

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 532 002**

51 Int. Cl.:

C12N 15/09 (2006.01)

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.08.2007** **E 07811226 (5)**

97 Fecha y número de publicación de la concesión europea: **25.02.2015** **EP 2049663**

54 Título: **Recombinación homóloga mediada por nucleasas con dedos de zinc**

30 Prioridad:

11.08.2006 US 837147 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.03.2015

73 Titular/es:

DOW AGROSCIENCES LLC (50.0%)
9330 Zionsville Road
Indianapolis IN 46268-1054, US y
SANGAMO BIOSCIENCES, INC. (50.0%)

72 Inventor/es:

CAI, QIHUA;
MILLER, JEFFREY;
AINLEY, WILLIAM MICHAEL;
GARRISON, ROBBY JANETTE;
PETOLINO, JOSEPH F.;
RUBIN-WILSON, BETH CORI;
SCHULENBERG, LISA LYNN y
WORDEN, ANDREW FREDERICK

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 532 002 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

Recombinación homóloga mediada por nucleasas con dedos de zinc

Campo técnico

5 La presente descripción está en el campo de la ingeniería genómica, reconocimiento génico, integración cromosómica dirigida y expresión de proteínas en plantas.

Antecedentes

10 Un campo de interés principal en la agricultura, en especial a la luz de la determinación de las secuencias de nucleótidos completas de una serie de genomas de plantas, es la alteración dirigida de secuencias genómicas. En particular, la capacidad para convertir secuencias endógenas de plantas facilitaría numerosas aplicaciones tales como, por ejemplo, la optimización de características de los cultivos que afectan al valor nutricional, rendimiento, tolerancia al estrés, resistencia a patógenos y resistencia a compuestos agroquímicos y/o la adaptación de plantas para usar como fábricas biológicas para la producción de compuestos farmacéuticos o productos químicos industriales.

15 En eucariotas, se han hecho intentos de alterar secuencias genómicas en células cultivadas aprovechando el fenómeno natural de la recombinación homóloga. Véase, por ejemplo, Capecchi (1989) *Science* 244:1288-1292; patentes de EE.UU. nº 6.528.313 y 6.528.314. Si un polinucleótido tiene suficiente homología con la región genómica que contiene la secuencia que se va a alterar, es posible que parte o toda la secuencia del polinucleótido sustituya la secuencia genómica por recombinación homóloga. Sin embargo, la frecuencia de la recombinación homóloga en estas circunstancias es extremadamente baja. Además, la frecuencia de inserción del polinucleótido exógeno en sitios genómicos que carecen de homología de secuencia supera la frecuencia de la recombinación homóloga en varios órdenes de magnitud.

20 La introducción de una rotura bicatenaria en el ADN genómico, en la región del genoma que contiene homología con un polinucleótido exógeno, se ha mostrado que simula la recombinación homóloga en este sitio en varios miles de veces en células cultivadas. Rouet et al. (1994) *Mol. Cell. Biol.* 14:8096-8106; Choulika et al. (1995) *Mol. Cell. Biol.* 15:1968-1973; Donoho et al. (1998) *Mol. Cell. Biol.* 18:4070-4078. Véase también Johnson et al. (2001) *Biochem. Soc. Trans.* 29:196-201; y Yanez et al. (1998) *Gene Therapy* 5:149-159. En estos métodos, la escisión del ADN en la región genómica deseada se llevaba a cabo por inserción de un sitio de reconocimiento para una meganucleasa (es decir, una endonucleasa cuya secuencia de reconocimiento es tan grande que no se produce, o se produce raramente en el genoma de interés) en la región genómica deseada.

30 Sin embargo, la recombinación homóloga estimulada por escisión por meganucleasa se basa en la presencia fortuita o en la inserción dirigida de un sitio de reconocimiento de la meganucleasa adecuado en la proximidad de la región genómica que se va a alterar. Puesto que los sitios de reconocimiento de las meganucleasas son raros (o no existen) en un genoma de planta típico, y la inserción de un sitio de reconocimiento de meganucleasa adecuado está lleno de las mismas dificultades asociadas con otras alteraciones genómicas, estos métodos no son ampliamente aplicables.

35 Por lo tanto, siguen siendo necesarias composiciones y métodos para la alteración dirigida de secuencias en cualquier genoma de planta, y composiciones y métodos para la introducción dirigida de secuencias exógenas en un genoma.

40 Durai et al., *Nucleic Acids Research*, 2005, Vol. 33, 5978-5990 se refiere a las nucleasas con dedos de zinc como tijeras moleculares diseñadas de forma personalizada para la ingeniería genómica de la planta y células de mamíferos.

Porteus y Carroll, *Nature Biotechnology*, 2005, Vol. 23, 967-973 se refiere al reconocimiento génico usando nucleasas con dedos de zinc.

Compendio

45 La presente invención se define mediante el conjunto de reivindicaciones adjuntas.

50 La presente descripción proporciona composiciones y métodos para la escisión dirigida de la cromatina celular en una región de interés y/o la recombinación homóloga en una región predeterminada de interés en células de plantas. Las células de plantas pueden ser de especies de plantas monocotiledóneas (monocotas) o dicotiledóneas (dicotas), y también se incluyen células cultivadas, células en una planta en cualquier etapa del desarrollo y células de plantas que se han separado de la planta entera y cuyas células (o sus descendientes) se devolverán a la planta.

Una región de interés en la cromatina celular puede ser, por ejemplo, una secuencia genómica o parte de la misma. Las composiciones incluyen polipéptidos de fusión que comprenden un dominio de unión con dedos de zinc genéticamente modificado (p. ej., un dominio de unión con dedos de zinc que tiene una especificidad nueva) y un dominio de escisión, y polipéptidos de fusión que comprenden un dominio de unión con dedos de zinc

genéticamente modificado y un semidominio de escisión. Los dominios de escisión y semidominios de escisión se pueden obtener, por ejemplo, de diferentes endonucleasas de restricción y/o endonucleasas de asentamiento.

5 En un aspecto, se describe en la presente memoria un vector que comprende una primera y segunda secuencias de ADN, en donde (i) la primera secuencia es homóloga con una tercera secuencia y la segunda secuencia es homóloga con una cuarta secuencia; (ii) la tercera y cuarta secuencias son secuencias de ADN cromosómico; y (iii) los bordes cercanos de la tercera y cuarta secuencias están separados por al menos 1 par de nucleótidos. En algunas realizaciones, la tercera y cuarta secuencias son secuencias endógenas.

10 En cualquiera de los vectores descritos en la presente memoria, al menos una de la primera o segunda secuencias tiene una longitud de 100 nucleótidos. Además, cualquier de los vectores descritos en la presente memoria puede comprender una quinta secuencia, en donde la quinta secuencia: (a) está interpuesta entre la primera y segunda secuencias; y (b) es una secuencia exógena. En algunas realizaciones, la quinta secuencia tiene un tamaño de al menos 1 par de bases, pero puede ser tan grande como de 22 kilopares de bases.

15 Los vectores (p. ej., la quinta secuencia) pueden comprender también secuencias que codifican una proteína o partes de una proteína. En algunas realizaciones, la secuencia que codifica proteína codifica un marcador seleccionable (p. ej., proteína verde fluorescente (GFP), β -glucuronidasa (GUS), fosfinotricina N-acetil-transferasa (PAT, BAR), neomicina fosfotransferasa, β -lactamasa, catecol dioxigenasa, α -amilasa, tirosinasa, β -galactosidasa, luciferasa, aequorina, EPSP sintasa, nitrilasa, acetolactato sintasa (ALS), dihidrofolato reductasa (DHFR), dalapon deshalogenasa y antranilato sintasa). En otras realizaciones, la secuencia que codifica proteína (p. ej., la quinta secuencia) codifica una proteína o parte de proteína, por ejemplo, una secuencia que es homóloga con secuencias cromosómicas.

20

En otras realizaciones más, los vectores (p. ej., la quinta secuencia) comprende una o más secuencias reguladoras transcripcionales. En otras realizaciones más, los vectores (p. ej., la quinta secuencia) puede comprender un equivalente natural de una secuencia cromosómica mutante, o alternativamente un equivalente mutante de una secuencia cromosómica natural.

25 En cualquiera de los vectores descritos en la presente memoria, la primera secuencia tiene una homología de al menos 90% a 95%, al menos 95%, 96%, 97%, 98%, 99% o 1000% con la cuarta secuencia.

30 En otro aspecto más, se describe en la presente memoria un método para introducir una secuencia exógena en el genoma de una célula de planta, comprendiendo el método las etapas de: (a) poner en contacto la célula con cualquiera de los vectores descritos antes; y (b) expresar una o más nucleasas en la célula, en donde la una o más nucleasas escinden el ADN cromosómico en el espacio entre 0,4 y 3 kilopares de bases de cualquiera de la tercera o cuarta secuencias; de modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación del vector director en el genoma por recombinación homóloga.

La una o más nucleasas son fusiones entre el dominio de escisión de una endonucleasa de restricción FokI y un dominio de unión con dedos de zinc genéticamente modificado.

35 También se describe en la presente memoria, un método para expresar una proteína en una célula de planta, comprendiendo el método las etapas de: (a) poner en contacto la célula con un vector como se describe en la presente memoria; y (b) expresar una o más nucleasas en la célula, en donde la una o más nucleasas escinden el ADN cromosómico en el espacio entre 0,1 y 3 kilopares de bases de cualquiera de la tercera o cuarta secuencias; de modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación del vector en el genoma por recombinación homóloga. La una o más nucleasas pueden ser fusiones entre el dominio de escisión de una endonucleasa de restricción de tipo IIS y un dominio de unión con dedos de zinc genéticamente modificado.

40

También se describe en la presente memoria, un vector director que comprende: (a) primera y segunda secuencias codificantes; y (b) primera, segunda y tercer secuencias codificantes; en donde la secuencias están dispuestas en el orden: primera secuencia diana, primera secuencia codificante, segunda secuencia diana, segunda secuencia codificante, tercera secuencia diana; en donde además la primera secuencia diana es homóloga con una primera secuencia cromosómica y la tercera secuencia diana es homóloga con una segunda secuencia cromosómica. La primera y/o segunda secuencia codificantes pueden codificar un marcador seleccionable o, alternativamente, la primera y/o segunda secuencia codificantes pueden codificar una proteína que no es un marcador seleccionable. La primera y segunda secuencias cromosómicas pueden ser secuencias cromosómicas endógenas. Además, los vectores pueden comprender una o más repeticiones de la primera, segunda y/o tercera secuencias diana.

45

50

Se describe además un método para introducir una secuencia exógena en el genoma de una célula de planta, comprendiendo el método las etapas de: (a) poner en contacto la célula con un vector director como se describe en el párrafo precedente; y (b) expresar una o más nucleasas en la célula, en donde la una o más nucleasas escinden el ADN cromosómico en el espacio entre 0,1 y 3 kilopares de bases de cualquiera de la primera o segunda secuencias; de modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación del vector director en el genoma por recombinación homóloga. La una o más nucleasas pueden comprender un semidominio de escisión; por ejemplo, la nucleasa es una fusión entre el dominio de escisión de una endonucleasa de restricción de tipo IIS y un dominio de unión con dedos de zinc genéticamente modificado.

55

- Se describe un método para expresar una proteína en una célula de planta, comprendiendo el método las etapas de: (a) poner en contacto la célula con un vector director como se describe dos párrafos antes; y (b) expresar una o más nucleasas en la célula, en donde la una o más nucleasas escinden el ADN cromosómico en el espacio entre 0,1 y 3 kilopares de bases de cualquiera de la primera o segunda secuencias; de modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación del vector director en el genoma por recombinación homóloga. Una o más nucleasas pueden ser fusiones entre el dominio de escisión de una endonucleasa de restricción de tipo IIS y un dominio de unión con dedos de zinc genéticamente modificado.
- También se describe una célula de planta transgénica obtenida de acuerdo con cualquiera de los métodos descritos en la presente memoria.
- Se describe además una planta que comprende una célula de planta transgénica obtenida como se describe en la presente memoria.
- Se describe además un método para eliminar secuencias del genoma de una célula de planta transgénica que comprende la primera y segunda secuencias codificantes; y la primera, segunda y tercera secuencias diana; en donde las secuencias están dispuestas en el orden: primera secuencia diana, primera secuencia codificante, segunda secuencia diana, segunda secuencia codificante, tercera secuencia diana, en donde el método comprende: (a) proporcionar una célula de planta transgénica como se describe en la presente memoria; y (b) expresar la primera y segunda nucleasas en la célula, en donde la primera nucleasa escinde en la primera secuencia diana y la segunda nucleasa escinde en la segunda secuencia diana.
- Se describe también en la presente memoria, un método para eliminar secuencias del genoma de una célula de planta transgénica que comprende la primera y segunda secuencias codificantes; y la primera, segunda y tercera secuencias diana; en donde las secuencias están dispuestas en el orden: primera secuencia diana, primera secuencia codificante, segunda secuencia diana, segunda secuencia codificante, tercera secuencia diana, en donde el método comprende: (a) proporcionar una célula de planta transgénica como se describe en la presente memoria; y (b) expresar la primera y segunda nucleasas en la célula, en donde la primera nucleasa escinde en la segunda secuencia diana y la segunda nucleasa escinde en la tercera secuencia diana.
- Se describe además en la presente memoria, un método para eliminar secuencias del genoma de una célula de planta transgénica que comprende la primera y segunda secuencias codificantes; y la primera, segunda y tercera secuencias diana; en donde las secuencias están dispuestas en el orden: primera secuencia diana, primera secuencia codificante, segunda secuencia diana, segunda secuencia codificante, tercera secuencia diana, en donde el método comprende: (a) proporcionar una célula de planta transgénica como se describe en la presente memoria; y (b) expresar la primera y segunda nucleasas en la célula, en donde la primera nucleasa escinde en la primera secuencia diana y la segunda nucleasa escinde en la tercera secuencia diana.
- Se describe un método para la recombinación homóloga intramolecular en el genoma de una célula (p. ej., célula de planta), comprendiendo el método las etapas de: (a) proporcionar un segmento de ADN que comprende una primera secuencia que es homóloga con una segunda secuencia; y (b) poner en contacto dicho segmento de ADN con una nucleasa, en donde la nucleasa escinde el segmento de ADN en la tercera secuencia. El segmento de ADN puede ser endógeno para la célula. La recombinación homóloga se puede producir en un cromosoma, por ejemplo, cuando se elimina del cromosoma el ADN entre la primera y segunda secuencias. Las secuencias eliminadas del cromosoma pueden codificar, por ejemplo, todo o parte de un marcador seleccionable. El ADN eliminado se puede sustituir por una secuencia exógena, por ejemplo, en donde el método comprende además: introducir un polinucleótido en la célula, en donde el polinucleótido comprende: (a) la cuarta y quinta secuencias, en donde la cuarta secuencia es homóloga con secuencias no eliminadas cercanas a la primera secuencia y la quinta secuencia es homóloga con secuencias no eliminadas cercanas a la segunda secuencia; y (b) la secuencia exógena (p. ej., un marcador seleccionable, una proteína o parte de una proteína distinta de un marcador seleccionable, un ARN tal como ARNip). En cualquiera de los métodos proporcionados en la presente memoria, el marcador seleccionable puede ser, por ejemplo, la proteína verde fluorescente (GFP), β -glucuronidasa (GUS), fosfotricina N-acetiltransferasa (PAT, BAR), neomicina fosfotransferasa, β -lactamasa, catecol dioxigenasa, α -amilasa, tirosinasa, β -galactosidasa, luciferasa, aequorina, EPSP sintasa, nitrilasa, acetolactato sintasa (ALS), dihidrofolato reductasa (DHFR), dalapon deshalogenasa y antranilato sintasa.
- En cualquiera de los métodos, la tercera secuencia (es decir, la secuencia escindida por la nucleasa) puede ser única en el genoma. La nucleasa es un par de proteínas de fusión, en donde cada proteína de fusión es una fusión entre el dominio de escisión de FokI y un dominio de unión con dedos de zinc genéticamente modificado. Además, la tercera secuencia puede estar entre la primera y segunda secuencias (es decir, las secuencias homólogas) y/o al menos 1 par de bases desde la primera y/o segunda secuencias.
- En cualquiera de los métodos descritos en la presente memoria, una o tanto la primera como la segunda secuencias pueden ser exógenas para el organismo.

Breve descripción de los dibujos

Las figuras 1A y 1B son representaciones esquemáticas de la construcción directora denominada pDAB1585. La

figura 1A es una representación lineal de varios elementos de la construcción. La figura 1B es una representación de la construcción circular.

Las figuras 2A y 2B representan ZFN de ejemplo y sus sitios diana. La figura 2A representa los ZFN y las regiones de unión. La figura 2B representa los sitios diana.

- 5 La figura 3 es una representación esquemática del plásmido pDAB1400.
La figura 4 es una representación esquemática del plásmido pDAB782.
La figura 5 es una representación esquemática del plásmido pDAB1582.
La figura 6 es una representación esquemática del plásmido pDAB354.
La figura 7 es una representación esquemática del plásmido pDAB1583.
- 10 La figura 8 es una representación esquemática del plásmido pDAB2407.
La figura 9 es una representación esquemática del plásmido pDAB1584.
La figura 10 es una representación esquemática del plásmido pDAB2418.
La figura 11 es una representación esquemática del plásmido pDAB4045.
La figura 12 es una representación esquemática del plásmido pDAB1575.
- 15 La figura 13 es una representación esquemática del plásmido pDAB1577.
La figura 14 es una representación esquemática del plásmido pDAB1579.
La figura 15 es una representación esquemática del plásmido pDAB1580.
La figura 16 es una representación esquemática del plásmido pDAB3401.
La figura 17 es una representación esquemática del plásmido pDAB1570.
- 20 La figura 18 es una representación esquemática del plásmido pDAB1572.
La figura 19 es una representación esquemática del plásmido pDAB4003.
La figura 20 es una representación esquemática del plásmido pDAB1571.
La figura 21 es una representación esquemática del plásmido pDAB7204.
La figura 22 es una representación esquemática del plásmido pDAB1573.
- 25 La figura 23 es una representación esquemática del plásmido pDAB1574.
La figura 24 es una representación esquemática del plásmido pDAB1581.
La figura 25 es una representación esquemática del plásmido pDAB1576.
- Las figuras 26A y 26B son representaciones esquemáticas del vector plasmídico director denominado pDAB1600. La figura 26A es una representación lineal de varios elementos del plásmido. La figura 26B es una representación del plásmido circular.
- 30 La figura 27 es una representación esquemática del plásmido pDAB7002.
La figura 28 es una representación esquemática del plásmido pDAB7025.
La figura 29 es una representación esquemática del plásmido pDAB1591.
La figura 30 es una representación esquemática del plásmido pcDNA3.1-IL1-L0-FokI.
- 35 La figura 31 es una representación esquemática del plásmido pcDNA3.1-SCD27-L0-FokI.
La figura 32 es una representación esquemática del plásmido pDAB1592.
La figura 33 es una representación esquemática del plásmido pDAB1594.
- Las figuras 34 A, B y C son representaciones esquemáticas de los plásmidos pDAB1596 y pDAB1598. La figura 34A es un esquema de los plásmidos linearizados. La figura 34B muestra pDAB1596. La figura 34C muestra pDAB1598.

La figura 35 es una representación esquemática del plásmido pDAB1577.

La figura 36 es una representación esquemática del plásmido pDAB1578.

Las figuras 37A y B son representaciones esquemáticas del plásmido pDAB1601. La figura 37A muestra varios elementos en forma lineal. La figura 37B muestra el plásmido circular.

5 La figura 38 es un esquema que representa la recombinación homóloga intercromosómica prevista estimulada por la fusión IL-1 ZFN-FokI.

La figura 39 es un esquema que representa la recombinación homóloga intercromosómica prevista estimulada por la fusión Scd27 ZFN-FokI.

10 La figura 40 muestra el análisis por PCR de los recombinantes. Las bandas designadas 1-20 muestran sucesos de recombinación homóloga a partir de la transformación de NT1-240 con la construcción de la proteína de fusión SCD27-FokI. Las bandas designadas 21 y 22 muestran sucesos de recombinación homóloga a partir de la transformación de NT1-240 con construcciones de la proteína de fusión SCD27-FokI. Se muestran bandas de control en las 3 bandas más a la izquierda.

15 La figura 41 muestra el análisis de transferencia Southern de los recombinantes. Las bandas designadas 1-20 muestran sucesos de recombinación homóloga a partir de la transformación de NT1-240 con la construcción de la proteína de fusión SCD27-FokI. Las bandas designadas 21 y 22 muestran sucesos de recombinación homóloga a partir de la transformación de NT1-240 con construcciones de la proteína de fusión SCD27-FokI. Se muestran bandas de control en las 2 bandas más a la izquierda.

20 La figura 42 es un esquema que representa la recombinación homóloga intracromosómica prevista estimulada por la fusión IL-1 ZFN-FokI.

La figura 43 es el análisis por PCR que confirma que la GFP es reconstituida en tejidos fluorescentes que expresan la proteína de fusión IL-1-FokI.

La figura 44 es una representación esquemática del plásmido pSB11.

La figura 45 es una representación esquemática del plásmido pSB1.

25 La figura 46 es una representación esquemática del plásmido pDAB3872.

La figura 47 es una representación esquemática del plásmido pDAB4365.

Descripción detallada de la invención

30 Se describen en la presente memoria métodos útiles para la escisión dirigida de cromatina celular de planta y para la alteración dirigida de una secuencia de nucleótidos celular de planta, p. ej., mediante la escisión dirigida seguido de recombinación homóloga intracromosómica, o mediante la escisión dirigida seguido de recombinación homóloga entre un polinucleótido exógeno (que comprende una o más regiones de homología con la secuencia de nucleótidos celular) y una secuencia genómica.

35 Las secuencias genómicas incluyen las presentes en cromosomas, episomas, genomas de orgánulos (p. ej., mitocondrias, cloroplastos), cromosomas artificiales y cualquier otro tipo de ácido nucleico presente en una célula, tal como, por ejemplo, secuencias amplificadas, cromosomas doble minuto y los genomas de bacterias y virus endógenos o infecciosos. Las secuencias genómicas pueden ser normales (es decir, naturales) o mutantes; las secuencias mutantes pueden comprender, por ejemplo, inserciones, eliminaciones, translocaciones, reordenaciones y/o mutaciones puntuales. Una secuencia genómica también puede comprender uno o una serie de alelos diferentes.

