

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 603 587**

51 Int. Cl.:

C12N 15/82

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.01.2010** **PCT/GB2010/000111**

87 Fecha y número de publicación internacional: **29.07.2010** **WO10084331**

96 Fecha de presentación y número de la solicitud europea: **25.01.2010** **E 10705401 (7)**

97 Fecha y número de publicación de la concesión europea: **24.08.2016** **EP 2389441**

54 Título: **Direccionamiento génico en plantas**

30 Prioridad:

26.01.2009 GB 0901249

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

28.02.2017

73 Titular/es:

ALGENTECH SAS (100.0%)
Pépinière Genepole Enterprise 4, rue Pierre
Fontaine
91058 Evry Cedex, FR

72 Inventor/es:

MALCUIT, ISABELLE y
SOROKIN, ALEXANDER

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 603 587 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

Direccionamiento génico en plantas

5 La presente invención se refiere a un método para alterar un genotipo de planta mediante la introducción de alteraciones genéticas a un genoma de célula vegetal, células vegetales que comprenden ADN nuclear genéticamente alterado, secuencias de ácidos nucleicos y polipéptidos generados como resultado de la introducción de cambios genéticos y plantas completas y/o partes de plantas. Particularmente, el método se refiere a un método más eficiente para introducir in situ mutaciones o deleciones en secuencias de ácido nucleico de interés, la reparación de secuencias de ácido nucleico en células vegetales, la introducción de nuevos ácidos nucleicos en sitios cromosómicos de interés y la generación de nuevos dominios dentro de genes de interés de una manera dirigida, el material genético requerido para ello, tal como ADN y ARN, vectores, células huésped, células vegetales que comprenden genomas nucleares genéticamente modificados, plantas derivadas de dichas células vegetales y usos de estas.

15 Una desventaja de los métodos de transformación nuclear de la técnica anterior es que la eficacia del direccionamiento génico para la reparación, alteración de los propósitos de la función o para la introducción de polinucleótidos de interés en sitios específicos dentro del genoma tiende a ser baja. Además, el direccionamiento preciso de los genes nucleares dentro de las poblaciones de células es ineficiente. Como consecuencia, el direccionamiento de genes dentro del genoma nuclear en las células es irregular, con un bajo porcentaje de absorción de material genético introducido en el sitio o locus correcto. Como consecuencia adicional, la eficacia de la reparación de los genes en sitios específicos dentro del genoma, de la introducción de alteraciones deseadas, o incluso de la introducción de genes de interés, es baja. Otra desventaja de los métodos de la técnica anterior es que la entrega de la información genética en el genoma tiende a ser errática puesto que los mecanismos de suministro empleados dependen de la oportunidad para la entrega satisfactoria de información genética, tal como ADN, en el genoma nuclear en los sitios deseados. La capacidad de dirigir eficientemente los genes de interés dentro de la célula vegetal necesita que se mejore de manera que se puedan crear mejores variedades de plantas con un mayor grado de certidumbre y con un mayor nivel de rendimiento de direccionamiento génico. Como tal, los procesos de la técnica anterior parecen carecer de un medio verdaderamente eficiente de direccionamiento génico. Estas y otras desventajas de la tecnología de direccionamiento génico de la técnica anterior resultarán evidentes a partir de la descripción.

30 Los presentes inventores han desarrollado una técnica robusta y reproducible para las modificaciones específicas del sitio dentro del genoma nuclear de la planta. La esencia del direccionamiento génico como se desarrolla en la presente descripción radica en la colocación eficiente de secuencias polinucleotídicas del donador, tales como ADN nativo o ADNc de elección, junto con el ADN de un sitio objetivo dentro de un gen o una pluralidad de secuencias de genes, e induciendo la recombinación entre la secuencia polinucleotídica del donador o las secuencias polinucleotídicas del donador y el sitio o sitios objetivo, pueden producirse eventos de transformación eficientes con una alta eficiencia. La técnica de direccionamiento génico usa dos componentes:

- (1) una proteína de recombinación de múltiples dominios (MDRP); y
- 40 (2) un sistema de vector de amplificación donador (DAV).

La MDRP de la invención posee al menos las dos primeras de las siguientes funcionalidades:

- (i) es capaz de unirse a un polinucleótido donador, tal como un ADN nativo introducido o un ADNc;
- 45 (ii) es capaz de unirse a un polinucleótido cromosómico;
- (iii) es capaz de situar secuencias de ácido nucleico objetivo específicas, tales como ADN que codifica para un gen nativo o una parte de este en el genoma nuclear; y
- (iv) es capaz de iniciar la hibridación de polinucleótidos homólogos, tales como secuencias de ADN, entre objetivo y donador.

50 El DAV es una herramienta molecular para la amplificación de un fragmento donador o de fragmentos donadores de polinucleótidos, tales como los derivados de secuencias nativas o de ADNc de interés en una célula vegetal, y es capaz de proporcionar un suministro permanente y abundante de secuencias polinucleotídicas del donador, tales como ADNc o moléculas de ADN nativo, para el direccionamiento génico eficiente.

55 De conformidad con un primer aspecto de la invención, se proporciona una secuencia de polinucleótido aislada como se define en la reivindicación 1.

60 El dominio o dominios que se unen al ADN donador pueden seleccionarse de la superfamilia CACTA de elementos transponibles, especialmente de los elementos transponibles encontrados en eucariotas tales como plantas, por ejemplo, el elemento transponible de maíz *En/Spm* (Pereira y otros, (1986) EMBO J 1986, 5:835-841), *Tpn1* de *Ipomoea nil* (Inagaki y otros, (1994) Plant Cell, 6:375-383); *Tdc1* de *Daucus carota* (Ozeki y otros, (1997) Mol Gen Genet, 254:407-416); *Psl* de *Petunia hybrida* (Snowden & Napoli (1998) Plant J, 14:43-54); *Candystripe1* (*Cs1*) de *Sorghum bicolor* (Chopra y otros, (1999) Proc Natl Acad Sci, 96:15330-15335); *Atenspm1* (o *CAC1*) de *Arabidopsis thaliana* (Miura A, y otros, (2001) Nature, 411:212-214); *Tpo1* de *Lolium perenne* (Langdon y otros, (2003) Genetics, 163:1097-1108); y *TamRSA1* de *A. majus* (Roccaro y otros, (2007) Mol Genet Genomics, 278:243-254), y dominios de

unión al ADN de endonucleasas de restricción bacterianas, tal como restrictasa EcoRV. Los dominios de unión al ADN adecuados pueden seleccionarse a partir de la proteína *tnpA* completa de *En/Spm*, o fragmentos seleccionados de esta, por ejemplo la secuencia polipeptídica entre las posiciones 122 a 427 de *tnpA*.

5 Los dominios de unión al ADN para el uso en la presente invención tienen típicamente una fuerte afinidad de unión para las denominadas regiones "subterminales". Las regiones subterminales adecuadas para el uso en los métodos de la presente invención comprenden motivos cortos repetitivos que son típicamente de longitud de aproximadamente 7 a
10 aproximadamente 25 pares de bases de nucleótidos. Tales motivos se sitúan típicamente en ambos extremos de un elemento de ADN transponible CACTA, tal como, por ejemplo, el transposón *En/Spm*. Los motivos repetitivos pueden situarse en cualquiera de los extremos o en ambos extremos de la familia CACTA de secuencias de ácidos nucleicos del elemento transponible. Los miembros de la familia CACTA de secuencias de ácidos nucleicos del elemento transponible incluyen *En/Spm* de maíz (Pereira y otros, (1986) EMBO J 1986, 5:835-841), *Tpn1* de *Ipomoea nil* (Inagaki y otros, (1994) Plant Cell, 6:375- 83); *Tdc1* de *Daucus carota* (Ozeki y otros, (1997) Mol Gen Genet, 254:407-416); *PsI* de *Petunia hybrida* (Snowden & Napoli (1998) Plant J, 14:43-54); *Candystripe1* (*Cs1*) de *Sorghum bicolor* (Chopra y otros, (1999) Proc Natl Acad Sci, 96:15330- 5335); *Atenspm1* (o *CAC1*) de *Arabidopsis thaliana* (Miura A, y otros, (2001) Nature, 411:212-214); *Tpo1* de *Lolium perenne* (Langdon y otros, (2003) Genetics, 163:1097-1108); y *TamRSA1* de *A. majus* (Roccaro y otros, (2007) Mol Genet Genomics, 278:243-254).

20 Las secuencias subterminales adecuadas para el uso en la invención incluyen los 6 motivos hiperreactivos de 12 pb en la región subterminal izquierda (STR1) y los 8 motivos hiperreactivos de 12 pb en la región subterminal derecha (STR2) de la secuencia de polinucleótidos de la secuencia de ADN *En/Spm*. Las secuencias subterminales pueden usarse como regiones flanqueantes para el fragmento donador en el vector de amplificación donador de la invención.

25 Las proteínas del elemento transponible, tal como *tnpA*, tienen típicamente una segunda función, la de mediar la asociación de las regiones subterminales con secuencias de ADN cromosómicas, y como resultado además pueden poseer una funcionalidad de unión a ADN cromosómico, tal como en el caso de la secuencia de proteína *tnpA* proporcionada en la presente descripción. Otras proteínas del elemento transponible pueden carecer de tal funcionalidad de unión al ADN cromosómico, y este tipo de funcionalidad puede incorporarse en ellas mediante la fusión de secuencias que tienen tal funcionalidad con elementos transponibles que carecen de funcionalidad, usando
30 procedimientos convencionales en la técnica.

Por lo tanto, el dominio de unión cromosómica (CBD) además puede seleccionarse a partir de elementos genéticos móviles tales como *Spm* de maíz (Pereira y otros, (1986) EMBO J 1986, 5:835-841), *Tpn1* de *Ipomoea nil* (Inagaki y otros, (1994) Plant Cell, 6:375-383); *Tdc1* de *Daucus carota* (Ozeki y otros, (1997) Mol Gen Genet, 254:407-416); *PsI* de *Petunia hybrida* (Snowden & Napoli (1998) Plant J, 14:43-54); *Candystripe1* (*Cs1*) de *Sorghum bicolor* (Chopra y otros, (1999) Proc Natl Acad Sci, 96:15330-15335); *Atenspm1* (o *CAC1*) de *Arabidopsis thaliana* (Miura A, y otros, (2001) Nature, 411:212-214); *Tpo1* de *Lolium perenne* (Langdon y otros, (2003) Genetics, 163:1097-1108); y *TamRSA1* de *A. majus* (Roccaro y otros, (2007) Mol Genet Genomics, 278:243-254) y dominios de
35 unión al ADN de las endonucleasas de restricción bacterianas, tal como la bien caracterizada restrictasa EcoRV.

40 La secuencia polinucleotídica aislada que codifica la MDRP comprende además un tercer dominio que es un dominio de unión al objetivo cromosómico y se diseña típicamente de conformidad con la secuencia del gen o fragmento génico de interés del sitio objetivo. El dominio de unión al objetivo de la MDRP puede seleccionarse de las proteínas de unión al ADN específicas de genes tales como los factores de transcripción o proteínas de dominio MADS como se describe, por ejemplo, en, Pellegrini y otros, (1995) Nature, 376, 490-498; Schwechheimer y otros, (1998) Ann Review Plant Physiol Plant Mol Biol., 49, 127-150), nucleasas de dedos de zincs (ZFN) o dominios de unión de meganucleasa (MN) (Bibikova y otros, (2003) Science, 300, 764-766; Epinat y otros, (2003) Nuc Acid Res, 31, 2952-2962).

50 La secuencia de polinucleótidos aislada que codifica una MDRP comprende un cuarto dominio que es un dominio inductor de la recombinación (RID). El apareamiento del ADN homólogo puede promoverse por dominios peptídicos cortos de proteínas de reparación de ADN. Típicamente, tales dominios peptídicos cortos tienen de aproximadamente 15 a aproximadamente 50 residuos de aminoácidos de longitud y como ejemplo de un péptido corto de uso en la invención se proporciona el dominio de 20 aminoácidos derivado de la proteína *recA* de *E. coli*. Estos dominios cortos pueden fusionarse a la proteína de unión al ADN, por ejemplo, *tnpA* como se proporciona en la presente descripción, resultando en *tnpAA*, y se usa como un RID de la MDRP. El RID puede seleccionarse de proteínas bien caracterizadas implicadas en la inducción de recombinación homóloga como se describe por Symington, (2002) Microbiology and Molecular Biology Reviews, 66, 630-670. Tales proteínas incluyen las proteínas rad 51 y rad 52 que se encuentran en plantas, hongos, algas y animales.

60 Naturalmente, el hombre con experiencia en la técnica apreciará que las secuencias polinucleotídicas de la MDRP pueden estar compuestas por secuencias de ADN genómico o ADNc o una combinación de ambas. La secuencia polinucleotídica que codifica para la proteína MRDP se suministra preferentemente en un vector como se describe en la presente descripción. Sin embargo, el hombre con experiencia en la técnica apreciará que las secuencias polinucleotídicas separadas que codifican para diferentes dominios de la proteína MDRP de la invención, pueden
65 administrarse en más de un vector.

En un segundo aspecto de la invención se proporciona un método para transformar una célula vegetal según se define en la reivindicación 3.

5 La segunda secuencia de ácido nucleico comprende además un origen de replicación viral situado en 5' y 3' de cada uno de dichos bordes izquierdo y derecho. El origen de replicación viral puede seleccionarse de *Mastreviruses* (tales como el virus de la raya del maíz (MSV), virus enano del trigo (WDV), *Curtoviruses* (tal como el virus de la rizomanía de la remolacha (BCTV), *Begomoviruses* (tales como virus del mosaico dorado del frijol (BGMV), virus del mosaico dorado del tomate (TGMV), virus del enrollamiento de la hoja de la calabaza (SqLCV) y virus del enrollamiento de la hoja amarilla del tomate (TYLCV) (Gutierrez, (1999) CMLS, Cell. Mol. Life Sci. 56 (1999) 313-329).

El promotor nuclear de la planta puede seleccionarse de cualquier promotor adecuado, tal como promotores constitutivos, inducibles, específicos de tejido, u otros promotores como se indica en la presente descripción.

15 El término "inducible", como se aplica a un promotor, es bien comprendido por los expertos en la técnica. En esencia, la expresión bajo el control de un promotor inducible es "activado" o aumentado en respuesta a un estímulo aplicado (que puede generarse dentro de una célula o proporcionarse exógenamente). La naturaleza del estímulo varía entre los promotores. Algunos promotores inducibles causan niveles de expresión poco o indetectables (o ninguna expresión) en ausencia del estímulo apropiado. Otros promotores inducibles causan una expresión constitutiva detectable en ausencia del estímulo. Sea cual sea el nivel de expresión en ausencia del estímulo, la expresión de cualquier promotor inducible aumenta en presencia del estímulo correcto. La situación preferida es cuando el nivel de expresión aumenta tras la aplicación del estímulo pertinente mediante una cantidad efectiva para alterar una característica fenotípica. Por lo tanto, puede usarse un promotor inducible (o "conmutable") que provoca un nivel básico de expresión en ausencia del estímulo cuyo nivel es demasiado bajo para producir un fenotipo deseado (y de hecho puede ser cero). Tras la aplicación del estímulo, la expresión se aumenta (o activa) hasta un nivel, lo que produce aproximadamente el fenotipo deseado. Un ejemplo de un promotor inducible es el interruptor génico inducible por etanol descrito en Caddick y otros (1998) Nature Biotechnology 16: 177-180. En la técnica se conocen varios promotores inducibles.

30 Se pueden usar promotores regulados químicamente para modular la expresión de un gen o una secuencia polinucleotídica de la invención en una planta mediante la aplicación de un regulador químico exógeno. Dependiendo del objetivo, el promotor puede ser un promotor químicamente inducible, donde la aplicación del producto químico induce la expresión del gen, o un promotor químicamente reprimible, donde la aplicación del producto químico reprime la expresión del gen. Los promotores químicamente inducibles son conocidos en la técnica e incluyen, pero no se limitan a, el promotor In2-2 del maíz, que se activa por los protectores herbicidas de la bencenosulfonamida, el promotor GST del maíz, que se activa por compuestos electrofílicos hidrofóbicos que se usan como herbicidas pre emergentes, y el promotor PR-1a del tabaco, que se activa por el ácido salicílico. Otros promotores de interés químicamente regulados incluyen promotores que responden a esteroides (ver, por ejemplo, el promotor inducible por glucocorticoides en Schena y otros (1991) Proc. Natl. Acad. Sci. USA 88:10421-10425 y McNellis y otros (1998) Plant J. 14(2):247-257) y promotores inducibles por tetraciclina y reprimibles por la tetraciclina (ver, por ejemplo, Gatz y otros (1991) Mol. Gen. Genet. 227:229-237, y Patentes de Estados Unidos Núms. 5,814,618 y 5,789,156), donde se desea una expresión mejorada en tejidos particularmente, se pueden utilizar promotores específicos de tejidos.

45 Los promotores específicos del tejido incluyen los descritos por Yamamoto y otros (1997) Plant J. 12(2):255-265; Kawamata y otros (1997) Plant Cell Physiol. 38(7):792-803; Hansen y otros (1997) Mol. Gen. Genet. 254(3):337-343; Russell y otros (1997) Transgenic Res. 6(2):157-168; Rinehart y otros (1996) Plant Physiol. 112(3):1331-1341; Van Camp y otros (1996) Plant Physiol. 112(2):525-535; Canevascini y otros (1996) Plant Physiol. 112(2):513-524; Yamamoto y otros (1994) Plant Cell Physiol. 35(5):773-778; Lam (1994) Results Probl. Cell Differ. 20:181-196; Orozco y otros (1993) Plant Mol Biol. 23(6):1129-1138; Matsuoka y otros (1993) Proc Natl. Acad. Sci. USA 90(20):9586-9590; y Guevara-Garcia y otros (1993) Plant J. 4(3):495-505.

50 Se pueden usar además los denominados promotores constitutivos en los métodos de la presente invención. Los promotores constitutivos incluyen, por ejemplo, el promotor ubiq3At, el promotor CaMV 35S (Odell y otros (1985) Nature 313:810-812); actina de arroz (McElroy y otros (1990) Plant Cell 2:163-171); ubiquitina (Christensen y otros (1989) Plant Mol. Biol. 12:619-632 y Christensen y otros (1992) Plant Mol. Biol. 18:675-689); pEMU (Last y otros (1991) Theor. Appl. Genet. 81:581-588); MAS (Velten y otros (1984) EMBO J. 3:2723-2730); promotor ALS (solicitud de Estados Unidos núm de serie 08/409,297), y similares. Otros promotores constitutivos incluyen aquellos en las Patentes de Estados Unidos núms. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; y 5,608,142.

