento está determinada por el grado de homología de la secuencia entre la secuencia de direccionamiento y la secuencia genómica.

Además, la presente invención no se limita a los procedimientos particulares que se utilizan en el presente documento para ejecutar la modificación de secuencias genómicas. De hecho, se contemplan varios procedimientos. Por ejemplo, los genes pueden direccionarse utilizando oligonucleótidos formadores de triple hélice (TFO). Los TFOs pueden generarse sintéticamente, por ejemplo, mediante PCR o mediante el uso de un aparato sintetizador de genes. Además, los TFO pueden aislarse del ADN genómico si se encuentran secuencias naturales adecuadas. Los TFOs pueden utilizarse de varias maneras, incluyendo, por ejemplo, mediante la unión a un mutágeno tal como, pero no limitado a, psoraleno o clorambucil (véase, por ejemplo, Havre et al., *Proc Nat'l Acad Sci, U.S.A.* 90:7879-7883, 1993 Havre et al., *J Virol* 67:7323-7331, 1993 Wang et al., *Mol Cell Biol* 15:1759-1768, 1995 Takasugi et al., *Proc Nat'l Acad Sci, U.S.A.* 88:5602-5606, 1991 Belousov et al., *Nucleic Acids Res* 25:3440-3444, 1997). Además, por ejemplo, los TFO pueden estar enlazados al ADN dúplex donante (véase, por ejemplo Chan et al., *J Biol Chem* 272:11541-11548, 1999). Los TFO también pueden actuar uniéndose con suficiente afinidad para provocar una reparación propensa a errores (Wang et al., *Science* 271:802-805, 1996).

Los procedimientos divulgados en el presente documento no se limitan a la naturaleza o al tipo de reactivo modificador del ADN que se utiliza. Por ejemplo, estos reactivos modificadores del ADN liberan radicales que provocan la rotura de la cadena de ADN. Alternativamente, los reactivos alquilan el ADN para formar aductos que bloquearían la replicación y la transcripción. En otra alternativa, los reactivos generan enlaces cruzados o moléculas que inhiben las enzimas celulares que conducen a la rotura de la cadena. Los ejemplos de reactivos modificadores del ADN que se han unido a los oligonucleótidos para formar TFO incluyen, pero no se limitan a, los indolocarbazoles, la diimida de naftalina (NDI), el transplatino, la bleomicina, los análogos del ciclopropapirrolindol y las fenantodihidrodioxinas. En particular, los indolocarbazoles son inhibidores de la topoisomerasa I. La inhibición de estas enzimas da lugar a roturas de la cadena y a la formación de aductos proteicos del ADN [Arimondo et al., *Bioorganic and Medicinal Chem.* 8, 777, 2000]. NDI es un fotooxidante que puede oxidar las guaninas, lo que podría causar mutaciones en sitios de residuos de guanina [Nunez, et al., *Biochemistry*, 39, 6190, 2000]. Se ha demostrado que el transplatino reacciona con el ADN en un objetivo triplex cuando el TFO está unido al reactivo. Esta reacción provoca la formación de aductos de ADN que serían mutagénicos [Columbier, et al., *Nucleic Acids Research*, 24: 4519, 1996]. La bleomicina es un rompedor del ADN, ampliamente utilizado como mimético de la radiación. Se ha unido a oligonucleótidos y se ha demostrado que es activo como rompedor en ese formato [Sergeyev, *Nucleic Acids Research* 23, 4400, 1995; Kane, et al., *Biochemistry*, 34, 16715, 1995]. Los análogos del ciclopropapirrolindol se han unido a los TFO y han demostrado que alquilan el ADN en una secuencia diana triplex. El ADN alquilado contendría entonces aductos químicos que serían mutagénicos [Lukhtanov, et al., *Nucleic Acids Research*, 25, 5077, 1997]. Las fenantodihidrodioxinas son quinonas enmascaradas que liberan especies radicales tras la fotoactivación. Se han enlazado a los TFO y se ha demostrado que introducen roturas en el ADN dúplex en la fotoactivación [Bendinskas et al., *Bioconjugate Chem.* 9, 555, 1998].

La presente invención contempla otros procedimientos para inducir modificaciones y/o recombinación. Por ejemplo, otra realización implica la inducción de la recombinación homóloga entre un fragmento de ADN exógeno y el gen diana (véase por ejemplo Capecchi et al., *Science* 244:1288-1292, 1989) o mediante el uso de ácidos nucleicos peptídicos (PNA) con afinidad por el sitio direccionado. Otros procedimientos incluyen el reconocimiento del ADN específico de la secuencia y la orientación mediante poliamidas (véase, por ejemplo, Dervan et al., *Curr Opin Chem*

Biol 3:688-693, 1999 Biochemistry 38:2143-2151, 1999) y el uso de nucleasas con actividad específica (por ejemplo, proteínas de dedo de zinc, TALEN, meganucleasas y/o CRISPR).

La presente invención no se limita a ninguna frecuencia particular de modificación y/o recombinación. Los procedimientos de la invención dan como resultado una frecuencia de modificación en la secuencia de nucleótidos diana de entre el 0,2 % y el 3 %. No obstante, cualquier frecuencia (es decir, entre el 0 % y el 100 %) de modificación y/o recombinación se contempla dentro del ámbito de la presente invención. La frecuencia de la modificación y/o recombinación depende del procedimiento utilizado para inducir la modificación y/o recombinación, el tipo de célula utilizada, el gen específico al que se dirige y el reactivo de mutación del ADN utilizado, si lo hay. Además, el procedimiento utilizado para detectar la modificación y/o recombinación, debido a las limitaciones del procedimiento de detección, puede no detectar todas las ocurrencias de modificación y/o recombinación. Además, algunos eventos de modificación y/o recombinación pueden ser silenciosos, no dando ninguna indicación detectable de que la modificación y/o recombinación ha tenido lugar. La incapacidad de detectar eventos de modificación y/o recombinación silenciosa da una estimación artificialmente baja de modificación y/o recombinación. Por estas y otras razones, la invención no se limita a ninguna modificación y/o frecuencia de recombinación en particular. En una realización, la frecuencia de modificación y/o recombinación está entre el 0,01 % y el 100 %. En otra realización, la frecuencia de modificación y/o recombinación está entre el 0,01 % y el 50 %. En otra realización, la frecuencia de modificación y/o recombinación está entre el 0,1 % y el 10 %. En otra realización más, la frecuencia de modificación y/o recombinación está entre el 0,1 % y el 5 %.

El término "frecuencia de mutación", tal y como se utiliza en el presente documento en referencia a una población de células que son tratadas con una molécula modificadora del ADN que es capaz de introducir una mutación en un sitio diana del genoma de las células, se refiere al número de células de la población tratada que contienen la mutación en el sitio diana en comparación con el número total de células que son tratadas con la molécula modificadora del ADN. Por ejemplo, con respecto a una población de células que es tratada con la molécula modificadora del ADN TFO unida al psoraleno que está diseñada para introducir una mutación en un sitio diana en el genoma de las células, una frecuencia de mutación del 5 % significa que de un total de 100 células que son tratadas con TFO-psoraleno, 5 células contienen una mutación en el sitio diana.

Aunque la presente invención no está limitada a ningún grado de precisión en la modificación y/o recombinación del ADN en la célula, se contempla que algunas realizaciones de la presente invención requieren mayores grados de precisión, dependiendo del resultado deseado. Por ejemplo, los cambios de secuencia específicos requeridos para la reparación del gen (por ejemplo, cambios de base particulares) requieren un mayor grado de precisión en comparación con la producción de una anulación del gen en el que sólo es necesaria la interrupción del gen. Con los procedimientos de la presente invención, el logro de mayores niveles de precisión en las técnicas de modificación y/o recombinación homóloga es mayor que con los procedimientos de la técnica anterior.