40 Las composiciones útiles para la escisión dirigida y recombinación incluyen proteínas de fusión que comprenden un dominio de escisión (o un semidominio de escisión) y un dominio de unión con dedos de zinc, polinucleótidos que codifican estas proteínas y combinaciones de polipéptidos y polinucleótidos que codifican polipéptidos. Un dominio de unión con dedos de zinc puede comprender uno o más dedos de zinc (p. ej., 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de zinc), y se puede modificar genéticamente para que se una a cualquier secuencia genómica. Por lo tanto, 45 identificando una región genómica diana de interés en la que se desea la escisión o recombinación, según los métodos descritos en la presente memoria, se pueden construir una o más proteínas de fusión que comprenden un dominio de escisión (o semidominio de escisión) y un dominio con dedos de zinc genéticamente modificado para reconocer una secuencia diana en dicha región genómica. La presencia de dicha proteína (o proteínas) de fusión en una célula dará como resultado la unión de la(s) proteína(s) de fusión en su(s) sitio(s) de unión y escisión dentro o 50 cerca de dicha región genómica. Además, si también está presente un polinucleótido exógeno homólogo con la región genómica en dicha célula, la recombinación homóloga se produce con una tasa mayor entre la región genómica y el polinucleótido exógeno.

General

La práctica de los métodos, así como la preparación y uso de las composiciones descritas en la presente memoria, usan, salvo que se indique otra cosa, técnicas convencionales de biología molecular, bioquímica, estructura y análisis de cromatina, química computacional, cultivo celular, ADN recombinante y campos relacionados, como conoce el experto en la técnica. Estas técnicas están explicadas con detalle en la bibliografía. Véase, por ejemplo, Sambrook et al. *Molecular Cloning: A Laboratory Manual*, Second edition, Cold Spring Harbor Laboratory Press, 1989 y tercera edición, 2001; Ausubel et al., *Current Protocols In Molecular Biology*, John Wiley & Sons, New York, 1987 y actualizaciones periódicas; la serie de *Methods In Enzymology*, Academic Press, San Diego; Wolffe, *Chromatin Structure And Function*, 3ª edición, Academic Press, San Diego, 1998; *Methods In Enzymology*, Vol. 304, "Chromatin" (P.M. Wassarman y A. P. Wolffe, eds.), Academic Press, San Diego, 1999; y *Methods In Molecular Biology*, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definiciones

Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan de forma intercambiable y se refieren a un polímero desoxirribonucleótido o ribonucleótido en una conformación lineal o circular, y en forma monocatenaria o bicatenaria. Para los propósitos de la presente descripción, estos términos no deben considerarse limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que son modificados en los restos de base, azúcar y/o fosfato (p. ej., cadenas principales de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases, es decir, un análogo de A dará emparejamiento de base con T.

Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable y se refieren a un polímero de restos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de un aminoácido natural correspondiente.

"Unión" se refiere a una interacción no covalente, específica de secuencia, entre macromoléculas (p. ej., entre una proteína y un ácido nucleico). No es necesario que todos los componentes de una interacción de unión sean específicos de la secuencia (p. ej., se ponen en contacto con restos de fosfato en una cadena principal de ADN), siempre que la interacción en conjunto sea específica de la secuencia. Dichas interacciones en general se caracterizan por una constante de disociación (K_d) de $10^{-6} M^{-1}$ o menor. La "afinidad" se refiere a la fuerza de unión: correlacionándose una mayor afinidad de unión con una menor K_d .

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

Una "proteína de unión a ADN con dedos de zinc" (o dominio de unión) es una proteína o un dominio dentro de una proteína mayor, que se une al ADN de una forma específica de secuencia a través de uno o más dedos de zinc, que son regiones de secuencia de aminoácidos dentro del dominio de unión, cuya estructura se estabiliza por coordinación de un ion zinc. La expresión proteína de unión a ADN con dedos de zinc a menudo se abrevia como proteína con dedos de zinc o ZFP.

Los dominios de unión con dedos de zinc se pueden "modificar genéticamente" para que se unan a una secuencia de nucleótidos predeterminada. Los ejemplos no limitantes de métodos para modificar genéticamente proteínas con dedos de zinc son diseño y selección. Una proteína con dedos de zinc diseñada es una proteína que no se encuentra en la naturaleza cuyo diseño/composición resultan principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y algoritmos computacionales para el procesamiento de información en una base de datos que almacena información de diseños de ZFP existentes y datos de unión. Véase, por ejemplo, las patentes de EE.UU. nº 6.140.081; 6.453.242; 6.534.261; y 6.785.613; véase también las publicaciones internacionales WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496; y patentes de EE.UU. 6.746.838; 6.866.997; y 7.030.215.

Una proteína con dedos de zinc "seleccionada" es una proteína que no se encuentra en la naturaleza cuya producción resulta principalmente de un procedimiento empírico tal como presentación en fago, trampa de interacción o selección de híbrido. Véase, p. ej., los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; US 6.733.970; US RE39.229; y WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197 y WO 02/099084.

El término "secuencia" se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser ADN o ARN; puede ser lineal, circular o ramificada y puede ser monocatenaria o bicatenaria. La expresión "secuencia donadora" se refiere a una secuencia de nucleótidos que se inserta en un genoma. La secuencia donadora puede ser de cualquier longitud, por ejemplo, entre 2 y 25.000 nucleótidos de longitud (o cualquier valor entero entre estos o por

encima de estos), preferiblemente entre aproximadamente 100 y 5.000 nucleótidos de longitud (o cualquier número entero entre estos), más preferiblemente entre aproximadamente 200 y 2.500 nucleótidos de longitud.

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

Se conocen en la materia técnicas para determinar la identidad de secuencia de ácidos nucleicos y aminoácidos. Típicamente, dichas técnicas incluyen determinar la secuencia de nucleótidos del ARNm para un gen y/o determinar la secuencia de aminoácidos codificada por el mismo, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta forma.

En general, la identidad se refiere a una correspondencia exacta de nucleótido a nucleótido o aminoácido a aminoácido de dos secuencias de polinucleótidos o polipéptidos, respectivamente. Dos o más secuencias (polinucleótido o aminoácidos) se pueden comparar determinando su porcentaje de identidad. El porcentaje de identidad de dos secuencias, sean secuencias de ácidos nucleicos o aminoácidos, es el número de correspondencias exactas entre dos secuencias alineadas dividido entre la longitud de las secuencias más cortas y multiplicado por 100. Se proporciona un alineamiento aproximado para secuencias de ácidos nucleicos mediante el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). Este algoritmo se puede aplicar a secuencias de aminoácidos usando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*. M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, y normalizada por Gribskov, *Nucl. Acids Res.* 14(6):6745- 6763 (1986). Una implementación de ejemplo de este algoritmo para determinar el porcentaje de identidad de una secuencia la proporciona el Genetics Computer Group (Madison, WI) en la aplicación de utilidad "BestFit". Los parámetros por defecto de este método se describen en el manual de Wisconsin Sequence Analysis Package Program Manual, Versión 8 (1995) (disponible en Genetics Computer Group, Madison, WI). Un método preferido de establecer el porcentaje de identidad en el contexto de la presente descripción es usar el paquete de programas MPSRCH protegido por derecho de autor por la University of Edinburgh, desarrollado por John F. Collins y Shane S. Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). A partir de este juego de paquetes se puede usar el algoritmo de Smith-Waterman cuando se usan los parámetros por defecto para la tabla de puntuación (por ejemplo, penalización por abertura de hueco de 12, penalización por extensión de hueco de 1 y un hueco de 6). A partir de los datos generados, el valor de "correspondencia" refleja la identidad de secuencia. Otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias son conocidos en general en la técnica, por ejemplo, otro programa de alineamiento es BLAST, usado con los parámetros por defecto. Por ejemplo, se pueden usar BLASTN y BLASTP usando los siguientes parámetros por defecto: código genético = estándar; filtro = ninguno; cadena = ambas; punto de corte = 60; esperado = 10; Matriz = BLOSUM62; Descripciónes = 50 secuencias; clasificación por = HIGH SCORE; Bases de datos = no redundantes, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Se pueden encontrar los detalles de estos programas en internet. Con respecto a las secuencias descritas en la presente memoria, el intervalo de grados de identidad de secuencia deseado es aproximadamente de 35% a 100% y cualquier valor entre los mismos. Típicamente, los porcentajes de identidades entre secuencias son al menos 35%-40%; 40%-45%; 45%-50%; 50%-60%; 60%-70%; 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, todavía más preferiblemente 95%, y lo más preferiblemente 98% de identidad de secuencia.

Alternativamente, el grado de similitud de secuencia entre polinucleótidos se puede determinar por hibridación de polinucleótidos en condiciones que permiten la formación de dúplex estables entre regiones homólogas, seguido de digestión con nucleasa(s) específicas de monocatenarios, y determinación del tamaño de los fragmentos digeridos. Dos secuencias de ácidos nucleicos o polipéptidos son sustancialmente homólogas entre sí cuando las secuencias presentan al menos aproximadamente 70%-75%, preferiblemente 80%-82%, más preferiblemente 85%-90%, incluso más preferiblemente 92%, todavía más preferiblemente 95%, y lo más preferiblemente 98% de identidad de secuencia a lo largo de una longitud definida de las moléculas, determinada usando los métodos anteriores. Como se usa en la presente memoria, sustancialmente homólogo también se refiere a secuencias que muestran identidad completa con una secuencia especificada de ADN o polipéptido. Las secuencias de ADN que son sustancialmente homólogas se pueden identificar en un experimento de hibridación Southern, por ejemplo, en condiciones restrictivas, como se define para ese sistema particular. Definir las condiciones de hibridación adecuadas depende del experto en la técnica. Véase, p. ej., Sambrook et al., véase antes, *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

La hibridación selectiva de dos fragmentos de ácido nucleico se puede determinar como sigue. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta a la eficacia y fuerza de los sucesos de hibridación entre dichas moléculas. Una secuencia de ácido nucleico parcialmente idéntica inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica con una molécula diana. La inhibición de la hibridación de la secuencia completamente idéntica se puede evaluar usando ensayos de hibridación que son bien conocidos en la técnica (p. ej., transferencia Southern (ADN), transferencia Northern (ARN), hibridación en solución o similares, véase Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, 2ª edición, (1989) Cold Spring Harbor, N.Y.). Dichos ensayos se pueden llevar a cabo usando diferentes grados de selectividad, por ejemplo, usando condiciones que varían desde restricción baja a alta. Si se usan condiciones de restricción baja, se puede evaluar la ausencia de unión no específica usando una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda que tiene menos de aproximadamente 30% de identidad de secuencia con la molécula diana), de modo que en ausencia de sucesos de unión no específicos, la sonda secundaria no hibridará con la diana.

Cuando se usa un sistema de detección basado en hibridación, se elige una sonda de ácido nucleico que es complementaria con una secuencia de ácido nucleico de referencia, y después mediante la selección de las condiciones adecuadas, la sonda y la secuencia de referencia hibridan o se unen selectivamente entre sí para formar una molécula dúplex. Una molécula de ácido nucleico que es capaz de hibridar selectivamente con una secuencia de referencia en condiciones de hibridación moderadamente restrictivas típicamente hibrida en condiciones que permiten la detección de una secuencia de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tiene al menos aproximadamente 70% de identidad de secuencia con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación restrictivas típicamente permiten la detección de secuencias de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencia mayor de aproximadamente 90-95% con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de la sonda/secuencia de referencia, cuando la sonda y la secuencia de referencia tienen un grado específico de identidad de secuencia, se pueden determinar como se conoce en la técnica (véase, por ejemplo, *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

Las condiciones para la hibridación son bien conocidas para los expertos en la técnica. La restricción de la hibridación se refiere al grado al que las condiciones de hibridación desfavorecen la formación de híbridos que contienen nucleótidos mal apareados, con una mayor restricción correlacionada con una menor tolerancia para los híbridos mal apareados. Los factores que afectan a la restricción de hibridación son bien conocidos para los expertos en la técnica e incluyen, pero no se limitan a temperatura, pH, fuerza iónica, y concentración de los disolventes orgánicos tales como, por ejemplo, formamida y dimetilsulfóxido. Como conocen los expertos en la técnica, la restricción de la hibridación aumenta con temperaturas más altas, menor fuerza iónica y menores concentraciones de disolvente.

Con respecto a las condiciones restrictivas para la hibridación, es bien conocido en la técnica que se pueden usar numerosas condiciones equivalentes para establecer una restricción particular variando, por ejemplo, los siguientes factores: la longitud y naturaleza de las secuencias, composición de bases de las diferentes secuencias, concentraciones de sales y otros componentes de la solución de hibridación, la presencia o ausencia de agentes de bloqueo en las soluciones de hibridación (p. ej., dextrano-sulfato y polietilenglicol), parámetros de temperatura y tiempo de reacción de hibridación, así como, diferentes condiciones de lavado. La selección de un conjunto particular de condiciones de hibridación se selecciona siguiendo métodos convencionales en la técnica (véase, por ejemplo, Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, 2ª edición, (1989) Cold Spring Harbor, N.Y.).

“Recombinación” se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, “recombinación homóloga (HR)” se refiere a la forma especializada en que tiene lugar dicho intercambio, por ejemplo, durante la reparación de roturas bicatenarias en células. Este proceso que requiere homología de secuencias de nucleótidos, usa una molécula “donadora” para la reparación del molde de una molécula “diana” (es decir, la que ha experimentado la rotura bicatenaria), y se conoce de modos diversos como “conversión génica no cruzada” o “conversión génica de tramo corto”, porque lleva a la transferencia de información genética del donador a la diana. Sin querer estar ligados por ninguna teoría particular, dicha transferencia puede implicar la corrección de apareamientos erróneos del ADN heterodúplex que se forma entre la diana rota y el donador, y/o “reasociación de cadenas dependiente de síntesis” en la que el donador se usa para resintetizar la información genética que formará parte de la diana y/o procedimientos relacionados. Dicha HR especializada a menudo da como resultado una alteración de la secuencia de la molécula diana, de modo que parte o toda la secuencia del polinucleótido donador se incorpora en el polinucleótido diana.

“Escisión” se refiere a la rotura de la cadena principal covalente de una molécula de ADN. La escisión se puede iniciar por una variedad de métodos incluyendo, pero sin limitar, la hidrólisis enzimática o química de un enlace fosfodiéster. Son posibles tanto la escisión monocatenaria como bicatenaria, y la escisión bicatenaria se puede producir como resultado de dos sucesos distintos de escisiones monocatenarias. La escisión del ADN puede producir extremos romos o extremos en bisel. En algunas realizaciones, se usan polipéptidos de fusión para la escisión dirigida de ADN bicatenario.

Un “dominio de escisión” comprende una o más secuencias de polipéptido que tienen actividad catalítica para la escisión del ADN. Un dominio de escisión puede estar contenido en una sola cadena de polipéptido o la actividad de escisión puede ser resultado de la asociación de dos (o más) polipéptidos.

5 Un “semidominio de escisión” es una secuencia de polipéptido que junto con un segundo polipéptido (idéntico o diferente) forma un complejo que tiene actividad de escisión (preferiblemente actividad de escisión bicatenaria).

10 La “cromatina” es la estructura de nucleoproteínas que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN y proteínas, incluyendo proteínas cromosómicas histonas y no histonas. La mayor parte de la cromatina celular eucariota existe en forma de nucleosomas, en donde un núcleo del nucleosoma comprende aproximadamente 150 pares de bases de ADN asociadas con un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y un ADN conector (de longitud variable dependiendo del organismo) se extiende entre los núcleos de los nucleosomas. En general una molécula de histona H1 está asociada con el ADN conector. Para el propósito de la presente descripción, el término “cromatina” se entiende que abarca todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye tanto cromatina cromosómica como episómica.

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

20 Un “episoma” es un ácido nucleico replicante, complejo nucleoproteico u otra estructura que comprende un ácido nucleico que no es parte del cariotipo cromosómico de una célula. Los ejemplos de episomas incluyen plásmidos y ciertos genomas víricos.

25 Una “región accesible” es un sitio en la cromatina celular en el que un sitio diana presente en el ácido nucleico se puede unir a una molécula exógena que reconoce el sitio diana. Sin querer estar ligados por ninguna teoría particular, se cree que una región accesible es una que no está empaquetada en una estructura de nucleosoma. La estructura diferente de una región accesible a menudo se puede detectar por su sensibilidad a sondas químicas y enzimáticas, por ejemplo nucleasas.

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

30 Una “molécula exógena” es una molécula que normalmente no está presente en una célula, pero se puede introducir en una célula mediante uno o más métodos genéticos, bioquímicos u otros métodos. La “presencia normal en la célula” está determinada con respecto a la etapa de desarrollo particular y a las condiciones del entorno de la célula. Así, por ejemplo, una molécula que está presente sólo durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular de adulto. Igualmente, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula que no ha sufrido choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión que funciona de una molécula endógena que funciona mal, o una versión que funciona mal de una molécula endógena que funciona normalmente.

35 Una molécula exógena puede ser, entre otras cosas, una pequeña molécula, tal como se genera mediante un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, hidrato de carbono, lípido, glucoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprenda una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser monocatenarios o bicatenarios; pueden ser lineales, ramificados o circulares; y pueden tener cualquier longitud. Los ácidos nucleicos incluyen aquellos capaces de formar dúplex, así como ácidos nucleicos que forman tríplex. Véanse, por ejemplo, las patentes de EE.UU. nº 5.176.996 y 5.422.251. Las proteínas incluyen, pero no se limitan a, proteínas que se unen a ADN, factores de transcripción, factores que remodelan la cromatina, proteínas que se unen a ADN metilado, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, quinasas, fosfatasa, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.

40 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína o ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma vírico infeccioso, cadena T de *Agrobacterium tumefaciens*, un plásmido o episoma introducido en una célula, o un cromosoma que normalmente no está presente en la célula. Los métodos para la introducción de moléculas exógenas en células son conocidos por los expertos en la técnica, e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo con partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano, y transferencia mediada por un vector vírico.

55 Por el contrario, una molécula “endógena” es aquella que está presente normalmente en una célula particular en una etapa particular del desarrollo, en condiciones medioambientales particulares. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, cloroplasto u otro orgánulo, o un ácido

nucleico episómico natural. Las moléculas endógenas adicionales pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

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

La expresión de una proteína de fusión en una célula puede ser resultado del suministro de la proteína de fusión a la célula o del suministro de un polinucleótido que codifica la proteína de fusión a una célula, en donde el polinucleótido es transcrito y el transcrito es traducido, para generar la proteína de fusión. También pueden estar implicados el empalme trans, escisión de polipéptido y ligado de polipéptido en la expresión de una proteína en una célula. Los métodos para el suministro de polinucleótidos y polipéptidos a células se presentan en otra parte en esta descripción.

Un "gen" para los propósitos de la presente descripción, incluye una región de ADN que codifica un producto génico (véase más adelante), así como todas las regiones de ADN que regulan la producción del producto génico, sean o no dichas secuencias reguladoras adyacentes a las secuencias codificantes y/o transcritas. Por consiguiente, un gen incluye, pero no se limita necesariamente a secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión al ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aisladores, elementos límite, orígenes de replicación, sitios de unión de matriz, y regiones de control de locus.

"Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto transcripcional directo de un gen (p. ej., ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural, o cualquier otro tipo de ARN), o una proteína producida por traducción de un ARNm. Los productos génicos también incluyen los ARN que están modificados mediante procesos tales como protección terminal, poliadenilación, metilación, y edición, y proteínas modificadas, por ejemplo, mediante metilación, acetilación, fosforilación, ubiquitinación, ribosilación con ADP, miristilación, y glicosilación.

"Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero no se limita a la activación génica y represión génica.

Las células de "planta" incluyen, pero no se limitan a células de plantas monocotiledóneas (monocotas) o dicotiledóneas (dicotas). Los ejemplos no limitantes de monocotas incluyen plantas de cereales tales como el maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, cebolla, plátano y coco. Los ejemplos no limitantes de dicotas incluyen tabaco, tomate, girasol, algodón, remolacha azucarera, patata, lechuga, melón, soja, canola (colza), y alfalfa. Las células de plantas pueden ser de cualquier parte de la planta y/o de cualquier etapa de desarrollo de la planta.

Una "región de interés" es cualquier región de la cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante dentro o adyacente a un gen, en el que es conveniente unir una molécula exógena. La unión puede ser para el propósito de la escisión dirigida de ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma de orgánulo (p. ej., mitocondria, cloroplasto) o un genoma vírico infeccioso, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de las regiones no codificantes transcritas tales como, por ejemplo, secuencias líder, secuencias remolque o intrones, o dentro de las regiones no transcritas, tanto en la dirección 5' como en la dirección 3' de la región codificante. Una región de interés puede ser tan pequeña como un solo par de nucleótidos o de hasta 25.000 pares de nucleótidos de longitud, o cualquier valor entero de pares de nucleótidos.

Las expresiones "unión operativa" y "operativamente unido" (o "unido operativamente") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en la que los componentes se ordenan de modo que ambos componentes funcionan con normalidad y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unido operativamente a una secuencia codificante si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores de regulación transcripcional. Una secuencia reguladora de la transcripción está en general unida operativamente en *cis* a una secuencia codificante, pero no necesita estar directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está unida operativamente a una secuencia codificante, incluso aunque no sean contiguas.

Respecto a los polipéptidos de fusión, la expresión "operativamente unido" se puede referir al hecho de que cada uno de los componentes realiza la misma función en la unión a otro componente, que si no estuviesen así unidos.

Por ejemplo, respecto a un polipéptido de fusión en el que el dominio de unión al ADN con ZFP está fusionado con un dominio de escisión, el dominio de unión al ADN con ZFP y el dominio de escisión están operativamente unidos si, en el polipéptido de fusión, la parte del dominio de unión al ADN con ZFP es capaz de unirse a su sitio diana y/o su sitio de unión, mientras que el dominio de escisión es capaz de escindir el ADN en la proximidad del sitio diana.

- 5 Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, aunque conserva la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede tener más, menos, o el mismo número de restos que la correspondiente molécula natural, y/o puede contener una o más sustituciones de aminoácidos o nucleótidos. Los métodos para determinar la función de un ácido nucleico (p. ej., función codificante, capacidad para hibridar con otro ácido nucleico) se conocen bien en la técnica. De igual forma, se conocen bien los métodos para determinar la función de la proteína. Por ejemplo, la función de unión al ADN de un polipéptido se puede determinar, por ejemplo, mediante unión a filtro, cambio de la movilidad electroforética o ensayos de inmunoprecipitación. La escisión del ADN se puede ensayar mediante electroforesis en gel. Véase Ausubel et al., más arriba. La capacidad de una proteína para interactuar con otra proteína se puede determinar, por ejemplo, mediante coinmunoprecipitación, ensayos de doble híbrido o de complementación, tanto genéticos como bioquímicos. Véase, por ejemplo, Fields et al. (1989) Nature 340: 245-246; patente de los EE.UU. nº 5.585.245 y publicación internacional PCT WO 98/44350.