60 Una 'secuencia de ácido nucleico donador' es una que comprende una secuencia de interés que es capaz de reemplazar un sitio objetivo cromosómico del receptor. Las secuencias de interés adecuadas incluyen aquellas que comprenden secuencias reparadoras de polinucleótidos reparadores, es decir secuencias que introducirán ganancia o pérdida de función en una secuencia objetivo que proporcione un beneficio deseado en términos de genotipo para un propósito final específico. Por ejemplo, la ganancia de secuencias funcionales puede incluir la introducción de mutaciones puntuales, por ejemplo en el gen de acetolactato sintasa (ALS) de *Arabidopsis* que resulta en la resistencia al herbicida sulfonilurea de plantas modificadas (Li y otros, (1992) Plant Physiol, 100: 662-668). Otras manipulaciones genéticas de la expresión génica que pueden realizarse usando métodos de la invención incluyen el aumento de la

expresión génica en la planta, por ejemplo al i) añadir un potenciador de la expresión a la región promotora de un gen deseado; ii) añadir dominios responsables de la expresión específica de tejido. Otros usos de los métodos de la invención incluyen la adición *en la planta* de dominios de etiquetas de epitopos para la detección del anticuerpo con la secuencia (s) de ácido(s) nucleico(s) (por ejemplo gen (es) de interés o la fusión de secuencia específica de ácido nucleico de interés, tales como genes, con marcadores visuales tales como GFP y luciferasa y similares. La pérdida de las secuencias funcionales puede incluir la introducción de mutaciones puntuales o deleciones en genes específicos, por ejemplo, los genes responsables de romper vainas en Brassica (hidrolasas de pared celular tales como poligalacturonasas, las regiones reguladoras de los correspondientes genes de plantas (WO 1997/013865 y WO 1997/000417). La secuencia de ácido nucleico donador puede comprender además secuencias que codifican para la producción de una o más proteínas de interés en una planta. Las secuencias adecuadas incluyen las que codifican proteínas heterólogas o exógenas de interés para propósitos de producción de proteínas, o secuencias heterólogas o exógenas de interés que se usan para introducir nuevas rutas genéticas en las plantas que no poseen tales rutas, tales como genes de la ruta de fijación de nitrógeno y/o secuencias de ADNc, genes para resistencia a la sequía, resistencia a insectos y similares. Adecuadamente, la producción de la proteína dirigida puede mejorarse empleando métodos de la invención sobre proteínas de la planta y proteínas farmacéuticas para uso en mamíferos, que incluyen el hombre, tales como insulina, preproinsulina, proinsulina, glucagón, interferones tales como interferón α , interferón β , interferón γ , factores de coagulación de la sangre seleccionados de Factor VII, VIII, IX, X, XI y XII, hormonas de fertilidad que incluyen hormona luteinizante, factores de crecimiento de la hormona estimulante del foliculo que incluyen factor de crecimiento epidérmico, factor de crecimiento derivado de plaquetas, factor estimulante de colonias de granulocitos y similares, prolactina, oxitocina, hormona estimulante del tiroides, hormona adrenocorticotrópica, calcitonina, hormona paratiroidea, somatostatina, eritropoyetina (EPO), enzimas tales como β -glucocerebrosidasa, hemoglobina, albúmina sérica, colágeno, proteína tóxica de insectos de *Bacillus thuringiensis*; proteína de resistencia a herbicidas (glifosato); proteínas de tolerancia a la sal; proteínas de mejora nutricional implicadas en la biosíntesis de compuestos fenólicos, almidones, azúcares, alcaloides, vitaminas y vacunas comestibles, y similares. Además, el método de la invención puede usarse para la producción de anticuerpos monoclonales específicos o fragmentos activos de estos y de enzimas industriales.

Todas las proteínas mencionadas anteriormente son de tipo vegetal y humano. Otras proteínas que se contemplan para la producción en la presente invención incluyen proteínas para el uso en cuidados veterinarios y pueden corresponder a homólogos en animales de proteínas humanas, tales como las proteínas humanas mencionadas anteriormente.

Las dos secuencias flanqueantes pueden seleccionarse de secuencias subterminales que son capaces de asociarse con el dominio de unión del donador de la MDRP y secuencias de ADN cromosómicas bajo la mediación de dominios de unión a ADN específicos, tales como *tnpA* como se menciona en la presente descripción. Las secuencias subterminales adecuadas incluyen las regiones subterminales de la superfamilia CACTA de elementos transponibles que comprenden motivos repetitivos con afinidad a dominio (s) de proteína similar a *tnpA* y similares.

El método descrito anteriormente, comprende además, típicamente, introducir una secuencia de ácido nucleico que comprende un promotor nuclear de la planta unido operativamente a una tercera secuencia de ácido nucleico que codifica para una proteína replicasa viral. Tales proteínas replicasa viral incluyen las proteínas Rep de Geminivirus (Gutierrez (1999) CMLS Cell Mol Life Sci., 56: 313-329) y estas deben ser capaces de unirse a cada uno de dichos orígenes de replicación viral proporcionados en cada extremo de las dos secuencias flanqueantes del dominio de ADN donador.

En una alternativa, el método de la invención proporciona dicha segunda secuencia de ácido nucleico que comprende un sitio de unión del iniciador (PBS) que se fusiona con el extremo 3' de una secuencia del borde flanqueante derecho como se describe en la presente descripción. En este aspecto alternativo de la invención, puede ser utilizada una secuencia de ácido nucleico adicional que comprende un promotor nuclear de planta unido operativamente a una tercera secuencia de ácido nucleico que codifica para una proteína de transcriptasa inversa, tal como RT-RH-tnt1, o un ADN que codifica retrotransposón de ARN o ADN que codifica el ARN retroviral. Las secuencias de ácido nucleico adecuadas incluyen la secuencia de ADN de la transcriptasa inversa del retrotransposón de virus del mosaico de la coliflor (CaMV), Ty1 de levadura, Ty1-copia de la planta, Ty3-gypsy (Kumar & Bennetzen, (1999) Annual Review of Genetics, 33: 479-532) y otros retrotransposones LTR. El sitio de unión del iniciador puede seleccionarse de un sitio de unión del iniciador específico de retrotransposón, y un sitio de unión del iniciador de un retrovirus. Preferentemente, el PBS es un sitio de unión del iniciador específico de retrotransposón.

Dicha tercera secuencia de ácido nucleico puede seleccionarse además de las secuencias de ácido nucleico de la transcriptasa inversa a partir de retrotransposón de la planta Ty1-copia, Ty3-gypsy (Kumar & Bennetzen, (1999) Annual Review of Genetics, 33: 479-532) y otros retrotransposones LTR, por ejemplo, transcriptasa inversa-RNasa H.

En un tercer aspecto de la invención se proporciona una célula vegetal según se define en la reivindicación 8.

Una célula vegetal de la invención comprende una proteína de recombinación de múltiples dominios que comprende un tercer dominio que es un dominio de unión a objetivo cromosómico y un cuarto dominio que es un dominio inductor de la recombinación.

- Una célula vegetal de la invención puede ser una seleccionada de tabaco (*Nicotiana tabacum*) y otras especies de *Nicotiana*, arabidopsis, patata, grano (maíz), canola (colza), arroz, trigo, cebada, brassica sp. tales como coliflor, brócoli (por ejemplo, brotes verdes y púrpuras), repollo (por ejemplo, coles rojas, verdes y blancas), col rizada, coles de Bruselas, algodón, algas (por ejemplo, especies de verde azul), lemnochora o musgo (por ejemplo, *physcomitrella patens*), tomate, pimiento, calabaza, girasol, soja, zanahoria, melones, vid, lechuga, fresa, remolacha azucarera, guisantes y sorgo. Especialmente preferidas son las células vegetales de la invención que se seleccionan de algodón, arroz, especies de semillas oleaginosas de Brassica tales como canola, grano (maíz) y soja.
- En un cuarto aspecto de la invención se proporciona un método para producir al menos una especie de ARN heteróloga o exógena en una planta que comprende:
- 1) introducir en una célula vegetal regenerable una secuencia de ácido nucleico que comprende un promotor nuclear de planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con el primer aspecto de la invención;
 - 2) introducir en dicha célula vegetal regenerable una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador; que es monocatenario; y
 - ii) al menos dos secuencias flanqueantes que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma; y
 - iii) dos orígenes de replicación viral situados 5' y 3' a dichos bordes izquierdo y derecho, en donde dicha segunda secuencia de ácido nucleico está operativamente unida a un promotor nuclear de la planta; e
 - 3) introducir en dicha célula vegetal regenerable una tercera secuencia de ácido nucleico que codifica para una proteína de replicasa viral, en donde la tercera secuencia de ácido nucleico está operativamente unida a un promotor nuclear de planta o un promotor viral nativo;
 - 4) cultivar dicha célula vegetal regenerable de las etapas 1) a 3);
 - 5) seleccionar una célula vegetal de 4);
 - 6) regenerar una planta a partir de la célula vegetal de 5); y
 - 7) cultivar la planta de 6.
- Naturalmente, el hombre con experiencia en la técnica apreciará que las regiones flanqueantes y la proteína replicasa serán capaces de unirse a un dominio de unión del donador de dicha proteína de recombinación multidominio, y al origen de las secuencias de replicación viral, respectivamente.
- En un quinto aspecto de la invención se proporciona un método para producir al menos una especie de ARN heteróloga o exógena en una planta que comprende:
- 1) introducir en una célula vegetal regenerable una secuencia de ácido nucleico que comprende un promotor nuclear de planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con el primer aspecto de la invención;
 - 2) introducir en dicha célula vegetal regenerable una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador que es monocatenario;
 - ii) al menos dos secuencias flanqueantes que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma, en donde dicha segunda secuencia de ácido nucleico está unida operativamente a un promotor nuclear de la planta; y
 - iii) un sitio de unión del iniciador que se fusiona con el extremo 3' de dicha secuencia de borde flanqueante derecho;
 - 3) introducir en dicha célula vegetal regenerable una tercera secuencia de ácido nucleico que codifica para una proteína de transcriptasa inversa, en donde la tercera secuencia de ácido nucleico está unida operativamente a un promotor nuclear de planta;
 - 4) cultivar dicha célula vegetal regenerable de las etapas 1) a 3);
 - 5) seleccionar una célula vegetal de 4);
 - 6) regenerar una planta a partir de la célula vegetal de 5); y
 - 7) cultivar la planta de 6.
- En un sexto aspecto de la invención se proporciona una secuencia polinucleotídica aislada que comprende una secuencia de ácido nucleico donador que es monocatenario, al menos dos secuencias flanqueantes situadas adyacentes a cada extremo de dicha secuencia de ácido nucleico donador formando los bordes izquierdo y derecho de la misma, y un origen de replicación viral situado 5' y 3' a dichos bordes izquierdo y derecho, respectivamente, adecuados para su uso en un método de conformidad con el segundo aspecto de la invención. Además, se proporciona una secuencia polinucleotídica aislada de conformidad con este sexto aspecto de la invención que comprende una secuencia de ácido nucleico adicional que comprende un promotor nuclear de planta o un promotor viral nativo unido operativamente a una secuencia de ácido nucleico que codifica para una proteína de replicasa viral adecuada para su uso en un método de conformidad con el segundo o cuarto aspecto de la invención.
- Se proporcionan además los vectores de ácidos nucleicos adecuados para la transformación de una célula vegetal o bacteriana y que incluyen un polinucleótido que codifica una MDRP de la invención. Tales vectores de ácido nucleico

son adecuados para la transformación de una célula vegetal o de una célula procariota, tal como una célula de *Agrobacterium*.

Las células huésped que contienen un polinucleótido heterólogo o vectores de ácido nucleico de la invención se abarcan además en el ámbito de la invención e incluyen células vegetales, y pueden comprenderse en una planta, una parte de planta o un propágulo vegetal o un extracto o derivado de una planta o en un cultivo de células vegetales. Por lo tanto, las plantas que comprenden células vegetales de la invención se abarcan además dentro del ámbito de la invención. Las plantas adecuadas de la invención incluyen plantas que se seleccionan del grupo que consiste en tabaco (*Nicotiana tabacum*) y otras *Nicotiana* especies, tales como *Nicotiana benthamiana*, zanahoria, aceite vegetal y semillas oleaginosas de *Brassica*, melones, pimientos, vid, lechuga, fresa, remolacha azucarera, trigo, cebada, (grano)maíz, arroz, soja, guisantes, sorgo, girasol, tomate, algodón y patata.

Las plantas preferidas de la invención son aquellas que se seleccionan del grupo que consiste en algodón, arroz, especies de semillas oleaginosas de *Brassica* tales como canola, grano(maíz) y soja.

Naturalmente, el hombre con experiencia en la técnica apreciará que las secuencias de ADN terminador estarán presentes en los constructos usados en la invención. Se contempla un terminador como una secuencia de ADN en el extremo de una unidad de transcripción que indica la terminación de la transcripción. Estos elementos son secuencias 3' no traducidas que contienen señales de poliadenilación, que actúan para provocar la adición de secuencias de poliadenilato al extremo 3' de los transcritos primarios. Para la expresión en células vegetales la secuencia del terminador transcripcional de la nopalina sintasa (A. Depicker y otros, 1982, J. of Mol. & Applied Gen. 1:561-573) sirve como señal de terminación transcripcional.

Aquellos con experiencia en la técnica son capaces de construir vectores y diseñar protocolos para secuencias de ácidos nucleicos recombinantes o expresión génica. [0081] Los vectores adecuados pueden elegirse o construirse, que contienen secuencias reguladoras adecuadas, que incluyen secuencias promotoras, fragmentos terminadores, secuencias de poliadenilación, secuencias potenciadoras, genes marcadores y otras secuencias según sea adecuado. Para más detalles ver, por ejemplo, *Molecular Cloning: a Laboratory Manual*, 2da edición, Sambrook y otros, 1989, Cold Spring Harbor Laboratory Press. Muchas técnicas y protocolos conocidos para la manipulación de los ácidos nucleicos, por ejemplo en la preparación de los constructos de ácido nucleico, mutagénesis, secuenciación, introducción de ADN en las células y expresión génica, y análisis de proteínas, se describen en detalle en *Current Protocols in Molecular Biology*, Ausubel y otros eds., John Wiley & Sons, 1992.

Procedimientos específicos y vectores previamente usados con gran éxito en las plantas son descritos por Bevan (Nucl. Acids Res. 12, 8711-8721 (1984)) y Guerineau y Mullineaux (1993) (Plant transformation and expression vectors. En *Plant Molecular Biology Labfax* (Croy RRD ed.) Oxford, BIOS Scientific Publishers, págs. 121-148).

Naturalmente, el destinatario experto apreciará que cada secuencia de ácido nucleico introducida estará típicamente bajo control regulador de su propio promotor exógeno y terminador. Cuando dos o más proteínas objetivo se destinan para producirse a partir de un único ARN portador, es preferible que puedan separarse fácilmente, por ejemplo, uniéndose a diferentes anticuerpos específicos de la proteína (monoclonal o policlonal) en la fase de recolección del sistema de cultivo de células vegetales.

Los marcadores genéticos seleccionables pueden facilitar la selección de plantas transgénicas y éstas pueden consistir en genes quiméricos que confieren fenotipos seleccionables tales como resistencia a antibióticos tales como espectinomycin, estreptomycin, kanamicina, neomicina, higromicina, puramicina, fosfinotricina, clorsulfurón, metotrexato, gentamicina, espectinomycin, imidazolinonas y glifosato.

Cuando se introducen secuencias de ácidos nucleicos seleccionadas de conformidad con la presente invención en una célula, deben tenerse en cuenta ciertas consideraciones, bien conocidas por los expertos en la técnica. El ácido nucleico que se inserta debe ensamblarse dentro de un constructo, que contiene elementos reguladores efectivos, que impulsarán la transcripción. Debe haber un método para transportar el constructo dentro de la célula. Una vez que el constructo está dentro de la membrana celular, la integración en el material cromosómico endógeno se producirá o no. Por último, en lo que a plantas se refiere el tipo de célula objetivo debe ser de manera que las células puedan regenerarse en plantas completas.

Las plantas transformadas con segmentos de ADN que contienen secuencias de interés tal como se proporcionan en la presente descripción pueden producirse mediante técnicas estándares, que ya son conocidas para la manipulación genética de plantas. El ADN puede transformarse en células vegetales usando cualquier tecnología adecuada, tal como un vector de plásmido Ti desarmado llevado por *Agrobacterium* que explota su capacidad natural de transferencia génica (EP-A-270355, EP-A-0116718, NAR 12(22) 8711-87215 1984), bombardeo de partícula o microproyectil (US 5100792, EP-A-444882, EP-A-434616) microinyección (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green y otros (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporación (EP 290395, WO 8706614) otras formas de absorción directa de ADN (DE 4005152, WO 9012096, US 4684611), absorción de ADN mediado por liposoma (por ejemplo (por ejemplo Freeman y otros *Plant Cell Physiol.* 29: 1353 (1984)), o el método con vórtex (por ejemplo Kindle,

PNAS U.S.A. 87: 1228 (1990d) Los métodos físicos para la transformación de células vegetales se revisan en Oard, 1991, *Biotech. Adv.* 9: 1-11.

Así, una vez que se ha identificado una secuencia de ácido nucleico o un gen, puede reintroducirse en células vegetales usando técnicas bien conocidas por aquellos con experiencia en la técnica para producir plantas transgénicas del fenotipo apropiado.