#### **Suministro de oligonucleobases reparadoras de genes en células vegetales**

Cualquier procedimiento comúnmente conocido utilizado para transformar una célula vegetal puede ser utilizado para suministrar las oligonucleobases de reparación de genes. A continuación se enumeran procedimientos ilustrativos. La presente invención contempla muchos procedimientos para transfectar las células con el reactivo o reactivos modificadores del ADN. De hecho, la presente invención no se limita a ningún procedimiento en particular. Los procedimientos para la introducción de reactivos modificadores del ADN en una célula o células son bien conocidos en la técnica e incluyen, pero no se limitan a, la microinyección, la electroporación, la adsorción pasiva, la coprecipitación de fosfato de calcio-ADN, la transfección mediada por DEAE-dextrano, la transfección mediada por polibreno, la fusión de liposomas, la lipofectina, la nucleofección, la fusión de protoplastos, la infección retroviral, la biolística (es decir, el bombardeo de partículas) y otros similares.

El uso de microportadores metálicos (microesferas) para introducir grandes fragmentos de ADN en células vegetales que tienen paredes celulares de celulosa mediante la penetración de proyectiles es bien conocido por los expertos en la técnica (de aquí en adelante, suministro biolístico). La Pat. de EE.UU. Nos. 4.945.050; 5.100.792 y 5.204.253 describen técnicas generales de selección de microportadores y dispositivos para proyectarlos.

Las condiciones específicas para el uso de microportadores en los procedimientos de la presente invención se describen en la Publicación Internacional WO 99/07865. En una técnica ilustrativa, se añaden microportadores fríos (60 mg/ml), oligonucleótidos dúplex mezclados (60 mg/ml) 2,5 M CaCl<sub>2</sub> y 0,1 M de espermidina en ese orden; la mezcla se agita suavemente, por ejemplo, mediante vórtex, durante 10 minutos y se deja reposar a temperatura ambiente durante 10 minutos, tras lo cual los microportadores se diluyen en 5 volúmenes de etanol, se centrifugan y se resuspenden en etanol al 100 %. Se pueden obtener buenos resultados con una concentración en la solución adherente de 8-10 µg/µl de microportadores, 14-17 µg/ml de oligonucleótido dúplex mixto, 1,1-1,4 M de CaCl<sub>2</sub> y 18-22 mM de espermidina. Los resultados óptimos se observaron en las condiciones de 8 µg/µl de microportadores, 16,5 µg/ml de oligonucleótido dúplex mixto, 1,3 M de CaCl<sub>2</sub> y 21 mM de espermidina.

Las oligonucleobases reparadoras de genes también pueden introducirse en las células vegetales para la práctica de la presente invención utilizando microfibras para penetrar la pared y la membrana celular. La Pat. de EE.UU. No. 5.302.523 de Coffee et al. describe el uso de fibras de carburo de silicio de 30.veces.0.5 µm y 10.veces.0.3 µm para

facilitar la transformación de cultivos de maíz en suspensión de Black Mexican Sweet. Cualquier técnica mecánica que pueda utilizarse para introducir ADN para la transformación de una célula vegetal utilizando microfibras puede utilizarse para suministrar oligonucleobases reparadoras de genes para la transmutación.

Una técnica ilustrativa para el suministro de microfibras de una oligonucleobase reparadora de genes es la siguiente:  
 5 Las microfibras estériles (2 µg) se suspenden en 150 µl de medio de cultivo vegetal que contiene aproximadamente 10 µg de un oligonucleótido dúplex mixto. Se deja sedimentar un cultivo en suspensión y se mezclan en vórtex volúmenes iguales de células empaquetadas y de la suspensión estéril de fibras/nucleótidos durante 10 minutos y se colocan en placas. Los medios selectivos se aplican inmediatamente o con un retraso de hasta aproximadamente 120 horas, según convenga para el rasgo concreto.

10 En una realización alternativa, las oligonucleobases reparadoras de genes pueden ser suministradas a la célula vegetal por electroporación de un protoplasto derivado de una parte de la planta. Los protoplastos se forman por tratamiento enzimático de una parte de la planta, en particular una hoja, según técnicas bien conocidas por los expertos en la técnica. Véase, por ejemplo Gallois et al., 1996, en *Methods in Molecular Biology* 55:89-107, Humana Press, Totowa, N.J. Kipp et al., 1999, en *Methods in Molecular Biology* 133:213-221, Humana Press, Totowa, N.J. No es necesario cultivar los protoplastos en medios de crecimiento antes de la electroporación. Las condiciones ilustrativas para la electroporación son 3. veces. 10. sup. 5 protoplastos en un volumen total de 0,3 mL con una concentración de oligonucleobase de reparación de genes de entre 0,6-4 µg/mL.

En una realización alternativa, los ácidos nucleicos son absorbidos por los protoplastos de las plantas en presencia del agente modificador de la membrana polietilenglicol, según técnicas bien conocidas por los expertos en la técnica.  
 20 En otra realización alternativa, las oligonucleobases reparadoras de genes pueden ser suministradas inyectándolas con un microcapilar en células vegetales o en protoplastos.

En una realización alternativa, los ácidos nucleicos son embebidos en microperlas compuestas de alginato de calcio y son absorbidos por los protoplastos de las plantas en presencia del agente modificador de la membrana polietilenglicol (véase, por ejemplo, Sone et al., 2002, Liu et al., 2004).

25 En una realización alternativa, los ácidos nucleicos congelados en agua e introducidos en las células vegetales por bombardeo en forma de micropartículas (véase, p. ej, Gilmore, 1991, Patente de EE.UU. 5.219.746; Brinegar et al.).

En una realización alternativa, los ácidos nucleicos unidos a las nanopartículas se introducen en células vegetales intactas mediante la incubación de las células en una suspensión que contiene la nanopartícula (véase, por ejemplo, Pasupathy et al., 2008) o suministrándolos en células intactas a través del bombardeo de partículas o en protoplastos mediante la coincubación (véase, por ejemplo, Torney et al., 2007).  
 30

En una realización alternativa, los ácidos nucleicos acomplejados con péptidos penetrantes y suministrados a las células por coincubación (véase, por ejemplo, Chugh et al, 2008, WO 2008148223 A1; Eudes y Chugh).

En una realización alternativa, los ácidos nucleicos se introducen en células intactas mediante electroporación (véase, por ejemplo, He et al, 1998, US 2003/0115641 A1, Dobres et al.).

35 En una realización alternativa, los ácidos nucleicos se suministran en las células de embriones secos empapándolos en una solución con ácidos nucleicos (empapando los embriones secos en (ver, por ejemplo, Töpfer et al., 1989, Senaratna et al., 1991).

### Selección de plantas

En diversas realizaciones, las plantas divulgadas en el presente documento pueden ser de cualquier especie de planta dicotiledónea, monocotiledónea o gimnosperma, incluyendo cualquier especie de planta leñosa que crezca como árbol o arbusto, cualquier especie herbácea, o cualquier especie que produzca frutos, semillas o verduras comestibles, o cualquier especie que produzca flores coloridas o aromáticas. Por ejemplo, la planta puede seleccionarse de una especie de planta del grupo que consiste en canola, girasol, maíz, tabaco, remolacha azucarera, algodón, maíz, trigo, cebada, arroz, alfalfa, sorgo, tomate, mango, melocotón, manzana, pera, fresa, plátano, melón, patata, zanahoria,  
 45 lechuga, cebolla, soja, soya spp, caña de azúcar, guisante, garbanzo, guisante de campo, haba, lentejas, nabo, colinabo, coles de Bruselas, altramuza, coliflor, col rizada, judías de campo, álamo, pino, eucalipto, uva, cítricos, triticale, alfalfa, centeno, avena, césped y hierbas forrajeras, lino, colza, mostaza, pepino, gloria de la mañana, bálsamo, pimienta, berenjena, caléndula, loto, col, margarita, clavel, tulipán, iris, lirio y plantas productoras de frutos secos, en la medida en que no se mencionen específicamente.

50 Las plantas y las células vegetales pueden someterse a pruebas de resistencia o tolerancia a un herbicida utilizando procedimientos comúnmente conocidos en la técnica, por ejemplo, cultivando la planta o la célula vegetal en presencia de un herbicida y midiendo la tasa de crecimiento en comparación con la tasa de crecimiento en ausencia del herbicida.

Tal como se utiliza en el presente documento, el crecimiento sustancialmente normal de una planta, órgano vegetal, tejido vegetal o célula vegetal se define como una tasa de crecimiento o tasa de división celular de la planta, órgano vegetal, tejido vegetal o célula vegetal que es al menos 35 %, al menos 50 %, al menos 60 % o al menos 75 % de la  
 55

tasa de crecimiento o tasa de división celular en una planta, órgano vegetal, tejido vegetal o célula vegetal correspondiente que expresa la proteína AHAS de tipo salvaje.