Sitios diana

- 20 Los métodos y composiciones descritos incluyen proteínas de fusión que comprenden un dominio de escisión (o un semidominio de escisión) y un dominio con dedos de zinc, en las que el dominio con dedos de zinc, mediante la unión a una secuencia en la cromatina celular (p. ej., un sitio diana o un sitio de unión), dirige la actividad del dominio de escisión (o semidominio de escisión) a las proximidades de la secuencia y, por lo tanto, induce la escisión en las proximidades de la secuencia diana. Como se ha expuesto en otra parte en esta descripción, un dominio con dedos de zinc se puede modificar genéticamente para que se una a prácticamente cualquier secuencia deseada. Por consiguiente, después de identificar una región de interés que contiene una secuencia en la que se desea la escisión o recombinación, se pueden modificar genéticamente uno o más dominios de unión con dedos de zinc para que se unan a una o más secuencias en la región de interés. La expresión de una proteína de fusión que comprende un dominio de unión con dedos de zinc y un dominio de escisión (o de dos proteínas de fusión, que comprende cada una un dominio de unión con dedos de zinc y un semidominio de escisión), en una célula, realiza la escisión en la región de interés.

- 30 La selección de una secuencia en la cromatina celular para la unión por un dominio de unión con dedos de zinc (p. ej., un sitio diana) se puede llevar a cabo, por ejemplo, de acuerdo con los métodos descritos en la patente en copropiedad de EE.UU. nº 6.453.242 (17 de septiembre, 2002), que también describe métodos para diseñar ZEP para la unión a una secuencia seleccionada. Estará claro para los expertos en la técnica que la simple inspección visual de una secuencia de nucleótidos también puede ser útil para la selección de un sitio diana. Por consiguiente, se puede usar cualquier medio de selección de sitios diana en los métodos reivindicados.

- 40 Los sitios diana en general están compuestos de una pluralidad de subsitios diana adyacentes. Un subsitio diana se refiere a la secuencia (normalmente un triplete de nucleótidos o un cuadruplete de nucleótidos que se puede solapar en un nucleótido con un cuadruplete adyacente) unida a un dedo de zinc individual. Véase, por ejemplo, la publicación internacional WO 02/077227. Si la cadena con la que una proteína con dedos de zinc hace la mayor parte de los contactos se designa la cadena diana "cadena de reconocimiento principal" o "cadena de contacto principal", algunas proteínas con dedos de zinc se unen a un triplete de tres bases en la cadena diana y a cuatro bases en la cadena no diana. Un sitio diana en general tiene una longitud de al menos 9 nucleótidos y, por consiguiente, se une con un dominio de unión con dedos de zinc que comprende al menos 3 dedos de zinc. Sin embargo, también es posible la unión, por ejemplo, de un dominio de unión con 4 dedos de zinc a un sitio diana de 12 nucleótidos, un dominio de unión con 5 dedos de zinc a un sitio diana de 15 nucleótidos o un dominio de unión con 6 dedos de zinc a un sitio diana de 18 nucleótidos. Como será evidente, también es posible la unión de dominios de unión mayores (p. ej., de 7, 8, 9 dedos de zinc o más) a sitios diana mayores.

- 50 No es necesario que un sitio diana sea un múltiplo de 3 nucleótidos. Por ejemplo, en casos en los que se producen interacciones de cadenas cruzadas (véase, p. ej., la patente de EE.UU. 6.453.242 y publicación internacional WO 02/077227), uno o más de los dedos de zinc individuales de un dominio de unión con múltiples dedos se puede unir a subsitios cuadrupletes que se solapan. Como resultado, una proteína con tres dedos se puede unir a una secuencia de 10 nucleótidos, en donde el décimo nucleótido es parte de un cuadruplete unido por un dedo terminal, una proteína de 4 dedos se puede unir a una secuencia de 13 nucleótidos, en donde el decimotercer nucleótido es parte de un cuadruplete unido por un dedo terminal, etc.

- 60 La longitud y naturaleza de las secuencias conectoras de ácidos nucleicos entre dedos de zinc individuales en un dominio de unión con múltiples dedos también afecta a la unión a una secuencia diana. Por ejemplo, la presencia de un llamado "conector no canónico", "conector largo" o "conector estructurado" entre dedos de zinc adyacentes en un dominio de unión con múltiples dedos puede permitir que estos dedos se unan a subsitios que no están inmediatamente adyacentes. Los ejemplos no limitantes de dichos conectores se describen, por ejemplo, en la

patente de EE.UU. 6.479.626 y publicación internacional WO 01/53480. Por consiguiente, uno o más subsitios en un sitio diana para un dominio de unión con dedos de zinc, pueden estar separados entre sí por 1, 2, 3, 4, 5 o más nucleótidos. Para proporcionar un ejemplo, un dominio de unión con 4 dedos se puede unir a un sitio diana de 13 nucleótidos que comprende, en orden, dos subsitios contiguos de 3 nucleótidos, un nucleótido intermedio y dos subsitios contiguos tripletes.

La distancia entre secuencias (p. ej., sitios diana) se refiere al número de nucleótidos o pares de nucleótidos que intervienen entre dos secuencias, medido desde los bordes de las secuencias más cercanos entre sí.

En algunas realizaciones en las que la escisión depende de la unión de dos moléculas de fusión de dominio con dedos de zinc/semidominio de escisión para separar sitios diana, los dos sitios diana pueden estar en cadenas de ADN opuestas. En otras realizaciones, ambos sitios diana están en la misma cadena de ADN.

Dominios de unión con dedos de zinc

Un dominio de unión con dedos de zinc comprende uno o más dedos de zinc. Miller *et al.* (1985) *EMBO J.* 4:1609-1614; Rhodes (1993) *Scientific American* Feb.:56-65; patente de EE.UU. nº 6.453.242. Típicamente, un solo dominio con dedos de zinc tiene aproximadamente 30 aminoácidos de longitud. Estudios estructurales han demostrado que cada dominio con dedos de zinc (patrón) contiene dos láminas beta (mantenidas en una vuelta beta que contiene los dos restos de cisteína invariantes) y una hélice alfa (que contiene los dos restos de histidina invariantes), que se mantienen en una conformación particular mediante la coordinación de un átomo de zinc con las dos cisteínas y las dos histidinas.

Los dedos de zinc incluyen tanto dedos de zinc C_2H_2 canónicos (es decir, aquellos en los que el ion zinc se coordina con las dos cisteínas y las dos histidinas) y dedos de zinc no canónicos, tales como por ejemplo, dedos de zinc C_3H (aquellos en los que el ion zinc se coordina con tres restos de cisteína y un resto de histidina) y dedos de zinc C_4 (aquellos en los que el ion zinc se coordina con cuatro restos de cisteína). Véase también la publicación internacional WO 02/057293.

Los dominios de unión con dedos de zinc se pueden diseñar para que se unan a una secuencia de elección. Véase, por ejemplo, Beerli *et al.*, (2002) *Nature Biotechnol.* 20: 135-141; Pabo *et al.* (2001) *Ann. Rev. Biochem.* 70: 313-340; Isalan *et al.* (2001) *Nature Biotechnol.* 19: 656-660; Segal *et al.*, (2001) *Curr. Opin. Biotechnol.* 12: 632-637; Choo *et al.*, (2000) *Curr. Opin. Struct. Biol.* 10: 411-416. Un dominio de unión con dedos de zinc genéticamente modificado puede tener una nueva especificidad de unión, en comparación con una proteína con dedos de zinc que se produce de forma natural. Los métodos de modificación por ingeniería genética incluyen, pero no se limitan al diseño racional y diversos tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos tripletes (o cuadrupletes) y secuencias de aminoácidos individuales de dedos de zinc, en las que cada secuencia de nucleótidos triplete o cuadruplete está asociada a una o más secuencias de aminoácidos de dedos de zinc que se unen a la secuencia triplete o cuadruplete particular. Véase, por ejemplo, las patentes de los EE.UU. en copropiedad 6.453.242 y 6.534.26. Se describen métodos de diseño adicionales, por ejemplo, en las patentes de EE.UU. 6.746.838; 6.785.613; 6.866.997; y 7.030.215.

Los métodos de selección de ejemplo, incluyendo los sistemas de presentación en fagos y los sistemas de doble híbrido, se describen en las patentes de EE.UU. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como en las publicaciones internacionales WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.

La potenciación de la especificidad de unión para los dominios de unión con dedos de zinc se ha descrito, por ejemplo, en la patente de EE.UU. en copropiedad nº 6.749.136.

Puesto que un dedo de zinc individual se une a una secuencia de tres nucleótidos (es decir, triplete) (o una secuencia de cuatro nucleótidos que se puede solapar, en un nucleótido, con el sitio de unión de cuatro nucleótidos de un dedo de zinc adyacente), la longitud de una secuencia para la cual se modifica genéticamente un dominio de unión con dedos de zinc para la unión (p. ej., una secuencia diana) determinará el número de dedos de zinc en un dominio de unión con dedos de zinc genéticamente modificado. Por ejemplo, para los ZFP en los que los motivos de los dedos no se unen a subsitios que se solapan, una secuencia diana de 6 nucleótidos se une por un dominio de unión de dos dedos; una secuencia diana de nueve nucleótidos se une por un dominio de unión de tres dedos, etc. Como se indica en la presente memoria, los sitios de unión para dedos de zinc individuales (es decir, subsitios) en un sitio diana no es necesario que sean contiguos, pero pueden estar separados por uno o varios nucleótidos dependiendo de la longitud y naturaleza de las secuencias de aminoácidos entre los dedos de zinc (es decir, los conectores entre dedos) en un dominio de unión con múltiples dedos.

En un dominio de unión con dedos de zinc con múltiples dedos, los dedos de zinc adyacentes pueden estar separados por secuencias conectoras de aminoácidos de aproximadamente 5 aminoácidos (llamadas conectores entre dedos "canónicos") o, alternativamente, por uno o más conectores no canónicos. Véase, p. ej., las patentes de EE.UU. en copropiedad nº 6.453.242 y 6.534.261. Para los dominios de unión con dedos de zinc genéticamente modificados que comprenden más de 3 dedos, se puede preferir la inserción de conectores entre dedos más largos ("no canónicos") entre algunos dedos de zinc, ya que puede aumentar la afinidad y/o especificidad de la unión por el

dominio de unión. Véase, por ejemplo, la patente de EE.UU. nº 6.479.626 y la publicación internacional WO 01/53480. Por consiguiente, los dominios de unión con dedos de zinc con múltiples dedos también se pueden caracterizar con respecto a la presencia y localización de conectores entre dedos no canónicos. Por ejemplo, un dominio de unión con dedos de zinc con 6 dedos que comprende 3 dedos (unidos por dos conectores entre dedos canónicos), un conector largo y tres dedos adicionales (unidos por 2 conectores entre dedos canónicos) se indica como una configuración 2x3. Igualmente, un dominio de unión que comprende 2 dedos (con un conector canónico entre ellos), un conector largo y 2 dedos adicionales (unidos por un conector canónico) se indica como una proteína 2x2. Una proteína que comprende 3 unidades de 2 dedos (en cada una de las cuales los dos dedos están unidos por un conector canónico) y en la que cada unidad de 2 dedos está unida a la unidad de 2 dedos adyacente mediante un conector largo, se denomina una proteína 3x2.

La presencia de un conector entre dedos largo o no canónico entre dos dedos de zinc adyacentes en un dominio de unión con múltiples dedos a menudo permite que los dos dedos se unan a subsitios que no están inmediatamente contiguos en la secuencia diana. Por consiguiente, puede haber huecos de uno o más nucleótidos entre subsitios en un sitio diana; es decir un sitio diana puede contener uno o más nucleótidos que no están conectados por un dedo de zinc. Por ejemplo, un dominio de unión con dedos de zinc 2x2 se puede unir a 2 secuencias de 6 nucleótidos separadas por un nucleótido, es decir, se une a un sitio diana de 13 nucleótidos. Véase también, Moore et al. (2001a) *Proc. Natl. Acad. Sci. USA* 98:1432-1436; Moore et al. (2001b) *Proc. Natl. Acad. Sci. USA* 98:1437-1441 y publicación internacional WO 01/53480.

Como se ha mencionado previamente, un subsitio diana es una secuencia de 3 o 4 nucleótidos al que se une un solo dedo de zinc. Para algunos propósitos, una unidad de 2 dedos se indica como un módulo de unión. Un módulo de unión se puede obtener, por ejemplo, seleccionando dos dedos adyacentes en el contexto de una proteína de múltiples dedos (en general 3 dedos) que se une a una secuencia diana de 6 nucleótidos particular. Alternativamente, se pueden construir módulos ensamblando dedos de zinc individuales. Véase también las publicaciones internacionales WO 98/53057 y WO 01/53480.

25 Dominios de escisión

La parte del dominio de escisión de las proteínas de fusión descritas en la presente memoria se puede obtener de cualquier endonucleasa o exonucleasa. Las endonucleasas de ejemplo de las cuales se puede obtener un dominio de escisión incluyen, pero no se limitan a endonucleasas de restricción y endonucleasas de asentamiento. Véase, por ejemplo, el catálogo 2002-2003 de *New England Biolabs*, Beverly, MA; y Belfort et al. (1997) *Nucleic Acids Res.* 25: 3379-3388. Se conocen otras enzimas que escinden el ADN (p. ej., nucleasa S1; nucleasa de la soja verde; ADNasa I pancreática; nucleasa microcócica; endonucleasa HO de levadura; véase también Linn et al. (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o fragmentos funcionales de las mismas) se pueden usar como fuente de dominios de escisión y semidominios de escisión.

De igual forma, un semidominio de escisión (p. ej., proteínas de fusión que comprenden un dominio de unión con dedos de zinc y un semidominio de escisión) puede proceder de cualquier nucleasa o parte de la misma, como se ha expuesto más arriba, que requiera la dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Alternativamente, se puede usar una única proteína que comprende dos semidominios de escisión. Los dos semidominios de escisión pueden proceder de la misma endonucleasa (o de fragmentos funcionales de la misma), o cada semidominio de escisión puede proceder de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios diana para las dos proteínas de fusión están preferentemente dispuestos, uno con respecto al otro, de modo que la unión de las dos proteínas de fusión a sus respectivos sitios diana coloca los semidominios de escisión en una orientación espacial de uno respecto al otro que permite que los semidominios de escisión formen un dominio de escisión funcional, p. ej., mediante dimerización. Por lo tanto, en algunas realizaciones, los bordes más cercanos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número entero de nucleótidos o pares de nucleótidos puede intercalarse entre dos sitios diana (p. ej., de 2 a 50 nucleótidos o más). En general, el punto de escisión se localiza entre los sitios diana.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unirse a secuencias específicas del ADN (en un sitio de reconocimiento), y escindir el ADN en o cerca del sitio de unión. Determinadas enzimas de restricción (p. ej., de tipo IIS) escinden el ADN en sitios eliminados del sitio de reconocimiento y tienen dominios de escisión y unión separables. Por ejemplo, la enzima de tipo IIS FokI cataliza la escisión bicatenaria del ADN a 9 nucleótidos de su sitio de reconocimiento en una cadena y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véase, por ejemplo, las patentes de los EE.UU. nº 5.356.802; 5.436.150 y 5.487.994; así como Li et al. (1992) *Proc. Natl. Acad. Sci. USA* 89: 4275-4279; Li et al. (1993) *Proc. Natl. Acad. Sci. USA* 90: 2764-2768; Kim et al., (1994a) *Proc. Natl. Acad. Sci. USA* 91: 883-887; Kim et al. (1994b) *J. Biol. Chem.* 269: 31978-31982. Las proteínas de fusión usadas según la invención, comprenden el dominio de escisión de FokI y son dominios de unión con dedos de zinc genéticamente modificados.

Una enzima de restricción de tipo IIS de ejemplo, cuyo dominio de escisión es separable del dominio de unión, es Fok I. Esta enzima particular es activa en forma de dímero. Bitinaite et al., (1998) *Proc. Natl. Acad. Sci. USA* 95:

10570-10575. Por consiguiente, para los propósitos de la presente descripción, la parte de la enzima Fok I usada en las proteínas de fusión descritas se considera un semidominio de escisión. Por lo tanto, para la escisión bicatenaria dirigida y/o el reemplazo dirigido de secuencias celulares usando fusiones de dedos de zinc-FokI, se pueden usar dos proteínas de fusión, comprendiendo cada una un semidominio de escisión de FokI, para reconstruir un dominio de escisión catalíticamente activo. Alternativamente, también se puede usar una sola molécula de polipéptido que contiene un dominio de unión con dedos de zinc y dos semidominios de escisión de FokI. En otra parte de esta descripción se proporcionan parámetros para la escisión dirigida y la alteración dirigida de secuencia usando fusiones de dedos de zinc-Fok I.

5

10 Un dominio de escisión o semidominio de escisión puede ser cualquier parte de una proteína que conserva la actividad de escisión, o que conserva la capacidad para multimerizarse (p. ej., dimerizarse) para formar un dominio de escisión funcional.

Las enzimas de restricción de tipo IIS de ejemplo se listan en la tabla 1. Las enzimas de restricción adicionales también contienen dominio de unión y escisión separables.

Véase, por ejemplo, Roberts et al. (2003) *Nucleic Acids Res.* 31:418-420.

15 Tabla 1: Algunos tipos de enzimas de restricción

Aar I	BsrB I	SspD5 I
Ace III	BsrD I	Sth132 I
Aci I	BstF5 I	Sts I
Alo I	Btr I	TspDT I
Bae I	Bts I	TspGW I
Bbr7 I	Cdi I	Tth111 II
Bbv I	CjeP I	UbaP I
Bbv II	Drd II	Bsa I
BbvC I	Eci I	BsmB I
Bcc I	Eco31 I	
Bce83 I	Eco57 I	
BceA I	Eco57M I	
Bcef I	Esp3 I	
Bcg I	Fau I	
BciV I	Fin I	
Bfi I	Fok I	
Bin I	Gdi II	
Bmg I	Gsu I	
Bpu10 I	Hga I	
BsaX I	Hin4 II	
Bsb I	Hph I	
BscA I	Ksp632 I	
BscG I	Mbo II	
BseR I	Mly I	
BseY I	Mme I	
Bsi I	Mnl I	
Bsm I	Pfl1108 I	
BsmA I	Ple I	
BsmF I	Ppi I	
Bsp24 I	Psr I	
BspG I	RleA I	
BspM I	Sap I	
BspNC I	SfaN I	
Bsr I	Sim I	

Fusiones de dominio con dedos de zinc-dominio de escisión

20 Los métodos para el diseño y construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos para los expertos en la técnica. Por ejemplo, se describen métodos para el diseño y construcción de proteínas de fusión que comprenden proteínas con dedos de zinc (y polinucleótidos que codifican las mismas) en las patentes de EE.UU. en copropiedad n° 6.453.242 y 6.534.261. En algunas realizaciones, se construyen los polinucleótidos que codifican dichas proteínas de fusión. Estos polinucleótidos se pueden insertar en un vector y el vector se puede introducir en una célula (véase más adelante para la descripción adicional en relación con vectores y métodos para introducir polinucleótidos en células).

25 En algunas realizaciones de los métodos descritos en la presente memoria, una proteína de fusión comprende un

dominio de unión con dedos de zinc y un semidominio de escisión de la enzima de restricción *Fok I*, y dos de dichas proteínas de fusión son expresadas en una célula. La expresión de dos proteínas de fusión en una célula puede resultar del suministro a la célula de las dos proteínas; el suministro a la célula de una proteína y un ácido nucleico que codifica una de las proteínas; el suministro a la célula de dos ácidos nucleicos, codificando cada uno una de las proteínas; o del suministro a la célula de un solo ácido nucleico que codifica ambas proteínas. En realizaciones adicionales, una proteína de fusión comprende una sola cadena polipeptídica que comprende dos semidominios de escisión y un dominio de unión con dedos de zinc. En este caso, es expresada una sola proteína de fusión en una célula y sin querer estar ligados por la teoría, se cree que se escinde el ADN como resultado de la formación de un dímero intramolecular de los semidominios de escisión.

En algunas realizaciones, los componentes de las proteínas de fusión (es decir, fusiones ZFP-*Fok I*) están dispuestos de modo que el dominio con dedos de zinc está más cercano al extremo amino de la proteína de fusión, y el semidominio de escisión está más cercano al extremo carboxi. Esto refleja la orientación relativa del dominio de escisión en dominios de escisión que dimerizan naturales tales como los procedentes de la enzima *Fok I*, en la que el dominio de unión al ADN es más cercano al extremo amino y el semidominio de escisión es más cercano al extremo carboxi. En estas realizaciones, la dimerización de los semidominios de escisión para formar una nucleasa funcional se lleva a cabo mediante la unión de las proteínas de fusión a sitios en cadenas de ADN opuestas, estando los extremos 5' de los sitios de unión próximos entre sí.

En realizaciones adicionales, los componentes de las proteínas de fusión (es decir, fusiones ZFP-*Fok I*) están dispuestos de modo que el semidominio de escisión está más cercano al extremo amino de la proteína de fusión, y el dominio con dedos de zinc está más cercano al extremo carboxi. En estas realizaciones, la dimerización de los semidominios de escisión para formar una nucleasa funcional se lleva a cabo mediante la unión de las proteínas de fusión a sitios en cadenas de ADN opuestas, estando los extremos 3' de los sitios de unión próximos entre sí.

En otras realizaciones adicionales, una primera proteína de fusión contiene el semidominio de escisión más cercano al extremo amino de la proteína de fusión, y el dominio con dedos de zinc más cercano al extremo carboxi, y una segunda proteína de fusión está dispuesta de modo que el dominio con dedos de zinc está más cercano al extremo amino de la proteína de fusión, y el semidominio de escisión está más cercano al extremo carboxi. En estas realizaciones, ambas proteínas de fusión se unen a la misma cadena de ADN, conteniendo el sitio de unión de la primera proteína de fusión el dominio con dedos de zinc más cercano al extremo carboxi situado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio con dedos de zinc más cercano al extremo amino.

En algunas realizaciones, en las proteínas de fusión descritas, la secuencia de aminoácidos entre el dominio con dedos de zinc y el dominio de escisión (o semidominio de escisión) se indica como el "conector ZC". El conector ZC debe distinguirse de los conectores entre dedos descritos antes. Véase, p. ej., las publicaciones de patente de EE.UU. 20050064474A1 y 20030232410, y publicación de patente internacional WO05/084190, para los detalles sobre la obtención de conectores ZC que optimizan la escisión.