La transformación de *Agrobacterium* es ampliamente usada por aquellos con experiencia en la técnica para transformar especies dicotiledóneas. La producción de plantas transgénicas estables, fértiles en casi todas las plantas monocotiledóneas económicamente relevantes es ahora además rutinario: (Toriyama, y otros (1988) *Bio/Technology* 6, 1072-1074; Zhang, y otros (1988) *Plant Cell Rep.* 7, 379-384; Zhang, y otros (1988) *Theor. Appl. Genet* 76, 835-840; Shimamoto, y otros (1989) *Nature* 338, 274-276; Datta, y otros (1990) *Bio/Technology* 8, 736-740; Christou, y otros (1991) *Bio/Technology* 9, 957-962; Peng, y otros (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, y otros (1992) *Plant Cell Rep.* 11, 585-591; Li, y otros (1993) *Plant Cell Rep.* 12, 250-255; Rathore, y otros (1993) *Plant Molecular Biology* 21, 871-884; Fromm, y otros (1990) *Bio/Technology* 8, 833-839; Gordon-Kamm, y otros (1990) *Plant Cell* 2, 603-618; D'Halluin, y otros (1992) *Plant Cell* 4, 1495-1505; Walters, y otros (1992) *Plant Molecular Biology* 18, 189-200; Koziel, y otros (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, y otros (1993) *Plant Physiology* 102, 1077-1084; Somers, y otros (1992) *Bio/Technology* 10, 1589-1594; WO92/14828). Particularmente, la transformación mediada por *Agrobacterium* es ahora un método de transformación alternativo altamente eficiente en monocotiledóneas (Hiei y otros (1994) *The Plant Journal* 6, 271-282).

La generación de plantas transgénicas fértiles se ha logrado en los cereales de arroz, maíz, trigo, avena y cebada (revisado en Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, y otros (1992) *Bio/Technology* 10, 667-674; Vain y otros, 1995, *Biotechnology Advances* 13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 página 702). Wan y Lemaux (1994) *Plant Physiol.* 104: 37-48 describen técnicas para la generación de gran número de plantas de cebada fértiles transformadas de manera independiente.

El bombardeo con microproyector, la electroporación y la absorción directa de ADN se prefieren cuando el *Agrobacterium* es ineficiente o ineficaz. Alternativamente, puede emplearse una combinación de diferentes técnicas para aumentar la eficiencia del proceso de transformación, por ejemplo bombardeo con micropartículas recubiertas con *Agrobacterium* (EP-A-486234) o bombardeo con microproyector para inducir la herida seguido de cultivo con *Agrobacterium* (EP-A-486233).

Después de la transformación, una planta puede regenerarse, por ejemplo a partir de células individuales, tejido de callos o discos foliares, como es estándar en la técnica. Casi cualquier planta puede regenerarse completamente a partir de células, tejidos y órganos de la planta. Las técnicas disponibles se revisan en Vasil y otros, *Cell Culture and Somatic Cell Genetics of Plants*, Vol. I, II y III, *Laboratory Procedures and Their Applications*, Academic Press, 1984, y Weiss Bach y Weiss Bach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

La elección particular de una tecnología de transformación se determinará por su eficiencia para transformar ciertas especies vegetales, así como la experiencia y preferencia de la persona que practica la invención con una metodología particular de elección. Será evidente para los expertos en la técnica que la elección particular de un sistema de transformación para introducir ácido nucleico en células vegetales no es esencial ni una limitación de la invención, ni tampoco la elección de la técnica para la regeneración de plantas.

La invención comprende además una célula huésped transformada con vectores o constructos como se ha expuesto anteriormente, especialmente una planta o una célula microbiana. Por lo tanto, se proporciona una célula huésped, tal como una célula vegetal, que incluye secuencias de nucleótidos de la invención como se indica en la presente descripción. Dentro de la célula, la secuencia de nucleótidos puede incorporarse dentro del cromosoma.

Además de conformidad con la invención se proporciona una célula vegetal que incorpora en su genoma al menos una secuencia de nucleótidos, particularmente secuencias de nucleótidos heterólogas, tal como se proporciona por la presente invención bajo el control operativo de secuencias reguladoras para el control de la expresión como se describe en la presente descripción. La secuencia codificante puede estar unida operativamente a una o más secuencias reguladoras que pueden ser heterólogas o extrañas a las secuencias de ácido nucleico empleadas en la invención, tales como aquellas que no están naturalmente asociadas con la secuencia o secuencias de ácido nucleico para esta (su) expresión. La secuencia de nucleótidos de conformidad con la invención se puede colocar bajo el control de un promotor inducible externamente para colocar la expresión bajo el control del usuario. Un aspecto adicional de la presente invención proporciona un método para preparar una célula vegetal de este tipo que implica la introducción de secuencia (s) de ácidos nucleicos contempladas para su uso en la invención o un vector adecuado que incluya la secuencia (s) contempladas para uso en la invención en una célula vegetal y que cause o permita la recombinación entre el vector y el genoma de la célula vegetal para introducir dichas secuencias en el genoma. La invención se extiende a células vegetales que contienen una secuencia de nucleótidos de conformidad con la invención como resultado de la introducción de la secuencia de nucleótidos en una célula antecesora.

El término "heterólogo" puede usarse para indicar que el gen/secuencia de nucleótidos en cuestión se ha introducido en

- dichas células de la planta o un antecesor de esta, usando ingeniería genética, es decir, mediante intervención humana. Puede proporcionarse una célula vegetal transgénica, es decir, transgénica para la secuencia de nucleótidos en cuestión. El transgén puede estar en un vector extragenómico o incorporarse, preferentemente de forma estable, al genoma. Un gen heterólogo puede reemplazar un gen endógeno equivalente, es decir, uno que normalmente realiza la misma función o una función similar, o la secuencia insertada puede ser adicional al gen endógeno u otra secuencia. Una ventaja de la introducción de un gen heterólogo es la capacidad de colocar la expresión de una secuencia bajo el control de un promotor de elección, para ser capaz de influir en la expresión de conformidad con la preferencia. Además, mutantes, variantes y derivados del gen silvestre, por ejemplo con mayor actividad que el silvestre, pueden usarse en lugar del gen endógeno. Las secuencias de nucleótidos heterólogas, o exógenas o extrañas, a una célula vegetal pueden no ser de origen natural en células de ese tipo, variedad o especie. Así, una secuencia de nucleótidos puede incluir una secuencia codificante o derivada de un tipo particular de célula vegetal o especie o variedad de planta, colocada dentro del contexto de una célula vegetal de un tipo o especie o variedad de planta diferente. Una posibilidad adicional es para una secuencia de nucleótidos que se coloca dentro de una célula en la que ésta o un homólogo se encuentra naturalmente, pero en donde la secuencia de nucleótidos está unida y/o adyacente al ácido nucleico que no se produce naturalmente dentro de la célula o células de ese tipo o especie o variedad de planta, tal como unida operativamente a una o más secuencias reguladoras, tales como una secuencia promotora, para el control de la expresión. Una secuencia dentro de una planta u otra célula huésped puede ser identificablemente heteróloga, exógena o extraña.
- Las plantas que incluyen una célula vegetal de conformidad con la invención se proporcionan además, junto con cualquier parte o propágulo de la misma, semilla, progenie y descendientes autofecundados o híbridos. Se proporcionan particularmente plantas de cultivo transgénico, que se han diseñado para llevar genes identificados como se ha indicado anteriormente. Los ejemplos de plantas adecuadas incluyen tabaco (*Nicotiana tabacum*) y otras especies *Nicotiana*, zanahoria, vegetales y semillas oleaginosas de *Brassica*, melones, pimiento, vid, lechuga, fresa, remolacha azucarera, trigo, cebada, (grano)maíz, arroz, soja, guisantes, sorgo, girasol, tomate, algodón y patata. Las plantas transgénicas especialmente preferidas de la invención incluyen algodón, arroz, especies de semillas oleaginosas de *Brassica* tales como canola, grano(maíz) y soja.
- Además de una planta, la presente invención proporciona cualquier clon de una planta, semilla, progenie y descendientes autofecundados o híbridos de ese tipo, y cualquier parte de cualquiera de estos, tales como recortes, semillas. La invención proporciona cualquier propágulo vegetal que es cualquier parte que se puede usar en la reproducción o propagación, sexual o asexual, que incluyen estacas, semillas y así sucesivamente. Se abarca además por la invención una planta que es un vástago, clon o descendiente de una planta propagado sexual o asexualmente de una planta o cualquier parte o propágulo de dicha planta, vástago, clon o descendiente de ese tipo.
- La presente invención abarca además el producto de expresión de polipéptido de una molécula de ácido nucleico de conformidad con la invención como se describe en la presente descripción o puede obtenerse de conformidad con la información y las sugerencias de la presente descripción. Se proporcionan además, métodos para preparar dicho producto de expresión mediante la expresión a partir de una secuencia de nucleótidos que codifica por lo tanto bajo condiciones adecuadas en células huésped adecuadas por ejemplo de *E.coli*. Aquellos con experiencia en la técnica son capaces de construir vectores y diseñar protocolos y sistemas para la expresión y recuperación de los productos de expresión génica recombinante.
- "Homología" con relación a una secuencia de aminoácidos o una secuencia polipeptídica producida por el método de la invención, puede usarse para referirse a la identidad o similitud, preferentemente a la identidad. Como se ha indicado anteriormente, un alto nivel de identidad de aminoácidos puede estar limitado a dominios o regiones funcionalmente significativas.
- En algunas modalidades, un alelo, variante, derivado, derivado mutante, mutante u homólogo de la secuencia específica pueden mostrar poca homología global, por ejemplo aproximadamente 20%, o aproximadamente 25%, o aproximadamente 30%, o aproximadamente 35%, o aproximadamente 40% o aproximadamente 45%, con la secuencia específica. Sin embargo, en dominios o regiones funcionalmente significativas, la homología de aminoácidos puede ser mucho mayor. Los dominios o regiones putativos funcionalmente significativos pueden identificarse usando los procesos de bioinformática, que incluyen la comparación de las secuencias de homólogos.
- Pueden combinarse los dominios o regiones funcionalmente significativos de diferentes polipéptidos para la expresión de ácido nucleico codificante como una proteína de fusión. Por ejemplo, las propiedades particularmente ventajosas o deseables de diferentes homólogos pueden combinarse en una proteína híbrida, de manera que el producto de expresión resultante, puede incluir si es apropiado, fragmentos de diversas proteínas parentales.
- La similitud de las secuencias de aminoácidos puede ser como se define y determina con el programa TBLASTN, de Altschul y otros (1990) *J. Mol. Biol.* 215: 403-10, que es de uso estándar en la técnica. Particularmente, TBLASTN 2.0 puede usarse con la Matrix BLOSUM62 y las penalizaciones GAP: existencia: 11, extensión: 1. Otro programa estándar que puede usarse es BestFit, que es parte del Paquete de Wisconsin, Versión 8, Septiembre de 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, Estados Unidos, Wisconsin 53711). BestFit fabrica una alineación óptima del mejor segmento de similitud entre dos secuencias. Las alineaciones óptimas se encuentran

insertando interrupciones para maximizar el número de coincidencias usando el algoritmo de homología local de Smith y Waterman (*Adv. Appl. Math.* (1981) 2: 482-489). Otros algoritmos incluyen GAP, que usa el algoritmo de Needleman y Wunsch para alinear dos secuencias completas que maximiza el número de coincidencias y minimiza el número de interrupciones. Como en cualquier algoritmo, generalmente se usan los parámetros por defecto, que para GAP son una penalización de creación de interrupción = 12 y la penalización de extensión de interrupción = 4. Alternativamente, puede usarse una penalización de creación de interrupción de 3 y una penalización de extensión de interrupción de 0,1. El algoritmo FASTA (que usa el método de Pearson y Lipman (1988) *PNAS USA* 85: 2444-2448) es una alternativa adicional.

El uso de cualquiera de los términos "homología" y "homóloga" en la presente descripción no implica ninguna relación evolutiva necesaria entre las secuencias comparadas, manteniendo por ejemplo, el uso estándar de términos tales como "recombinación homóloga", que requiere simplemente que dos secuencias de nucleótidos sean suficientemente similares para recombinarse en las condiciones apropiadas. Se describe más abajo, una discusión adicional de polipéptidos de conformidad con la presente invención, que puede codificarse por ácido nucleico de conformidad con la presente invención.

La invención se ilustra ahora en los siguientes ejemplos no limitantes y figuras.

Figuras

Figura 1: Estructura de la proteína de recombinación de múltiples dominios usada para el direccionamiento génico en células vegetales. CBD = dominio de unión cromosómica; TBD = dominio de unión al objetivo; DBD = dominio de unión del donador; RID = dominio inductor de la recombinación.

Figura 2: Sistemas vectoriales adecuados para la amplificación de los fragmentos de ADN donadores en las células vegetales.

- (A) Sistema para la amplificación del ADN viral: VOR - origen de la replicación viral; STR1, STR2 - regiones subterminales con afinidad al dominio de unión del donador de la proteína de recombinación multidominio; DONADOR - fragmento donador del ADN para el direccionamiento génico; REP -proteína de replicasa viral que se une a VOR e induce la replicación del ADN donador.
- (B) Sistema de amplificación de ADN basado en retrotransposón de ARN. El fragmento donador se fusiona a un sitio de unión del iniciador (PBS) y se expresa a partir de un promotor fuerte específico de planta. El ARN generado se usa posteriormente para una reacción de transcripción inversa usando transcriptasa inversa específica del retrotransposón. El sitio de unión del iniciador - PBS usado para la hibridación del iniciador (tRNA-Met usado como un iniciador) y la transcripción inversa del ARN donador en el ADN; RT - RH-transcriptasa inversa específica del retrotransposón-proteína RNasa H que se une a PBS e inicia la transcripción inversa.

Figura 3: Presentación esquemática del direccionamiento génico usando proteína de recombinación multidominio.

Figura 4: Constructo deletéreo ALG125 usado para la transformación de *Arabidopsis*.

Figura 5: Un conjunto de constructos para el direccionamiento génico usando el sistema de amplificación del donador viral.

Figura 6: Conjunto de constructos para el direccionamiento génico usando el sistema de amplificación de donadores basados en retrotransposón de ARN.

Figura 7: Constructo ALG460 para la mutagénesis del gen acetolactato sintasa del tabaco (ALS).

Figura 8: Un conjunto de constructos para el direccionamiento de la quelatosa de magnesio en el arroz.

Sección experimental

Concepto

Se desarrolla una técnica robusta y reproducible para la modificación específica del sitio en sitios dentro del genoma nuclear de la planta. La esencia del direccionamiento génico tal como se desarrolla en la presente descripción radica en la colocación eficiente del ADN donador junto al ADN de un sitio de direccionamiento y la inducción de la recombinación entre estos dos fragmentos de ADN. Nuestra técnica de direccionamiento génico se basa en dos componentes esenciales:

- (1) una proteína de recombinación de múltiples dominio (MDRP); y
- (2) vector de amplificación donador (DAV).

La MDRP tiene las siguientes funcionalidades:

- (i) capaz de unirse al ADN donador;
- (ii) capaz de unirse al ADN cromosómico;
- (iii) capaz de buscar y situar el ADN objetivo; y
- (iv) capaz de iniciar la hibridación de cadenas de ADN homólogas entre el ADN objetivo y donador.

5

El DAV es una herramienta para la amplificación de un fragmento donador o fragmentos donadores del ADN de interés en una célula vegetal, que proporciona un suministro permanente y abundante de moléculas de ADN donador para un direccionamiento génico eficiente.

10 Proteína de recombinación de múltiples dominios (MDRP)

La proteína de recombinación de múltiples dominios es el núcleo de nuestra tecnología de direccionamiento génico y funciona como un dispositivo eficiente para la inducción de la recombinación entre las moléculas de ADN objetivo y donador. Comprende cuatro dominios principales (1) un dominio de unión al ADN donador; (2) un dominio de unión al ADN cromosómico; (3) un dominio de unión al objetivo; y (4) un dominio inductor de la recombinación (ver Figura 1).

15

El dominio de unión al ADN donador (DBD) se requiere para la captura eficiente de las moléculas de ADN donador y la formación de un complejo proteína ADN donador (DPC).

20

El dominio de unión cromosómica (CBD) se requiere para la unión del DPC al ADN cromosómico de interés, y sirve como un motor para impulsar el DPC donador a lo largo del ADN cromosómico facilitando una búsqueda activa de un sitio objetivo.

25

El Dominio de Unión al Objetivo (TBD) facilita el reconocimiento de un sitio objetivo en el cromosoma y aumenta la eficiencia de la alineación de los fragmentos del ADN donador y fragmentos del ADN objetivo.

El dominio inductor de la recombinación (RID) se requiere para la primera etapa de recombinación homóloga: la hibridación de un ADN donador con el ADN de un sitio de direccionamiento.

30

Sistemas vectoriales adecuados para la amplificación del donador

Para proporcionar un exceso de moléculas de ADN donador para el direccionamiento génico, se diseñaron vectores de amplificación de donadores. Se han desarrollado dos sistemas para la amplificación del donador:

35

- (1) un sistema de amplificación del ADN viral; y
- (2) un sistema de amplificación de ADN basado en retrotransposón de ARN (ver Figura 2).

40

(1) El sistema de amplificación de ADN viral se basa en la colocación de un origen de replicación viral (VOR) a partir de un virus de ADN (por ejemplo a partir de geminivirus de planta tales como el virus de la rizomanía de la remolacha BCTV, virus del mosaico dorado del tabaco (TGMV) virus de la raya del maíz (MSV), etc.) sobre ambos flancos o extremos de un fragmento de ADN donador que comprende el fragmento de ADN donador y secuencias de ADN subterminal con alta afinidad al dominio de unión del donador de la MDRP fusionado a ambos extremos de estos. Por lo tanto, el fragmento de ADN donador unió a cada flanco o extremo, unas secuencias de ADN subterminal y se fusionó a cada una de ellas un VOR como se representa en la Figura 2. La coexpresión de una replicasa específica de VOR genera la replicación del fragmento de ADN de interés que se coloca o localiza entre los dos orígenes virales flanqueantes de la replicación. Como resultado, se producen en el núcleo de la planta una gran abundancia o exceso de fragmentos de ADN monocatenario (ADNmc) o ADN bicatenario (ADNbc) del ADN donador. Las dos secuencias subterminales del ADN adicionales se denominan "STR1" y "STR2" en la Figura 2.

50

(2) El sistema de amplificación de ADN basado en retrotransposón de ARN se basa en la generación del ARN a partir de un casete transgénico que contiene un fragmento de ADN donador (flanqueado por STR1 y STR2) fusionado a un sitio de unión del iniciador específico del retrotransposón (PBS). El PBS es capaz de capturar el ARNt-Met celular y lo usa como un iniciador para una reacción de transcripción inversa. La coexpresión de una secuencia de transcriptasa inversa específica de retrotransposón o de transcriptasa inversa-RNasa H (RT-RH), tal como un ADNc o secuencia génica de longitud completa, facilita la transcripción inversa del transcripto a partir del casete donador en el ADN monocatenario que posteriormente se usa después, como un donador de ADN para el direccionamiento génico.