5 Tal como se utiliza en el presente documento, el desarrollo sustancialmente normal de una planta, órgano vegetal, tejido vegetal o célula vegetal se define como la ocurrencia de uno o más eventos de desarrollo en la planta, órgano vegetal, tejido vegetal o célula vegetal que son sustancialmente iguales a los que ocurren en una planta, órgano vegetal, tejido vegetal o célula vegetal correspondiente que expresa la proteína de tipo salvaje.

10 En ciertas realizaciones, los órganos vegetales proporcionados en el presente documento incluyen, pero no se limitan a, hojas, tallos, raíces, yemas vegetativas, yemas florales, meristemos, embriones, cotiledones, endosperma, sépalos, pétalos, pistilos, carpelos, estambres, anteras, microsporas, polen, tubos de polen, óvulos, ovarios y frutos, o secciones, rodajas o discos tomados de los mismos. Los tejidos vegetales incluyen, pero no se limitan, los tejidos de callo, los tejidos del suelo, los tejidos vasculares, los tejidos de almacenamiento, los tejidos meristemáticos, los tejidos de las hojas, los tejidos de los brotes, los tejidos de las raíces, los tejidos de las agallas, los tejidos de los tumores vegetales y los tejidos reproductivos. Las células vegetales incluyen, pero no se limitan a, células aisladas con paredes celulares, agregados de las mismas de diferentes tamaños y protoplastos.

15 Las plantas son sustancialmente "tolerantes" a un herbicida relevante cuando son sometidas a este y proporcionan una curva dosis/respuesta que se desplaza hacia la derecha cuando se compara con la proporcionada por una planta similar no tolerante sometida. Estas curvas de dosis/respuesta tienen la "dosis" trazada en el eje X y el "porcentaje de muerte", el "efecto herbicida", etc., trazado en el eje Y. Las plantas tolerantes necesitarán más herbicida que las plantas similares no tolerantes para producir un efecto herbicida determinado. Las plantas que son sustancialmente  
20 "resistentes" al herbicida presentan pocas, si hay, lesiones necróticas, líticas, cloróticas o de otro tipo, cuando se las somete al herbicida en las concentraciones y dosis que suelen emplear los agroquímicos para matar las malezas en el campo. Las plantas que son resistentes a un herbicida son también tolerantes al herbicida.

### Generación de plantas

25 Se conoce el cultivo de tejidos de diversas especies vegetales y la regeneración de plantas a partir de ellos. Por ejemplo, la propagación de un cultivar de canola por cultivo de tejidos se describe en cualquiera de los siguientes, pero no se limita a ninguno de ellos: Chuong et al., "A Simple Culture Method for Brassica hypocotyls Protoplasts", Plant Cell Reports 4:4-6, 1985 Barsby, T. L., et al., "A Rapid and Efficient Alternative Procedure for the Regeneration of Plants from Hypocotyl Protoplasts of Brassica napus," Plant Cell Reports (Spring, 1996) Kartha, K., et al., "In vitro Plant Formation from Stem Explants of Rape," Physiol. Plant, 31:217-220, 1974 Narasimhulu, S., et al., "Species  
30 Specific Shoot Regeneration Response of Cotyledonary Explants of Brassicas," Plant Cell Reports (Spring 1988) Swanson, E., "Microspore Culture in Brassica," Methods in Molecular Biology, Vol. 6, Chapter 17, p. 159, 1990.

La reproducción adicional de la variedad puede producirse mediante el cultivo de tejidos y la regeneración. El cultivo de tejidos de diversos tejidos de soja y la regeneración de plantas a partir de los mismos es bien conocido y ampliamente publicado. Por ejemplo, se puede hacer referencia a Komatsuda, T. et al., "Genotype X Sucrose Interactions for Somatic Embryogenesis in Soybeans", Crop Sci. 31:333-337, 1991 Stephens, P. A., et al., "Agronomic  
35 Evaluation of Tissue-Culture-Derived Soybean Plants", Theor. Appl. Genet. 82:633-635, 1991 Komatsuda, T. et al., "Maturation and Germination of Somatic Embryos as Affected by Sucrose and Plant Growth Regulators in Soybeans Glycine gracilis Skvortz and Glycine max (L.) Merr." Plant Cell, Tissue and Organ Culture, 28:103-113, 1992 Dhir, S. et al., "Regeneration of Fertile Plants from Protoplasts of Soybean (Glycine max L. Merr.); Genotypic Differences in Culture Response," Plant Cell Reports 11:285-289, 1992 Pandey, P. et al., "Plant Regeneration from Leaf and Hypocotyl Explants of Glycine wightii (W. and A.) VERDC. var. longicauda," Japan J. Breed. 42:1-5, 1992y Shetty, K., et al., "Stimulation of In Vitro Shoot Organogenesis in Glycine max (Merrill.) by Allantoin and Amides," Plant Science 81:245-251, 1992. Las divulgaciones de Pat. de EE.UU. No. 5,024,944 otorgada el 18 de junio de 1991 a Collins et al. y Pat. de EE.UU. No. 5,008,200 otorgada el 16 de abril de 1991 a Ranch et al.

### 45 EJEMPLOS

#### Ejemplo 1: Longitud de la GRON

Sommer et al., (Mol Biotechnol. 33:115-22, 2006) describen un sistema informador para la detección de la conversión génica *in vivo* que se basa en un único cambio de nucleótido para convertir entre fluorescencia azul y verde las variantes de la proteína verde fluorescente (GFP). Este sistema informador se adaptó para su uso en los siguientes  
50 experimentos utilizando *Arabidopsis thaliana* como especie modelo con el fin de evaluar la eficiencia de la conversión de la GRON tras la modificación de la longitud de la GRON.

En resumen, para este y los siguientes ejemplos se creó una línea de *Arabidopsis* con múltiples copias de un gen de proteína azul fluorescente mediante procedimientos conocidos por los expertos en la técnica (véase, por ejemplo, Clough y Brent, 1998). Con esta línea se establecieron cultivos de tejidos meristemáticos derivados de la raíz, que se  
55 utilizaron para el aislamiento y cultivo de protoplastos (véase, por ejemplo, Mathur et al., 1995). El suministro de GRON a los protoplastos se logró mediante la captación de GRON por parte del polietilenglicol (PEG) en los protoplastos. Se utilizó un procedimiento en formato de 96 pocillos, similar al descrito por Fujiwara y Kato (2007). A continuación se

describe brevemente el protocolo. Los volúmenes indicados son los aplicados a pocillos individuales de una placa de 96 pocillos.

1. Mezclar 6,25 µl de la mezcla GRON (80 µM) con 25 µl de protoplastos derivados de tejido meristemático de raíces transgénicas de *Arabidopsis* a 5×10<sup>6</sup> células/ml en cada pocillo de una placa de 96 pocillos.
- 5 2. Se añadieron 31, 25 µl de una solución de PEG al 40 % y se mezclaron los protoplastos.
3. Las células tratadas se incubaron en hielo durante 30 minutos.
4. A cada pocillo se añadieron 200 µl de solución W5 y se mezclaron las células.
5. Las placas se dejaron incubar en hielo durante 30 minutos para que los protoplastos se depositaran en el fondo de cada pocillo.
- 10 6. Se retiraron 200 µl del medio por encima de los protoplastos asentados.
7. Se añadieron 85 µl de medio de cultivo (MSAP, véase Mathur et al., 1995).
8. Las placas se incubaron a temperatura ambiente en la oscuridad durante 48 horas. La concentración final de GRON tras añadir el medio de cultivo es de 8 µM.