Métodos para la escisión dirigida

Los métodos descritos se pueden usar para escindir ADN en una región de interés en la cromatina celular (p. ej., en un sitio deseado o predeterminado en un genoma, por ejemplo, en un gen mutante o natural). Para dicha escisión dirigida del ADN, se modifica genéticamente un dominio de unión con dedos de zinc para que se una a un sitio diana en o cerca del sitio de escisión predeterminado, y se expresa en una célula una proteína de fusión que comprende el dominio de unión con dedos de zinc genéticamente modificado y un dominio de escisión. Tras la unión de la parte de dedos de zinc de la proteína de fusión al sitio diana, el ADN es escindido cerca del sitio diana por el dominio de escisión. El sitio exacto de escisión puede depender de la longitud del conector ZC.

Alternativamente, se expresan en una célula dos proteínas de fusión, que comprende cada una un dominio de unión con dedos de zinc y un semidominio de escisión, y se unen a sitios diana que se yuxtaponen de modo que se reconstituye un dominio de escisión funcional y el ADN es escindido en la proximidad de los sitios diana. En una realización, la escisión se produce entre los sitios diana de dos dominios de unión con dedos de zinc. Uno o ambos dominios de unión con dedos de zinc pueden ser genéticamente modificados.

Para la escisión dirigida usando un polipéptido de fusión de dominio de unión con dedos de zinc-dominio de escisión, el sitio de unión puede abarcar el sitio de escisión, o el borde cercano del sitio de unión puede estar a 1, 2, 3, 4, 5, 6, 10, 25, 50 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) del sitio de escisión. La situación exacta del sitio de unión, con respecto al sitio de escisión, dependerá del dominio de escisión particular y la longitud del conector ZC. Para los métodos en los que se usan dos polipéptidos de fusión, que comprende cada uno un dominio de unión con dedos de zinc y un semidominio de escisión, los sitios de unión en general se solapan con los sitios de escisión. Por lo tanto, el extremo cercano del primer sitio de unión puede estar a 1, 2, 3, 4, 5, 6, 10, 25, 50 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en un lado del sitio de escisión, y el borde cercano del segundo sitio de unión puede estar a 1, 2, 3, 4, 5, 6, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en el otro lado del sitio de escisión. Los métodos para la cartografía de sitios de escisión in vitro e in vivo son conocidos para los expertos en la técnica.

Por lo tanto, los métodos descritos en la presente memoria pueden usar un dominio de unión con dedos de zinc genéticamente modificado fusionado con un dominio de escisión. En estos casos, el dominio de unión se modifica genéticamente para que se una a una secuencia diana, en o cerca de la que se desea la escisión. Se introduce en una célula de planta la proteína de fusión, o un polinucleótido que codifica la misma. Una vez introducida o expresada en la célula, la proteína de fusión se une a la secuencia diana y escinde en o cerca de la secuencia diana. El sitio exacto de la escisión depende de la naturaleza del dominio de escisión y/o la presencia y/o naturaleza de las secuencias conectoras entre los dominios de unión y escisión. En el caso donde se usen dos proteínas de fusión, comprendiendo cada una un semidominio de escisión, la distancia entre los bordes cercanos a los sitios de unión puede ser de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos). Los niveles óptimos de escisión también pueden depender tanto de la distancia entre los sitios de escisión de las dos proteínas de fusión (véase, por ejemplo, Smith et al. (2000) *Nucleic Acids Res.* 28:3361-3369; Bibikova et al. (2001) *Mol. Cell. Biol.* 21:289-297) como de la longitud del conector ZC en cada proteína de fusión. Véase, también, la publicación de patente de EE.UU. 20050064474A1 y publicaciones de patente internacionales WO05/084190, WO05/014791 y WO03/080809.

En algunas realizaciones, el dominio de escisión comprende dos semidominios de escisión, los cuales son ambos parte de un solo polipéptido que comprende un dominio de unión, un primer semidominio de escisión y un segundo semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o diferentes secuencias de aminoácidos, con la condición de que funcionen para escindir el ADN.

Los semidominios de escisión también se pueden proporcionar en moléculas separadas. Por ejemplo, se pueden introducir dos polipéptidos de fusión en una célula, en donde cada polipéptido comprende un dominio de unión y un semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o diferentes secuencias de aminoácidos, con la condición de que funcionen para escindir el ADN. Además, los dominios de unión se unen a secuencias diana que están típicamente dispuestas de modo que tras la unión de los polipéptidos de fusión, los dos semidominios de escisión están presentes en una orientación espacial entre sí que permite la reconstitución de un dominio de escisión (p. ej., por dimerización de los semidominios), posicionando de esta forma los semidominios uno con respecto al otro para formar un dominio de escisión funcional, que produce la escisión de la cromatina celular en una región de interés. En general, la escisión por el dominio de escisión reconstituido se produce a un sitio localizado entre las dos secuencias diana. Una o ambas de las proteínas pueden ser diseñadas para unirse a su sitio diana.

Las dos proteínas de fusión se pueden unir en la región de interés en la misma polaridad o la opuesta, y sus sitios de unión (es decir, sitios diana) pueden estar separados por cualquier número de nucleótidos, p. ej. de 0 a 200 nucleótidos o cualquier valor entero entre estos. En algunas realizaciones, los sitios de unión para las dos proteínas de fusión, que comprende cada una un dominio de unión con dedos de zinc y un semidominio de escisión, pueden estar localizados separados entre 5 y 18 nucleótidos, por ejemplo, separados por 5-8 nucleótidos, o separados por 15-18 nucleótidos, o separados por 6 nucleótidos, o separados por 16 nucleótidos, medido desde el borde de cada sitio de unión más cercano al otro sitio de unión, y la escisión se produce entre los sitios de unión.

El sitio en el que el ADN es escindido, en general está entre los sitios de unión para las dos proteínas de fusión. La rotura bicatenaria del ADN a menudo resulta de dos roturas monocatenarias, o "mellas", desplazadas en 1, 2, 3, 4, 5, 6 o más nucleótidos (por ejemplo, la escisión del ADN bicatenario por *Fok I* natural es resultado de las roturas monocatenarias desplazadas en 4 nucleótidos). Por lo tanto, la escisión no se produce necesariamente exactamente en sitios opuestos en cada cadena de ADN. Además, la estructura de las proteínas de fusión y la distancia entre los sitios diana puede influir en si la escisión se produce adyacente a un solo par de nucleótidos, o si la escisión se produce en varios sitios. Sin embargo, para muchas aplicaciones, incluyendo la recombinación dirigida y la mutagénesis dirigida (véase más adelante) la escisión dentro de un intervalo de nucleótidos en general es suficiente, y no se requiere la escisión entre pares de bases particulares.

Como se ha indicado antes, la o las proteínas de fusión se pueden introducir como polipéptidos y/o polinucleótidos, Por ejemplo, se pueden introducir en una célula dos polinucleótidos, comprendiendo cada uno secuencias que codifican uno de los polipéptidos mencionados antes, y cuando los polipéptidos son expresados y cada uno se une a su secuencia diana, se produce la escisión en o cerca de la secuencia diana. Alternativamente, se introduce en una célula un solo polinucleótido que comprende secuencias que codifican ambos polipéptidos de fusión. Los polinucleótidos pueden ser ADN, ARN o cualquier forma modificada o análogos de ADN y/o ARN.

Para potenciar la especificidad de la escisión, también se pueden usar composiciones adicionales en los métodos descritos en la presente memoria. Por ejemplo, los semidominios de escisión individuales pueden presentar actividad de escisión bicatenaria limitada. En métodos en los que se introducen en una célula dos proteínas de fusión, conteniendo cada una un dominio con dedos de zinc con 3 dedos y un semidominio de escisión, cualquiera de las proteínas especifica un sitio diana de aproximadamente 9 nucleótidos. Aunque la secuencia diana agregada de 18 nucleótidos es probable que sea única en un genoma de mamífero, cualquier sitio diana dado de 9 nucleótidos aparece, como media, aproximadamente 23.000 veces en el genoma humano. Por lo tanto, puede ocurrir la escisión no específica, debida a la unión específica de sitio de un solo semidominio. Por consiguiente, los métodos descritos en la presente memoria contemplan el uso de un mutante dominante negativo de un semidominio de escisión tal como *Fok I* (o un ácido nucleico que codifica el mismo) que es expresado en una célula junto con las dos proteínas

de fusión. El mutante dominante negativo es capaz de dimerizar pero no es capaz de escindir, y también bloquea la actividad de escisión de un semidominio con el que dimeriza. Proporcionando el mutante dominante negativo en exceso molar respecto a las proteínas de fusión, solo las regiones en las que están unidas ambas proteínas de fusión tendrán una concentración local suficientemente alta de semidominios de escisión funcionales para que se produzca la dimerización y escisión. En los sitios en los que solo está unida una de las dos proteínas de fusión, su semidominio de escisión forma un dímero con el semidominio del mutante dominante negativo, y no se produce la escisión no específica no deseable.

Se han identificado tres restos de aminoácidos catalíticos en el semidominio de escisión de *Fok I*: Asp 450, Asp 467 y Lys 469. Bitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 10.570-10.575. Por lo tanto, se pueden usar una o más mutaciones en uno de estos restos para generar una mutación negativa dominante. Además, se conocen muchos de los restos de aminoácidos catalíticos de otras endonucleasas de tipo IIS y/o se pueden determinar, por ejemplo, mediante alineamiento con secuencia de *Fok I* y/o por generación y ensayo de mutantes para la actividad catalítica.

Mutaciones del dominio de dimerización en el semidominio de escisión

Los métodos para la escisión dirigida que implica el uso de fusiones entre ZFP y un semidominio de escisión (tal como, p. ej., una fusión de ZFP/*Fok I*) requiere el uso de dos de dichas moléculas de fusión, en general cada una dirigida a una secuencia diana distinta. Las secuencias diana para las dos proteínas de fusión se pueden seleccionar de modo que la escisión dirigida se dirige a un sitio único en el genoma, como se ha descrito previamente. Una fuente potencial de especificidad de escisión reducida podría resultar de la homodimerización de una de las dos fusiones de ZFP/semidominio de escisión. Esto puede ocurrir, por ejemplo, debido a la presencia, en un genoma, de repeticiones invertidas de las secuencias diana para una de las dos fusiones de ZFP/semidominio de escisión, localizadas de modo que permitan que se unan dos copias de la misma proteína de fusión con una orientación y separación que permitan la formación de un dímero funcional.

Un procedimiento para reducir la probabilidad de este tipo de escisión aberrante en secuencias distintas del sitio diana previsto, implica la generación de variantes del semidominio de escisión que minimizan o previenen la homodimerización. Preferiblemente, uno o más aminoácidos en la región del semidominio implicado en su dimerización, son alterados. En la estructura cristalina del dímero de proteína de *Fok I*, se describe que la estructura de los semidominios de escisión es similar a la disposición de los semidominios de escisión durante la escisión del ADN por *Fok I*. Wah et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:10564-10569. Esta estructura indica que los restos de aminoácidos en las posiciones 483 y 487 tienen una función clave en la dimerización de los semidominios de escisión de *Fok I*. La estructura también indica que los restos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 están todas suficientemente cerca de la interfase de dimerización para influir en la dimerización. Por consiguiente, las alteraciones de la secuencia de aminoácidos en una o más de las posiciones mencionadas antes probablemente alterarán las propiedades de dimerización del semidominio de escisión. Dichos cambios se pueden introducir, por ejemplo, construyendo una biblioteca que contiene (o codifica) diferentes restos de aminoácidos en estas posiciones y seleccionando variantes con las propiedades deseadas, o diseñando racionalmente mutantes individuales. Además de prevenir la homodimerización, también es posible que algunas de estas mutaciones puedan aumentar la eficacia de la escisión por encima de la obtenida con los dos semidominios de escisión naturales.

Por consiguiente, la alteración de un semidominio de escisión de *Fok I* en cualquier resto de aminoácido que afecta a la dimerización se puede usar para prevenir que uno de un par de fusiones de ZFP/*Fok I* experimente homodimerización que puede conducir a la escisión en secuencias indeseadas. Por lo tanto, para la escisión dirigida usando un par de fusiones de ZFP/*Fok I*, una o ambas proteínas de fusión puede comprender una o más alteraciones de aminoácidos que inhiben la autodimerización, pero permiten que se produzca la heterodimerización de las dos proteínas de fusión de modo que la escisión se produce en el sitio diana deseado. En algunas realizaciones, están presentes alteraciones en ambas proteínas de fusión, y las alteraciones tienen efectos aditivos, es decir, se minimiza o anula la homodimerización de cualquiera de las proteínas de fusión que conduce a la escisión aberrante, mientras que se facilita la heterodimerización de las dos proteínas de fusión comprado con la obtenida con semidominios de escisión naturales.

Métodos para la alteración dirigida de secuencias genómicas y recombinación dirigida

También se describe en la presente memoria métodos para sustituir una secuencia genómica (p. ej., una región de interés en la cromatina celular) por una secuencia homóloga no idéntica (es decir, recombinación dirigida). Los intentos previos de sustituir secuencias particulares han implicado poner en contacto una célula con un polinucleótido que comprende secuencias que contienen homología con una región cromosómica (es decir, un ADN donador), seguido de la selección de células en las que la molécula de ADN donador ha experimentado recombinación homóloga en el genoma. La tasa de éxito de este método es baja, debido a la poca eficacia de la recombinación homóloga y una alta frecuencia de inserción no específica del ADN donador en regiones del genoma distintas del sitio diana.

La presente descripción proporciona métodos de alteración de secuencia dirigida caracterizada por una mayor

eficacia de la recombinación dirigida y una menor frecuencia de sucesos de inserción no específicos. Los métodos implican hacer y usar dominios de unión con dedos de zinc genéticamente modificados fusionados con dominios de escisión (o semidominios de escisión) para hacer una o más roturas bicatenarias en el ADN celular. Debido a que las roturas bicatenarias en el ADN celular estimulan mecanismos de reparación celular en varios miles de veces en la proximidad del sitio de escisión, dicha escisión dirigida permite la alteración o sustitución (mediante reparación dirigida por homología) de secuencias prácticamente en cualquier sitio en el genoma.

Además de las moléculas de fusión descritas en la presente memoria, la sustitución dirigida de una secuencia genómica seleccionada también requiere la introducción de la secuencia de sustitución (o donadora). La secuencia donadora se puede introducir en la célula antes, simultáneamente o después de la expresión de la(s) proteína(s) de fusión. El polinucleótido donador contiene suficiente homología con una secuencia genómica para soportar la recombinación homóloga (o reparación dirigida por homología) entre este y la secuencia genómica con la que tiene homología. Aproximadamente 25, 50, 100, 200, 500, 750, 1.000, 1.500, 2.000 nucleótidos o más de homología de secuencia entre un donador y una secuencia genómica (o cualquier valor entero entre 10 y 2.000 nucleótidos, o más) soportarán la recombinación homóloga entre ellos. Las secuencias donadoras pueden tener una longitud en el intervalo de 10 a 5.000 nucleótidos (o cualquier valor entero de nucleótidos entre estos) o más larga. Será fácilmente evidente que la secuencia donadora típicamente no es idéntica a la secuencia genómica que sustituye. Por ejemplo, la secuencia del polinucleótido donador puede contener uno o más cambios, inserciones, eliminaciones, inversiones o reordenaciones de bases individuales con respecto a la secuencia genómica, con la condición de que haya suficiente homología con la secuencia cromosómica. Alternativamente, una secuencia donadora puede contener una secuencia no homóloga flanqueada por dos regiones de homología. Además, las secuencias donadoras pueden comprender una molécula vector que contiene secuencias que no son homólogas con la región de interés en la cromatina celular. En general, la o las regiones homólogas de una secuencia donadora tendrán una identidad de secuencia de al menos 50% con una secuencia genómica con la que se desea la recombinación. En algunas realizaciones, hay presente una identidad de secuencia de 60%, 70%, 80%, 90%, 95%, 98%, 99% o 99,9%. Puede haber presente cualquier valor entre 1% y 100% de identidad de secuencia, dependiendo de la longitud del polinucleótido donador.

Una molécula donadora puede contener varias regiones discontinuas de homología con la cromatina celular. Por ejemplo, para la inserción dirigida de secuencias que normalmente no están presentes en una región de interés, dichas secuencias pueden estar presentes en una molécula de ácido nucleico donadora y flanqueada por regiones de homología con la secuencia en la región de interés.

Para simplificar los ensayos (p. ej., hibridación, PCR, digestión con enzimas de restricción) para determinar la inserción satisfactoria de la secuencia donadora, pueden estar presentes algunas diferencias de secuencia en la secuencia donadora comparada con la secuencia genómica. Preferiblemente, si se encuentra en una región codificante, dichas diferencias de secuencia de nucleótidos no cambiarán la secuencia de aminoácidos, o harán cambios de aminoácidos silenciosos (es decir, cambios que no afectan a la estructura o función de la proteína). El polinucleótido donador opcionalmente puede contener cambios en secuencias correspondientes a los sitios de unión del dominio con dedos de zinc en la región de interés, para prevenir la escisión de las secuencias donadoras que se han introducido en la cromatina celular por recombinación homóloga.

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

Un polinucleótido se puede introducir en una célula como parte de una molécula vector que tiene secuencias adicionales tales como, por ejemplo, orígenes de replicación, promotores y genes que codifican resistencia a antibióticos. Además, los polinucleótidos donadores se pueden introducir como ácido nucleico solo, como ácido nucleico en complejo con un agente tal como un liposoma o poloxámero, o se puede suministrar mediante bacterias o virus (p. ej., *Agrobacterium*, *Rhizobium sp.* NGR234, *Sinorhizobium meliloti*, *Mesorhizobium loti*, virus del mosaico del tabaco, virus X de la patata, virus del mosaico de la coliflor y virus del mosaico de la vena de la yuca. Véase, p. ej., Chung *et al.* (2006) *Trends Plant Sci.* 11(1):1-4.

Sin querer estar ligado por una teoría, parece que la presencia de una rotura bicatenaria en una secuencia celular, acoplada con la presencia de una molécula de ADN exógeno que tiene homología con una región adyacente o alrededor de la rotura, activa los mecanismos celulares que reparan la rotura por transferencia de información de secuencia de la molécula donadora a la secuencia celular (p. ej., genómica o cromosómica), es decir, por un proceso de reparación dirigida por homología, también conocido como "conversión génica". Los métodos de los autores de la invención combinan ventajosamente las potentes capacidades de dirigir de los ZFP genéticamente modificados con un dominio de escisión (o semidominio de escisión) para dirigir específicamente una rotura

bicatenaria a la región del genoma en la inserción de la secuencia exógena que se desea.

Para la alteración de una secuencia cromosómica, no es necesario que la secuencia entera del donador sea copiada en el cromosoma, con la condición de que sea copiada suficiente secuencia del donador para realizar la alteración de secuencia deseada.

5 La eficacia de la inserción de secuencias donadoras por recombinación homóloga está inversamente relacionada con la distancia, en el ADN celular, entre la rotura bicatenaria y el sitio en el que se desea la recombinación. En otras palabras, se observan eficacias de recombinación homóloga mayores cuando la rotura bicatenaria está más cerca del sitio en el que se desea la recombinación. En los casos en los que no está predeterminado un sitio preciso de recombinación (p. ej., el suceso de recombinación deseado se puede producir a lo largo de un intervalo de
10 secuencia genómica), la longitud y la secuencia del ácido nucleico donador, junto con el o los sitios de escisión, se seleccionan para obtener el suceso de recombinación deseado. En los casos en los que el suceso deseado está concebido para cambiar la secuencia de un solo par de nucleótidos en una secuencia genómica, la cromatina celular se escinde en el espacio de 10.000 nucleótidos en cualquiera de los lados de ese par de nucleótidos. En algunas realizaciones, la escisión se produce en el espacio de 1.000, 500, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, o 2
15 nucleótidos, o cualquier valor entero entre 2 y 1.000 nucleótidos, en cualquiera de los lados del par de nucleótidos cuya secuencia se va a cambiar.

Como se ha detallado antes, los sitios de unión para dos proteínas de fusión, que comprende cada una un dominio de unión con dedos de zinc y un semidominio de escisión, pueden encontrarse a 5-8 o 15-18 nucleótidos de separación, medido desde el borde de cada sitio de unión más cercano al otro sitio de unión, y la escisión se
20 produce entre los sitios de unión. El que la escisión se produzca en un solo sitio o en múltiples sitios entre los sitios de unión es irrelevante, puesto que las secuencias genómicas escindidas son sustituidas por las secuencias donadoras. Por lo tanto, para la alteración eficaz de la secuencia de un solo par de nucleótidos por recombinación dirigida, el punto medio de la región entre los sitios de unión está en el espacio de 10.000 nucleótidos de ese par de nucleótidos, preferiblemente en el espacio de 10.000 nucleótidos de este par de nucleótidos, preferiblemente en el
25 espacio de 1.000 nucleótidos, o 500 nucleótidos, o 200 nucleótidos, o 100 nucleótidos, o 50 nucleótidos, o 20 nucleótidos, o 10 nucleótidos, o 5 nucleótidos, o 2 nucleótidos, o 1 nucleótido, o en el par de nucleótidos de interés.

En algunas realizaciones, un cromosoma homólogo puede servir como el polinucleótido donador. Así, por ejemplo, la corrección de una mutación en un heterocigoto se puede lograr mediante la modificación genética de proteínas de fusión que se unen y escinden la secuencia mutante en un cromosoma, pero no escinden la secuencia natural en el
30 cromosoma homólogo. La rotura bicatenaria en el cromosoma que lleva mutación estimula un proceso de "conversión génica" basada en homología, en el que la secuencia natural del cromosoma homólogo es copiada en el cromosoma escindido, restaurando así dos copias de la secuencia natural.