55

60

En resumen, el sistema vector de amplificación del donador genera una cantidad o exceso constantemente abundante del ADN donador. El dominio de unión del donador (DBD) de la proteína de recombinación de múltiples dominios se une a las regiones STR1 y STR2 de las moléculas del ADN donador. Los dominios de unión cromosómica (CBD) y los dominios de unión al objetivo (TBD) realizan una búsqueda de un sitio de ADN objetivo y colocan una molécula de ADN donador junto al sitio de direccionamiento. El dominio inductor de la recombinación (RID) de la proteína de recombinación de múltiples dominios (MDRP) inicia la hibridación entre el ADN donador y ADN objetivo con la posterior

65

participación de la maquinaria de reparación génica del huésped que lleva a cabo el proceso de reemplazo del fragmento de ADN objetivo con el ADN donador (ver Figura 3).

5 Materiales y Métodos

Diseño de la proteína de recombinación de múltiples dominios (MDRP)

Para construir una proteína de recombinación de múltiples dominios (MDRP) una secuencia parcial de la proteína *tnpA* (1-543 aa) del elemento transponible de maíz *En/Spm* que comprende un dominio de unión al ADN (entre la posición 122 a 427) se usó como un núcleo. Este dominio de ADN tiene una fuerte afinidad con los motivos de 12pb en los subterminales del transposón *En/Spm* (Gierl y otros, (1988) EMBO J, 7, 4045-4053) y representa un dominio de unión del donador (DBD) de MDRP. Se detectaron seis motivos hiperreactivos de 12 pb en la región subterminal izquierda (STR1) y ocho en la región subterminal derecha (STR2) de la secuencia de ADN *En/Spm*, y estas regiones subterminales se usan como regiones flanqueantes para el fragmento donador en el vector de amplificación donador. La proteína *tnpA* media además, la asociación de regiones de ADN de STR1-STR2 con las secuencias de ADN cromosómicas (Masson y otros, (1991) Plant Cell, 3, 73-85), y como resultado puede cumplir la segunda función como un dominio de unión al ADN cromosómico (CBD) de MDRP.

Se demostró que el apareamiento del ADN homólogo - la reacción clave de la recombinación homóloga - puede promoverse por un péptido de 20 aminoácidos derivado de la proteína *recA* de *E. coli* (Voloshing y otros, (1996) Nature, 272, 868-872). Se fusionó el péptido *recA* con la proteína *tnpA* que resulta en la proteína *tnpAA*, y se utilizó como el dominio inductor de la recombinación (RID) de la MDRP. Estos tres dominios fusionados representan la cadena principal de nuestra MDRP. El dominio de unión del donador (DBD) y el dominio de unión cromosómica (CBD) pueden seleccionarse también de otros elementos genéticos móviles tales como transposones *TAM3*, *Ac-Dc*, *mutador* (Hehl y otros, (1991) Plant Mol Biol, 16, 369-371; Kunze y otros, (1987) EMBO J, 6, 1555-1563, Brown y otros, (1992) Genetics, 130, 889-898), grupos Ty1-copia, Ty3-gypsy de los retrotransposones (Kumar & Bennetzen (1999) Ann Review Genet, 33, 479-532), y los dominios de unión al ADN de las endonucleasas de restricción bacterianas (Shulze y otros, (1998) EMBO J, 17, 6757-6766; Kessler & Manta (1990) Gene, 92, 1-248).

El dominio inductor de la recombinación (RID) puede seleccionarse de las proteínas bien caracterizadas que participan en la inducción de la recombinación homóloga tales como *rad51*, *rad52* etc. (Symington, (2002) Microbiology and Molecular Biology Reviews, 66, 630-670).

El dominio de unión al objetivo (TBD) de la MDRP debe diseñarse de conformidad con la secuencia del sitio de direccionamiento. Los dominios de reconocimiento de objetivo ideales pueden representarse por dominios de unión a nucleasas con dedos de zinc (ZFN) y meganucleasa (MN) (Bibikova y otros, (2003) Science, 300, 764-766; Epinat y otros, (2003) Nuc Acid Res, 31, 2952-2962). El TBD puede seleccionarse también a partir de las proteínas de unión a ADN específicas de genes tales como factores de transcripción, proteínas de dominio de MADS, etc (Pellegrini y otros, (1995) Nature, 376, 490-498; Schwechheimer y otros, (1998) Ann Review Plant Physiol Plant Mol Biol., 49, 127-150). Como ejemplo, se utiliza un dominio de unión al ADN de la proteína del dominio MADS de *Arabidopsis*, APETALA1, que reconoce la secuencia de ADN del gen APETALA3 como un dominio de reconocimiento de objetivo.

Diseño del sistema vectorial para la amplificación del donador

Típicamente, un fragmento de ADN de 500-2000 pb con homología al ADN de un sitio de direccionamiento deseable se usa como donador para el direccionamiento génico. Para proporcionar la afinidad de las moléculas de ADN del donador al dominio de unión del donador de la MDRP, las secuencias subterminales STR1 y STR2 del transposón *En/Spm* de maíz que contienen los motivos repetitivos de 12bp se fusionaron con el fragmento donador diseñado (Figura 2). Estas secuencias juegan un papel importante en el reconocimiento y reclutamiento de las moléculas del ADN donador por la MDRP.

Se pueden usar dos sistemas diferentes para la amplificación de moléculas de ADN donador en células vegetales. Para el sistema viral de la amplificación del donador, se utiliza el origen de replicación viral (VOR) a partir del virus de la rizomanía de la remolacha (BCTV) (Stanley y otros, (1986) EMBO J, 5, 1761-1767), que se colocó en ambos extremos del fragmento de ADN donador (Figura 2A). El fragmento de ADN viral que contiene el promotor y el gen *repA* de la replicasa (REP) se administró junto con el vector de amplificación del donador para iniciar la replicación del fragmento donador en las células vegetales (Figura 5). Para las monocotiledóneas se usó un origen de replicación del virus de la raya del maíz de manera similar al origen de replicación del BCTV. La secuencia viral completa puede usarse también como fuente de la replicasa.

El sistema de amplificación de ADN basado en retrotransposón de ARN (Figura 2B) se basa en la generación del ARN a partir de un casete transgénico que contiene un fragmento donador fusionado a un sitio de unión del iniciador específico del retrotransposón (PBS). Se ha usado PBS del retrotransposón *tnt1* de tabaco (Casacuberta y otros, (1993), Nuc Acids Res., 21, 2087-2093), que recluta tRNA-Met citoplásmico como iniciador para la transcripción inversa.

Las secuencias parciales del polipéptido expresado por *tnt1* que representan los dominios de la endonucleasa (En) y

transcriptasa inversa (RT), o combinación de los dominios de En, RT y RNasa H (RH) (En-RT-RH) se administró bajo un promotor específico de la planta para generar la enzima transcriptasa inversa (Figura 6).

Transformación de las plantas de Arabidopsis

5

La transformación de las plantas de *Arabidopsis* se realizó como se describe en Clough & Bent (1998) Plant Journal 16:735-743). *Agrobacterium tumefaciens* cepa GV3101 (Koncz & Schell (1986) Mol Gen Genet 204:383-396) se usó para la transformación.

10 Transformación de explantes de hojas de tabaco con la cepa AGL1 de Agrobacterium

Todos los artículos se esterilizan en autoclave antes de su uso.

15 Esterilizar por filtro los antibióticos para evitar el crecimiento de hongos, mantener los antibióticos para el cultivo de tejidos vegetales en caja separada.

20 Esterilizar el material vegetal: tomar plantas de aproximadamente 9 cm de altura que no han comenzado a florecer. Las hojas cortadas que tienen una cutícula (4-6 hojas por constructo, suficiente para cortar 100 explantes), se sumergen en etanol al 70% y se sumergen inmediatamente en hipoclorito de Na al 1% (núm de cat. 01032500; usar una botella de lejía que tenga ≤ 3 meses porque el gas de cloro se evapora), mantener las hojas con una pinza y agitarla durante 20 min. Evitar dañar la cutícula, de cualquier otra manera el cloro entrará en el sistema vascular. Enjuagar brevemente en agua estéril 5-6 veces y dejar en agua hasta que esté listo para cortarse.

25 Cocultivo de agrobacterium con explantes de tabaco: cultivar AGL1 en caldo LB o L con antibióticos apropiados durante toda la noche a 28-30°C, al día siguiente resuspender agro en la solución de cocultivo de manera que la concentración final sea de alrededor de 0,4-0,6 de DO_{600nm}. Colocar las hojas de tabaco en caldo cocultivado y cortar cuadrados de 1-1,5cm x 1-1,5cm con un bisturí estéril redondeado usando una acción rodante. Sumergir los explantes foliares en la solución agro con pinzas estériles (almacenados en etanol al 100%, flamear y dejar enfriar antes de tocar el tejido foliar) secar sobre papel Whatman estéril y transferir sobre placas TSM no selectivas (6 explantes por placa) necesarias para preparar aproximadamente 15 placas por constructo. Repetir este procedimiento para cada constructo, asegurándose de que el bisturí y la pinza se sumergen en etanol y se flameen entre cada constructo para evitar la contaminación cruzada. Dejar durante 2 días sólo para AGL1 (3-4 días para otras cepas agro)

35 Transferir en placas TSM selectivas: usar pinzas estériles para recoger y lavar los explantes en un caldo de cocultivo de 100 ml suplementado con 320mg/l de timentina (un recipiente por constructo), agitar bien, secar con papel estéril whatman y colocar los explantes lavados en placas TSM selectivas suplementadas con antibióticos selectivos apropiados y 320mg/l de timentina para matar el agrobacterium.

40 Regeneración del brote: tarda alrededor de 1 mes ver que los brotes aparecen, los explantes deben transferirse en las placas frescas cada 10-14 días. Tener cuidado con el crecimiento recurrente de AGL1, si la timentina no es suficiente para matar el agrobacterium, agregar cefotaxima a 250mg/l.

45 Regeneración de la raíz: Tarda alrededor de 1 semana. Los brotes se cortan de los explantes y se colocan en las cajas de crecimiento que contienen TRM suplementado con los antibióticos selectivos apropiados y 320mg/l de timentina + 250mag/l de cefotaxima para prevenir el crecimiento recurrente de agrobacterium.

Mantenga las plantas en cajas de TRM: sustituirlas cada dos semanas hasta que estén listas para ser transferidas al invernadero

50 Adaptación a las condiciones de invernadero: empapar los sedimentos en agua estéril hasta que se hinchen a un tamaño normal y plantar cuidadosamente una planta por sedimento, incubar las plantas bajo condiciones de 100% de humedad en un propagador, abriendo gradualmente las ventanas hasta que las plantas se adapten a la atmósfera normal.

55 Recetas:
Cocultivo: MS con vitaminas y MES + 0,1mg/l de NAA + 1mg/l de BA + 3% sacarosa, pH 5,7
TSM: MS con vitaminas y MES + 0,1mg/l de NAA + 1mg/l de BA + 3% de sacarosa, pH5,7, 0,2% de gelrite
TRM: ½ sales de MS con vitaminas y MES + 0,5% de sacarosa, pH5,7, 0,2% de gelrite. Autoclave.

60 Concentración de antibióticos
Para cultivos de agrobacterium LB o L:
Cultivar AGL1 que porta pGreen/pSOUP: 100mg/l de Carbenicilina, 5mg/ml de Tetraciclina, 50mg/ml de Rifampicina, 50mg/ml de Kanamicina AGL1 que porta pSOUP: 100mg/l de Carbenicilina, 5mg/ml de Tetraciclina, 50mg/ml de Rifampicina.
65 AGL1 vacío: 100mg/l de Carbenicilina, 50mg/ml de Rifampicina.

- Para el cultivo de la planta:
 Kanamicina: 300mg/l (100mg/l si se usa bentamiana)
 Higromicina: 30mg/l (10mg/l si se usa bentamiana)
 PPT: 20mg/l (2mg/l si se usa bentamiana)
- 5 Timentina: 320mg/l. Se usa para matar agrobacterium, se prepara una pequeña cantidad de materia prima bastante inestable, se almacena en el congelador por hasta 1 mes después de que el antibiótico no es más eficiente.
 Cefotaxima: 250mg/l. También se usa para matar agrobacterium, se añade a TS.
- Transformación de embriones inmaduros de arroz.
- 10 Escisión del embrión inmaduro
- Día 1:
- 15 Eliminar las semillas inmaduras de la fase lechosa/post lechosa de las panículas (se desean embriones inmaduros de 1-2 mm de tamaño).
- Esterilizar semillas inmaduras: hipoclorito de sodio al 50% (12%) + 1 gota de tween 20. Agitar 10 min.
- 20 Enjuagar 3-5x en agua estéril desionizada. Evacuar el exceso de agua. Formar alícuotas con las semillas (alrededor de 40) en placas de Petri estériles.
- 25 Preparar una placa de Petri de 60 x 15 mm que contiene una solución de hipoclorito de sodio al 50% y cerca de ella un vaso estéril sobre su lado con un papel de filtro estéril en él. Usar pinzas estériles para eliminar asépticamente las glumas de la primera semilla. Sumergir esta semilla en hipoclorito de sodio al 50%. Eliminar las glumas de una segunda semilla y sumergir la segunda semilla en la solución de hipoclorito de sodio mientras se elimina la primera semilla y se almacena esta semilla descascarillada/estéril en el papel de filtro en el vaso. Continuar.
- 30 Después se eliminan todas las glumas:
 Esterilizar las semillas descascarilladas: hipoclorito de sodio al 50%: 5 min. con agitación.
 Enjuagar: 5-7 x en agua estéril desionizada, escurrir.
- 35 Colocar todas las semillas en una gran placa de Petri estéril. Formar alícuotas con para la escisión del embrión (para evitar que las semillas se sequen, trabajar con sólo 50-100 en la placa a la vez dejando el resto en la placa maestra).
- Retirar el embrión de cada semilla y colocar el embrión, escutelo hacia arriba, en una placa de Petri de 90 x 15 mm que contiene medio de proliferación (40-50 embriones / placa). Cultivar a 28°C en la oscuridad durante 2 días antes del bombardeo
- 40 Día 3:
- Comprobar cada embrión para la contaminación antes de explotarse.
- 45 Eliminar los embriones del medio de proliferación. Distribuir 35-40 embriones escutelo hacia arriba en un área de 1 cm² en el centro de una placa objetivo de 60 x 15 mm que contiene 10 ml de medio de proliferación + osmótico (0,6 M). Comprobar cada placa objetivo de manera que el escutelo sea recto. Dejar espacio suficiente para que los escutelos no se sombreen entre sí.
- Bombardeo:
- 50 Pistola Biolística 14 kV (PDS-1000/He instrumento de vacío (Biorad) : 25 pulgadas de Hg
- 1^{er} bombardeo 4 horas después del tratamiento osmótico
- 55 2^{do} bombardeo 4 horas después 1^{er} bombardeo
- Día 4:
- 4-16 horas después de la 2da transferencia blástica de embriones inmaduros en el medio de proliferación sin osmótico.
- 60 Cultivar en la oscuridad a 28°C durante 2 días.
- Selección:
- Día 5:
- 65

Cortar asépticamente con tijeras el brote de germinación. Transferir 16 - 20 embriones inmaduros al medio de proliferación fresco que contiene 30-50 mg/l de Higromicina (dependiendo del genotipo); cultivar en la oscuridad a 28°C; registrar el número total de embriones.

5 Después de 10 días, retirar cuidadosamente el callo del escutelo rompiéndolo en 2-10 pequeñas piezas; subcultivar en el medio de proliferación fresco + higromicina. No subcultivar el tejido marrón y el embrión inmaduro remanente que puede inhibir el crecimiento adicional del callo sano.

10 Subcultivar cada 10 días seleccionando el tejido sano: (embriogénico si está presente) y transferirlo al medio de proliferación fresco + higromicina. Eliminar el callo marrón ya que puede inhibirse el callo embriogénico.

30 a 40 días después del bombardeo cambiar el procedimiento de selección. En vez de eliminar tejidos de aspecto malo, sólo se mantiene el tejido embriogénico (eliminar tejido sano no embriogénico)

15 Regeneración:

Después de 40 a 60 días, transferir los callos embriogénicos establecidos que muestran crecimiento diferencial en el medio de proliferación + higromicina para el medio de regeneración + higromicina. Cultivar a 28°C con poca luz durante 10 días y después, bajo luz intensa durante 10 días adicionales. Comprobar las placas periódicamente en la luz para el desarrollo de embriones y brotes verdes. A medida que los brotes se desarrollan, a veces es beneficioso mover con suavidad el brote en desarrollo lejos del callo que lo originó y eliminar cualquier tejido muerto del propio brote para prevenir la inhibición del crecimiento.

25 Germinación:

Transferir los embriones compactos blancos y brotes verdes que inician las raíces en el medio de germinación bajo luz intensa a 28°C durante 1 a 2 semanas. Comprobar las placas periódicamente. Eliminar el tejido necrótico y dividir los embriones en germinación si es necesario.

30 Resultados

Para evaluar la eficacia del direccionamiento génico en *Arabidopsis* un conjunto de constructos se preparó para el direccionamiento de genes exógenos y endógenos (Figuras 5 - 6). Un gen exógeno *uidA* a partir de *E. coli* y el gen endógeno de la acetolactato sintasa (ALS) de *Arabidopsis* se usaron para el direccionamiento.

35 Se introdujo la delección de dos pares de bases en el gen *uidA* de *E. coli* y el fragmento mutado se fusionó con el dominio AP3.