15 Cuarenta y ocho horas después del suministro de GRON, las muestras se analizaron por citometría de flujo para detectar protoplastos cuya fluorescencia verde y amarilla es diferente de la de los protoplastos de control (BFP0 indica GRONs no-objetivo con ningún cambio en comparación con el objetivo BFP; C es el diseño de la cadena codificante y NC es el diseño de la cadena no-codificante). Una única diferencia de nucleótidos de C a T (cadena codificante) o una mutación dirigida de nucleótidos de G a A (cadena no codificante) en el centro de las moléculas de BFP4. La fluorescencia verde se debe a la introducción de una mutación dirigida en el gen BFP, que da lugar a la síntesis de GFP. Los resultados se muestran en la Figura 1

20 La siguiente tabla muestra la secuencia de GRONs ejemplares de 101-mer y 201-mer de BFP4/NC 5'-3PS/ 3'-3PS diseñados para la conversión de un gen de proteína azul fluorescente (BFP) a fluorescencia verde. (3PS indica 3 enlaces de fosfotioato en cada uno de los extremos del oligo 5' y 3').

Tabla 1:

Nombre de la GRON	Secuencia de nucleótidos de GRON
BFP4/NC 101-mer	G*T*C*GTGCTGCTTCATGTGGCGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTGAAGGTGTCACGAGGGTG GGCCAGGCACGGGCAGCTTGCCGG*T*G*G
BFP0/NC 101-mer	G*T*C*GTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTGGTGAAGGTGTCACGAGGGT GGCCAGGCACGGGCAGCTTGCCGG*T*G*G
BFP4/C 101-mer	C*C*A*CCGGCAAGCTGCCGTGCCCTGGCCACCCCTCGTGACCACTTCACCTACGGCGTGCAGTGCTTCAGCCGC TACCCCGACCATGAAGCAGCAC*G*A*C
BFP0/C 101-mer	C*C*A*CCGGCAAGCTGCCGTGCCCTGGCCACCCCTCGTGACCACTTCACCCACGGCGTGCAGTGCTTCAGCCGC TACCCCGACCATGAAGCAGCAC*G*A*C
BFP4/NC 201-mer	A*A*G*ATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTTCATGTGGTCGG GGTAGCGGCTGAAGCACTGCACGCCGTAGGTGAAGGTGTCACGAGGGTGGGCCAGGGCAGGGCAGCTTGCCG GTGGTGCAAGTGAACCTCAGGGTCAAGCTTGCCGTAGGTGGCATCGCCCTCG *C*C*C
BFP0/NC 201-mer	A*A*G*ATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTTCATGTGGTCGG GGTAGCGGCTGAAGCACTGCACGCCGTGGTGAAGGTGTCACGAGGGTGGGCCAGGGCAGGGCAGCTTGCCG GTGGTGCAAGTGAACCTCAGGGTCAAGCTTGCCGTAGGTGGCATCGCCCTCG *C*C*C
BFP4/C 201-mer	G*G*G*CGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTTCATCTGCACCACGGCAAGCTGCCCGTGCCC TGGCCACCCCTCGTACCACTTCACTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGAAGCAGCACGA CTTCTTCAAGTCCGCATGCCGAAGGCTACGTCCAGGAGCGCACCAT*C*T*T
BFPo/C 201-mer	G*G*G*CGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTTCATCTGCACCACGGCAAGCTGCCCGTGCCC TGGCCACCCCTCGTACCACTTCACTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGAAGCAGCACGA CTTCTTCAAGTCCGCATGCCGAAGGCTACGTCCAGGAGCGCACCAT*C*T*T

\* = enlace PS (fosfotioato)

25

### Ejemplo 2: Tasas de conversión utilizando GRONs marcadas con 5'Cy3/ 3'idC

El propósito de esta serie de experimentos es comparar las eficiencias de los GRONs marcados con fosfotioato (PS) (que tienen 3 moléculas PS en cada extremo de la GRON) con los GRONs marcados con 5'Cy3/ 3'idC. Los GRONs marcados con 5'Cy3/ 3'idC tienen un fluoróforo 5' Cy3 (amidita) y una base inversa 3' idC. La eficacia se evaluó mediante la conversión de la proteína azul fluorescente (BFP) en fluorescencia verde

30

En los tres experimentos, realizados mediante la administración de PEG de GRONs en protoplastos en tubos Falcon individuales (etiquetados como "Tubos") o en placas de 96 pocillos (etiquetadas como "Placa de 96 pocillos"), no hubo diferencias significativas entre las diferentes químicas de GRONs en la eficiencia de conversión de BFP a GFP determinada por citometría (Fig. 1).

5 **Ejemplo 3: Comparación entre la GRON 5'-3PS/ 3'-3PS de BFP4/NC de 41-meres y la GRON de fragmentos de Okazaki**

10 El propósito de esta serie de experimentos es comparar las eficiencias de conversión de las GRONs etiquetadas con fosfotoiato (PS) con fracciones de 3PS en cada extremo de la GRON a "GRONs de fragmentos Okazaki" en presencia y ausencia de un miembro de la familia de la bleomicina, Zeocina™ (1 mg/ml) para inducir roturas de ADN. El diseño de estas GRON se representa en la Fig. 2. Las GRONs se introdujeron en protoplastos BFP de *Arabidopsis* mediante un tratamiento con PEG y la conversión de BFP a GFP se determinó a las 24 h del tratamiento mediante citometría. Las muestras tratadas con zeocina (1 mg/ml) se incubaron con zeocina durante 90 minutos en hielo antes del tratamiento con PEG.

15 En general, la presencia de zeocina (1 mg/ml) incrementó la conversión de BFP a GFP según se determinó por citometría (Tabla 2). Tanto en presencia como en ausencia de zeocina, la GRON NC Okazaki que contenía un grupo 2'-O Me en la primera base de ARN en el extremo 5' de la GRON fue más eficaz en la conversión de BFP a GFP en comparación con la GRON NC Okazaki que contenía un grupo 2'-O Me en cada una de las primeras nueve bases de ARN 5' (Fig. 2 y Tabla 2).

20 En todos los experimentos, no hubo diferencia significativa entre el BFP4/NC 5'3PS/ 3'3PS de 41-mer y el fragmento de Okazaki BFP4/NC GRON de 71-mer que contiene un grupo 5' 2'-O me en la primera base de ARN 5' (denotado como BFP4 71-mer (1) NC) en la conversión de BFP a GFP tanto en presencia como en ausencia de 1 mg/ml de zeocina como se determinó por citometría (Fig. 2 y Tabla 2). Es importante señalar que en presencia de zeocina (y lo que se espera para la bleomicina, la fleomicina, la tallysomicina, la pepleomicina y otros miembros de esta familia de antibióticos) esa conversión se vuelve independiente de la cadena (es decir, tanto las GRONs C como los NC con los  
25 diseños probados en estos experimentos muestran una actividad aproximadamente igual).

Tabla 2: Comparación de un diseño GRON estándar con diseños GRON de fragmentos Okazaki en presencia y ausencia de un antibiótico glicopéptido zeocina.

Zeocina (+)												
Exp. Nombre	BFP4 41-mer			BFP4 71-mer (0)			BFP4 71-mer (1)			BFP4 71-mer (9)		
	NC	c		NC	c		NC	c		NC	c	
APT043	0,13:	0,0875:		0,2275:	0,2075:		0,355:	0,2275:		0,2325:	0,195:	
APT066	1,9:	0,713:		0,762:	0,683:		1,318:	0,7103:		0,769:	0,883:	
Media	1,015:	0,40025:		0,49475:	0,44525:		0,8365:	0,4689:		0,50075:	0,539:	
Desviación estándar	1,251579:	0,442295:		0,377949:	0,336229:		0,680944:	0,341391:		0,379363:	0,486489:	
SE	0,885134:	0,312797:		0,26729:	0,237786:		0,481573:	0,241436:		0,268291:	0,344052:	
Zeocina (-)												
Exp. Nombre	BFP4 41-mer			BFP4 71-mer (0)			BFP4 71-mer (1)			BFP4 71-mer (9)		
	NC	C		NC	C		NC	C		NC	C	
APT043	nd	nd		0,1875:	0,0175:		0,21:	0,025:		0,1:	0,0225:	
APT066	0,109:	0,007:		0,112:	0,005:		0,141:	0,023:		0,065:	0,021:	
Media	0,109:	0,007:		0,14975:	0,01125:		0,1755:	0,024:		0,0825:	0,02175:	
Desviación estándar	na	na		0,053387:	0,008839:		0,04879:	0,001414:		0,024749:	0,001061:	
SE	na	na		0,037756:	0,006251:		0,034505:	0,001:		0,017503:	0,00075:	
BFP4 71-mer (0)			los primeros 10bp de 5' son de ARN y la GRON no tiene protección									
BFP4 71-mer (1)			los primeros 10 pb de 5' son de ARN y el primer pb del extremo 5' tiene un 2'O-Me									
BFP4 71-mer (9)			5' los primeros 10 pb son ARN y los primeros nueve pb del extremo 5' tienen un 2'O-Me									
NC	C											