También se proporcionan métodos y composiciones que pueden potenciar los niveles de recombinación dirigida, incluyendo, pero no limitado al uso de fusiones de ZFP-dominio funcional adicionales para activar la expresión de genes implicados en la recombinación homóloga, tales como, por ejemplo, miembros del grupo de epistasia RAD52 (p. ej., *Rad50*, *Rad51*, *Rad51B*, *Rad51C*, *Rad51D*, *Rad52*, *Rad54*, *Rad54B*, *Mre11*, *XRCC2*, *XRCC3*), genes cuyos productos interactúan con los productos génicos mencionados antes (p. ej., BRCA1, BRCA2) y/o genes en el complejo NBS1. Véase, p. ej., Boyko et al. (2006) *Plant Physiology* 141:488-497 y LaFarge et al. (2003) *Nucleic Acids Res* 31(4): 1148-1155. Igualmente se pueden usar fusiones de ZFP-dominio funcional en combinación con los métodos y composiciones descritos en la presente memoria, para reprimir la expresión de genes implicados en la unión de extremos no homólogos (p. ej., Ku70/80, XRCC4, poli(ADP ribosa) polimerasa, ADN ligasa 4. Véase, por ejemplo, Riha et al. (2002) *EMBO* 21:2819- 2826; Freisner et al. (2003) *Plant J.* 34:427-440; Chen et al. (1994) *European Journal of Biochemistry* 224:135-142. Se describen métodos para la activación y represión de la expresión de genes usando fusiones entre un dominio de unión con dedos de zinc y un dominio funcional, por ejemplo, en las patentes de EE.UU. en copropiedad 6.534.261; 6.824.978 y 6.933.113. Los métodos de represión adicionales incluyen el uso de oligonucleótidos antisentido y/o ARN interferente pequeño (ARNip o ARNi) dirigido a la secuencia del gen que se va a reprimir.

Como una alternativa o además de la activación de la expresión de productos génicos implicados en la recombinación homóloga, las fusiones de estas proteínas (o fragmentos funcionales de las mismas) con un dominio de unión con dedos de zinc dirigido a la región de interés, se puede usar para reclutar estas proteínas (proteínas recombinantes) en la región de interés, aumentado así su concentración local y estimulando más los procesos de recombinación homóloga. Alternativamente, un polipéptido implicado en la recombinación homóloga como se ha descrito antes (o un fragmento funcional del mismo) puede ser parte de una proteína de fusión triple que comprende un dominio de unión con dedos de zinc, un dominio de escisión (o semidominio de escisión) y la proteína recombinante (o fragmento funcional de la misma). Las proteínas adicionales implicadas en la conversión génica y el remodelado de cromatina relacionado con recombinación, que se pueden usar en los métodos y composiciones mencionados antes, incluyen histona acetiltransferasas (p. ej., Esa1p, Tip60), histona metiltransferasas (p. ej., Dot1p), histona quinasas e histona fosfatasa. Véase, también Bhat et al. (1999) *Plant J.* 33:455-469.

Se logran aumentos adicionales de la eficacia de la recombinación dirigida, en células que comprenden una molécula de fusión de dedos de zinc/nucleasa y una molécula de ADN donadora, bloqueando las células en la fase
60

G₂ del ciclo celular, cuando los procesos de reparación dirigidos por homología están activos al máximo. Dicha detención se puede lograr mediante una serie de formas. Por ejemplo, las células se pueden tratar, p. ej., con fármacos, compuestos y/o moléculas pequeñas que influyen en el avance del ciclo celular para así detener las células en la fase G₂. Las moléculas de ejemplo de este tipo incluyen, pero no se limitan a compuestos que afectan a la polimerización de microtúbulos (p. ej., vinblastina, nododazol, taxol), compuestos que interaccionan con el ADN (p. ej., dicloruro de *cis*-platino(II)-diamina, cisplatino, doxorubicina) y/o compuestos que afectan a la síntesis de ADN (p. ej., timidina, hidroxiaurea, L-mimosina, etopósido, 5-fluorouracilo). Se logran aumentos adicionales en la eficacia de la recombinación mediante el uso de inhibidores de la histona desacetilasa (HDAC) (p. ej., butirato sódico, tricotatina A) que alteran la estructura de la cromatina para hacer el ADN genómico más accesible a la maquinaria de la recombinación celular.

Métodos adicionales para la detención del ciclo celular incluyen el exceso de expresión de proteínas que inhiben la actividad de las quinasas del ciclo celular CDK, por ejemplo, introduciendo una ADNc que codifica la proteína en la célula o introduciendo en la célula un ZFP genéticamente modificado que activa la expresión del gen que codifica la proteína. La detención del ciclo celular también se logra inhibiendo la actividad de las ciclinas y CDK, por ejemplo, usando métodos de ARNi (p. ej., patente de EE.UU. n° 6.506.559) o introduciendo en la célula un ZFP genéticamente modificado que reprime la expresión de uno o más genes implicados en el avance del ciclo celular, tales como, por ejemplo, genes de ciclina y/o CDK. Véase, p. ej., la patente de EE.UU. en copropiedad n° 6.534.261 para métodos de síntesis de proteínas con dedos de zinc diseñadas para la regulación de la expresión génica.

Alternativamente, en algunos casos, la escisión dirigida se lleva a cabo en ausencia de un polinucleótido donador (preferiblemente en fase S o G₂) y la recombinación se produce entre cromosomas homólogos.

Vectores de expresión

Un ácido nucleico que codifica uno o más ZFP o proteínas de fusión de ZFP se pueden clonar en un vector para la transformación en células procariontas o eucariotas para la replicación y/o expresión. Los vectores pueden ser vectores procariontas, p. ej., plásmidos o vectores lanzadera, vectores de insecto o vectores eucariotas. Un ácido nucleico que codifica un ZFP también se puede clonar en un vector de expresión, para la administración a una célula de planta.

Para expresar los ZFP o proteínas de fusión de ZFP, las secuencias que codifican los ZFP o fusiones de ZFP típicamente se subclonan en un vector de expresión que contiene un promotor para dirigir la transcripción. Los promotores bacterianos y eucariotas adecuados son bien conocidos en la técnica y se describen, p. ej., en Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2ª ed. 1989; 3ª ed., 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y *Current Protocols in Molecular Biology* (Ausubel et al., véase antes). Los sistemas de expresión bacterianos para expresar el ZFP están disponibles, p. ej., en *E. coli*, *Bacillus sp.* y *Salmonella* (Palva et al., *Gene* 22:229-235 (1983)). Los kits para dichos sistemas de expresión están disponibles en el comercio. Los sistemas de expresión eucariotas para células de mamífero, levadura, y células de insecto son bien conocidos para los expertos en la técnica, y también están disponibles en el comercio.

El promotor usado para la expresión directa de un ácido nucleico que codifica el ZFP depende de la aplicación particular. Por ejemplo, normalmente se usa un promotor constitutivo fuerte adecuado para la célula hospedante para la expresión y purificación de ZFP.

En cambio, cuando se administra un ZFP in vivo para la regulación génica de una planta (véase, "Suministro de ácidos nucleicos a células de plantas" en la siguiente sección), se usa un promotor constitutivo o uno inducible, dependiendo del uso particular del ZFP. Los ejemplos no limitantes de los promotores de plantas incluyen secuencias de promotores derivadas de ubiquitina-3 de *A. thaliana* (ubi-3) (Callis, et al., 1990, *J. Biol. Chem.* 265:12486-12493); manopina sintasa de *A. tumifaciens* (Δ mas) (Petolino et al., patente de EE.UU. n° 6.730.824); y/o virus del mosaico de la vena de yuca (CsVMV) (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139). Véase también, Ejemplos.

Además del promotor, el vector de expresión típicamente contiene una unidad de transcripción o casete de expresión que contiene todos los elementos adicionales requeridos para la expresión del ácido nucleico en células hospedantes, procariontas o eucariotas. Un casete de expresión típico contiene por lo tanto, un promotor operativamente unido, p. ej., a una secuencia de ácido nucleico que codifica la ZFP, y señales requeridas, p. ej., para la poliadenilación eficaz del transcrito, terminación transcripcional, sitios de unión al ribosoma o terminación de la traducción. Elementos adicionales del casete pueden incluir, p. ej., potenciadores y señales de empalme heterólogo.

El vector de expresión particular usado para transportar la información genética a la célula se selecciona con respecto al uso previsto del ZFP, p. ej., expresión en plantas, animales, bacterias, hongos, protozoos, etc. (véase los vectores de expresión descritos más adelante). Se conocen en la técnica vectores de expresión bacterianos y de animales convencionales y se describen con detalle, por ejemplo, en la publicación de patente de EE.UU. 20050064474A1 y publicaciones de patente internacional WO05/084190, WO05/014791 y WO03/080809.

Se pueden usar métodos de transfección convencionales para producir líneas de células bacterianas, de mamífero,

levadura o insecto que expresan cantidades grandes de proteína, que después se pueden purificar usando técnicas convencionales (véase, p. ej., Colley et al., *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification, in Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). La transformación de células eucariotas y procariontes se lleva a cabo de acuerdo con técnicas convencionales (véase, p. ej., Morrison, *J. Bact.* 132:349-351 (1977); Clark-Curtiss y Curtiss, *Methods in Enzymology* 101:347-362 (Wu et al., eds., 1983).

Se puede usar cualquiera de los procedimientos bien conocidos para introducir secuencias de nucleótidos extrañas en dichas células hospedantes. Estos incluyen el uso de transfección con fosfato de calcio, polibrenno, fusión de protoplastos, electroporación, métodos ultrasónicos (p. ej., sonoporación), liposomas, microinyección, ADN solo, vectores plasmídicos, vectores víricos, tanto episómicos como integrativos, y cualquiera de los otros métodos bien conocidos para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético extraño en una célula hospedante (véase, p. ej., Sambrook et al., véase antes). Solo es necesario que el procedimiento particular de modificación genética usado sea capaz de introducir satisfactoriamente al menos un gen en la célula hospedante capaz de expresar la proteína de elección.

Suministro de ácidos nucleicos en células de plantas

Como se ha indicado antes, se pueden introducir construcciones de ADN (p. ej., en el genoma de) un hospedante de planta deseado mediante una variedad de técnicas convencionales. Para revisiones de dichas técnicas, véase, por ejemplo, Weissbach y Weissbach *Methods for Plant Molecular Biology* (1988, Academic Press, N.Y.) Sección VIII, pág. 421-463; y Grierson y Corey, *Plant Molecular Biology* (1988, 2ª Ed.), Blackie, London, Ch. 7-9.

Por ejemplo, la construcción de ADN se puede introducir directamente en el ADN genómico de la célula de la planta usando técnicas tales como la electroporación y microinyección de protoplastos de célula de planta, o las construcciones de ADN se pueden introducir directamente en el tejido de la planta usando métodos biobalísticos, tales como bombardeo de partículas de ADN (véase, p. ej., Klein et al. (1987) *Nature* 327:70-73). Alternativamente, las construcciones de ADN se pueden combinar con regiones flanqueadoras de T-ADN adecuadas e introducir en un vector hospedante de *Agrobacterium tumefaciens* convencional. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens* incluyendo desarmado y uso de vectores binarios, están bien descritos en la bibliografía científica. Véase, por ejemplo, Horsch et al. (1984) *Science* 233:496-498, y Fraley et al. (1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803.

Además, la transferencia de genes se puede lograr usando bacterias no *Agrobacterium* o virus tales como *Rhizobium sp.* NGR234, *Sinorhizobium meliloti*, *Mesorhizobium loti*, virus X de la patata, virus del mosaico de la coliflor y virus del mosaico de la vena de la yuca y/o virus del mosaico del tabaco. Véase, p. ej., Chung et al. (2006) *Trends Plant Sci.* 11(1):1-4.

Las funciones de virulencia del hospedante *Agrobacterium tumefaciens* dirigirán la inserción de la construcción y el marcador adyacente en el ADN de la célula de planta cuando la célula es infectada por la bacteria usando el vector binario de ADN-T (Bevan (1984) *Nuc. Acid Res.* 12:8711-8721) o el procedimiento de cocultivo (Horsch et al. (1985) *Science* 227:1229-1231). En general, el sistema de transformación de *Agrobacterium* se usa para la modificación genética de plantas dicotiledóneas (Bevan et al. (1982) *Ann. Rev. Genet.* 16:357-384; Rogers et al. (1986) *Methods Enzymol.* 118:627-641). El sistema de transformación de *Agrobacterium* también se puede usar para transformar, así como transferir, ADN a plantas monocotiledóneas y células de plantas. Véase, la patente de EE.UU. n° 5.591.616; Hernalsteen et al. (1984) *EMBO J* 3:3039-3041; Hooykass-Van Slogteren et al. (1984) *Nature* 311:763-764; Grimsley et al. (1987) *Nature* 325:1677-179; Boulton et al. (1989) *Plant Mol. Biol.* 12:31-40; y Gould et al. (1991) *Plant Physiol.* 95:426-434.

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

Los métodos y composiciones descritos se pueden usar para insertar secuencias exógenas en una localización predeterminada en un genoma de célula de planta. Esto es útil considerando que la expresión de un transgén introducido en un genoma de planta, depende críticamente de su sitio de integración. Por consiguiente, se pueden insertar genes que codifican, p. ej., nutrientes, antibióticos o moléculas terapéuticas, mediante recombinación dirigida, en regiones de un genoma de planta favorable para su expresión.

Las células de planta transformadas que se producen por cualquiera de las técnicas de transformación anteriores, se pueden cultivar para regenerar una planta entera que tiene el genotipo transformado y por lo tanto el fenotipo deseado. Dichas técnicas de regeneración se basan en la manipulación de algunas fitohormonas en un medio de crecimiento de cultivo tisular, típicamente basándose en un marcador biocida y/o herbicida que se ha introducido

junto con las secuencias de nucleótidos deseadas. La regeneración de plantas desde protoplastos cultivados se describe en Evans, et al., "*Protoplasts Isolation and Culture*" en *Handbook of Plant Cell Culture*, pág. 124-176, Macmillan Publishing Company, New York, 1983; y *Binding, Regeneration of Plants, Plant Protoplasts*, pág. 21-73, CRC Press, Boca Raton, 1985. La regeneración también se puede obtener de callos de plantas, explantes, órganos, polen, embriones o partes de los mismos. Dichas técnicas de regeneración se describen en general en Klee et al (1987) *Ann. Rev. of Plant Phys.* 38:467-486.

Los ácidos nucleicos introducidos en una célula de planta se pueden usar para conferir las características deseadas esencialmente en cualquier planta. Se pueden modificar genéticamente una variedad de plantas y sistemas de células de plantas para las características fisiológicas y agronómicas descritas en la presente memoria, usando las construcciones de ácido nucleico de la presente descripción y los diferentes métodos de transformación mencionados antes. En realizaciones preferidas, las plantas diana y las células de planta para la genomaniplación incluyen, pero no se limitan a plantas monocotiledóneas y dicotiledóneas, tales como cosechas que incluyen cosechas de grano (p. ej., trigo, maíz, arroz, mijo, cebada), cultivos de frutas (p. ej., tomate, manzana, pera, fresa, naranja), cultivos forrajeros (p. ej., alfalfa), cultivos de hortalizas de raíz (p. ej., zanahoria, patata, remolacha azucarera, batata), cultivos de hortalizas de hoja verde (p. ej., lechuga, espinaca); plantas con flores (p. ej., petunia, rosa, crisantemo), coníferas y árboles de pino (p. ej., picea, abeto); plantas usadas en fitorremedio (p. ej., plantas que acumulan metales pesados); cultivos de aceite (p. ej., girasol, semillas de colza) y plantas usadas con fines experimentales (p. ej., *Arabidopsis*). Por lo tanto, los métodos y composiciones descritos tienen uso en una amplia variedad de plantas, incluyendo, pero no limitado a especies de los géneros *Asparagus*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucurbita*, *Daucus*, *Glycine*, *Gossypium*, *Hordeum*, *Lactuca*, *Lycopersicon*, *Malus*, *Manihot*, *Nicotiana*, *Oryza*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Solanum*, *Sorghum*, *Triticum*, *Vitis*, *Vigna*, y *Zea*.

Un experto en la técnica reconocerá que después de que el casete de expresión se ha incorporado de forma estable en plantas transgénicas y se ha confirmado que es operable, se puede introducir en otras plantas por cruce sexual. Se puede usar cualquiera de una serie de técnicas de reproducción convencionales, dependiendo de las especies que se van a cruzar.

Una célula de planta, callo, tejido o planta transformados se puede identificar y aislar por selección o cribado del material de la planta genéticamente modificado según las características codificadas por los genes de marcadores presentes en el ADN transformante. Por ejemplo, la selección se puede llevar a cabo cultivando el material de la planta genéticamente modificado en medio que contiene una cantidad inhibidora del antibiótico o herbicida al que la construcción génica transformante confiere resistencia. Además, las plantas y células de plantas transformadas también se pueden identificar por cribado según las actividades de cualquier gen de marcador visible (p. ej., la β -glucuronidasa, luciferasa, genes B o C1) que pueden estar presentes en construcciones de ácidos nucleicos recombinantes. Dichas metodologías de selección y cribado son bien conocidas para los expertos en la técnica.

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

Los efectos de la manipulación genética usando los métodos descritos en la presente memoria se pueden observar, por ejemplo, por transferencias Northern del ARN (p. ej. ARNm) aislado de tejidos de interés. Típicamente, si la cantidad de ARNm ha aumentado, se puede suponer que el correspondiente gen endógeno está siendo expresado con una tasa mayor que antes. Se pueden usar otros métodos para medir la actividad de genes y/o CYP74B. Se pueden usar diferentes tipos de ensayos enzimáticos, dependiendo del sustrato usado y el método de detección del aumento o disminución de un producto o subproducto de reacción. Además, los niveles de proteína CYP74B expresada se pueden medir por métodos inmunoquímicos, es decir, ELISA, RIA, EIA y otros ensayos basados en anticuerpos bien conocidos para los expertos en la técnica, tales como por ensayos de detección electroforética (con tinción o transferencia western). El transgén se puede expresar selectivamente en algunos tejidos de la planta o en algunas etapas del desarrollo, o el transgén se puede expresar en sustancialmente todos los tejidos de la planta, sustancialmente a lo largo de todo su ciclo de vida. Sin embargo, también es aplicable cualquier modo de expresión combinatoria.

También se describen semillas de las plantas transgénicas descritas antes, en donde la semilla tiene el transgén o construcción génica. La presente descripción abarca además la progenie, clones, líneas celulares o células de las plantas transgénicas descritas antes, en donde dicha progenie, clon, línea celular o célula tiene el transgén o construcción génica.

Los ZFP y vectores de expresión que codifican los ZFP se pueden administrar directamente a la planta para la escisión dirigida y/o recombinación.

5 La administración de las cantidades eficaces es por cualquiera de las vías usadas normalmente para introducir ZFP en contacto íntimo con la célula de la planta que se va a tratar. Los ZFP se administran de cualquier forma adecuada, preferiblemente con vehículos farmacéuticamente aceptables. Los métodos adecuados de administración de dichos moduladores están disponibles y son bien conocidos para los expertos en la técnica, y aunque se puede usar más de una vía para administrar una composición particular, una vía particular a menudo puede proporcionar una reacción más inmediata y más eficaz que otra vía.

10 Se pueden usar vehículos y se determinan en parte por la composición particular que se va a administrar, así como por el método particular usado para administrar la composición. Por consiguiente, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas que están disponibles (véase, p. ej., *Remington's Pharmaceutical Sciences*, 17^a ed. 1985)).

Aplicaciones

15 Los métodos descritos para la escisión dirigida se pueden usar para inducir mutaciones en una secuencia genómica. La escisión dirigida también se puede usar para generar inactivaciones de genes (p. ej., para genómica funcional o validación de dianas) y para facilitar la inserción dirigida de una secuencia en un genoma (es decir, introducción génica). La inserción puede ser por medios de sustituciones de secuencias cromosómicas por recombinación homóloga o por integración dirigida, en la que una nueva secuencia (es decir, una secuencia no presente en la región de interés), flanqueada por secuencias homólogas con la región de interés en el cromosoma, se inserta en un sitio diana predeterminado. También se pueden usar los mismos métodos para sustituir una secuencia natural con una secuencia mutante, o para convertir un alelo en un alelo diferente.

20 La escisión dirigida de patógenos de plantas infecciosos o integrados se puede usar para tratar infecciones patógenas en un hospedante de planta, por ejemplo, escindiendo el genoma del patógeno de modo que su patogenicidad se reduce o elimina. Además, se puede usar la escisión dirigida de genes que codifican receptores para virus de plantas para bloquear la expresión de dichos receptores, previniendo así la infección vírica y/o propagación vírica en la planta.

25 Los ejemplos de patógenos de planta de ejemplo incluyen, pero no se limitan a virus de plantas tales como *Alfamoviruses*, *Alphacryptoviruses*, *Badnaviruses*, *Betacryptoviruses*, *Bigeminiviruses*, *Bromoviruses*, *Bymoviruses*, *Capilloviruses*, *Carlaviruses*, *Carmoviruses*, *Caulimoviruses*, *Closteroviruses*, *Comoviruses*, *Cucumoviruses*, *Cytorhabdoviruses*, *Dianthoviruses*, *Enamoviruses*, *Fabaviruses*, *Fijiviruses*, *Furoviruses*, *Hordeoviruses*, *Hybrigeminiviruses*, *Idaeoviruses*, *Illaviruses*, *Ipomoviruses*, *Luteoviruses*, *Machlomoviruses*, *Macluraviruses*, *Marafiviruses*, *Monogeminiviruses*, *Nanaviruses*, *Necroviruses*, *Nepoviruses*, *Nucleorhabdoviruses*, *Oryzaviruses*, *Ourmiaviruses*, *Phytoreoviruses*, *Potexviruses*, *Potyvirus*, *Rymoviruses*, *ARN satélites*, *satelliviruses*, *Sequiviruses*, *Sobemoviruses*, *Tenuiviruses*, *Tobamoviruses*, *Tobraviruses*, *Tombusviruses*, *Tospoviruses*, *Trichoviruses*, *Tymoviruses*, *Umbraviruses*, *Varicosaviruses* y *Waikaviruses*; patógenos fúngicos tales como carbones (p. ej. *Ustilaginales*), royas (*Uredinales*), ergots (*Claviceps pupurea*) y mildiú; mohos (*Oomycetes*) tales como *Phytophthora infestans* (tizón de la patata); patógenos bacterianos tales como *Erwinia* (p. ej., *E. herbicola*), *Pseudomonas* (p. ej., *P. aeruginosa*, *P. syringae*, *P. fluorescense* y *P. putida*), *Ralstonia* (p. ej., *R. solanacearum*), *Agrobacterium* y *Xanthomonas*; ascárides (Nematoda); y *Phytopmyxa* (*Polymyxa* y *Plasmodiophora*).

30 Los métodos descritos para la recombinación dirigida se pueden usar para sustituir cualquier secuencia genómica por una secuencia homóloga no idéntica. Por ejemplo, una secuencia genómica mutante se puede sustituir por su equivalente natural, proporcionando así métodos para el tratamiento de enfermedades de plantas; proporcionar resistencia frente a patógenos de la planta; aumentar los rendimientos de los cultivos, etc. De la misma manera, un alelo de un gen se puede sustituir por un alelo diferente usando los métodos de recombinación dirigida descritos en la presente memoria.