Se introdujo el genmodificado *uidA* en *Arabidopsis* bajo el promotor CMV 35S y el terminador *nos* (constructo ALG125, Figura 4). Las líneas transgénicas se ensayaron para las actividades GUS como se describe por McCabe y otros, (Bio/Technol, 1988, 6, 923-926). No se detectó actividad GUS en las líneas transgénicas debido al marco de lectura en el OFR del gen *uidA*. Las plantas de *Arabidopsis* que portan el gen *uidA* deletéreo se cotransformaron después con los constructos donadores de reparación ALG358 y ALG363 que portan la fuente de replicasa y MDRP (Figura 5). Se evaluó la generación subsiguiente para las actividades GUS y se observaron grandes sectores azules en las hojas de aproximadamente el 40% de las líneas transgénicas, confirmando la reparación del gen *uidA* en la planta usando nuestro sistema de direccionamiento génico. Se obtuvo datos similares con el vector de amplificación del donador basado en el retrotransposón de ARN. Se introdujo una mutación puntual en el fragmento donador para el gen endógeno de la acetolactato sintasa (ALS) de *Arabidopsis* y se prepararon los constructos ALG242 y ALG362 para el direccionamiento de ALS (Figuras 5 - 6). Se demostró que la sustitución del residuo C por el residuo T en la posición 589 del gen confiere resistencia al herbicida de clorsulfurón en las plantas de *Arabidopsis* (Haughn & Somerville, (1986) Mol Gen Genet, 240, 430-434). El constructo ALG242 se cotransformó con ALG 363 y las plantas transgénicas se seleccionaron en el medio suplementado con 10 nm de clorsulfurón. Se recuperó una serie de plantas resistentes, lo que confirma la introducción de la mutación en el gen endógeno de *Arabidopsis* usando nuestro sistema de direccionamiento génico. El análisis por PCR y por secuenciación confirmó la presencia de mutación.

55 Se observaron datos similares en el tabaco para la reparación del gen *uidA* y la mutagénesis de ALS usando el constructo ALG460 (Figura 7)

60 Para evaluar la eficiencia del direccionamiento génico en el arroz, se prepararon un conjunto de constructos basados en el origen de replicación viral y la replicasa del virus de la raya del maíz (MSV) (Figura 8). Un codón de parada prematuro en el fragmento de la quelatata de magnesio de arroz (núm. de acceso EF065607) se introdujo y se clonó en un vector con orígenes de replicación de MSV (ALG462). La fuente de replicasa se administró usando el vector ALG463. La mutación en el gen está afectando el color del tejido de la hoja de verde a verde claro o blanco. Se observaron numerosos sectores de colores verdes claros o blancos en líneas específicas, lo que confirma que el método es eficiente en el arroz. El análisis de la secuenciación confirmó mutaciones en los sectores foliares afectados.

Dominio tnpA

El dominio tnpA del elemento transponible En/Spm (núm. de acceso M25427) se amplificó a partir de la genoteca de ADNc de maíz usando los siguientes iniciadores:

```

5
    AS900  gctcgagaacatgttcaggatggattcgtct      SEQ ID NO. 1
    AS101  cgcggatccgcgcatttgccctcctatcacac      SEQ ID NO. 2

10
    atgttcaggatggattcgtctggcaggagatccagatcacgaaggctcgaggggttcc
    agtgggtgctcctaacatgtttgagggcaccacaacaagcagaagcaggcaagagcag
    ttgcttgccctcccttgagcagatgcggggtagcagcggaccctccaatacagaaggg
15
    actacttcgcgtgcccgtgatttggtggcacctactatggcacctactgcgaggca
    gctgtggacgcggaggcagctgtggacgcggaggctgaggaagctgccgcagaactg
    gacgatggagaggagacttcgggagcagacgcttctacagaggaggctgccacacag
20
    gctccgccacggcgtgccatcagatatcgctcggtcctcacattgaagccttctaaa
    cccttcgatcaacggaggggttattgaaccgaaaggcaccagggcgtggaaggaggctc
    agttgggacggaacagggcatcgcacaccaatcctcacagagttggggatttggttg
25
    cgattcgcctaccgcggcatgggtgacagaggcggtcaagagatagctgctcattac

    tgggctcattgggatctcaagccgtacgggaacgacgggacacacacaagcaaagtg
30
    tgggatttggtttggggccaatttcgcgtgtgcgacccctataccttggtatgattca
    tatgtgctgaggtattcaatggctcagcagatagagctgtaaagggaatgatgtat
    aaagcacgattgagggccgtcacgggtgtacaaaagcgtcaggggaactattgtgac
35
    gcgaatatggcaaaggaaattcacctgacggctcagcagtacaaagagagtgaggtg
    gattggttgagccaccattcggtgcttgggcgtggatgtgagagattgggcttct
    gaggagtcttgccatatctaatagaatcgatgaatcggttgagcaagccaggg
40
    gtccacttcttcggggcggtgacacgcttggaaggctgcacgtatggcggtcga
    aatggagtggagcctactttgcttcaagtcttcgtcgaggacacaagggtcccgat
    ccgaatcatccggaaatactgaatgacagtaatgcgacagagaaaactggctcgctac
45
    attgacaacgtgagagaaaaaatggtccagacacggattggctcactggagagttc
    gatacggaggcggttataaggctggtggaggcgtaccacatgggaggttagcaatt
    ggtgacgggtgttgctccccgcgtcgagttacactagacgggtccaatttctcagctggg
50
    tcgaatcgacccccgacgacctctgcacgtgaaggagagctgcttgagaagatgact
    cagatggaggagagtatggcgcaatataagcaacaagtgcagcaacagatgcaaaa
    atgcaaaaattggatgctacatcagatgtatggaggtgccggaactcagttcggcatg
55
    ccgccttttcaacaacccctatcatcacacatccggtgtctggacaatcatcggac
    cgctccactgcagcggcagatggatcacagggttctgcaacttctgtccaagaccaa
    ttgatgccattgggtgtgatcggcgggcaaatgcgcggatcc

60
    SEQ ID NO. 3

65

```

ES 2 603 587 T3

dominio *recA* del *gen recA*

ggtaacctgaagcagtcctcaacacgctgctgatcttcatcaaccagatccgatatgaaa
attggtgtgatgttcggtaaccggaaccactaccggtggtaacgcgctgaaattc
tacgcctctgttcgtctcgacatccgctcgttac

SEQ ID NO. 4

El dominio *recA* (SEQ ID NO.:4) del *gen recA* se amplificó a partir del ADN genómico de la cepa JM101 de *E. coli* (New England Biolabs) usando los siguientes iniciadores:

AS380 cgggatccggttaacctgaagcagtcctcaaca SEQ ID NO. 5

AS381 cggctctagattttaacgacggatgtcgagacga SEQ ID NO. 6

Secuencia de localización nuclear *virD2* (*virD2NLS*)

La secuencia de localización nuclear de un *gen virD2* se amplificó de la cepa GV3101 de *Agrobacterium* usando los siguientes iniciadores:

AS380 cgggatccggttaacctgaagcagtcctcaaca
SEQ ID NO. 5

AS881 ctctcgcggtttgcgttcaactcggttctccatcatcatcttcaacgcggac
SEQ ID NO. 7

AS882 cactagttcatctctcgcggtttg SEQ ID NO. 8

AS899 ccatcatcatcttcaacgcggacgcttgtaacgacggatgtcgagacga
SEQ ID NO. 9

Secuencia de localización nuclear de *virD2*

aagcgtccgcgtgaagatgatgatggagaaccgagtgaaacgcaaacgcgagagatga
SEQ ID NO. 10

dominio de fusión *tnpA-recA-virD2NLS*

El dominio *recA* está subrayado; la secuencia de localización nuclear del *gen virD2* está en cursiva.

atgttcaggatggattcgtctggcaggagatccagatcacgaaggtcgaggggttcc
 5 agtggtgctcctaacaatgtttgagggcaccacaacaagcagaagcaggcaagagcag
 ttgcttgccctcccttgagcagatgcggggtagcagcggaccctccaatacagaaggg
 actacttcgcgtgccgctgatttgggtggcacctactatggcacctactgcggaggca
 10 gctgtggacgcggaggcagctgtggacgcggaggctgaggaagctgccgcagaactg
 gacgatggagaggagacttcgggagcagacgcttctacagaggaggctgccacacag
 gctccgccacggcgtgccatcagatatcgtcggtcctcacattgaagccttctaaa
 15 cccttcgatcaacggaggggttattgaaccgaaaggcaccagggcgtggaaggaggtc
 agttgggacggaacagggcatcgcacaccaatcctcacagagttggggatttgtttg
 cgattcgcctacccggcgatggtgacagaggcggtcaagagatagctgctcattac
 20 tgggctcattgggatctcaagccgtacgggaacgcgggacacacacaagcaaagtg
 tgggatttgttttggggccaatttcgcgtgtgcgacccctataccttggtatgattca
 25 tatgtgcgtgaggtattcaatggctcagcagatagagctgtaaaggggaatgatgtat
 aaagcacgattgagggccgtcacggtgtacaaaagcgtcaggggaactattgtgac
 gcgaatatggcaaaggaaattcacctgacggctcagcagtacaaagagagtgaggtg
 30 gattgggttgagccaccattcggatgcttgggcgtggatgtgcgagtattgggcttct
 gaggagttcttggccatatctaatagaatcgtatgaatcggttgagcaagccaggg
 gtccacttcttcggggcggtggacacgcttggaaggctgcacgtatggcggctcga
 35 aatggagtgagagcctactttgcttcaagtcttcgtcgagggaacacaaggggtcccgat
 ccgaatcatccggaaataactgaatgacagtaatgcgacagagaaactggctcgctac
 attgacaacgtgagagaaaaaatgggtccagacacggattgggtcactggagagttc
 40 gatacggaggcggttataaggctggtggaggcgtagccacatgggaggttagcaatt
 ggtgacggtgttgtccccgcgtcgaggttacactagacgggtccaatttctcagctggg
 45 tcgaatcgaccccgacgaccatctgcacgtgaaggagagctgcttgagaagatgact
 cagatggaggagagtatggcgcaatataagcaacaagtgcagcaacagatgcaacaa
 atgcaaaaattggatgctacatcagatgtatggaggtgccggaactcagttcggcatg
 50 ccgccttttcaacaacccctatcatcacacatccggtgtctggacaatcatcggac
 cgctccactgcagcggcagatggatcacaggggttctgcaacttctgtccaagaccaa
 ttgatgccattgggtgtgatcggcgggcaaattgcgcggatccggtaacctgaagcag
 55 tccaacacgctgctgatcttcatcaaccagatccgtatgaaaattgggtgtgatgttc
ggtaacccggaaaccactaccgggtggtaacgcgctgaaattctacgcctctgttcgt
ctcgacatccgtcggttacaagcgtccgcgtgaagatgatgatggagaaccgagtgaa
 60 cgcaaacgcgagagatga SEQ ID NO. 11

65

Promotor de la ubiquitina 3 de Arabidopsis

Se amplificó la secuencia del promotor del gen de la ubiquitina 3 de Arabidopsis (At5g03250) a partir del ADN genómico de la cepa Arabidopsis Col-0 usando los siguientes iniciadores:

5
AS724 cggtacctaccggatttggagccaagtc SEQ ID NO. 12
AS726 gctcgagtggttggtagacctgaaataaaacaatagaacaagt
10
SEQ ID NO. 13
taccggatttggagccaagtcataaacgccattgtggaagaaagtcttgagttgg
15
tggtaatgtaacagagtagtaagaacagagaagagagagagtgtagatacatgaat
tgtcgggcaacaaaaatcctgaacatcttatttttagcaaagagaaagagttccgagt
ctgtagcagaagagtgaggagaaatttaagctcttggacttgtgaattgttccgcct
20
cttgaatacttcttcaatcctcatatattcttcttctatgttacctgaaaaccggca
tttaatctcgcgggtttattccggttcaacatttttttgttttgagttattatctg
ggcttaataacgcaggcctgaaataaattcaaggcccaactgttttttttttaaga
25
agttgctgttaaaaaaaaaaaaaagggaattaacaacaacaacaaaaaaagataaaga
aaataataacaattactttaattgttagactaaaaaacatagattttatcatgaaaa
30
aaagagaaaaagaaataaaaaacttgatcaaaaaaaaaacatacagatcttctaatta
ttaacttttcttaaaaaattaggtcctttttcccaacaattaggttttagagttttgga
attaaacaaaaagattgttctaataaaataactcaaatttggtagataagtttcctta
35
ttttaattagtgcaatggtagatacttttttttcttttctttattagagtagattaga
atcttttatgccaagtattgataaattaaatcaagaagataaactatcataatcaac
atgaaattaaaaagaaaaatctcatatatagtagtatttctctatatatatattatga
40
ttgcttattcttaaatgggttgggttaaccaagacatagtccttaattggaaagaatctt
ttttgaactttttccttattgattaaattcttctatagaaaagaaagaaattatttg
aggaaaagtatatatacaaaaagaaaaatagaaaaatgtcagtgaagcagatgtaatgg
45
atgacctaatccaaccaccaccataggatgtttctacttgagtcggtcttttaaaaa
cgcacggtggaaaatatgacacgtatcatatgattccttcctttagtttcgtgataa
taatcctcaactgatatcttccttttttgttttggttaagatatattttattctcat
50
taatagaaaagacggttttgggcttttgggttgcgatataaagaagaccttcgtgtg
gaagataataattcatcctttcgtctttttctgactcttcaatctctcccaaagcct
aaagcgatctctgcaaatctctcgcgactctctctttcaaggtatatattttctgattc
55
tttttggttttgattcgtatctgatctccaatttttggttatgtggattattgaatct
tttgatataaattgcttttgacaatattgttcgtttcgtcaatccagcttctaaattt
tgtcctgattactaagatatcgattcgtagtgtttacatctgtgtaatttcttgctt
60
gattgtgaaattaggattttcaaggacgatctattcaatttttggtgttttctttggt
cgattctctctgttttaggtttcttatgttttagatccgtttctctttgggtgttgtt
65

tgattttctcttacggccttttgatttggtatatgttcgctgattgggttctacttgtt
ctattgttttatttcagggtcaccaaaca SEQ ID NO. 14

5

terminador ags

10 Se sintetizó la señal de poliA de agropina sintasa (*ags* terminador) basada en la secuencia de banco de genes EU181145.

15 Gaattaacagaggtggatggacagaccggttcttacaccggactgggcgcgggtag
gataattcagattgggatgggattgagcttaaagccggcgctgagaccatgctcaagg
taggcaatgtcctcagcgtcgagcccgcatctatgtcgagggcattgggtggagcgc
gcttcggggataccgtgcttgtaactgagaccggatatgaggccctcactccgcttg
20 atcttgggcaaagataatttgacgcatttattagtagtggttaattttcatttgacgtg
cagtattttctattcgatctttatgtaattcggttacaattaataaatattcaaata
gattattgactgtcatttgatcaaatcgtgtttaatggatatttttattataatat
25 tgatgat

SEQ ID NO. 15

30 secuencia BCTVi

El gen de la replicasa repA del virus de la rizomanía de la remolacha (BCTV, Acc. Núm. BCTCG028A) y los genes responsables de la regulación de la replicación viral se amplificaron usando el siguiente conjunto de iniciadores:

35 **AS710** aagccttgggaagtatgaagacgtgaag SEQ ID NO. 16
AS711 tttttttcgggccatccggtttatgaaagttgtactaatatat
SEQ ID NO. 17
40 **AS712** accggatggcccgaaaaaaatg SEQ ID NO. 18
AS713 ggtaccttataattacatatacatgta SEQ ID NO. 19

45 El bucle de tallo en la región del promotor del gen repA se deletionó parcialmente mediante los iniciadores superpuestos AS711 y AS712 para evitar la replicación de este casete en las células vegetales.

Se insertaron dos intrones del gen de *Arabidopsis* At2g29890 en la secuencia para mejorar la eficiencia de la expresión en plantas y para evitar la replicación del casete de reparación en células de *Agrobacterium* (subrayado). Se usaron los siguientes iniciadores superpuestos para introducir estos intrones en la secuencia BCTVi:

50 **AS795** aggccaaacaaagatgaaaattacctgggtggggataccagggtcgaat
SEQ ID NO. 20
55 **AS796** attcgacctgggtatccccaaccaggtaattttcatctttgtttggcct
SEQ ID NO. 21
AS797 atgtttggatggaaatgtgctgacctgcaatctcataagtacaatgga
60 SEQ ID NO. 22
AS798 tccattgtaacttatgagattgcagggtcagcacatttccatccaaacat
SEQ ID NO. 23
65

para el intrón 1, y

AS799 tcaagatgggaaaagatttcttaccttcaagaaacggggtccagtcaa

SEQ ID NO. 24

AS800 ttgactggaccccggtttcttgaaggtaagaaatcttttcccatcttga

SEQ ID NO. 25

AS801 tatctctagtctctttttcaagacctgcaatgtttccacgcatattat

SEQ ID NO. 26

AS802 ataatatgcgtggaaacattgcaggtcttgaaaaagagactagagata

SEQ ID NO. 27

para el intrón 2.

ttataattacatatacatgtaaaaataacgtatataaaacgacataattcaaaactt

gctgagcaagtttctgtggggcacacttactttacttttactaaagtaaagtact

ttaataaagttgtccccttcttttgaccagtcacaaatcttatctgaaaggaatct

tcttcaggaagtttcccgaggaatcttcttcaggaagtttcccgctcaaaaacgctc

acatcattcaacagtccccttccacgtgtcacgttttgattgggtgcccattttttt

cgggccatccggtttatgaaagttgtactaatatatataataactccaatactccaat

agatatgttaatatgctaagccacgtcatcggagtacaatgcctcctactaaaagat

ttcgtattcaagcaaaaaacatatttcttacatatacctcagtggttctctttcaaaag

aagaagctcttgagcaaattcaaagaatacaactttcatctaataaaaaatatatta

aaattgccagagagctacacgaagatgggcaacctcatctccacgtcctgcttcaac

tcgaaggaaaagttcagatcacaatatcagattattcgacctgggtatccccaacca

ggtaattttcatctttgtttggccttccaagtgccttttttgcctgtttacgggtgga

acttcagtaaaaatgggatcaaaacatcatatggcataaataaattttaagaatggc

gaactcgggggttacccaatatggcttcccttttccagtgtttcttagtccattgtact

tatgagattgcagggtcagcacatttccatccaaacattcagagagctaaatccagct

ccgacgtcaagtcctacgtcgacaaggacggagacacaattgaatggggagaattcc

agatcgacggtagaagtgcctagaggaggtcaacagacagctaacgactcatatgcc

aggcgttaaacgcaacttctcttgaccaagcacttcaaatattgaaggaagaacaac

caaaggattacttcccttcaacatcacaatcttttgaacaatgctcaaaagatatttc

agaggccacctgatccatggactccactatttccctctgtcctcattcacaacgcttc

ctgaggaaatgcaagaatgggctgatgcataatttcgggggttgatgccgctgcgcggc

ctttaagatataatagtatcatagtagagggtgattcaagaacaggggaagactatgt

5 gggctagatcttttaggggcccacaattacatcacagggcacttagatttttagcccta
 gaacgtattatgatgaagtggaatacaacgtcattgatgaogtagatcccacttact
 taaagatgaaacactggaaacaccttattggagcacaaaaggagtggcagacaaact
 10 taaagtatggaaaaccacgtgtcattaaagggtggtatcccctgcattatattatgca
 atccaggacctgagagctcataccaacaatttcttgaaaaaccagaaaatgaagccc
 ttaagtccctggacattacataattcaaccttctgcaaactccaaggtccgctcttta
 15 ataaccaagcagcagcatcctcgcaaggtgactctaccctgtaactgccacttcaca
 atacaccatgaatgtaatagaggatttttcgcacagggggaacctattactctccatca
 ggcaacaaattccgctcgaattcgagaatgtaccgaatccactgtatatgaaactcct
 atggttcgagagatacgggccaatctatcaactgaagatacaaatcagattcaacta
 caacctccggagagcgttgaatcttcacaagtgtctggatagagctgacgataactgg
 20 atcgaacaggatattgactggaccccgttttcttgaaggtaagaaatcttttcccatc
ttgaagtcacctcaaaccgaacgttaggaaattccaaaatgttttgatagtagtcta
cttagtttcaagttttgggtttgtgtatactttcactaataatatgcgtggaaacat
 25 tgcaggctcttgaaaaagagactagagatatacttggataatctaggtttaatttgta
 ttaataatgtaattagagggtttaaatcatgtcctgtatgaagaatttacttttgtat
 30 caagtgtaattcagaaccagagtgttgcaatgaacttgtactaatttcattattaat
 aataaatattattaataaaaaatagcatctacaattgccaaataatgtggcatacata
 ttagtattatccgtattatcattaacaacaacatagagtaaagcattctccttcacg
 35 tcttcataacttccca

SEQ ID NO. 28

40 Gen mutado *uidA*

45 Se introdujo una delección de dos pares de bases (delección GT) en el gen *uidA* de *E. coli*, generando un cambio de marco en el ORF de *uidA*. La posición de la delección está entre dos nucleótidos, y se muestra en mayúscula y subrayada.