**Ejemplo 4: Comparación entre las GRONs de 41, 101 y 201 marcadores de BFP4/NC 5'-3PS/ 3'-3PS**

El propósito de esta serie de experimentos fue comparar las eficiencias de conversión (en presencia y ausencia de zeocina) de las GRONs etiquetadas con fosfotioato (PS) con fracciones de 3PS en cada extremo de la GRON de diferentes longitudes: 41-mer, 101-mer y 201-mer que se muestran en la Tabla 1. Una vez más, la presencia de zeocina (1 mg/ml) aumentó las tasas de conversión de BFP a GFP, según se determinó por citometría (Tabla 3). La tendencia general en los tres experimentos fue lineal con el aumento de la longitud del NC GRON tanto en presencia como en ausencia de zeocina. Excepto el BFP-4/NC/101 y el BFP-4/C/101 en presencia de zeocina, éste tenía tasas de conversión casi iguales pero inferiores a las de la GRON de 41-meres NC. Esto contrasta con todos los experimentos anteriores en los que se utilizaron las GRON codificantes y no codificantes de BFP-4/41, en los que el no codificante fue siempre muy superior a la GRON codificante. Esta asimetría en la frecuencia de conversión también se aplica a las GRON BFP-4/201 utilizadas en esta serie experimental.

Tabla 3:

Zeocina (+)						
Exp. Nombre	BFP4 41-mer		BFP4 101-mer		BFP4 201-mer	
	NC	C	NC	C	NC	C
APT038	0,2425:	0,1275:	0,3025:	0,2575:	0,97:	0,245:
APT043	0,13:	0,0875:	0,185:	0,2275:	0,66:	0,1875:
APT047	0,3975:	0,145:	0,19:	0,125:	0,235:	0,085:
APT052	0,3275:	nd	0,17:	0,21:	0,585:	0,225:
APT052	nd	nd	0,3225:	0,3175:	0,5075:	0,3125:
APT058	1,4275:	nd	1,2:	nd	1,9:	nd
APT066	1,9:	0,713:	0,992:	1,05:	1,7:	0,916:
Media	0,7375:	0,26825:	0,480286:	0,364583:	0,936786:	0,3285:
Desviación estándar	0,73828:	0,297475:	0,428968:	0,341634:	0,630943:	0,297412:
SE	0,301462:	0,148738:	0,162119:	0,139499:	0,238452:	0,121442:
Zeocina (-)						
Exp. Nombre	BFP4 41-mer		BFP4 101-mer		BFP4 201-mer	
	NC	C	NC	C	NC	C
APT038	0,05:	0,01:	0,1025:	0,025:	0,5725:	0,025:
APT066	0,109:	0,007:	0,214:	0,047:	0,566:	0,035:
Media	0,0795:	0,0085:	0,15825:	0,036:	0,56925:	0,03:
Desviación estándar	0,041719:	0,002121:	0,078842:	0,015556:	0,004596:	0,007071:
SE	0,029504:	0,0015:	0,055758:	0,011002:	0,00325:	0,005001:

**Ejemplo 5: CRISPRs combinadas con GRONs para mejorar la conversión en las plantas.**

Se deben considerar tres componentes de diseño al ensamblar un complejo CRISPR: Cas9, gRNA (ARN guía) y la región diana (protoespaciador en el gen diana endógeno).

**Cas 9**

- Expresión transitoria del gen Cas9 de *Streptococcus pyogenes* con codón optimizado para *Arabidopsis* o maíz impulsado por 35S o ubiquitina de maíz respectivamente. Genes optimizados sintetizados por Genewiz o DNA 2.0. NB debe garantizar que no se creen intrones crípticos.

- Terminador RBCSE9 según G1155

-NLS de SV40 único (PKKRKV) como fusión C-terminal

- La columna vertebral del vector será la misma que la de todos nuestros sistemas de expresión transitoria: G1155.

**ARNg**

- Proponer el uso de un ARNtracr quimérico - pre-creRNA según Le Cong *et al.*, 2013 y Jinek *et al.*, 2013. Obsérvese que LeCong *et al.* demostraron que el complejo nativo de tracr + pre-crRNA se escinde de forma mucho más eficiente que la versión quimérica. Por lo tanto, una opción sería hacer una quimera utilizando el ARNtracr de longitud completa (89bp).

5 - Secuencia del ARNg ((N)<sub>20</sub> representa la secuencia guía). La secuencia entre corchetes comprende la forma de 89 pb de longitud completa

```

NNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATA
AGGCTAGTCCG(TTATGTTCTTGAAAAAGTGAGTGGCACCGAGTCGGTGGTG
CTTTTT)
    
```

La siguiente figura de Cong *et al.* muestra el complejo nativo y la quimera



10 - El ARNg se expresaría bajo el promotor de la RNA pol III AtU6 en *Arabidopsis* (secuencia indicada a continuación). En el maíz se podría utilizar el promotor ZmU6 RNA pol III. Estas selecciones se basan en Wang *et al.* 2008.

- Terminador RBCSE9 según G1155 o una cadena de T según Wang *et al.* 2013 y el enfoque de un solo componente que se muestra a continuación.

**En la secuencia del promotor U6 de Wang *et al***

15 **Región diana**

- La especificidad de la secuencia guía está definida por la secuencia de la región objetivo. Independientemente de la selección del organismo modelo, éste será el locus Y66H de BFP. Una secuencia PAM (NGG) en las proximidades de Y66H es la única restricción de diseño. Además, la inclusión de la posición Y66H en los 3' 12bp de la secuencia guía ("secuencia semilla") significaría que una vez lograda la reparación el sitio no será recortado.

Tc	gtg	acc	acc	ttc	acc	cac	ggc
	V	T	T	F		T	Y
G							
	61:	62:	63:	64:	65:	66:	67:

20 **Se necesitará una columna vertebral de vector distinta a la de G1155 para permitir el co-suministro de Cas9 y ARNg. Este problema se evitará con el enfoque de un solo componente:**

**Enfoque de un componente**



5 Cas9 purificada (50-500 nM) y el dúplex ARNtracr:crRNA (50-500 nM, 1:1) en un tampón de escisión del plásmido Cas9 (20 mM HEPES pH 7,5, 150 mM KCl, 0,5 mM DTT, 0,1 mM EDTA) con o sin 10 mM MgCl<sub>2</sub>. Las reacciones se detuvieron con un tampón de carga de ADN 5X que contenía 250 mM de EDTA, se resolvieron mediante electroforesis en gel de agarosa al 0,8 o al 1 % y se visualizaron mediante tinción con bromuro de etidio. Para los ensayos de escisión del mutante Cas9, las reacciones se detuvieron con tampón de carga 5X SDS (30 % de glicerol, 1,2 % de SDS, 250 mM de EDTA) antes de cargarlas en el gel de agarosa.

**Objetivos de los rasgos en los cultivos**

Dada la flexibilidad de la secuencia de reconocimiento CRISPR no es difícil encontrar potenciales secuencias protospacer definidas por una secuencia PAM 3' NGG.

10 **ZmEPSPS**

El ejemplo siguiente muestra una secuencia protospacer adecuada (amarillo) y PAM (azul) para crear una rotura de DS en el sitio catalítico de ZmEPSPS donde se sabe que las mutaciones en el T97 y P101 causan tolerancia al glifosato. La reparación posterior de la rotura, mediada por el oligo (ODM), daría lugar a los cambios deseados.