35 En muchos de estos casos, una región de interés comprende una mutación, y el polinucleótido donador comprende la correspondiente secuencia natural. Igualmente, una secuencia genómica natural se puede sustituir por una secuencia mutante, si se desea. Por ejemplo, el exceso de expresión de un oncogén se puede invertir por mutación del gen o sustituyendo sus secuencias de control por secuencias que aguanten un nivel no patógeno menor de expresión. Realmente, cualquier patología dependiente de una secuencia genómica particular, de cualquier manera se puede corregir o aliviar usando los métodos y composiciones descritos en la presente memoria.

40 La escisión dirigida y recombinación dirigida también se pueden usar para alterar secuencias no codificantes (p. ej., secuencias reguladoras tales como promotores, potenciadores, iniciadores, terminadores, sitios de empalme) para alterar los niveles de expresión de un producto génico. Dichos métodos se pueden usar, por ejemplo, para propósitos terapéuticos, genómica funcional y/o estudios de validación de dianas.

45 La modificación dirigida de la estructura de cromatina, como se describe en la publicación internacional en copropiedad WO 01/83793, se puede usar para facilitar la unión de proteínas de fusión a la cromatina celular.

En realizaciones adicionales, se pueden usar una o más fusiones entre un dominio de unión con dedos de zinc y una recombinasa (o fragmento funcional de la misma), además de o en lugar de fusiones de dedos de zinc-dominio de escisión descritas en la presente memoria, para facilitar la recombinación dirigida. Véase, por ejemplo, la patente de EE.UU. en copropiedad 6.534.261 y Akopian et al. (2003) *Proc. Natl. Acad. Set USA* 100:8688-8691.

5 Los métodos descritos se pueden usar para proporcionar fusiones de dominios de unión con ZFP con dominios de activación o represión transcripcional que requieren dimerización (homodimerización o heterodimerización) para su actividad. En estos casos, un polipéptido de fusión comprende un dominio de unión con dedos de zinc y un monómero de dominio funcional (p. ej., un monómero de un dominio de activación o represión transcripcional dimérico). La unión de dos de dichos polipéptidos de fusión a sitios diana adecuadamente situados, permite la
10 dimerización de modo que se reconstituye un dominio de activación o represión transcripcional funcional.

Además, como se ha descrito antes, los métodos expuestos en la presente memoria se pueden usar para la integración dirigida de secuencias exógenas en una región de interés en el genoma de una célula, por ejemplo, en el que la escisión potencia la inserción por mecanismos dependientes de homología (p. ej., inserción de una secuencia donadora que comprende una secuencia exógena junto con una o más secuencias que son idénticas u homólogas pero no idénticas, con una secuencia genómica predeterminada (es decir, un sitio diana)).
15

La secuencia donadora típicamente contiene suficiente homología en las regiones que flanquean la secuencia exógena, para soportar la reparación dirigida por homología de una rotura bicatenaria en una secuencia genómica, insertando así la secuencia exógena en el sitio diana genómico. Por lo tanto, el ácido nucleico donador puede ser de cualquier tamaño suficiente para soportar la integración de la secuencia exógena por mecanismos de reparación dependientes de homología (p. ej., recombinación homóloga). Sin querer estar ligados por ninguna teoría en particular, las regiones de homología que flanquean la secuencia exógena se cree que proporcionan los extremos de cromosoma roto con un molde para la resíntesis de la información genética en el sitio de la rotura bicatenaria. En algunas realizaciones están presentes dos de las secuencias idénticas o dos de las secuencias homólogas pero no idénticas (o una de cada), flanqueando la secuencia exógena. Una secuencia exógena (o ácido nucleico exógeno o polinucleótido exógeno) es uno que contiene una secuencia de nucleótidos que normalmente no está presente en la región de interés.
20
25

Las secuencias exógenas de interés incluyen, pero no se limitan a ADNc, secuencias promotoras, secuencias potenciadoras, marcadores de epítomos, genes marcadores, sitios de reconocimiento de enzimas de escisión y diferentes tipos de construcciones de expresión. Véase, por ejemplo, la patente de EE.UU. nº 6.833.252. Las endonucleasas de asentamiento de ejemplo adicionales incluyen I-CeuI, PI-PspI, PI-SceI, I-ScelV, I-Csml, I-PanI, I-Scell, I-Ppol, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII. Sus secuencias de reconocimiento son conocidas. Véase también la patente de EE.UU. nº 5.420.032; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol* 280:345-353 y el catálogo de New England Biolabs.
30
35

Los genes de marcadores incluyen, pero no se limitan a secuencias que codifican proteínas que median la resistencia a antibiótico (p. ej., resistencia a ampicilina, resistencia a neomicina, resistencia a G418, resistencia a puomicina), secuencias que codifican proteínas coloreadas o fluorescentes o luminiscentes (p. ej., proteína verde fluorescente, proteína roja fluorescente potenciada, proteína roja fluorescente, luciferasa), y proteínas que median el crecimiento celular potenciado y/o amplificación génica (p. ej., dihidrofolato reductasa). Por lo tanto, los genes de marcadores de ejemplo incluyen, pero no se limitan a β -glucuronidasa (GUS), fosfinotricina N-acetil-transferasa (PAT, BAR), neomicina fosfotransferasa, β -lactamasa, catecol dioxigenasa, α -amilasa, tirosinasa, β -galactosidasa, luciferasa, aequorina, EPSP sintasa, nitrilasa, acetolactato sintasa (ALS), dihidrofolato reductasa (DHFR), dalapon deshalogenasa y antranilato sintasa). En algunas realizaciones, la integración dirigida se usa para insertar una construcción de expresión de ARN, p. ej., secuencias responsables de la expresión regulada de micro ARN o ARNip. También se pueden incorporar promotores, potenciadores y secuencias reguladoras de la transcripción adicionales, como se ha descrito antes, en una construcción de expresión de ARN.
40
45

Ejemplos

Ejemplo 1 - Diseño y generación de vector diana

50 A. Estructura general de la secuencia diana

La construcción diana para el tabaco (una dicota) incluía los siguientes 7 componentes como se muestra en la figura 1: i) un casete de expresión de higromicina fosfotransferasa (HPT) que comprende un promotor de ubiquitina-3 de *A. thaliana* (ubi-3) (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493) que dirige el gen de HPT de *E. coli* (Waldron et al., 1985, *Plant Mol. Biol.* 18:189-200) terminado por marco de lectura abierto-24 (orf-24) de *A. tumifaciens* región no traducida 3' (UTR) (Gelvin et al., 1987, documento EP222493); ii) secuencia homóloga 1, que comprende la región de unión de la matriz RB7 de *N. tabacum* (MAR) (Thompson et al., 1997, publicación internacional WO9727207); iii) un fragmento de gen de la proteína verde fluorescente (GFP) 5' (Evrogen Joint Stock Company, Moscow, Rusia) dirigido por un promotor de manopina sintasa de *A. tumifaciens* (Δ mas) (Petolino et al., patente de EE.UU. nº
55

6.730.824); iv) un casete de expresión de β -glucuronidasa (GUS) que comprende un promotor del virus del mosaico de la vena de la yuca (CsVMV) (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139) que dirige un gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) terminado por la sintasa nopalina de *A. tumifaciens* (nos) 3' UTR (DePicker et al., 1982, *J. Mol. Appl. Genet.* 1:561-573); v) un fragmento de gen de GFP 3' (Evrogen Joint Stock Company, Moscow, Rusia) terminado por un orf-1 de *A. tumifaciens* 3' UTR (Huang et al., *J. Bacteriol.* 172:1814-1822); vi) secuencia homóloga 2, que comprende el intrón-1 de la 4-cumaroil-CoA sintasa (4-CoAS) de *A. thaliana* (Locus At3g21320, GenBank NC 003074) y; vii) un fragmento de gen de la fosfinotricina fosfotransferasa (PAT) de *S. viridochromogenes* 3' (Wohlleben et al., 1988, *Gene* 70:25-37) terminado por un ORF-25/26 de *A. tumifaciens* 3' UTR (Gelvin et al., 1987, documento EP222493).

10 Se insertó un sitio de unión de la proteína de fusión de dedos de zinc-FokI (IL-1-L0-FokI) (Urnov et al., 2005, documento US 2005/0064474) en la dirección 3' del promotor del CsVMV (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139) y se fusionó con la secuencia que codifica la GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) en el extremo N-terminal. Dos copias de un sitio de unión de una segunda proteína de fusión de dedos de zinc-FokI (Scd27-L0-FokI) (Urnov et al., 2005, documento US 2005/0064474) flanqueaban los fragmentos del gen de GFP 5' y 3' (Evrogen Joint Stock Company, Moscow, Rusia). Cada sitio de unión contenía 4 repeticiones en tándem de la secuencia de reconocimiento de la proteína de fusión de dedos de zinc-FokI particular, de modo que cada sitio de unión tenía un tamaño de ~200 pb (figura 2a). Esto se diseñó para asegurar que las secuencias de reconocimiento serían accesibles a la proteína de fusión de dedos de zinc-FokI en el entorno del complejo de cromatina. Cada secuencia de reconocimiento incluía una secuencia de repetición invertida a la que se unía una sola proteína de fusión de dedos de zinc-FokI como un homodímero y escindía el ADN bicatenario (figura 2b). Los fragmentos del gen de GFP 5' y 3' se solapan en 540 pb proporcionando homología dentro de la secuencia diana y se insertó un codón de parada en el extremo 3' del fragmento de GFP 5' para asegurar que no había traducción de GFP funcional desde la secuencia diana. El vector de transformación que comprendía la secuencia diana se generó mediante un proceso de clonación de múltiples etapas como se describe más adelante.

25 B. Construcción del vector binario de HPT (pDAB1584)

El vector pDAB1584, que contenía un casete de expresión de GUS, que comprende un promotor de ubi-3 de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493) que dirige el gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) terminado por el orf-1 de *A. tumifaciens* UTR (Huang et al., *J. Bacteriol.* 172:1814-1822), se usó como la construcción base de partida (figura 3).

30 Para evitar elementos reguladores repetidos innecesarios en la construcción diana, el orf-1 de *A. tumifaciens* UTR (Huang et al., *J. Bacteriol.* 172:1814-1822) en pDAB1400 se sustituyó por un orf-24 de *A. tumifaciens* UTR (Gelvin et al., 1987, documento EP222493), que se escindió de pDAB782 (figura 4) como un fragmento SacI/XbaI y se clonó en los mismos sitios en pDAB1400. La construcción resultante contenía un promotor de ubi-3 de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493) que dirige el gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) terminado por un orf-24 de *A. tumifaciens* UTR (Gelvin et al., 1987, EP222493) y se denominó pDAB1582 (figura 5).

40 La secuencia que codificaba HPT (Waldron et al., 1985, *Plant Mol. Biol.* 18:189-200) se amplificó por PCR a partir del plásmido pDAB354 (figura 6) usando los cebadores P1 y P2. Se añadió un sitio BbsI en el extremo 5' del cebador P1 y el sitio SacI se retuvo en el extremo 3' del cebador P2. El fragmento de la PCR de HPTII se digirió con BbsI/SacI y se clonó en pDAB1582 digerido con NcoI-SacI para sustituir el gen de GUS por el gen de HPT del fragmento de la PCR. El plásmido resultante se denominó pDAB1583 (figura 7).

45 Después se escindió el fragmento de ubi-3 de *A. thaliana*/HPT/orf-24 de *A. tumifaciens* de pDAB1583 por digestión con NotI y se trató con ADN polimerasa de T4 para generar extremos romos. El casete de expresión de HPT tratado con extremos romos se clonó en pDAB2407 (figura 8), un vector base binario, en el sitio de PmeI produciendo un plásmido pDAB1584 (figura 9).

C. Construcción del vector que comprende las secuencias homólogas y el sitio de unión de la proteína de fusión de dedos de zinc-FokI Scd27 (pDAB1580)

50 El orf-1 de *A. tumifaciens* UTR (Huang et al., *J. Bacteriol.* 172:1814-1822) en pDAB2418 (figura 10) se sustituyó por el orf25/26 de *A. tumifaciens* UTR (Gelvin et al., 1987, documento EP222493) para evitar las secuencias reguladoras repetidas en el vector diana. Para hacer el intercambio de UTR, el orf25/26 de *A. tumifaciens* UTR (Gelvin et al., 1987, documento EP222493) se amplificó por PCR a partir del plásmido pDAB4045 (figura 11) usando los cebadores P3 y P4. Se añadieron los sitios SmaI y AgeI al extremo 3' del fragmento de la PCR, y se retuvo el sitio SacI en el extremo 5'. El ADN del plásmido pDAB2418, que contenía un casete de expresión del gen de PAT que comprendía el promotor de ubiquitina-10 (ubi-10) de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493) que dirige el gen de PAT (Wohlleben et al., 1988, *Gene* 70:25-37), terminado por el orf-1 de *A. tumifaciens* UTR (Huang et al., *J. Bacteriol.* 172:1814-1822) y la secuencia de MAR RB7 de *N. tabacum* (Thompson et al., 1997, publicación internacional WO9727207), se digirió con SacI y AgeI y se recuperaron los dos fragmentos mayores. Estos fragmentos se ligaron con el producto de la PCR digerido con SacI y AgeI, orf25/26 de *A. tumifaciens* UTR (Gelvin et al., 1987, documento EP222493). El plásmido resultante se denominó pDAB1575 (figura 12). MAR RB7 de

N. tabacum (Thompson et al., 1997, W09727207) sirve como secuencia homóloga 1 en el vector diana.

El intrón-1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) se seleccionó para servir como la secuencia homóloga 2 en el vector diana. La secuencia que codifica el gen de PAT (Wohlleben et al., 1988, *Gene* 70:25-37) se analizó y las 299/300 pb en la dirección 3' del codón de inicio se identificaron como el sitio para la inserción del intrón, de modo que se podían formar los sitios de empalme 5' y 3' adecuados. El intrón de longitud completa después se fusionó con 253 pb de la secuencia codificante de PAT parcial 3' por síntesis de ADN (Picoscript Ltd., LLP, Houston, Texas). Se añadieron sitios NotI y SacI en el extremo 5' y 3' del fragmento de ADN, respectivamente. El fragmento de ADN sintetizado después se digirió con NotI/SacI y se insertó en pDAB1575 en los mismos sitios para sustituir la secuencia codificante de PAT de longitud completa. La construcción resultante se denominó pDAB1577 (figura 13).

Se sintetizó un fragmento de ADN de 241 pb que contenía 4 repeticiones en tándem de Scd27-L0-FokI (figura 2) (Picoscript Ltd., LLP, Houston, Texas) con un sitio SmaI añadido tanto en el extremo 5' como 3' del fragmento. El fragmento que contenía el sitio de unión de dedos de zinc-FokI sintetizado, después se digirió con SmaI y se insertó en pDAB1577 en el sitio MscI. El vector resultante se denominó pDAB1579 (figura 14). Un segundo fragmento que contenía el sitio de unión de dedos de zinc-FokI digerido con SmaI se insertó después en pDAB1579 en el sitio SmaI. La construcción resultante se denominó pDAB1580 (figura 15). Este vector contiene las secuencias homólogas 1 y 2 (MAR RB7 de *N. tabacum* e intrón I de 4-CoAS de *A. thaliana*, respectivamente) y dos sitios de unión de dedos de zinc-FokI Scd27 sintetizados, que contiene cada uno 4 repeticiones en tándem de los sitios de reconocimiento Scd27-L0-FokI.

D. Construcción del vector que contiene dos fragmentos de GFP no funcionales parcialmente duplicados (pDAB1572)

El gen de GFP, CopGFP, se adquirió en Evrogen Joint Stock Company (Moscow, Rusia) y la secuencia codificante de longitud completa se amplificó por PCR usando los cebadores P5 y P6. Se añadieron los sitios BbsI y SacI en los extremos 5' y 3' del producto de la PCR, respectivamente. El producto de la PCR CopGFP después se digirió con BbsI/SacI y se clonó en pDAB3401 (figura 16) que comprende el promotor de Δ mas de *A. tumefaciens* modificado (Petolino et al., documento US6730824) que dirige el gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) y terminado por orf-1 de *A. tumefaciens* UTR (Huang et al., *J. Bacteriol.* 172:1814-1822) en los sitios NcoI/SacI para sustituir el gen de GUS. El vector resultante se denominó pDAB1570 (figura 17).

Para hacer los dos fragmentos de GFP no funcionales parcialmente duplicados, un fragmento de ADN que contenía la mayoría de la secuencia codificante de CopGFP con una eliminación de 47 pb en el extremo 5', se amplificó por PCR usando los cebadores P9 y P10. Se añadió un sitio Apal tanto al extremo 5' como al 3' y se añadió un sitio Stul adicional en el extremo 5' en la dirección 3' del sitio Apal. Después el producto de la PCR se digirió con Apal y se insertó en pDAB1570 en el sitio Apal, creando así dos fragmentos de GFP no funcionales en el mismo vector con una secuencia duplicada de 540 pb. La construcción resultante se denominó pDAB1575 (figura 18).

E. Construcción del vector que contiene la fusión de sitio de unión de la proteína de fusión de IL-1 dedos de zinc-FokI/gen de GUS (pDAB1573)

Un fragmento de ADN de 233 pb que contenía 4 repeticiones en tándem del sitio de reconocimiento IL-1_L0-FokI (figura 2) fue sintetizado por Picoscript Ltd., LLP, (Houston, Texas) con sitios NcoI y AflIII añadidos a los extremos 5' y 3', respectivamente. El fragmento sintetizado se digirió con NcoI/AflIII y se insertó en pDAB4003 (figura 19), que contenía un gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) dirigido por un promotor de CsVMV (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139) terminado por el orf-1 de *A. tumefaciens* 3' UTR (Huang et al., *J. Bacteriol.* 172:1814-1822) en el sitio NcoI. Después se generó una fusión N-terminal entre el sitio de unión de IL-1_Lo-FokI y la secuencia que codifica GUS. El vector resultante se denominó pDAB1571 (figura 20).

Para evitar elementos 3' UTR que se repiten en el vector diana, la 3' UTR de nos de *A. tumefaciens* (DePicker et al., 1982, *J. Mol. Appl. Genet.* 1:561-573) se cortó de pDAB7204 (figura 21) como un fragmento SacI/PmeI y se clonó en pDAB1571, que se digirió con SacI/NaeI, para sustituir el orf-1 de *A. tumefaciens* 3' UTR (Huang et al., *J. Bacteriol.* 172:1814-1822). El plásmido resultante se denominó pDAB1573 (figura 22).

F. Construcción del vector diana final (PDAB1585)

Para hacer el vector diana final, el casete de expresión de GUS con la inserción del sitio diana de la proteína de fusión de IL-1-FokI se cortó de pDAB1573 por digestión con NotI, se trató para generar extremos romos y se insertó en pDAB1572 en el sitio de Stul. El vector intermedio resultante se denominó pDAB1574 (figura 23). El casete entero que contenía el promotor Δ mas modificado (Petolino et al., documento US6730824), una secuencia de GFP 5' parcialmente duplicada (Evrogen Joint Stock Company, Moscow, Rusia), el promotor de CsVMV (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139), una secuencia diana de la proteína de fusión IL-1-FokI, la región codificante del gen de GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405), una 3' UTR de nos de *A. tumefaciens* (DePicker et al., 1982, *J. Mol. Appl. Genet.* 1:561-573), una GFP 3' parcialmente duplicada (Evrogen Joint Stock Company, Moscow, Rusia) y el orf-1 de *A. tumefaciens* 3' UTR (Huang et al., *J. Bacteriol.* 172:1814-1822), se cortó de pDAB1574 y se insertó en pDAB1580 en el sitio de NotI. El plásmido resultante se denominó pDAB1581 (figura

24). El fragmento de Agel de pDAB1581 después se insertó en pDAB1584 en el sitio de Agel creando así la construcción diana final, pDAB1585 (figura 1).

Ejemplo 2 - Generación de líneas celulares transgénicas con secuencias diana integradas

5 Se usaron dos cultivos de suspensión de células de tabaco diferentes en los que se integraron de forma estable secuencias diana del ejemplo 1 mediante transformación en *Agrobacterium*. El primer cultivo, denominado NT1, se obtuvo de Arnold Bendich de la University of Washington, Seattle, WA, EE.UU. Este cultivo prolifera como células de 15-20 µm de diámetro en grupos de 20-30 células con un tiempo de duplicado de aproximadamente 48 h. Los cultivos de suspensión de células NT1 se mantuvieron en medio que contenía sales basales MS (PhytoTechnology Labs M524), K₂HPO₄ 137,4 mg/l, sacarosa 30 g/l, 2,4-D 2,22 mg/l, tiamina-HCl 1 mg/l, mio-inositol 100 mg/l y MES 10 0,5 g/l a un pH de 5,7. Las células de NT1 se subcultivaron cada 7 días por adición de 40 ml de medio basado en MS de nueva aportación a 1 ml de volumen de células empaquetadas (PCV).

15 El segundo cultivo de células de tabaco usado, denominado BY2, se obtuvo de Jun Ueki de Japan Tobacco, Iwata, Shizuoka, Japón. Este cultivo prolifera como células de 5-10 µm de diámetro en grupos de 100-150 células con un tiempo de duplicado de aproximadamente 18 h. Los cultivos de suspensión de células BY2 se mantuvieron en medio que contenía sales basales LS (PhytoTechnology Labs L689), K₂HPO₄ 170 mg/l, sacarosa 30 g/l, 2,4-D 0,2 mg/l y tiamina-HCl 0,6 mg/l, a un pH de 6,0. Las células BY2 se subcultivaron cada 7 días por adición de 50 ml de medio basado en LS a 0,25 ml de PCV. Los cultivos de suspensiones celulares tanto NT1 como BY2 se mantuvieron en matraces de 250 ml en un agitador rotatorio a 25°C y 125 rpm.

20 Con el fin de generar cultivos celulares NT1 y BY2 transgénicos con secuencias diana integradas, un matraz de una suspensión de tabaco después de subcultivo de 4 días se dividió en cuatro partes alícuotas de 10-12 ml que se cocultivaron en placas Petri de 100x25 mm con 100 µl de la cepa de *Agrobacterium* LBA4404 que alberga pDAB1585 cultivada durante la noche hasta una DO₆₀₀ ~1,5. Las placas se envolvieron con Parafilm y se incubaron a 25°C sin agitación durante 3 días, después de lo cual se separó el exceso de líquido y se sustituyó con 100 ml de medio basal (basado en MS o LS para NT1 y BY2, respectivamente) que contenía carbenicilina 500 mg/ml.