50 atgttacgtcctgtagaaaccccaacccgtgaaatcaaaaaactcgacggcctgtgg
 gcatttcagtctggatcgcgaaaactgtggaattgatcagcgttggtgggaaagcgcg
 ttacaagaaagccgggcaattgctgtgccaggcagttttaacgatcagttcgccgat
 gcagatattcgtaattatgcgggcaacgtctggtatcagcgcaagtctttataccg
 55 aaagggttgggcaggccagcgtatcgtgctgcgttttcgatgcggtcactcattacggc
 aaagtgtgggtcaataatcaggaagtgtgagcatcagggcgggtatacgcattt
 gaagccgatgtcacgccgtatgttattgcccgggaaaagTAcgtatcacccgtttgtgt
 60 gaacaacgaactgaactggcagactatcccgcgggaatggtgattaccgacgaaaa
 cggcaagaaaaagcagcttacttccatgatttctttaactatgccggaatccatcg
 cagcgtaatgctctacaccacgccgaacacctgggtggacgatatacccggttgac
 65

gcatgtcgcgcaagactgtaaccacgcgtctgttgactggcaggtggtggccaatgg
 tgatgtcagcgttgaactgcgtgatgcggatcaacaggtggttgcaactggacaagg
 5 cactagcgggactttgcaagtgggtgaatccgcacctctggcaaccgggtgaaggtta
 tctctatgaactgtgcgtcacagccaaaagccagacagagtgtgatatctacccgct
 10 tcgcgtcggcatccggtcagtggcagtgaaagggcgaacagttcctgattaaccacaa
 accgtttctactttactggcgtttggcgtcatgaagatgcggacttacgtggcaaagg
 attcgataacgtgctgatggtgcacgaccacgcattaatggactggattggggccaa
 15 ctctaccgtacctcgcattacccttacgctgaagagatgctcgactgggcagatga
 acatggcatcgtgggtgattgatgaaactgctgctgctcggcctttaacctctcttttagg
 20 cattggtttctgaagcgggcaacaagccgaaagaactgtacagcgaagaggcagtcaa
 cggggaaactcagcaagcgcacttacaggcgattaaagagctgatagcgcgtgacaa
 aaaccacccaagcgtgggtgatgtggagtattgccaacgaaccggataaccgctccgca
 25 agtgcacgggaatatcttcgccactggcggaagcaacgcgtaaactcgacccgacgcg
 tccgatcacctgcgtcaatgtaatgttctgcgacgctcacaccgataccatcagcga
 30 tctctttgatgtgctgtgcctgaaccggttattacggatgggtatgtccaaagcggcga
 tttggaaacggcagagaaggtactggaaaaagaacttctggcctggcaggagaaact
 gcatcagccgattatcatcacccaatacggcgtggatacgttagccgggctgcactc
 35 aatgtacaccgacatgtggagtgaagagtatcagtgtgcatggctggatatgtatca
 ccgcgtctttgatcgcgtcagcgccgtcgtcgggtgaacaggtatggaatttcgccga
 40 ttttgcgacctcgcaaggcatattgcgcggttggcggtaacaagaaagggatcttcac
 tcgcgacccgcaaaccgaagtcggcggttttctgctgcaaaaacgctggactggcat
 gaacttcggtgaaaaaccgcagcagggaggcaacaatga
 45

SEQ ID NO. 29

Promotor CMV 35S

El gen *uidA* mutado se colocó bajo control del promotor 35S del virus del mosaico de la coliflor (CMV), que se sintetizó en base a la secuencia disponible del banco de genes (Acc Núm. AF502128).

caatcccacaaaaatctgagcttaacagcacagttgctcctctcagagcagaatcgg
 gtattcaacaccctcatatcaactactacgttgtgtataacgggtccacatgccggta
 5 tatacgatgactgggggtgtacaaaggcggcaacaaacggcggtcccggagttgcac
 acaagaaatgtgccactattacagaggcaagagcagcagctgacgcgtacacaacaa
 gtcagcaaacagacaggttgaacttcatccccaaggagaagctcaactcaagccca
 10 agagctttgctaaggccctaacaagcccaccaaagcaaaaagcccactgggtcacgc
 taggaaccaaaggccagcagtgatccagcccaaaagagatctcctttgccccgg
 agattacaatggacgatttctctatctttacgatctaggaaggaagttcgaagggtg
 15 aagtagacgacactatgttcaccactgataatgagaagggttagcctcttcaatttca
 gaaagaatgctgaccacagatgggttagagaggcctacgcagcaggtctcatcaaga
 cgatctacccgagtaacaatctccaggagatcaaataccttcccaagaaggttaaag
 atgcagtcaaaagattcaggactaattgcatcaagaacacagagaaagacatatctt
 tcaagatcagaagtactattccagtatggacgattcaaggcttgcttcataaaccaa
 25 ggcaagtaatagagattggagttctctaaaaaggtagttcctactgaatctaaggcca
 tgcattggagtttaagattcaaatacgaggatctaacagaactcgccgtgaagactggc
 gaacagttcatacagagtccttttacgactcaatgacaagaagaaaatcttcgtcaac
 30 atgggtggagcacgacactctgggtctactccaaaaatgtcaaagatacagttctcagaa
 gaccaaagggtattgagactttttcaacaaaggataatttcgggaaacctcctcgga
 ttccattgcccagctatctgtcacttcatcgaaaaggacagtagaaaaggaagggtggc
 35 tctacaaatgccatcattgagataaaggaaagggtatcattcaagatctctctgcc
 gacagtggtcccaaagatggacccccaccacgaggagcatcgtggaaaaagaagac
 gttccaaccacgtcttcaaagcaagtggattgatgtgacatctccactgacgtaagg
 40 gatgacgcacaatccactatccttcgcaagacccttctctatataaggaagttca
 tttcatttgagaggacacg

45 SEO ID NO. 30

Nos Terminador

50 El fragmento *nos* terminador se sintetizó en base a la secuencia de acceso de GenBank EU048864.

gtcaagcagatcggttcaaacatttggcaataaagtttcttaagattgaatcctggtg
 ccggctcttgcatgattatcatataatttctgttgaattacgtgaagcatgtaataa
 55 ttaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtcccg
 aattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaat
 60 tatcgcgcgcggtgtcatctatgttactagatcgac

SEQ ID NO. 31

Origen de replicación viral BCTV (VOR)

65

ES 2 603 587 T3

El origen de replicación viral se amplificó a partir del ADN viral (Acc. Núm. BCTCG028A) con los siguientes iniciadores:

AS252 tgtactccgatgacgtggctt SEQ ID NO. 32

AS253 ttcaggaagtttcccqctca SEO ID NO. 33

tgtactccgatgacgtggcttagcatattaacatatctattggagtagttggagtatt
atatatattagtagacaactttcataagggccatccgttataatattaccggatggccc
gaaaaaaatgggcacccaatcaaaacgtgacacgtggaaggggactgttgaatgatg
tgacgttttttgacgqggaaacttcctgaa

SEO ID NO. 34

Región Subterminal 1 del elemento transponible de maíz En/Spm (STR1)

La secuencia de STR1 se amplificó a partir de ADN genómico de maíz usando los siguientes iniciadores:

AS150 aggagtgtcagttaattaaaga SEQ ID NO. 35

AS151 gtcgactgtatactggacg SEQ ID NO. 36

Gagtgtcagtttaattaaagagtgtcggggccgacactcttaatcgaagtaaaagtgt
gggttttgcgtgcaccgacactcttaatttaagagtgtcggggcccgatgaaaccga
cgcttttaatttaagagtgtgggtttttccacaccgacactcttatgaatgttacc
taaattccccaatctattctacagccgtcgtgcttctctctctcttctccctgcc
cgccgtccagtataca

SEQ ID NO. 37

Región Subterminal 2 del elemento transponible de maíz En/Spm (STR2)

La secuencia para STR2) se amplificó a partir de ADN genómico de maíz usando los siguientes iniciadores:

AS152 agtgtcggccaaaacccac SEQ ID NO. 38

AS153 cacactcttatatgcgccag SEQ ID NO. 39

ccactcttatatgcgcccaggtagcttactgatgtgcgcgcagtaagagtgacggcc
acgggtactggccgacacttttaacataagagtgctcggttgcttggtgaaccgacact
tttaacataagagcgtcgggtccccacacttctatacgaataagagcgtccatttttag
agtgcggtgaagagtgctcgggtcaaccgacactcttatacttagagtgctcggcttat
ttcagtaagagtggtgggggttttgccgacact

SEO ID NO. 40

secuencia de donador de reparación uidA (uidA-R)

La secuencia de donador para la reparación del gen *uidA* se amplificó a partir de la cepa de ADN genómico de *E. coli* DH10B usando los siguientes iniciadores:

ES 2 603 587 T3

AS866 ctgcgtttcgatgcggtcac SEQ ID NO. 41

AS867 accaatgcctaaagagaggtt SEQ ID NO. 42

5 Ctgcgtttcgatgcggtcactcattacggcaaagtgtgggtcaataatcaggaagtg
atggagcatcagggcggctatacgccatttgaagccgatgtcacgccgtatgtttatt
gccgggaaaagtgtacgtatcaccgtttgtgtgaacaacgaactgaactggcagact
10 atcccgccgggaatggtgattaccgacgaaaacggcaagaaaaagcagtcttacttc
catgatttctttaactatgccggaatccatcgacgcgtaatgctctacaccacgccg
aacacctgggtggacgatatcaccggtgacgcgatgtcgcgcaagactgtaaccac
15 gcgctctgttgactggcaggtggtggccaatggtgatgtcagcggtgaactgcgtgat
gcggatcaacaggtggttgcaactggacaaggcactagcgggactttgcaagtgtg
aatccgcacctctggcaaccgggtgaaggttatctctatgaactgtgcgtcacagcc
20 aaaagccagacagagtgtgatctaccgccttcgcgtcggcatccggtcagtggca
gtgaagggcgaaacagttcctgattaaccacaaaccggttctactttactggcttgggt
cgatcatgaagatgcggacttgcggtggcaaaggattcgataacgtgctgatggtgcac
25 gaccacgcattaatggactggattggggccaactcctaccgtacctcgcattaccct
tacgctgaagagatgctcgactgggcagatgaacatggcatcggtggtgattgatgaa
actgctgctgtcggctttaacctctctttaggcattggt
30 SEQ ID NO. 43

Secuencia de mutagénesis ALS de Arabidopsis (ALS-M)

35 Se preparó un donador mutado (mutación C a T está en capital y subrayado) para direccionar el gen de la acetolactato sintasa endógena de *Arabidopsis* usando el siguiente conjunto de iniciadores:

AS803 gcggctgcagggcggcggaacaacaacaaca

40 SEQ ID NO. 44

AS804 atctgtaccaatcatagcagagagacttgtcctgtgattgctacaag

SEQ ID NO. 45

45 **AS805** cttgtagcaatcacaggacaagtctctcgtcgtatgattggtacagat

SEQ ID NO. 46

AS806 accatctagaccaatctcagccgagtcaatatca

50 SEQ ID NO. 47

gcggcggaacaacaacaacaacacatcttcttcgatctccttctccacaaacca
tctccttctcctccaaatcaccattaccaatctccagattctccctcccattctcc
55 ctaaaccccaacaatcatcctcctcctcccgccgcggtatcaaatecagctct
cctcctccatctccgctgtcacaacaaccaccaatgtcacaaccaactccctct
ccaaccaaacctaccaaacccgaaacattcatctcccgattcgctccagatcaaccc
60 cgcaaaggcgctgatatcctcgtcgaagctttagaacgtcaaggcgtagaaaccgta

65

ES 2 603 587 T3

ttcgcttaccctggaggtgcatcaatggagattcaccaagccttaacccgctcttcc
tcaatccgtaacgtccttcctcgtcacgaacaaggaggtgtattcgcagcagaagga
tacgctcgatcctcaggtaaacagggtatctgtatagccacttcagggtcccggagct
5 acaaatctcgttagcggattagccgatgcggtttagatagtgttcctctttagca
atcacaggacaagtcTctcgtcgtatgattgggtacagatgcgtttcaagagactccg
attggttgaggtaacgcgttcgattacgaagcataactatcttgtgatggatggtgaa
10 gatatccctaggattattgaggaagctttcttttttagctacttctggtagacctgga
cctgtttttggttgatgttcctaaagatattcaacaacagcttgcgattcctaattgg
gaacaggctatgagattacctgggttatatgtctaggatgcctaaacctccggaagat
15 tctcatttgaggagcagattgttaggttgatttctgagtctaagaagcctgtgttgat
gttggtggtggttgtttgaattctagcgatgaattgggtagggttgttgagcttacg
gggatccctgttgcgagtacgttgatggggctgggatcttatccttgtgatgatgag
20 ttgtcgttacatatgcttggaatgcatgggactgtgtatgcaaattacgctgtggag
catagtgatttgttgttggtggttgggtaagggttgatgatcggtgcacgggtaag
cttgaggcttttgctagtagggctaagattgttcatattgatattgactcggctgag
25 attgg

SEQ ID NO. 48

dominio de ADN APETALA 3 (AP3)

El dominio AP3 se amplificó usando los siguientes iniciadores:

AS690

ctcgagactttccatttttagtaactaccaccatgttacgtcctgtagaaaccca

SEQ ID NO. 49

AS300 aattgtttgcctccctgctggcctt

SEQ ID NO. 50

agttactaaaaatggaaagt

SEQ ID NO. 51

Dominio de unión al ADN APETALA1 (AP1)

El dominio Ap1 se amplificó a partir del ADN genómico de Arabidopsis usando los siguientes iniciadores:

AS429 ggctcgagatgggaaggggtaggg

SEQ ID NO. 52

AS430 catcgatttcggcgtaagagtacct

SEQ ID NO. 53

atgggaaggggtagggttcaattgaagaggatagagaacaagatcaatagacaagt
acattctcgaaaagaagagctgggtcttttgaagaaagctcatgagatctctgttctc
tgtgatgctgaagttgctcttgttgtcttctcccataagggaaaactcttcgaatac

tccactgattcttgtatggagaagatacttgaacgctatgagaggtactcttacgcc
gaa

SEQ ID NO. 54

Secuencia de Unión del Iniciador

La secuencia de unión del iniciador a partir de retrotransposón *tnt1* de tabaco se amplificó con los siguientes iniciadores:

5 **AS912** gccgcggctttattaccgtgaatatta SEQ ID NO. 55

10 **AS913** cgcgcccgctctgataagtgcaacctgatt SEQ ID NO. 56

GccgcggctttattaccgtgaatattatTTTTGGTAAGGGGTTATTCCCAACAAC
15 gtatcagagcacaggttctgctcgttcactgaaatactattcactgtcggtagtact
atacttggtgaaaaataaaaaatgtctggagtaaagtagcaggtagcaaaattcaatg
gagataacggtttctcaacatggcaaagaaggatgagagatctgctcatccaacaag
20 gattacacaagggttctagatgtttgattccaaaaagcctgataccatgaaagctgagg
attgggctgacttggatgaaagagctgctagtgcattcaggttgacttatcagagc
ggccgcg

25 SEQ ID NO. 57

Transcriptasa inversa-RNasaH (RT-RH)

RT-RH se amplificó a partir de retrotransposón tnt1 de tabaco usando los siguientes iniciadores:

30 **AS885** tctcgagaacatgcagcatagagtctcatttc SEQ ID NO. 58

35 **Tnt-RT-R** ctagtttgaatgcattccgacaa SEQ ID NO. 59

atgtcagaaaaggtgaagaatggtataattcctaactttgttactattccttctact
40 tctaacaatcccacaagtgacagaaagtagcaccgacgaggtttccgagcagggggag
caacctggtgaggttattgagcagggggagcaacttgatgaagggtgtcgaggaagtg
gagcaccctcactcagggagaagaacaacatcaacctctgaggagatcagagaggcca
45 agggtagagtcacgcaggtacccttccacagagtatgtcctcatcagtgatgagggg
gagccagaaagtccttaaggaggtggtgtcccatccagaaaagaaccagtggatgaaa
gctatgcaagaagagatggaatctctccagaaaaatggcacatacaagctggttgaa
50 cttccaaagggtaaaagaccactcaaatgcaaatgggtcctttaaactcaagaaagat
ggagatggcaagctgggtcagatacaaagctcgattggtggttaaaggcttcgaacag
aagaaaggtattgattttgacgaaatcttccccgttggttaaaatgacttctatt
55 cgaacaattttgagcttagcagctagcctagatcttgaagtggagcagttggatgtg
aaaactgcatttcttcatggagatttggaagaggagatttatatggagcaaccagaa
ggatttgaagtagctggaaaagaacacatgggtgtgc aaattgaataagagtcctttat
60 ggattgaagcaggcaccaaggcagtggtacatgaagtttgattcattcatgaaaagt
caaacatacctaaagacctatttctgatccatgtgtatacttcaaaagattttctgag

aataactttattatattgttgttgtatgtggatgacatgctaattgtaggaaaagac
 aaggggttgatagcaaagttgaaaggagatctgtccaagtcatttgatatgaaggac
 5 ttgggcccagcacaacaaattctagggatgaagatagttcgagagagaacaagtaga
 aagttgtggctatctcaggagaagtacattgaacgtgtactagaacgcttcaacatg
 aagaatgctaagccagtcagcacacctcttgctggatcatctaaagttgagtaaaaag
 10 atgtgtcctacaacagtggaagagaaaagggaaacatggctaaagttccttattcttca
 gcagtcggaagccttgatgtatgcaatggatgtactagacctgatattgctcacgca
 gttgggtgttgcagcaggtttcttgaaaatcctggaaaggaacattgggaagcagtc
 15 aagtggatactcaggtacctgagaggtaccacgggagattgtttgtgctttggagga
 tctgatccaatcttgaagggctatacagatgctgatatggcaggtgacattgacaac
 agaaaatccagttactggatatttgtttacattttcagggggagctatatcatggcag
 20 tctaagttgcaaaagtgcgttgcactttcaacaactgaagcagagtagattgctgct
 acagaaactggcaaggagatgatatggctcaagcgattccttcaagagccttgattg
 catcagaaggagtagtgcgtctattgtgacagtcgaaagtgaatagaccttagcaag
 25 aactctatgtaccatgcaaggaccaaacacattgatgtgagatatcattggattcga
 gaaatggtagatgatgaatctctaaaagtcttgaagatttctacaaatgagaatccc
 30 gcagatatgctgaccaaggtggtaccaaggaacaagttcgagctatgcaaagaactt
 gtcggaatgcattcaaactag

SEQ ID NO.60

35 **Secuencia de mutagénesis ALS de tabaco**

Se preparó un donador mutado para direccionar el gen de acetatolactato sintasa (ALS) endógeno del tabaco usando el siguiente conjunto de iniciadores:

40 AS1031 gcccggtcctccaccctcctccctagat SEQ ID NO.61
 AS1034 atcctacgtgccacttgaccggttatagcaa SEQ ID NO.62
 45 AS1035 cgggtcaagtggcacgtaggatgatcggtact SEQ ID NO.63
 AS1036 atctagaacctggatagcatat SEQ ID NO.64

50 **La mutación que induce la resistencia al clorsulfurón está capitalizada (ver más abajo) y subrayada.**

gcccggtcctccaccctcctccctagatccaccttccctttccccaccaccccca
 55 caaaaccacccaccacccctccacctacccccaccacattcacagccaacgccg
 tcgtttcaccatctccaatgtcattttccactacccaaaaagtttccgagacccaaaa
 agccgaaaactttcggtttcccggtttgcccctgacgaaccagaaaggggttcgacgt
 60 tctcgtggaggccctcgaaagagaaggggttacggacgtttttgcgtagccaggcgg
 cgcttccctcgagattcaccaagctttgacgcgctcaagcatcatccgcaacgtgct
 accacgtcacgagcaggggtggtgtcttcgcgctgaggggttacgcacgcgccaccgg
 65 cttccccggcggtttgcattgccacctccggccctggcgccaccaatctcgtcagtgg

cctcgcggacgcctactggatagcgtccccattgttgctataaccgggtcaagtGc
 acgtaggatgatcgggtactgatgcttttcaggaaactccgattgttgaggtaactag
 5 atcgattaccaagcataattatctcgttatggacgtagaggatattcctagggttgt
 acgtgaggcttttttccttgcgagatcggggccggcctggccctgttttgattgatgt
 10 acctaaggatattcagcaacaattggtgatacctgactgggatcagccaatgaggtt
 gcctgggttacatgtctaggttacctaattgcccaatgagatgcttttagaacaat
 tgttaggcttattttctgagtcaaagaagcctgttttgatgtggggggtgggtgttc
 15 gcaatcgagtgaggagttgagacgattcgtggagctcaccgggtatccccgtggcaag
 tactttgatgggtcttgagcttttccaactggggatgagctttccctttcaatgtt
 gggatgcatgggtactgtttatgctaattatgctgtggacagtagtgatttattgct
 20 cgcatttgggggtgagggttgatgatagagttactggaaagttagaagcttttgctag
 ccgagcgaaaattgttcacattgatattgattcagctgagattggaaagaacaagca
 gcctcatgtttccatttgtgcggatatcaagttggcgttacagggtttgaattcgat
 25 attggagagtaaggaaggtaaactgaagttggattttttctgcttggaggcaggagtt
 gacggtgcagaaagtgaagtacccgttgaatttttaaaacttttggtgatgctattcc
 tccgcaatatgctatccaggttctagat SEQ ID NO.65

30

Se amplificó el origen de replicación del virus de la raya de maíz (MSV) para el direccionamiento génico en monocotiledóneas con los siguientes iniciadores:

35 AS1020 gggtaccccgacgacggaggttgaggct SEQ ID NO.66

AS1021 cgaattcatgggctgctttagataaatc SEQ ID NO.67

40 y

AS1022 ttctagaccgacgacggaggttgaggct SEQ ID NO.68

45 AS1023 cccgcggatgggctgctttagataaatc SEQ ID NO.69

Ccgacgacggaggttgaggctgagggatggcagactgggagctccaaactctatagt
 50 ataccctgctgccttcgaaatccgccgctcccttgctttatagtgggttgcaaatggg
 ccggaccggggccggcccagcaggaaaagaaggcgcgcactaatattaccgctcttc
 ttttcctgagggcccggtagggcccgagcgatttgatgtaaagtttggtcctgct
 55 ttgtatgatttatctaaagcagcccat

SEQ ID NO.70

La replicasa de MSV se amplificó usando los siguientes iniciadores:

60

65

	AS965 ggaattcatgaatgaatcgacttgtag	SEQ ID NO.71
	AS966 actaatattacgagggcccggtagggcccgag	SEQ ID NO.72
5	AS967 accgggcccctcgtaatattagtgcgcgcttc	SEQ ID NO.73
	AS968 cggtagcgagtgttggaaccagtaatga	SEQ ID NO.74
10	AS1039 gtgtttgagagaggtatgtaaatatgaactgtat	SEQ ID NO.75
	AS1040 catatttacatacctctctcaaacacagccagag	SEQ ID NO.76
15	AS1041 gtataacttgcaggtactttcattcctaggaagt	SEQ ID NO.77
	AS1042 ggaatgaaagtagctgcaagttatacgatcagga	SEQ ID NO.78
20	Se eliminó el lado derecho del bucle del tallo en la región promotora para evitar la recombinación con el origen de replicación y se insertó un intrón de arroz de alfa tubulina en el gen RepA. La secuencia del intrón está en cursiva y subrayada.	
25	ggaattcatgaatgaatcgacttgtaggcaattttatagtgaccgggaccggattc tttagaatgggctgcttttagataaatcatacaaagcaggaccaaactttacatcaaa tcgctcggggccctaccggggccctcgtaatattagtgcgcgcttcttttcctgctgg gccggcccgggtccggcccatttgcaaccactataagacaaggagcggcggtttcg aaggcgcaagggtatactatagagtttgaggtcccagtcctgccatccctcagcctc 30 aacctccgtcgtcggcaatggcctcctcctcatccaaccgtcagttctcacaccgga acgctaacacgttcttaacctatccaaagtgtccagaaaaatcctgaaatcgctgtc agatgatctgggagctcgtttgttcgttggattcccaaatacattctatgtgcccgag aggcacataaagatggaagtttgcatttacatgcattgcttcagacagagaagccgg 35 taaggatatctgactcaagggttctttgatataaatgggtttcacccaaatattcaga gtgccaagtcagtaaacagggtagggattacattctcaaggaacctctggctgtgt ttgagagaggtatgtaaatatgaactgtatatgacaagtctcactgccatgggttac gtttcaatcaagtcacacaagtgataaatattgtcctgatcgtataacttgcaggtac 40 tttcattcctaggaagtcctccctcctaggaataatctgattcagaggtaaaggaaaa aaagccttctaaagatgaaataatgcgagacattatctcacacgctacttccaaaga agagtacctctcatgatccagaaagagcttccctttgattgggtccacaaaattgca gtattttgaatactctgcaataagctttttcctgagattcaggaagagttcaccaa 45 tctcatccaccctcatcacctgatttactttgtaatgagtcaatcaatgattggct ccagcctaacatcttccagggttagtcccgaagcttacatgctccttcaacctacctg ttataccctcgaggatgcaatttctgacctccaatggatggattctgtatccagtca tcagatgaaagatcaagaaagcagagcctctacatcgctcgcccaacaagaaccgga 50 aaatctacttgggcccagaagcctaggggttcataaataactggcaaaaataatgttgat tggtcttcataacaacgaagacgcaatctataacatcgtagatgatattccgtttaaa ttctgtccttgttggaacagttagttggctgtcagagggatttcattgtaaacccc aagtatggtaaaaaagaaaaaggtgcagaagaagtctaagcctacaataatcctcgcc 55 aactcggatgaagattggatgaaggaaatgactccagggcagctggagtatttcgag gcaaactgcacatctttacattatgtcgccgggggagaaatggattctccccctgag ctgcctcctacggaggcagtagattcagatagatcttgatttttcgatgttctgccc gccgagcaccacccttagccgcccgcgactacccccctgtattgattgtgtgtgtt 60 tttcgtgccatcgcacgacataattaatgtaagctttcagcattcatcagatataata aaaacggcggttttattcattactgggttgccaacactcggtagcg	
		SEQ ID NO.79
65		

ES 2 603 587 T3

Se usó un fragmento del gen de quelatasa de magnesio (MgChe) de arroz para el direccionamiento génico en arroz. Se amplificó a partir de arroz japonica (Nipponbare) usando los siguientes iniciadores:

5	AS1011 gctgcag atcgtggggcaggacgagatga	SEQ ID NO.80
	AS1012 ggcgcgccgtgggagatggagataccctc	SEQ ID NO.81
10	AS1013 ggcgcgccccctgctcggttcacctcat	SEQ ID NO.82
	AS1014 gactagt aagacttcataaaaacttctcaa	SEQ ID NO.83

15 Se introdujo un codón de parada prematuro en el marco para generar un fenotipo mutado.

Gctgcagatcgtggggcaggacgagatgaagctgtgcctgctgctcaacgtcatcga
ccctaagatcggcgggtgtcatgatcatgggagaccgtggcaccggcaaattccaccac
cgtccgctcgtcgtcgcacctgctcccgatatccgcgtcgttgttggcgacccttt
caattccgaccctgacgatcccgaggtcatgggccctgaggtccgggaacgcgtgct
ggaggggtgagaagcttcctgttgtcacggccaagatcacctatggtagatcttcccct
tggtgccactgaggatagagtctgtggcaccattgatattgagaaggcgctcaccga
tggtgtcaaggcgttcgagcctggtttgcttgccaaggccaacagggggattcttta
tgtggatgaggtcaatttgttgatgaccatctagtagatgtgcttctggattctgc
tgcgtcaggatggaacaccgtggagagagaggggtatctccatctcccacggcgcgcc
ccctgctcggttcacctcattgggtctggtaaccccgaggaaggggagctccggcc
acagctgcttgaccggtttggcatgcacgcgcaggttggtactgtcagggatgctga
actcaggggtgaaaattgttgagagagagctcggttcgacagggatccaaaggcgtt
ccgtgagtcctacttggaggaacaagacaagctccagcagcagatttcacatctgctcg
gagtaacottggtgctgtgcagattgaccatgatcttcgtgttaagatttctaaagt
gtgtgcagagttgaatgttgatggattaagaggggacattgtgactaacagggctgc
caaggcgttggcagcactcaaaggcagggacactgtcactgtagaggacattgccac
tgttatccccaactgcttgaggcatcggcttcggaaggaccacttgaatcaattga
ctcaggattgctcgtgggttgagaagttttatgaagtcttactagtg

50
SEQ ID NO.84

55

60

65

REIVINDICACIONES

1. Una secuencia polinucleotídica aislada que codifica para una proteína de recombinación de múltiples dominios y es capaz de iniciar la hibridación entre el ADN donador monocatenario y el ADN objetivo, la secuencia que comprende:
 - i) una secuencia polinucleotídica que codifica para un dominio de unión al ADN donador;
 - ii) una secuencia polinucleotídica que codifica para un dominio de unión cromosómico;
 - iii) una secuencia polinucleotídica que codifica para un tercer dominio que es un dominio de unión al objetivo cromosómico; y
 - iv) una secuencia polinucleotídica que codifica para un cuarto dominio que es un dominio inductor de la recombinación.
2. Una secuencia polinucleotídica aislada de conformidad con la reivindicación 1, que comprende ADN genómico o ADNc.
3. Un método para transformar una célula vegetal que comprende:
 - 1) introducir en dicha célula vegetal una secuencia de ácido nucleico que comprende un promotor nuclear de la planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con la reivindicación 1 o la reivindicación 2; e
 - 2) introducir en dicha célula vegetal una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador que es monocatenario;
 - ii) al menos dos secuencias flanqueantes que son capaces de unirse a un dominio de unión del donador de dicha proteína de recombinación de múltiples dominios que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma; y
 - iii) dos orígenes de replicación viral situados 5' y 3' a cada uno de dichos bordes izquierdo y derecho.
4. Un método de conformidad con la reivindicación 3, que comprende además introducir una secuencia de ácido nucleico que comprende un promotor nuclear de la planta o un promotor nativo viral unido operativamente a una tercera secuencia de ácido nucleico que codifica para una proteína de replicasa viral.
5. Un método para transformar una célula vegetal que comprende:
 - 1) introducir en dicha célula vegetal una secuencia de ácido nucleico que comprende un promotor nuclear de la planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con la reivindicación 1 o la reivindicación 2; e
 - 2) introducir en dicha célula vegetal una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador que es monocatenario;
 - ii) al menos dos secuencias flanqueantes que son capaces de unirse a un dominio de unión del donador de dicha proteína de recombinación de múltiples dominios que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma, en donde dicha segunda secuencia de ácido nucleico está operativamente unida a un promotor nuclear de la planta y comprende además un sitio de unión del iniciador que se fusiona al extremo 3' de una secuencia de borde flanqueante derecho; e
 - iii) introducir una secuencia de ácido nucleico que comprende un promotor nuclear de planta unido operativamente a una tercera secuencia de ácido nucleico que codifica para una proteína de transcriptasa inversa.
6. Un método de conformidad con la reivindicación 5, en donde el sitio de unión del iniciador se selecciona de un sitio de unión del iniciador específico del retrotransposón, y un sitio de unión del iniciador de un retrovirus.
7. Un método de conformidad con la reivindicación 5 o la reivindicación 6, en donde dicha tercera secuencia de ácido nucleico se selecciona del ADN que codifica el ARN retrotransposon o el ADN que codifica el ARN retroviral, tal como la transcriptasa inversa-RNAasa H.
8. Una célula vegetal que comprende una proteína de recombinación de múltiples dominios que comprende
 - i) un dominio de unión al ADN donador;
 - ii) un dominio de unión cromosómico
 - iii) un tercer dominio que es un dominio de unión al objetivo cromosómico; y
 - iv) un cuarto dominio que es un dominio inductor de la recombinación.

9. Una célula vegetal de conformidad con la reivindicación 8, en donde la célula vegetal se selecciona de tabaco (*Nicotiana tabacum*) y otras especies de *Nicotiana* arabis, patata, grano (maíz), canola (colza), arroz, trigo, cebada, brassica sp. seleccionada de coliflor, brócoli (por ejemplo, brotes verdes y púrpuras), repollo (por ejemplo, coles rojas, verdes y blancas), col rizada, coles de Bruselas, algodón, algas (por ejemplo, especies de verde azul), lemnospora o musgo (por ejemplo, *physcomitrella patens*), tomate, pimiento, calabaza, girasol, soja, zanahoria, melones, vid, lechuga, fresa, remolacha azucarera, guisantes y sorgo.
10. Un método para producir al menos una especie de ARN heterólogo o exógeno en una planta que comprende:
 - 1) introducir en una célula vegetal regenerable una secuencia de ácido nucleico que comprende un promotor nuclear de la planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con la reivindicación 1 o la reivindicación 2;
 - 2) introducir en dicha célula vegetal regenerable una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador que es monocatenario; y
 - ii) al menos dos secuencias flanqueantes que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma; y
 - iii) dos orígenes de replicación viral situados 5' y 3' a dichos bordes izquierdo y derecho, en donde dicha segunda secuencia de ácido nucleico está operativamente unida a un promotor nuclear de la planta;
 - 3) introducir en dicha célula vegetal regenerable una tercera secuencia de ácido nucleico que codifica para una proteína de replicasa viral, en donde la tercera secuencia de ácido nucleico está unida operativamente a un promotor de planta nuclear o viral nativo;
 - 4) cultivar dicha célula vegetal regenerable de las etapas 1) a 3);
 - 5) seleccionar una célula vegetal de 4);
 - 6) regenerar una planta a partir de la célula vegetal de 5); y
 - 7) cultivar la planta de 6).
11. Un método para producir al menos una especie de ARN heterólogo o exógeno en una planta que comprende:
 - 1) introducir en una célula vegetal regenerable una secuencia de ácido nucleico que comprende un promotor nuclear de la planta unido operativamente a una primera secuencia de ácido nucleico que comprende una secuencia polinucleotídica de conformidad con la reivindicación 1 o la reivindicación 2;
 - 2) introducir en dicha célula vegetal regenerable una segunda secuencia de ácido nucleico que comprende
 - i) una secuencia de ácido nucleico donador que es monocatenario;
 - ii) al menos dos secuencias flanqueantes que se sitúan adyacentes a cada extremo de la secuencia polinucleotídica del donador formando bordes izquierdo y derecho a la misma, en donde dicha segunda secuencia de ácido nucleico está unida operativamente a un promotor nuclear de la planta; y
 - iii) un sitio de unión del iniciador que se fusiona con el extremo 3' de dicha secuencia de borde flanqueante derecho;
 - 3) introducir en dicha célula vegetal regenerable una tercera secuencia de ácido nucleico que codifica para una proteína de transcriptasa inversa, en donde la tercera secuencia de ácido nucleico está unida operativamente a un promotor nuclear de la planta;
 - 4) cultivar dicha célula vegetal regenerable de las etapas 1) a 3);
 - 5) seleccionar una célula vegetal de 4);
 - 6) regenerar una planta a partir de la célula vegetal de 5); y
 - 7) cultivar la planta de 6).
12. Una secuencia polinucleotídica aislada que comprende una secuencia de ácido nucleico donador que es monocatenario, al menos dos secuencias flanqueantes situadas adyacentes a cada extremo de dicha secuencia de ácido nucleico donador que forman bordes izquierdo y derecho a la misma y un origen de replicación viral situado 5' y 3' a dichos bordes izquierdo y derecho, respectivamente, adecuados para su uso en un método de conformidad con la reivindicación 3.
13. Una secuencia polinucleotídica aislada de conformidad con la reivindicación 12, que comprende una secuencia de ácido nucleico adicional que comprende un promotor nuclear o nativo viral de la planta unido operativamente a una secuencia de ácido nucleico que codifica para una proteína de replicasa viral adecuada para su uso en un método de conformidad con la reivindicación 3 o reivindicación 10.
14. Un vector de ácido nucleico adecuado para la transformación de una célula vegetal o célula bacteriana tal como una célula de *Agrobacterium* y que incluye un polinucleótido de conformidad con la reivindicación 1 o la reivindicación 2.
15. Una célula huésped tal como una célula vegetal que contiene un polinucleótido heterólogo o un vector de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1, 2, 12, 13 y 14 o una célula huésped tal como una célula vegetal que contiene un polinucleótido heterólogo o un vector de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1,2, 12, 13 y 14 y comprendido en una planta, una parte de la planta, un propágulo de planta o un cultivo de célula vegetal.