T A M R P L T V A V act gca atg cgg cca ttg aca gca gct gtt act gct gct gg

15 La siguiente tabla muestra las secuencias de protospacer de los genes de interés en los cultivos de interés:

Cultivo	Gen	Secuencia Protospacer
Canola	EPSPS 2-22 P101	ccgctgccgttactgctgca
	EPSPS - 2-23 P101	cggctgcagttactgctgct
	EPSPS 2-25 P101	ccgctgcagttactgctgca
	EPSPS 2-28 P101	ccgctgcagttacagctgca
Lino	EPSPS P101	cagctgctgtaacagccgct
Patata	EPSPS - 2.1/2.2/2.3 P101	cagcagcagttgctgtagct
	Gen EPSPS 1 P101	cagcagcagttacagtagct
Patata	PPX R144	tgcgctcgccttctt
	PPX A220	atttacaggtgtttacgcc

20 Una limitación de las restricciones de diseño es que a menudo es difícil encontrar una secuencia NGG dentro de 12 pb del nucleótido que está siendo alterado por ODM. Esto es significativo porque si este fuera el caso, el éxito de la ODM significaría que el corte posterior no sería posible porque la secuencia de semillas del protospacer estaría alterada. Jinek et al. (2012) mostraron que esto era perjudicial para la eficiencia del corte.

**Referencias**

25 LeCong et al 2013 Science : vol. 339 no. 6121 pp. 819-823.  
 Jinek et al 2012 Science. 337, 816-21  
 Wang et al 2008 ARN 14: 903-913  
 Zhang et al 2013. Plant Physiol. 161: 20-27

Un experto en la técnica aprecia fácilmente que la presente invención está bien adaptada para llevar a cabo los objetos y obtener los fines y ventajas mencionados, así como los inherentes a los mismos. Los ejemplos proporcionados en el presente documento son representativos de las realizaciones preferidas, son ejemplares y no pretenden ser limitaciones al ámbito de la invención.

30 **LISTADO DE SECUENCIAS**

<110> CIBUS US LLC CIBUS EUROPE B.V.

<120> PROCEDIMIENTOS Y COMPOSICIONES PARA AUMENTAR LA EFICIENCIA DE LA MODIFICACIÓN GENÉTICA DIRECCIONADA UTILIZANDO LA REPARACIÓN GENÉTICA MEDIADA POR OLIGONUCLEÓTIDOS

35 <130> CIBUS-024-PCT

<140> PCT/US2014/029566

<141> 2014-03-14  
 <150> 61/801,333  
 <151> 2013-03-15  
 <160> 38  
 5 <170> PatentIn version 3.5  
 <210> 1  
 <211> 99  
 <212> ADN  
 <213> Secuencia Artificial  
 10 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 <220>  
 <221> características\_misc  
 <222> (1)..(4)  
 15 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 <221> características\_misc  
 <222> (96)..(99)  
 <223> Enlace de fosforotioato entre nucleótidos  
 20 <400> 1  
 gtcgtgctgc ttcattgtggc ggggtagcgg ctgaagcact gcacgccgta ggtgaagggtg 60  
 gtcacgaggg tgggccaggc acgggcagct tgccggtgg 99  
 <210> 2  
 <211> 101  
 <212> ADN  
 25 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <220>  
 <221> características\_misc  
 30 <222> (1)..(4)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 <221> características\_misc  
 <222> (98)..(101)  
 35 <223> Enlace de fosforotioato entre nucleótidos

ES 2 921 207 T3

<400> 2  
 gtcgtgctgc ttcattgtgt cgggtagcgc gctgaagcac tgcacgccgt ggggtgaaggt 60  
 ggtcacgagg gtgggccagg gcacgggcag cttgccggtg g 101  
 <210> 3  
 <211> 101  
 5 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <220>  
 10 <221> características\_misc  
 <222> (1)..(4)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 <221> características\_misc  
 15 <222> (98)..(101)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <400> 3  
 ccaccggcaa gctgcccggt ccctggccca ccctcgtgac caccttcacc tacggcgtgc 60  
 agtgcttcag ccgctacccc gaccacatga agcagcacga c 101  
 <210> 4  
 20 <211> 101  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 25 <220>  
 <221> características\_misc  
 <222> (1)..(4)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 30 <221> características\_misc  
 <222> (98)..(101)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <400> 4

ES 2 921 207 T3

ccaccggcaa gctgcccggtg ccctggccca ccctcgtgac caccttcacc cacggcgtgc 60  
 agtgcttcag ccgctacccc gaccacatga agcagcacga c 101

<210> 5  
 <211> 201  
 <212> ADN  
 5 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <220>  
 <221> características\_misc  
 10 <222> (1)..(4)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 <221> características\_misc  
 <222> (198)..(201)  
 15 <223> Enlace de fosforotioato entre nucleótidos  
 <400> 5

aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc 60  
 ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt aggtgaaggt ggtcacgagg 120  
 gtgggccagg gcacgggacag cttgccggtg gtgcagatga acttcagggt cagcttgccg 180  
 taggtggcat cgcctcggc c 201

<210> 6  
 <211> 201  
 20 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <220>  
 25 <221> características\_misc  
 <222> (1)..(4)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <220>  
 <221> características\_misc  
 30 <222> (198)..(201)  
 <223> Enlace de fosforotioato entre nucleótidos  
 <400> 6

## ES 2 921 207 T3

```

    --- -
aagatggtgc gctcctggac gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc      60
ttcatgtggt cggggtagcg gctgaagcac tgcacgccgt gggagaaggt ggtcacgagg      120
gtgggccagg gcacgggcag cttgccggtg gtgcagatga acttcagggt cagcttgccg      180
taggtggcat cgcctcggc c                                     201
<210> 7
<211> 201
<212> ADN
5 <213> Secuencia Artificial
<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético
<220>
<221> características_misc
10 <222> (1)..(4)
<223> Enlace de fosforotioato entre nucleótidos
<220>
<221> características_misc
<222> (198)..(201)
15 <223> Enlace de fosforotioato entre nucleótidos
<400> 7
gggagagggc gatgccacct acggcaagct gacctgaag ttcatctgca ccaccggcaa      60
gctgcccgtg ccctggccca ccctcgtgac caccttcacc tacggcgtgc agtgcttcag      120
ccgctacccc gaccacatga agcagcacga cttcttcaag tccgccatgc ccgaaggcta      180
cgtccaggag cgcaccatct t                                     201
<210> 8
<211> 201
20 <212> ADN
<213> Secuencia Artificial
<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético
<220>
25 <221> características_misc
<222> (1)..(4)
<223> Enlace de fosforotioato entre nucleótidos
<220>
<221> características_misc
30 <222> (198)..(201)

```

ES 2 921 207 T3

<223> Enlace de fosforotioato entre nucleótidos  
 <400> 8  
 gggcgagggc gatgccacct acggcaagct gaccctgaag ttcactctgca ccaccggcaa 60  
 gctgcccgtg cctcggccca cctcgtgac caccttcacc cacggcgtgc agtgcttcag 120  
 ccgctacccc gaccacatga agcagcacga cttcttcaag tccgcatgc ccgaaggcta 180  
 cgtccaggag cgcacccatct t 201  
 <210> 9  
 5 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
 10 <400> 9  
 Pro Lys Lys Arg Lys Val  
 1 5  
 <210> 10  
 <211> 110  
 <212> ADN  
 15 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <220>  
 <221> base\_modificada  
 20 <222> (1)..(20)  
 <223> a, c, t, g, desconocido u otro  
 <400> 10  
 nnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatgttc ttgaaaaaag tgagtggcac cgagtcggtg gtgctttttt 110  
 <210> 11  
 25 <211> 48  
 <212> ARN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 30 <220>  
 <221> base\_modificada  
 <222> (3)..(32)

# ES 2 921 207 T3

<223> a, c, u, g, desconocido u otro  
 <400> 11  
 acnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnguuiuuga gcuaugcu 48  
 <210> 12  
 5 <211> 67  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 10 <220>  
 <223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
 <400> 12  
 agcauagcaa guaaaaaua ggctaguccg uuaucaacuu gaaaagugg caccgagucg 60  
 gugcuuu 67  
 <210> 13  
 15 <211> 62  
 <212> ARN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 20 <220>  
 <221> base\_modificada  
 <222> (1)..(20)  
 <223> a, c, u, g, desconocido u otro  
 <400> 13  
 nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaaagc aaguuaaaau aaggcuaguc 60  
 25 cg 62  
 <210> 14  
 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 <400> 14  
 togtgaccac ctcaccac ggc 23  
 <210> 15