25 Después de volver a suspender las células de tabaco, se dispuso 1 ml de suspensión sobre placas de 100x25 mm de medio base adecuado que contenía carbenicilina 500 mg/l e higromicina 200 mg/l solidificado con agar TC 8 g/l, y se incubaron sin envolver a 28°C en la oscuridad. Esto produjo 120-144 placas de selección para un solo tratamiento. Los aislados resistentes a la higromicina individuales aparecieron 10-14 días después del cultivo en placa (tabla 1) y se transfirieron a placas individuales de 60x20 mm (un aislado por placa) donde se mantuvieron 30 como callo en un programa de subcultivo de 14 días hasta que fueron necesarios para el análisis y posteriores experimentos de retransformación.

Tabla 1. Resumen de la generación de cultivo de células diana transgénicas

Cultivo de células de tabaco	nº de placas de selección	nº de sucesos transgénicos
NT1	360	305
BY2	720	551

Ejemplo 3 - Cribado y caracterización de sucesos transgénicos diana

35 Los sucesos transgénicos resistentes a la higromicina generados a partir de la transformación del vector diana en cultivos de células de tabaco BY2 o NT1 (como se describe en el ejemplo 2) se analizaron como sigue.

40 Los análisis iniciales llevados a cabo para el cribado de estos sucesos transgénicos incluían el análisis de la expresión de GUS para indicar la accesibilidad de la secuencia diana, análisis por PCR de la secuencia diana parcial y de longitud completa para confirmar la presencia e integridad del vector diana y análisis de transferencia Southern para determinar el número de copias de secuencia diana integradas. Un subconjunto de los sucesos transgénicos que mostraban expresión de GUS contenían una sola copia de la secuencia diana de longitud completa; estos se seleccionaron para el restablecimiento de cultivos celulares para generar las líneas dianas para la posterior retransformación. Estas líneas diana reestablecidas también se sometieron a posterior caracterización, que incluía 45 análisis de transferencia Southern más minucioso, confirmación de secuenciación del inserto diana entero y análisis de secuencias genómicas flanqueadoras.

50 En los cultivos de suspensión o tejido de callo de tabaco transgénico iniciados a partir de los sucesos seleccionados se analizó la actividad de GUS incubando 50 mg de muestras en 150 µl de tampón de ensayo durante 24-48 h a 37°C. El tampón de ensayo consistía en fosfato sódico 0,2 M a pH 8,0, ferricianuro potásico y ferrocianuro potásico 0,1 mM cada uno, EDTA sódico 1,0 mM, 5-bromo-4-cloro-3-indoil-β-glucuronido 0,5 mg/ml y Triton X-100 al 0,6% (v/v) (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405). La aparición de regiones de color azul se usó como indicador de la expresión del gen de GUS, que indicaba que la inserción de la secuencia diana era transcripcionalmente activa y por lo tanto accesible en el entorno genómico local.

Los sucesos transgénicos que expresaban GUS se ensayaron por PCR usando la pareja de cebadores P15/P16 que

condujeron a la amplificación de un fragmento de ADN de 10 kb que se extendía desde la 3' UTR del casete de expresión de HTP en el extremo 5' de la secuencia diana a la 3' UTR del casete del gen de PAT parcial en el extremo 3' de la secuencia diana. Puesto que todos los sucesos se obtuvieron con selección con higromicina, se dio por supuesto que el casete de expresión de HPT estaba intacto en todos los sucesos diana. Por lo tanto, solo se cubrió la 3' UTR del casete de expresión de HPT en el análisis de PCR de longitud completa. También se ensayó por PCR un subconjunto de sucesos, usando las parejas de cebadores P15/P17 y P18/P19 para determinar la integridad de los extremos 5' y 3' de la secuencia diana, respectivamente. Todos los sucesos diana confirmados por análisis de PCR se ensayaron después por análisis de transferencia Southern para determinar el número de copias de la secuencia diana integradas.

El análisis de transferencia Southern se llevó a cabo para todos los sucesos diana que pasaron el cribado de la expresión de GUS y PCR de longitud completa. Se digirieron 10 µg de ADN genómico con Nsil, que era un cortado único dentro de la secuencia diana. El ADN genómico digerido se separó en un gel de agarosa al 0,8% y se transfirió a una membrana de nailon. Después de reticulación, el ADN transferido sobre la membrana se hibridó con una sonda de gen de HPT para determinar el número de copias del extremo 5' de la secuencia diana. Después, se separó y rehibridó la misma transferencia con una sonda de gen de PAT para determinar el número de copias del extremo 3' de la secuencia diana.

Tres sucesos que mostraban expresión de GUS y contenían una sola copia de la secuencia diana de longitud completa se seleccionaron para la caracterización adicional, que incluía análisis de transferencia Southern más minucioso, confirmación de la secuencia diana entera y análisis de la secuencia genómica flanqueadora (tabla 2). Estos tres sucesos eran BY2-380, NT1-240 y NT1-260. Se restablecieron cultivos de suspensión a partir de estos tres sucesos para la posterior retransformación con vectores que comprendían ADN de donador y genes de la proteína de fusión de dedo de zinc-Fok1.

Tabla 2. Caracterización de cultivos de células diana transgénicas seleccionados

Línea celular	Nº de sucesos	Resistencia a higromicina	Expresión de GUS		nº copias de PAT	nº copias de HPTII	PCR de longitud completa	5' PCR	3' PCR
			callo	suspensión					
BY2	380	+	+	+	1	1	+	+	+
NT1	240	+	+	+	1	1	+	+	+
NT1	260	+	+	+	1	1	+	+	+

* con una inserción en el extremo 3' de la diana

Para asegurar que los tres cultivos de suspensiones establecidos a partir de los sucesos diana BY2-380, NT1-240 y NT1-260 contenían la secuencia diana intacta como se esperaba, la secuencia diana principal desde 3' UTR del casete de expresión de HTP en el extremo 5' de la secuencia diana a la 3' UTR del casete del gen de PAT parcial en el extremo 3' de la secuencia diana, se amplificó por PCR usando la pareja de cebadores P15/P16 y se clonó en el vector TOPO pCR2.1 (Invitrogen, Carlsbad, California). Los productos de la PCR insertados en el vector TOPO fueron secuenciados por Lark technology, Inc. (Houston, Texas). Los resultados de la secuencia indicaban que mientras que tanto BY2-380 como NT1-240 tenían la secuencia diana completa como se esperaba, NT1-260 tenía una inserción de ADN de 5475 pb. La inserción se localizó 27pb en la dirección 5' del extremo 3' de la secuencia de PAT parcial 3'. Es interesante que había un orf de 2883 pb en esta inserción. Una búsqueda por BLAST de este orf frene a la base de datos de NCBI mostró que se correspondía con un transposón de *Agrobacterium*. Aunque esta línea diana, NT1-260, puede no ser adecuada para los experimentos de recombinación homóloga intercromosómica debido a la secuencia de inserción extra dentro del gen del marcador seleccionable PAT, todavía se podría usar para ensayar la recombinación homóloga intracromosómica usando el sistema de indicador de GFP diseñado en el vector diana.

Las tres líneas se analizaron más para obtener las secuencias genómicas flanqueadoras usando el kit Universal GenomeWalker Kit (Clontech, Mountain View, California). Brevemente, se digirieron 2,5 µg de ADN genómico con tres enzimas de restricción de extremos romos, EcoRV, Dral y Stul en reacciones separadas. El ADN digerido se purificó por extracción con fenol/cloroformo y se ligó con BD GenomeWalker Adaptor. Se llevó a cabo la amplificación por PCR anidada con ligado como molde y el cebador P20 (paseo en la dirección 5' del extremo 5' de la inserción de la secuencia diana) y P21 (paseo en la dirección 3' del extremo 3' de la inserción de la secuencia diana) para la primera reacción de PCR y cebador P22 (paseo en la dirección 5' del extremo 5' de la inserción de la secuencia diana) y P23 (paseo en la dirección 3' del extremo 3' de la inserción de la secuencia diana) para la segunda reacción de PCR anidada. Los fragmentos amplificados de las segundas reacciones de PCR se clonaron en el vector pCR2.1 TOPO o pCR Blunt II TOPO (Invitrogen, Carlsbad, California) y se secuenciaron usando un kit de secuenciación Dye Terminator Cycle Sequencing Kit (Beckman Coulter, Fullerton, CA). Se obtuvieron las secuencias genómicas flanqueadoras de las tres líneas diana mediante este procedimiento. Después se diseñaron los cebadores basándose en las secuencias genómicas flanqueadoras para amplificar la secuencia diana entera. Los fragmentos amplificados obtenidos a partir de estas líneas diana eran del tamaño esperado. Ambos extremos de los fragmentos amplificados se confirmaron por secuenciación.

Ejemplo 4: Diseño y generación del vector de ADN donador

La construcción de ADN donadora incluía la secuencia homóloga 1 (RB7 MAR de *N. tabacum*) (Thompson et al., 1997, publicación internacional WO9727207), un promotor de ubi10 de longitud completa de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493), 299 pb de la secuencia que codifica el gen de PAT 5' parcial (Wohlleben et al., 1988, *Gene* 70:25-37) y la secuencia homóloga 2 (intrón 1 de 4-CoAS de *A. thaliana*) (Locus At3g21320, GenBank NC 003074). Tanto la secuencia homóloga 1 como la secuencia 2 en el vector donador eran idénticas a la secuencia homóloga 1 y secuencia 2 correspondientes en el vector diana (pDAB1585).

La construcción del vector donador, la secuencia que codifica PAT parcial 5', se fusionó con el intrón 1 de 4-CoAS de *A. thaliana* de longitud completa (Locus At3g21320, GenBank NC 003074) por síntesis de ADN por Picoscript Ltd., LLP, (Houston, Texas). Se añadieron los sitios NcoI y XhoI en el extremo 5' y 3' del fragmento, respectivamente. Este fragmento de ADN sintetizado después se digirió con NcoI/XhoI y se insertó en pDAB1575 en los mismos sitios para sustituir la secuencia que codifica el gen de PAT de longitud completa y su 3' UTR. La construcción resultante se denominó pDAB1576 (figura 25).

Después se digirió pDAB1576 con AgeI y el fragmento entero que contenía el casete de expresión de PAT parcial 5' flanqueado por la secuencia homóloga 1 y la secuencia homóloga 2, se insertó en pDAB2407, el vector base binario, en el mismo sitio. La construcción resultante se denominó pDAB1600 y era la versión binaria del vector donador para la retransformación de las células de la planta (figura 26).

Ejemplo 5 - Diseño y generación del vector de expresión de la nucleasa con dedos de zinc

El gen de la proteína de fusión de dedos de zinc-FokI se dirigió mediante un promotor de CsVMV y 5' UTR (Verdaguer et al., 1996, *Plant Molecular Biology* 31:1129-1139). También estaban incluidos en el casete las 5' y 3' UTR de la osmotina de *N. tabacum* (Merlo et al., documento US2005102713). Para hacer estos vectores, el fragmento de HindIII/SacI que comprendía el promotor de CsVMV y la 5' UTR que dirige PAT en pDAB7002 (figura 27) se sustituyó por un fragmento que comprendía el promotor de CsVMV y la 5' UTR y la 5' UTR de *N. tabacum* que dirige GUS, que se cortó de pDAB7025 (figura 28) con HindIII/SacI. El plásmido resultante se denominó pDAB1591 (figura 29).

Las secuencias codificantes de IL1-L0-FokI y Scd27-L0-FokI se amplificaron por PCR a partir de sus vectores originales, pCDNA3.1-IL1-L0-FokI (figura 30) y pCDNA3.1-SCD27a-L0-FokI (figura 31) usando la pareja de cebadores P11/P12 y P13/P14, respectivamente. Se añadieron los sitios BbsI y SacI en el extremo 5' y 3' de los fragmentos de PCR, respectivamente. El gen de PAT en pDAB1591 se sustituyó por el fragmento de la PCR del gen de la proteína de fusión con dedos de zinc por clonación por SacI/NcoI. Los plásmidos resultantes se denominaron pDAB1592 (figura 32) y pDAB1594 (figura 33) para IL-1-FokI y Scd27-FokI, respectivamente.

Las versiones binarias de estos vectores se construyeron cortando los casetes de expresión de la proteína de fusión con dedos de zinc de pDAB1592 y pDAB1594 como fragmentos PmeI/XhoI, llenando los extremos y clonando en pDAB2407 en el sitio de PmeI (figura 34A). Los plásmidos resultantes se denominaron pDAB1596 (figura 34B) y pDAB1598 (figura 34C) para IL-1 ZFN y Scd27 ZFN, respectivamente y eran la versión binaria de los vectores del gen de la proteína de fusión con dedos de zinc de la retransformación de la célula de planta.

Ejemplo 6 - Diseño y generación del vector de control positivo

Para calcular la frecuencia de recombinación ilegítima y que sirva como un control positivo, se usó un vector que contenía el casete de expresión del gen de PAT. Con el fin de ser comparable con los recombinantes finales, se insertó el intrón 1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) en las 299/300 pb de la secuencia que codifica PAT (Wohlleben et al., 1988, *Gene* 70:25-37). Para hacer esta construcción, se ligó el fragmento Swal/Clal de 2559 pb de pDAB1576 con el fragmento de la cadena principal de pDAB1577 (La figura 35) que se digirió con las mismas enzimas de restricción. El vector resultante contenía el casete de expresión del gen de PAT con las 1743 pb de la inserción del intrón 1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) (Locus At3g21320, GenBank NC 003074) en medio de la secuencia codificante de PAT (Wohlleben et al., 1988, *Gene* 70:25-37). Este vector se denominó pDAB1578 (figura 36).

Para hacer la versión binaria de pDAB1578, el casete de expresión del gen de PAT con el intrón 1 de *A. thaliana* 1 (Locus At3g21320, GenBank NC 003074) se cortó de pDAB1578 con PmeI/XhoI. Después de tratar el extremo 3' del fragmento para generar extremo romo, se insertó en pDAB2407, el vector base binario, en el sitio PmeI. El vector resultante se denominó pDAB1601 (figura 37) comprendía el gen de PAT (Wohlleben et al., 1988, *Gene* 70:25-37) que contenía la secuencia del intrón 1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) dirigida por el promotor de ubi10 de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493) y terminado por el orf25/26 de *A. tumefaciens* 3' UTR (Gelvin et al., 1987, documento EP222493).

Ejemplo 7: Retransformación de los cultivos de células diana con genes de nucleasa con dedos de zinc y secuencias de ADN donadoras

Se seleccionaron 3 cultivos de células transgénicas, resistentes a la higromicina, independientes (NT1-240, NT1-260

y BY2-380), que contenía cada uno una sola copia integrada de longitud completa de la secuencia diana, y se usaron para reiniciar los cultivos en suspensión poniendo ~ 250-500 mg de tejido de callo en 40-50 ml de medio basal (basado en MS o LS para NT1 y BY2, respectivamente) que contenía higromicina 100 mg/l y se subcultivaron cada 7 días como se ha descrito antes. Antes de la retransformación, los cultivos de suspensiones se transfirieron a medio basal sin higromicina.

La transformación mediada por *Agrobacterium* de los cultivos de células diana se llevó a cabo como se ha descrito antes. Para cada experimento, se generaron 10 placas de cocultivo como sigue: una placa comprendía células cocultivadas con 100 µl de una cepa de *Agrobacterium* que albergaba pDAB1600 (ADN donador); una placa comprendía células cocultivadas con 100 µl de una cepa de *Agrobacterium* que albergaba pDAB1601 (marcador seleccionable PAT); 4 placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que albergaba pDAB1600 (ADN donador) y 250 µl de una cepa de *Agrobacterium* que albergaba pDAB1596 (IL-1 ZFP-FokI); y 4 placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que albergaba pDAB1600 (ADN donador) y 250 µl de una cepa de *Agrobacterium* que albergaba pDAB1598 (Scd 27a ZFP-FokI). Después de cocultivo usando los métodos descritos antes, las células se cultivaron en placa en medio basal (basado en MS o LS para NT1 y BY2, respectivamente) que contenía carbenicilina 500 mg/l y Bialaphos® 10 mg/l o 15 mg/l, respectivamente para NT1 o BY2. Aparecieron aislados individuales resistentes a Bialaphos® 2-4 semanas después de cultivo en placa (tabla 3) y se transfirieron a placas individuales de 60x20 mm (un aislado por placa) donde se mantuvieron como callo en un programa de subcultivo de 14 días hasta que fueron necesarios para el análisis.

Tabla 3. Resumen de la retransformación de los cultivos de células diana con genes de proteína de fusión de dedos de zinc-FokI y ADN donador

Cultivo celular diana	Tratamiento	nº de placas de selección	nº de sucesos transgénicos	nº medio de sucesos por placa de selección
NT1-240	pDAB1601 (marcador seleccionable PAT)	59	1.490	25,3
	pDAB1600 (solo ADN donador)	35	0	0
	pDAB 1600 + pDAB 1596 (ADN donador + ZFP-Fok1)	251	293	1,2
	pDAB 1600 + pDAB 1598 (ADN donador + Scd27a ZFP-Fok1)	251	247	1,0
NT1-260	pDAB1601 (marcador seleccionable PAT)	35	427	12,2
	pDAB1600 (solo ADN donador)	35	0	0
	pDAB 1600 + pDAB 1596 (ADN donador + IL-1 ZFP-Fok1)	251	35	0,1
	pDAB 1600 + pDAB 1598 (ADN donador + Scd27a ZFP-Fok1)	251	76	0,3
BY2-380	pDAB1601 (marcador seleccionable PAT)	46	536	11,7
	pDAB1600 (solo ADN donador)	46	0	0
	pDAB 1600 + pDAB 1596 (ADN donador + IL-1 ZFP-Fok1)	214	43	0,2
	pDAB 1600 + pDAB 1598 (ADN donador + Scd27a ZFP-Fok1)	214	47	0,2

Ejemplo 8 - Confirmación de la recombinación homóloga

A. Recombinación homóloga inter cromosómica

Se desarrollaron dos estrategias y se ensayaron para la recombinación inter cromosómica facilitada por la proteína de fusión de dedos de zinc-FokI en el sistema de tabaco de ejemplo descrito en los ejemplos 1 a 7.

En la estrategia 1, el sitio de unión para la proteína de fusión de dedos de zinc-FokI (IL-1-L0-FokI) se incluyó en medio de la construcción diana (figura 37). En esta estrategia, el sitio de unión estaba flanqueado por ~3 kb de secuencias no homólogas en ambos lados seguido de la secuencia homóloga 1 (MAR RB7 de *N. tabacum*) y la secuencia homóloga 2 (intrón 1 de 4-CoAS de *A. thaliana*) en la dirección 5' y dirección 3', respectivamente (figura 38). Se pensó que en presencia de la proteína de fusión de dedos de zinc-FokI IL-1, las secuencias de unión IL-1-L0-FokI serían reconocidas y se induciría una rotura bicatenaria en este sitio específico, lo cual estimularía el proceso de reparación de ADN endógeno. En presencia del ADN donador, que contenía secuencias homólogas idénticas a las de la secuencia diana, el gen de PAT parcial 5' junto con su promotor, sustituiría el fragmento de ADN de ~6 kb entero entre las secuencias homólogas en la diana mediante recombinación homóloga. Mediante este proceso, las dos secuencias del gen de PAT parciales, con el intrón 1 de 4-CoAS de *A. thaliana* interpuesto entre ellas, reconstituirían un gen de PAT funcional, dando como resultado la expresión de PAT y un fenotipo de

resistencia a herbicida.

En la estrategia 2, se incluyeron dos sitios de unión de dedos de zinc-FokI (Scd27-L0-FokI) en el vector diana: uno directamente en la dirección 3' de MAR RB7 de *N. tabacum* y el otro directamente en la dirección 5' del intrón 1 de 4-CoAS de *A. thaliana* (figura 39). Entre los dos sitios de unión de la proteína de fusión de dedos de zinc-FokI había ~6 kb de secuencia, que incluían el fragmento de GFP 5', un casete de expresión de GUS y el fragmento de GFP 3'. Se pensó que en presencia de la proteína de fusión de dedos de zinc-FokI Scd27, las dos secuencias de unión serían reconocidas y se inducirían roturas del ADN bicatenarias en ambas localizaciones, lo que eliminaría el fragmento de ADN de ~6 kb entre estas dos secuencias de unión, y estimularía el proceso de reparación de ADN endógeno. De forma similar a la estrategia 1, en presencia del ADN donador, que contenía secuencias homólogas idénticas a la de la secuencia diana, el gen de PAT parcial 5', sería insertado en la secuencia diana mediante recombinación homóloga en el sitio donde se habían inducido las roturas de ADN bicatenarias. Mediante este proceso, las dos secuencias del gen de PAT parciales, con el intrón 1 de 4-CoAS de *A. thaliana* interpuesto entre ellas, reconstituirían un gen de PAT funcional, dando como resultado la expresión de PAT y un fenotipo de resistencia a herbicida.

Todos los aislados obtenidos después de la selección con herbicida (Bialaphos®) se analizaron primero por PCR usando el par de cebadores P24/P25, que amplificaba un fragmento de ADN que abarcaba el gen de PAT reconstituido. El cebador P24 era homólogo al extremo 5' de la secuencia que codifica PAT en el ADN donador y el cebador P25 era homólogo al extremo 3' de la secuencia que codifica PAT en el ADN diana. Resultaría un fragmento de la PCR de 2,3 kb si las dos secuencias que codifican PAT se unieran por recombinación homóloga. Como se muestra en la figura 40, se obtuvo un producto de la PCR de 2,3 kb de muchos de los aislados analizados. Estos aislados se obtuvieron de la cotransformación tanto del gen de la proteína de fusión de dedos de zinc-FokI IL-1/ADN donador como del gen de la proteína de fusión de dedos de zinc-FokI Scd27/ADN donador. Los productos de la PCR de 2,3 kb de múltiples aislados independientes representativos de los derivados de las transformaciones tanto del gen de la proteína de fusión de dedos de zinc-FokI IL-1 como de dedos de zinc-FokI Scd27, se purificaron en geles de agarosa y se clonaron en el vector pCR2.1 TOPO vector (Invitrogen, Carlsbad, California). El producto de la PCR de 2,3 kb insertado en el vector TOPO después se secuenció usando el kit de secuenciación Dye Terminator Cycle Sequencing Kit (Beckman Coulter). Los resultados de la secuenciación confirmaron que todos los productos de la PCR clonados en el vector TOPO contenían la secuencia recombinada prevista, incluyendo las secuencias del gen de PAT parciales 5' y 3' con el intrón 1 de 4-CoAS de *A. thaliana* intermedio. Estos resultados confirmaron la recombinación intercromosómica prevista para ambas estrategias ensayadas y el direccionamiento génico ilustrado a través de la expresión del gen de la proteína de fusión de dedos de zinc-FokI.