16. Un método para producir una célula de conformidad con la reivindicación 15, el método que incluye incorporar dicho polinucleótido o vector de ácido nucleico en la célula por medio de transformación.
- 5 17. Un método de conformidad con la reivindicación 16 que incluye regenerar una planta a partir de una célula de conformidad con la reivindicación 15 a partir de una o más células transformadas.
- 10 18. Una planta que comprende una célula vegetal de conformidad con la reivindicación 15, que se selecciona del grupo que consiste en tabaco (*Nicotiana tabacum*) y otras especies de *Nicotiana* tal como *Nicotiana benthamiana*, zanahoria, vegetal y semillas oleaginosas de *Brassica*, melones, pimiento, vid, lechuga, fresa, remolacha azucarera, trigo, cebada, (grano)maíz, arroz, soja, guisantes, sorgo, girasol, tomate, algodón y patata.

Figura 1

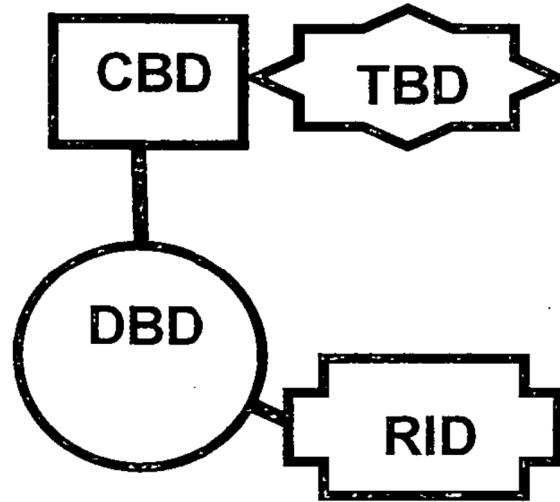
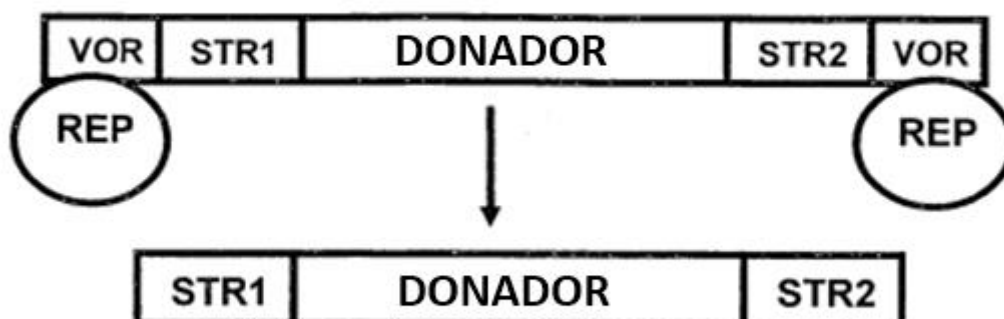


Figura 2

(A)



(B)

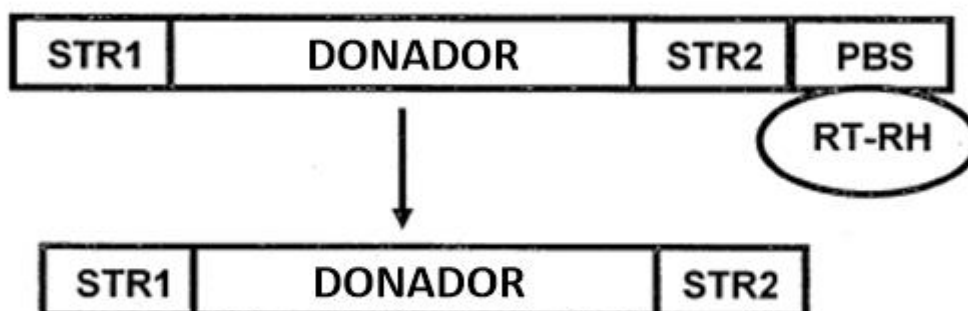


Figura 3

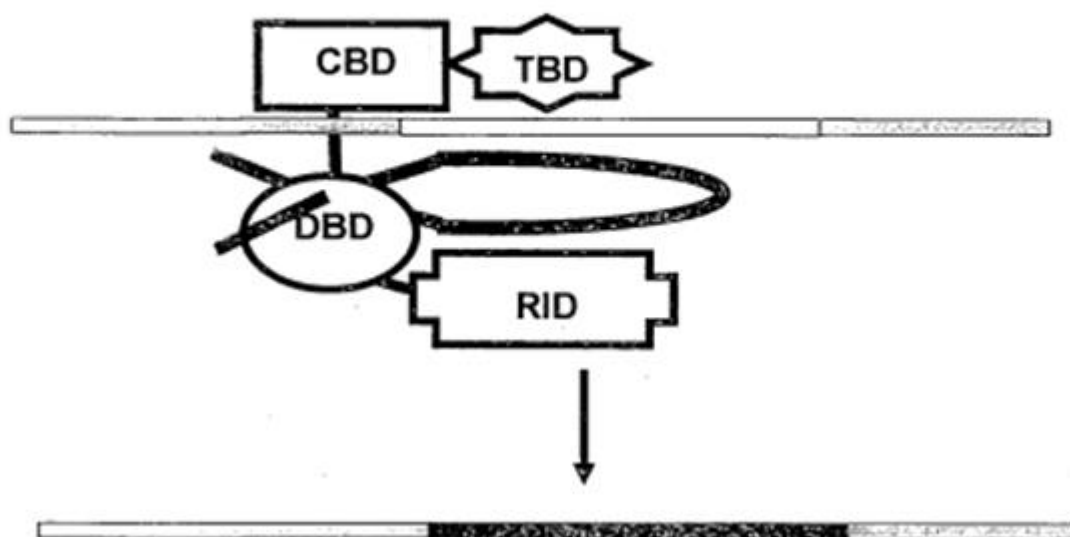


Figura 4

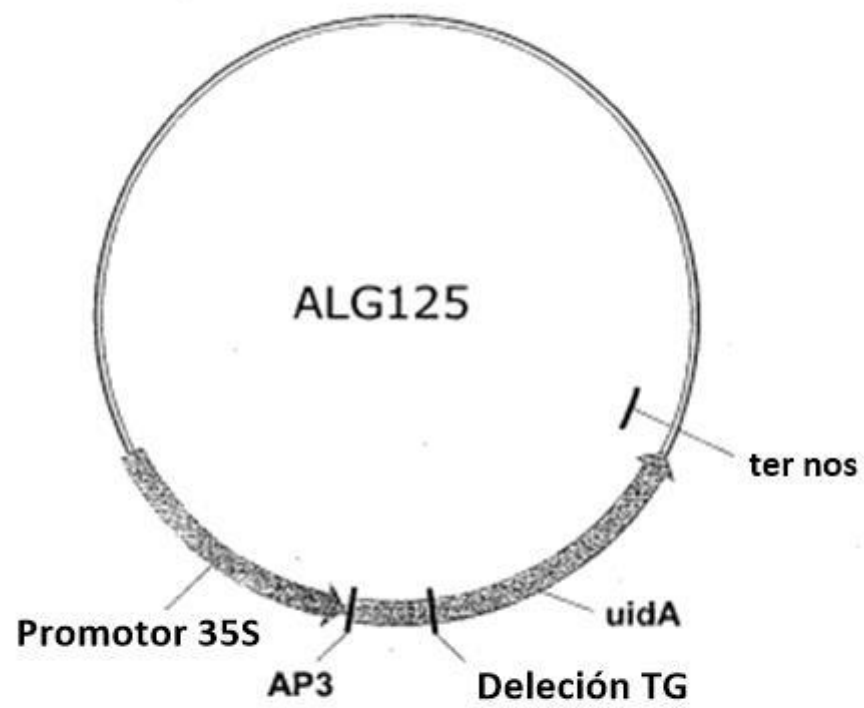


Figura 5

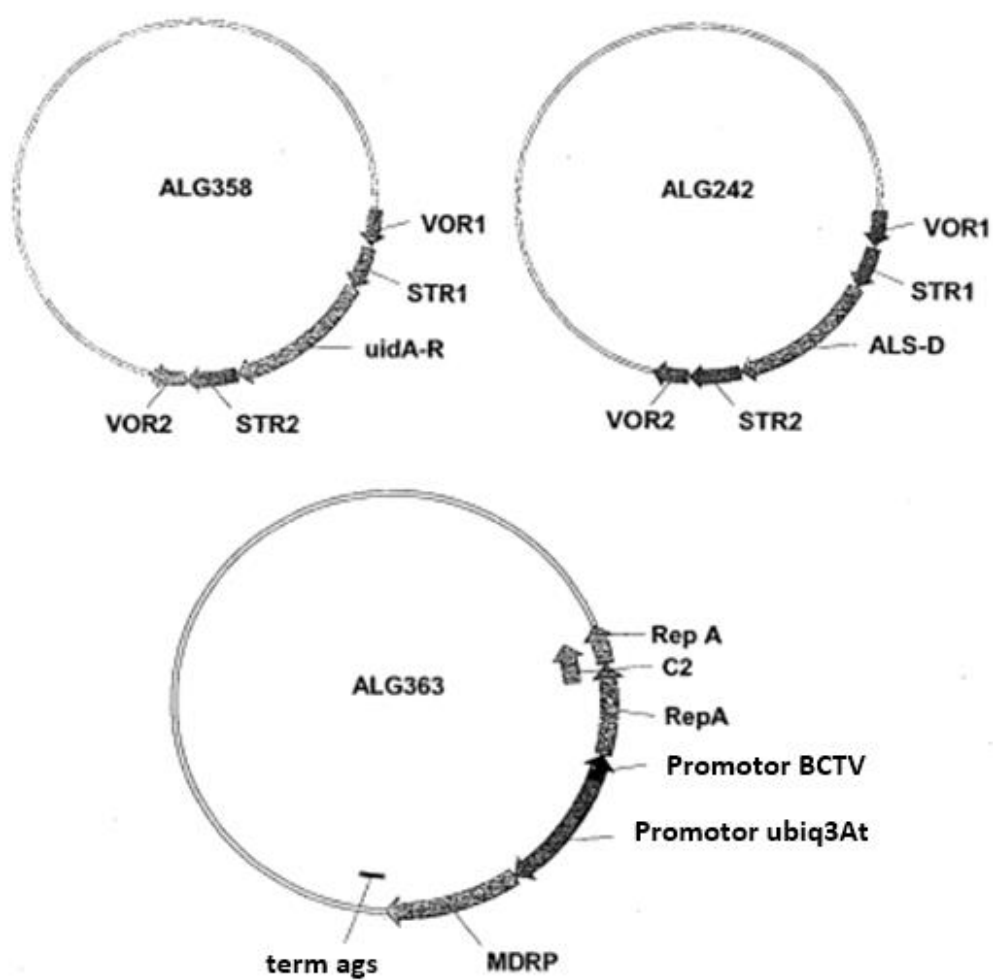


Figura 6

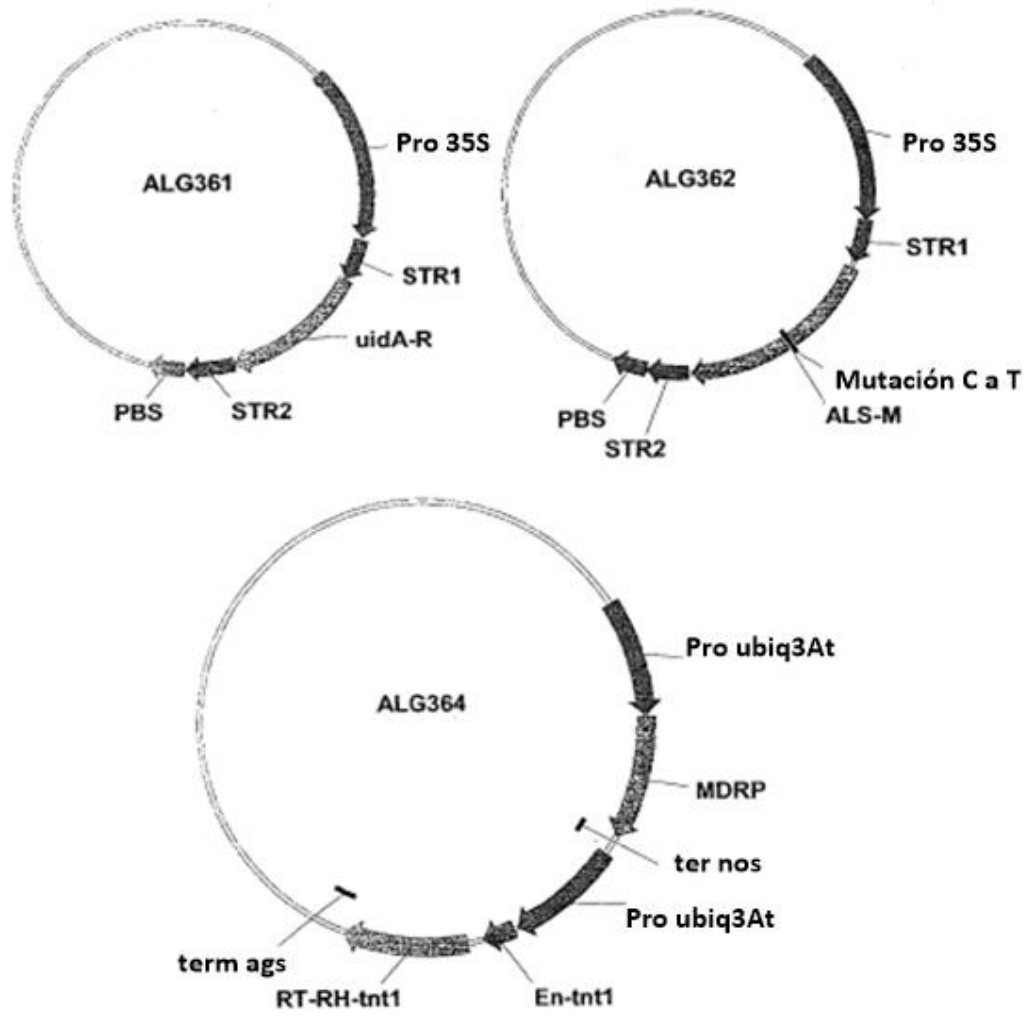


Figura 7

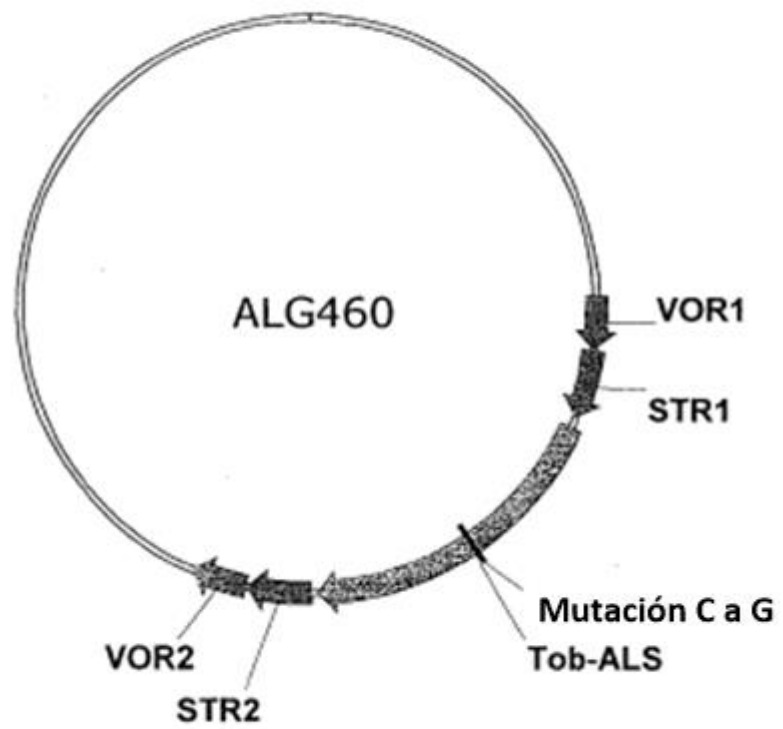


Figura 8