## ES 2 921 207 T3

ccgctgccgt tactgctgca 20  
<210> 23  
<211> 20  
<212> ADN  
5 <213> Secuencia Artificial  
<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 23  
cggctgcagt tactgctgct 20  
10 <210> 24  
<211> 20  
<212> ADN  
<213> Secuencia Artificial  
<220>  
15 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 24  
ccgctgcagt tactgctgca 20  
<210> 25  
<211> 20  
20 <212> ADN  
<213> Secuencia Artificial  
<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 25  
25 ccgctgcagt tacagctgca 20  
<210> 26  
<211> 20  
<212> ADN  
<213> Secuencia Artificial  
30 <220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 26  
cagctgctgt aacagccgct 20  
<210> 27  
35 <211> 20  
<212> ADN  
<213> Secuencia Artificial

## ES 2 921 207 T3

<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 27  
cagcagcagt tgctgtagct 20  
5 <210> 28  
<211> 20  
<212> ADN  
<213> Secuencia Artificial  
<220>  
10 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 28  
cagcagcagt tacagtagct 20  
<210> 29  
<211> 20  
15 <212> ADN  
<213> Secuencia Artificial  
<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 29  
20 tgcgcctcgc tttgtcttgt 20  
<210> 30  
<211> 20  
<212> ADN  
<213> Secuencia Artificial  
25 <220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<400> 30  
atttacagg tgttacgcc 20  
<210> 31  
30 <211> 71  
<212> ADN  
<213> Secuencia Artificial  
<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
35 <220>  
<223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
<220>

# ES 2 921 207 T3

<221> base\_modificada  
 <222> (1)..(1)  
 <223> nucleótido modificado 2'-O-Me  
 <220>

5

<221> base\_modificada  
 <222> (11)..(71)  
 <223> nucleótido modificado 2'-O-Me  
 <400> 31  
 uucauguggu cggggtagcg gctgaagcac tgcacgccgt aggtgaaggt ggtcacgagg 60  
 gtgggccagg g 71

10

<210> 32  
 <211> 71  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>

15

<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 <220>  
 <223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
 <220>

20

<221> base\_modificada  
 <222> (1)..(1)  
 <223> nucleótido modificado 2'-O-Me  
 <220>

25

<221> base\_modificada  
 <222> (11)..(71)  
 <223> nucleótido modificado 2'-O-Me  
 <400> 32  
 uucauguggu cggggtagcg gctgaagcac tgcacgccgt gggtgaaggt ggtcacgagg 60  
 gtgggccagg g 71

30

<210> 33  
 <211> 71  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 <220>

ES 2 921 207 T3

<223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
<220>  
<221> base\_modificada  
<222> (1)..(1)  
5 <223> nucleótido modificado 2'-O-Me  
<220>  
<221> base\_modificada  
<222> (11)..(71)  
<223> nucleótido modificado 2'-O-Me  
10 <400> 33  
gcugcccgug ccctggccca ccctcgtgac caccttcacc tacggcgtgc agtgcttcag 60  
ccgctacccc g 71  
<210> 34  
<211> 71  
<212> ADN  
15 <213> Secuencia Artificial  
<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<220>  
<223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
20 <220>  
<221> base\_modificada  
<222> (1)..(1)  
<223> nucleótido modificado 2'-O-Me  
<220>  
25 <221> base\_modificada  
<222> (11)..(71)  
<223> nucleótido modificado 2'-O-Me  
<400> 34  
gcugcccgug ccctggccca ccctcgtgac caccttcacc cacggcgtgc agtgcttcag 60  
30 ccgctacccc g 71  
<210> 35  
<211> 71  
<212> ADN  
35 <213> Secuencia Artificial

ES 2 921 207 T3

<220>  
<223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<220>  
<223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
5 <220>  
<221> base\_modificada  
<222> (1)..(9)  
<223> nucleótido modificado 2'-O-Me  
<220>  
10 <221> base\_modificada  
<222> (11)..(71)  
<223> nucleótido modificado 2'-O-Me  
<400> 35  
uucauguggu cggggtagcg gctgaagcac tgcacgccgt aggtgaaggt ggtcacgagg 60  
gtgggccagg g 71  
15 <210> 36  
<211> 71  
<212> ADN  
<213> Secuencia Artificial  
<220>  
20 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
<220>  
<223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
<220>  
<221> base\_modificada  
25 <222> (1)..(9)  
<223> nucleótido modificado 2'-O-Me  
<220>  
<221> base\_modificada  
<222> (11)..(71)  
30 <223> nucleótido modificado 2'-O-Me  
<400> 36  
uucauguggu cggggtagcg gctgaagcac tgcacgccgt gggtgaaggt ggtcacgagg 60  
gtgggccagg g 71  
<210> 37  
<211> 71

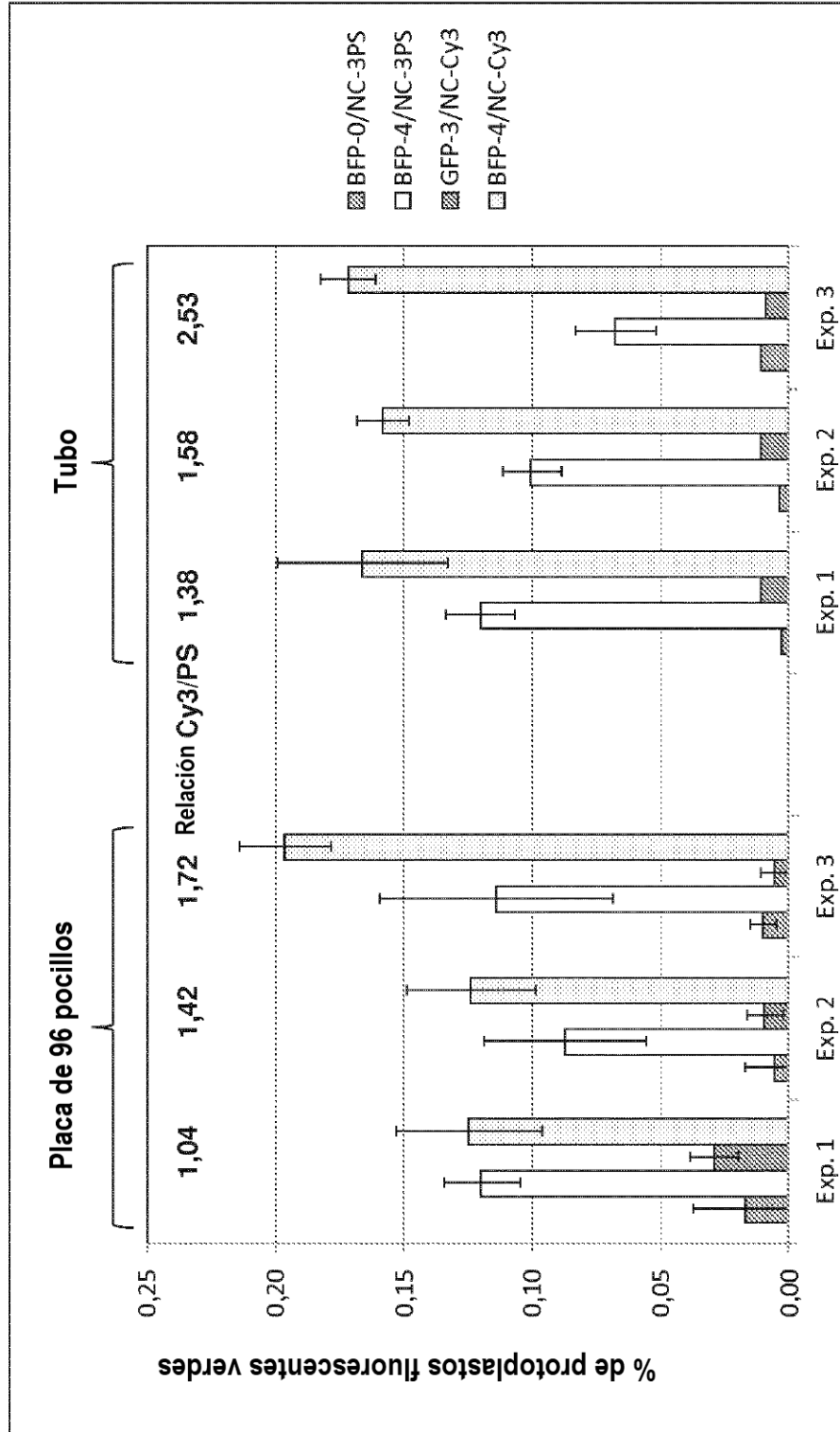
# ES 2 921 207 T3

<212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 5 <220>  
 <223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
 <220>  
 <221> base\_modificada  
 <222> (1)..(9)  
 10 <223> nucleótido modificado 2'-O-Me  
 <220>  
 <221> base\_modificada  
 <222> (11)..(71)  
 <223> nucleótido modificado 2'-O-Me  
 15 <400> 37  
 gcugcccgug ccctggccca ccctcgtgac caccttcacc tacggcgtgc agtgcttcag 60  
 ccgctacccc g 71  
 <210> 38  
 <211> 71  
 <212> ADN  
 20 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido Sintético  
 <220>  
 <223> Descripción de la molécula combinada de ADN/ARN: oligonucleótido sintético  
 25 <220>  
 <221> base\_modificada  
 <222> (1)..(9)  
 <223> nucleótido modificado 2'-O-Me  
 <220>  
 30 <221> base\_modificada  
 <222> (11)..(71)  
 <223> nucleótido modificado 2'-O-Me  
 <400> 38  
 gcugcccgug ccctggccca ccctcgtgac caccttcacc cacggcgtgc agtgcttcag 60  
 ccgctacccc g 71

## REIVINDICACIONES

1. Un procedimiento para introducir una mutación mediada por una oligonucleobasa de reparación de genes (GRON) en una secuencia de ácido desoxirribonucleico (ADN) diana en una célula vegetal, que comprende:
- 5 suministro de una GRON en la célula vegetal; y  
 suministro de una endonucleasa específica de sitio en la célula vegetal, en la que la endonucleasa específica de sitio se selecciona entre nucleasas de dedos de zinc, nucleasas efectoras similares a activadores de la transcripción (TALEN) y complejos de Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas con Regularidad (complejos CRISPR);  
 10 en el que la GRON comprende una o más alteraciones de los nucleótidos convencionales de ARN y ADN seleccionadas entre uno o más nucleótidos 2'O-metilos en el extremo 5' o 3' del mismo, una tapa en el terminal 5', uno o más tintes fluorescentes unidos covalentemente al mismo, una o más modificaciones de fosfotioato, una base inversa en el extremo 3' del mismo; y  
 en el que la endonucleasa específica del sitio induce una rotura de cadena simple o doble en la vecindad del sitio diana para la conversión por GRON.
- 15 2. El procedimiento de la reivindicación 1, en el que la GRON comprende además una o más de las siguientes características:
- la GRON tiene una longitud superior a 55 bases, la GRON comprende opcionalmente dos o más sitios de mutación para su introducción en el ADN diana;  
 la GRON comprende uno o más nucleótidos abásicos  
 20 la GRON comprende uno o más nucleótidos 8'oxo dA y/o 8'oxo dG;  
 la GRON comprende uno o más nucleótidos de ARN 2'O-metilo en su extremo 5';  
 la GRON comprende al menos dos nucleótidos de ARN 2'-O-metilo en su extremo 5';  
 la GRON comprende un tinte intercalado;  
 25 la GRON comprende una modificación de la columna vertebral seleccionada del grupo que consiste en una modificación de fosfonato de metilo, una modificación de ácido nucleico bloqueado (LNA), una modificación de O - (2-metoxietil) (MOE), una modificación de di PS y una modificación de ácido nucleico peptídico (PNA);  
 la GRON comprende uno o más enlaces cruzados intracadena;  
 la GRON comprende una o más bases que aumentan la energía de hibridación.
3. El procedimiento de la reivindicación 1 o 2, en el que el procedimiento comprende además sintetizar todo o una parte de la GRON utilizando multímeros de nucleótidos.
4. El procedimiento de cualquiera de las reivindicaciones 1 a 3, en el que la endonucleasa específica del sitio es un complejo de Repeticiones Palindrómicas Cortas Agrupadas Interespaciadas con Regularidad (complejo CRISPR).
5. El procedimiento de cualquiera de las reivindicaciones 1 a 3, en el que la endonucleasa de sitio específico es una nucleasa efectora similar al activador de la transcripción (TALEN).
- 35 6. El procedimiento de cualquiera de las reivindicaciones 1 a 3, en el que la endonucleasa de sitio específico es una nucleasa de dedo de zinc.
7. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que la secuencia de ácido desoxirribonucleico (ADN) diana está dentro del genoma de la célula vegetal.
- 40 8. El procedimiento v, en el que la célula es una célula vegetal de una especie seleccionada del grupo que consiste en canola, girasol, maíz, tabaco, remolacha azucarera, algodón, maíz, trigo, cebada, arroz, alfafa, cebada, sorgo, tomate, mango, melocotón, manzana, pera, fresa, plátano, melón, patata, zanahoria, lechuga, cebolla, soja, soya spp, caña de azúcar, guisante, garbanzo, guisante de campo, haba, lentejas, nabo, colinabo, coles de Bruselas, altramuz, coliflor, col rizada, judía de campo, álamo, pino, eucalipto, uva, cítricos, triticale, alfalfa, centeno, avena, césped y hierbas forrajeras, lino, colza, mostaza, pepino, gloria de la mañana, bálsamo, pimienta, berenjena, caléndula, loto, col, margarita, clavel, tulipán, iris y lirio.
- 45 9. El procedimiento de una de las reivindicaciones 1-4, en el que la célula es transgénica.
10. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que la secuencia de ADN diana es un gen endógeno de la célula.
11. El procedimiento de cualquiera de las reivindicaciones anteriores, que comprende además la regeneración de una planta a partir de la célula vegetal.
- 50 12. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que la célula vegetal es un protoplasto.

FIG. 1



A. Okasaki Fragmentos GRON (BFP4 o 0/C o NC/71/5' 2'-0-Me (1))  
 BFP4/NC: U UCA UGU GGU CCG GGT ACC GGC TGA AGC ACT GCA CGC CGT AGG TGA AGG TGG TCA CGA GGG TGG GCC AGG G (71-mer)  
 BFP0/NC: U UCA UGU GGU CCG GGT ACC GGC TGA AGC ACT GCA CGC CGT GGG TGA AGG TGG TCA CGA GGG TGG GCC AGG G (71-mer)  
 BFP4/C: G CUG CCC GUG CCC TGG CCC ACC CTC GTG ACC ACC TTC ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC G (71-mer)  
 BFP0/C: G CUG CCC GUG CCC TGG CCC ACC CTC GTG ACC ACC TTC ACC CAC GGC GTG CAG TGC TTC AGC CGC TAC CCC G (71-mer)  
bases de ARN

Grupo 2'-0-Me en todas las bases de ARN excepto en la base más cercana a la primera base de ADN de la GRON I  
 (la base de ARN sin el grupo 2'-0-Me está en verde)

B. Okasaki Fragmentos GRON (BFP4 o 0/C o NC/71/5' 2'-0-Me (9))  
 BFP4/NC: U UCA UGU GGU CCG GGT ACC GGC TGA AGC ACT GCA CGC CGT AGG TGA AGG TGG TCA CGA GGG TGG GCC AGG G (71-mer)  
 BFP0/NC: U UCA UGU GGU CCG GGT ACC GGC TGA AGC ACT GCA CGC CGT GGG TGA AGG TGG TCA CGA GGG TGG GCC AGG G (71-mer)  
 BFP4/C: G CUG CCC GUG CCC TGG CCC ACC CTC GTG ACC ACC TTC ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC G (71-mer)  
 BFP0/C: G CUG CCC GUG CCC TGG CCC ACC CTC GTG ACC ACC TTC ACC CAC GGC GTG CAG TGC TTC AGC CGC TAC CCC G (71-mer)  
bases de ARN

Grupo 2'-0-Me en todas las bases de ARN excepto en la base más cercana a la primera base de ADN de la GRON I  
 (la base de ARN sin el grupo 2'-0-Me está en verde)

FIG. 2