Un par de muestra se analizaron más por PCR usando el par de cebadores P26/P25 que amplificó un fragmento de ADN a lo largo de toda la región recombinada entera. El cebador P26 era homólogo del extremo 3' de la región que codifica el gen de HPT en la secuencia diana y el cebador P25 era homólogo del extremo 3' de la región que codifica el gen de PAT en la secuencia diana. Se obtendría un producto de la PCR de 5,2 kb previsto si se produjera recombinación homóloga entre la secuencia diana y el ADN donador. La diana no recombinada daría un producto de la PCR de ~10 kb. Se obtuvo un producto de la PCR de 5,2 kb de ambas muestras analizadas. El producto de la PCR de 5,2 kb de una de las muestras analizadas se purificó en un gel de agarosa, se clonó en el vector pCR2.1 TOPO (Invitrogen, Carlsbad, California). El producto de la PCR después se secuenció usando el kit de secuenciación Dye Terminator Cycle Sequencing Kit (Beckman Coulter, Fullerton, CA). Los resultados de la secuenciación confirmaron que el producto de la PCR contenía la secuencia recombinada, incluyendo (de 5' a 3'): el extremo 3' de la región que codifica HPT (desde la secuencia diana), el orf-24 de *A. thaliana* 3' UTR (desde la secuencia diana), MAR RB7 de *N. tabacum* (secuencia homóloga 1), promotor de ubi-10 de *A. thaliana* (desde ADN donador), gen de PAT parcial 5' (desde ADN donador), intrón 1 de 4-CoAS de *A. thaliana* (secuencia homóloga 2), gen de PAT parcial 3' (desde la secuencia diana). Este resultado confirmaba que los productos de la PCR resultaban de la recombinación homóloga intercromosómica.

Para analizar más los productos recombinantes a nivel genómico, se llevó a cabo el análisis de transferencia Southern con un conjunto de 22 aislados seleccionados de las transformaciones tanto del gen de la proteína de fusión de dedos de zinc-FokI IL-1/ADN donador como del gen de la proteína de fusión de dedos de zinc-FokI Scd27/ADN donador. Se confirmó que todas las muestras eran recombinantes por análisis por PCR. Se digirieron 10 µg de ADN genómico de cada muestra con BanII. El ADN genómico digerido se separó en un gel de agarosa al 0,8% y se transfirió a una membrana de nailon. Después de entrecruzamiento sobre la membrana, el ADN se hibridó con una sonda de PAT 3'. Los productos recombinantes esperados darían una banda de 2079 pb mientras que la secuencia diana no recombinada daría una banda de 3018 pb. Los resultados mostraron que de las 22 muestras analizadas, 18 muestras tenían una banda con el tamaño esperado de ~2 kb, que indicaba que estos sucesos derivaban de la recombinación homóloga en la proximidad de los dos fragmentos de PAT parciales (figura 41). Dos muestras presentaban una banda de tamaño no esperado (una mayor y otra menor que el control de la diana) lo que sugería que la recombinación en estas dos muestras no era totalmente dependiente de homología o que se producían reordenaciones de secuencia adicionales durante o después de recombinación. Otras 2 muestras presentaban una banda idéntica al control de la diana (~3 kb), indicando que estas muestras eran escapes de la selección con herbicida o una población mixta de células siendo solo una proporción pequeña derivada de la recombinación homóloga por debajo del nivel de detección para el análisis de transferencia Southern. El último caso es el más probable puesto que estas muestras presentaban el producto de amplificación esperado, correspondiente a la recombinación homóloga cuando se analizaron por PCR. De las 18 muestras que presentaban la banda de

recombinación esperada, 4 muestras también tenían una banda adicional con el tamaño idéntico al control de la diana, indicando que estas muestras representaban una población mixta de células con algunas células no recombinadas. Otras 4 muestras tenían bandas adicionales de diferentes tamaños, indicando que estas muestras eran quimeras genéticas que comprendían algunas células que habían experimentado sucesos no dependientes de homología. En total, 10 de las 22 muestras se confirmaron por análisis de transferencia Southern que habían experimentado recombinación homóloga como se había previsto, al menos en la región que implicaba los dos fragmentos de PAT parciales.

B. Recombinación homóloga intracromosómica

Para ensayar la recombinación homóloga intracromosómica facilitada por dedos de zinc-FokI, se incluyeron dos fragmentos de GFP no funcionales con secuencias que se solapaban 540 pb, en el vector diana como se describe en la figura 42. Entre estos dos fragmentos había un casete de expresión del gen de GUS. La secuencia de unión de la proteína de fusión IL-1-FokI se fusionó con la secuencia que codifica GUS en su extremo N-terminal. Se pensó que en presencia de la proteína de fusión IL-1-FokI, las secuencias de unión de ZFN IL-1 serían reconocidas y se induciría una rotura de ADN bicatenaria, lo que estimularía el proceso de reparación de ADN endógeno. Sin la presencia de ADN donador, los dos fragmentos de GFP parcialmente homólogos experimentarían un proceso de recombinación homóloga intracromosómica y se reconstituiría un gen de GFP funcional.

Las dos líneas diana, BY2-380 y NT1-260 se transformaron con el plásmido pDAB1596 (el vector binario del gen de la proteína de fusión IL-1-FokI) mediante transformación mediada por *Agrobacterium* como se ha descrito antes. Tanto el ADN donador (pDAB1600) como el ADN de control de PAT (pDAB1601) se incluyeron como tratamientos de control separados. Las células se cultivaron en placa en medio no selectivo después de la transformación. La expresión apreciable del gen GFP funcional constituido produjo fluorescencia visible aproximadamente a los 5-8 días de transformación. Como se resume en la tabla 4, se observaron ~50 locus fluorescentes en cada placa transformada con la construcción del gen de la proteína de fusión IL-1-FokI (pDAB1596). No se observaron diferencias significativas entre las dos líneas diana ensayadas. No se observó fluorescencia apreciable más allá de una ligera de fondo, en los controles negativos transformados con las construcciones de ADN donador o gen de PAT.

Tabla 4. Constitución de GFP funcional mediante recombinación homóloga intracromosómica estimulada por la proteína de fusión con dedos de zinc-FokI IL-1

Línea diana	Tratamiento de transformación	nº de locus fluorescentes por placa
BY2-380	IL-1-FokI (pDAB1596)	51,4
	Gen de PAT (pDAB1601)	2,0
	ADN donador solo (pDAB1600)	0
NT1-260	IL-1-FokI (pDAB1596)	53,0
	Gen de PAT (pDAB1601)	0
	Solo ADN donador (pDAB1600)	0

Para confirmar que los locus de proteína fluorescentes verdes resultaban de la reconstitución de un gen de la GFP funcional, se llevó a cabo el análisis molecular del tejido que expresa GFP. Puesto que todas las células se cultivaron en placa en medio no selectivo, era difícil obtener una población de células que fuera homogénea con respecto a la expresión de la GFP. Se aislaron una serie de segmentos de tejidos fluorescentes de la placa (con ayuda de un microscopio de disección) y se enriquecieron mediante varios pasos de subcultivo selectivo. Se aisló ADN genómico de estos tejidos enriquecidos visualmente y se ensayó por PCR usando el par de cebadores P27/P28. El cebador P27 era homólogo del extremo 5' del fragmento de GFP parcial 5' en la secuencia del ADN diana y el cebador P28 era homólogo del extremo 3' del fragmento de GFP parcial 3' en la secuencia del ADN diana. Se obtendría un producto de la PCR de 0,6 kb previsto si el gen de GFP se hubiera reconstituido por recombinación homóloga intracromosómica entre estos dos fragmentos de GFP parciales. La diana no recombinada daría un producto de la PCR de 4,1 kb. Como se muestra en la figura 43, todas las muestras que se habían enriquecido a partir de los tejidos fluorescentes tenían el producto de la PCR de 0,6 kb previsto, lo que indicaba que se había reconstituido un gen de GFP funcional en estos tejidos. También se observó un segundo producto de la PCR de 4,1 kb en todas estas muestras enriquecidas, lo que indicaba la presencia de la población de células no recombinadas. Esto no se esperaba, puesto que estas muestras solo se enriquecieron por selección visual usando la fluorescencia como indicador. Los productos de la PCR se separaron en un gel de agarosa al 0,8%, se transfirieron a una membrana de nailon y se hibridaron con sonda con la secuencia que codifica el gen de la GFP. Los resultados indicaban que los dos productos de la PCR, de 0,6 kb y 4,1 kb, contenían la secuencia de GFP, confirmando de esta forma que la fluorescencia resultaba de la expresión de un gen de GFP reconstituido.

Ejemplo 9. Diseño de la proteína de fusión de dedos de zinc-FokI dirigida contra un gen en una monocotiledónea

Las nucleasas con dedos de zinc para facilitar la reparación dirigida por homología en monocotiledóneas (p. ej., maíz, sorgo, trigo, cebada, arroz) se diseñan y sintetizan como sigue. Se selecciona un gen de interés, cuyo gen preferiblemente incluye al menos un codón que puede ser dirigido por una sustitución de aminoácidos. La parte

relevante del gen seleccionado se clona y se determina la secuencia de nucleótidos del clon.

La secuencia así obtenida se examina, opcionalmente usando un programa de ordenador que contiene una lista de dedos de zinc individuales y sus sitios diana y/o una lista de módulos de dos dedos y sus sitios diana, para encontrar un par de secuencias diana, separadas por 5-6 pares de nucleótidos, en donde cada secuencia diana puede estar unida por una proteína con dedos de zinc con 3, 4, 5 o 6 dedos. Véase, por ejemplo, la patente de EE.UU. nº 6.785.613; WO 98/53057; WO 01/53480 y publicación de solicitud de patente de EE.UU. nº 2003/0092000. Se describen métodos adicionales para diseñar ZFP, por ejemplo, en las patentes de EE.UU. nº 5.789.538; 6.013.453; 6.410.248; 6.733.970; 6.746.838; 6.785.613; 6.866.997; 7.030.215; WO 01/088197; WO 02/099084; y publicaciones de solicitudes de patentes de EE.UU. 2003/0044957; 2003/0108880; 2003/0134318 y 2004/0128717.

Para cada secuencia diana identificada en la etapa previa, se sintetiza un gen que codifica una fusión entre un semidominio de escisión de FokI y una proteína con dedos de zinc que se une a la secuencia diana. Véase, por ejemplo, la patente de EE.UU. nº 5.436.150; WO 2005/084190 y publicación de solicitud de patente de EE.UU. nº 2005/0064474. Después se ensaya en cada proteína de fusión la afinidad con la que se une a su secuencia diana, usando un ensayo ELISA, como describen, por ejemplo, Bartsevich et al. (2003) *Stem Cells* 21:632-637. Las proteínas que tienen afinidades de unión con la secuencia diana que superan un valor umbral predeterminado se someten a posterior ensayo en un ensayo de indicador basado en célula.

Opcionalmente, la especificidad de una o más proteínas de fusión descritas antes, se puede evaluar y, si es necesario, mejorar usando métodos descritos en la patente de EE.UU. nº 6.794.136.

El ensayo basado en células se describe, por ejemplo, en Umov et al. (2005) *Nature* 435:646-651 y la publicación de solicitud de patente de EE.UU. nº 2005/0064474. Brevemente, un par de secuencias diana, identificadas como antes, se inserta en un gen de proteína verde fluorescente (GFP) cromosómico defectuoso, bajo el control de transcripción de un promotor inducible por doxociclina, en una línea celular adecuada. Las células se transfectan con ácidos nucleicos que codifican dos proteínas de fusión de dedos de zinc/FokI (cada una de las cuales se une a una de las secuencias diana) y con un ácido nucleico que contiene las secuencias que, si sirven como molde para la reparación dirigida por homología del gen de GFP cromosómico defectuoso, reconstituirán un gen de GFP funcional. Las células en las que se ha producido la reparación dirigida por homología se pueden identificar y cuantificar por separación de células activadas por fluorescencia, seguido de inducción con doxociclina.

Ejemplo 10: Diseño y generación del vector de ADN donador para monocotiledóneas

La construcción del ADN donador incluye la secuencia codificante (CDS) para el gen seleccionado y la secuencia genómica en la dirección 5' y/o dirección 3' de la CDS. La CDS y/o las secuencias genómicas se obtienen de la base de datos del Centro nacional para la información biotecnológica (NCBI) y/o la base de datos de genoma de plantas. Se pueden usar correspondencias de "cántigos" entre secuencias parciales y secuencias conocidas para confirmar que las secuencias derivan del mismo locus en el genoma de monocotiledónea seleccionado. Para evitar que la secuencia de unión de la nucleasa con dedos de zinc se escinda de forma repetida después de recombinación, también se pueden hacer dos mutaciones de un solo nucleótido, que no producen un cambio en la secuencia de aminoácidos codificada, dentro de la secuencia de unión de la nucleasa con dedos de zinc (ZFN) en el ADN donador. Una o ambas mutaciones pueden crear un sitio de enzimas de restricción, que facilita la caracterización molecular en la dirección 3'.

La construcción del vector de ADN donador, un fragmento de ADN, que cubre algo de la CDS y secuencia en la dirección 3', se amplifica por PCR a partir del ADN genómico de la monocotiledónea seleccionada usando cebadores adecuados. Los sitios de restricción (p. ej., sitios de SacI y BamHI) preferiblemente se añaden en el extremo 5' y 3' de los fragmentos de la PCR, respectivamente. El producto de la PCR se clona en el vector TOPO pCR Blunt II (Invitrogen, Carlsbad, California).

Se puede introducir una mutación en la localización deseada sustituyendo uno o más nucleótidos seleccionados usando el kit de mutagénesis dirigida al sitio QuickChange® (Stratagene, La Jolla, California). Como se ha indicado antes, también se pueden introducir mutaciones silenciosas individuales (p. ej., 2, 3 o más mutaciones individuales) en la secuencia de unión de dedos de zinc-FokI de una forma similar usando el kit de mutagénesis dirigida al sitio QuickChange® (Stratagene, La Jolla, California). El fragmento de la PCR con la mutación de sustitución y mutaciones individuales silenciosas adicionales se aísla del vector pCR Blunt II TOPO mediante digestión adecuada y se clona en pSB11 (figura 44) en los mismos sitios. El ADN plasmídico resultante después se transforma en una cepa de *Agrobacterium* que alberga el plásmido pSB1 (figura 45) para formar el vector donador superbinario para la transformación de la célula de la planta mediante recombinación homóloga.

Ejemplo 11: Diseño y generación de un vector del gen de la proteína de fusión de dedos de zinc-FokI de monocotiledónea

El gen de la proteína de fusión de dedos de zinc-FokI es dirigido por un promotor de monocotiledónea adecuado (p. ej., promotor de ubi1 de *Z. mays* (Quail et al., documento US5614399)). El casete también puede incluir la peroxidasa-5 de *Z. mays* (per5) 3' UTR (Ainley et al., documento WO9856921). Para hacer este vector, el gen de la proteína de fusión de dedos de zinc-FokI se aísla de su vector fuente original por digestión con NcoI/SacI y se

clona en pDAB3872 (figura 46) en los mismos sitios. El casete de expresión que incluye el promotor de ubi-1 de *Z. mays*, el gen de la proteína de fusión de dedos de zinc-FokI y per-5 de *Z. mays* 3' UTR, se aisló del vector intermedio hecho antes por digestión con HindIII/MscI y se insertó en pSB11 en el sitio HindIII/PmeI. El plásmido resultante después se transforma en una cepa de *Agrobacterium* que alberga el plásmido pSB1 para formar el vector superbinario, pDAB4365 mediante recombinación homóloga (figura 47). pDAB4365 es la versión superbinaria del vector de expresión del gen de la proteína de fusión de dedos de zinc-FokI para la transformación de la célula de la planta.

Ejemplo 12. Transformación de células de maíz con genes de nucleasa con dedos de zinc y ADN donador y generación de cultivos celulares

La semilla del cruce F₁ de "High II" (Armstrong, et al., 1991, *Maize Genet. Coop. News Lett.* 65:92-93) se puede plantar directamente en macetas de 19 litros (5 galones) que contienen mezcla de suelo comercial (Conrad Fafard, Inc., Springfield; mezcla de suelo n° 3). Las plantas se cultivan en el invernadero con un fotoperiodo de 16 h complementado por una combinación de lámparas de sodio de alta presión y de haluro metálico con 15.000 lux (1.500 ft-candle) de iluminación. La temperatura de día y de noche se mantiene a 27/20 ± 2°C. Las plantas se riegan según sea necesario con la mezcla de tanque de fertilización convencional.

Para obtener embriones inmaduros, se pueden llevar a cabo polinizaciones controladas (hermanos o autopolinizaciones). Las plantas en antesis se preparan para la polinización recortando las barbas un día antes de la polinización, produciendo así un cepillo de barbas completo para la fertilización máxima y producción de semilla. Un día antes de la polinización, los flecos que se sueltan activamente se ponen en una bolsa, y se recoge el polen reciente y se aplica con cuidado en las barbas. Cuando los embriones en desarrollo alcanzan 1,0-1,2 mm de tamaño (9-10 días después de la polinización), la mazorca se puede cortar y esterilizar la superficie. Brevemente, las mazorcas se someten a inmersiones en etanol al 70% durante 2-5 minutos, y lejía comercial al 20% (hipoclorito sódico al 0,1%) durante 30-45 minutos seguido de 3 aclarados en agua destilada estéril. Después de esterilización, se pueden aislar los embriones maduros.

Se usan dos cepas diferentes de *Agrobacterium* para la transformación. La primera alberga el gen de la nucleasa con dedos de zinc descrita en los ejemplos 9-11, y la segunda alberga la secuencia de ADN donadora que comprende la mutación de sustitución. Se puede usar el sistema de vector "superbinario" de Japan Tobacco descrito en la patente de EE.UU. n° 5.591.616.

Para preparar las suspensiones de *Agrobacterium*, 1-2 bucles de bacterias de una placa rayada (que contiene extracto de levadura 5 g/l, Bacto-peptona 10 g/l, cloruro sódico 5 g/l, espectinomicina 50 mg/l, tetraciclina 10 mg/l y agar Bacto 15 g/l) en 5 ml de medio de "infiltración". El medio de infiltración consiste en sales basales de LS (Linsmaier et al., 1965, *Physiol. Plant.*), vitaminas N6 (Chu et al., 1975, *Sci. Sinica* 18:659-668), 2,4-D 1,5 mg/l, sacarosa 68,5 g/l, glucosa 36 g/l, prolina 6 mM ajustado a pH 5,2 antes de la esterilización por filtración. La mezcla se mezcla en mezcladora vorticial hasta que se logra una suspensión uniforme. La concentración bacteriana se puede determinar usando un colorímetro fotoeléctrico Klett-Summerson mediante lectura de la densidad de la solución. La solución se ajusta a una concentración de Klett 200 (~1x10⁹ ufc/ml) y se añade acetosiringona para alcanzar una concentración final 100 µM.

Los embriones inmaduros se aíslan directamente en un tubo de microfuga que contiene 2 ml de medio de "infiltración". Cada tubo que contiene ~100 embriones, se mezcla en mezcladora vorticial durante 3-5 s. Se separa el medio y se sustituye por medio líquido de nueva aportación de la misma composición y se repite la mezcla vorticial. El medio líquido se vuelve a separar y se sustituye por 1 ml de solución de *Agrobacterium* (800 µl de cepa con nucleasa con dedos de zinc y 200 µl de la cepa con ADN donador) a la concentración Klett 200. La mezcla de *Agrobacterium* y embriones se mezcla en mezcladora vorticial durante 30 s. Después de 5 min de incubación a temperatura ambiente, los embriones se pueden transferir y poner el eje del embrión hacia abajo en medio de "cocultivo" durante 5 días en la oscuridad a 25°C. El medio de "cocultivo" consistía en sales basales de LS (Linsmaier et al., 1965, *Physiol. Plant.*), vitaminas N6 (Chu et al., 1975, *Sci. Sinica* 18:659-668), 2,4-D 1,5 mg/l, sacarosa 30 g/l, prolina 6 mM, nitrato de plata 0,85 mg/l, acetosiringona 100 µM, GELRITE® 3 g/l ajustado a pH 5,8 antes de la esterilización por filtración.

Ejemplo 13: Selección de recombinantes homólogos

Después de cocultivo con una mezcla 5:1 de dos cepas de *Agrobacterium* que albergaban el casete del gen de la nucleasa con dedos de zinc y el ADN donador, respectivamente, los embriones se movieron al medio de "callo" que puede incluir componentes que detienen el crecimiento adicional de *Agrobacterium* (p. ej., Cefotaxima 250 mg/l y/o Pursuit® 500 nM). El medio de "callo" consistía en sales basales de LS (Linsmaier et al., 1965, *Physiol. Plant.*), vitaminas N6 (Chu et al., 1975, *Sci. Sinica* 18:659-668), 2,4-D 1,5 mg/l, MES 0,5 g/l, sacarosa 30 g/l, prolina 6 mM, nitrato de plata 1 mg/l, agar TC 8 g/l (PhytoTechnology Laboratories, Shawnee Mission, KS) ajustado a pH 5,8 antes de tratamiento en autoclave. A lo largo de la fase de selección, los embriones se cultivan en la oscuridad a 28°C.

Para la regeneración de la planta, los cultivos de callo se transfieren a medio de "inducción" y se incuban a 27°C con un fotoperiodo de luz/oscuridad de 16/8 en luz baja (13 µE/m²/s) durante 1 semana, seguido de 1 semana con luz

- alta ($40 \mu\text{E}/\text{m}^2/\text{s}$) proporcionada mediante lámparas fluorescentes blanco frío. El medio de "inducción" está compuesto de sales basales de MS y vitaminas (Murashige et al., 1962, *Physiol. Plant.* 15:473-497), sacarosa 30 g/l, 6-bencilaminopurina 5 mg/l, 2,4-D 0,025 mg/l, GELRITE[®] 2,5 g/l con el pH ajustado a 5,7 antes de tratamiento en autoclave. Después de este periodo de inducción de 2 semanas, el callo se transfiere a medio de "regeneración" y se incuba con luz alta ($40 \mu\text{E}/\text{m}^2/\text{s}$) a 27°C. El medio de "regeneración" es idéntico al medio de "inducción" excepto que carece de hormonas. El callo se puede subcultivar con medio de "regeneración" de nueva aportación cada 2 semanas hasta que aparecen brotes.
- 5
- Cuando las plántulas alcanzan aproximadamente 3-5 cm de longitud, se transfieren a tubos de cultivo de 150x25 mm que contienen sales basales de SH y vitaminas (Schenk et al., 1972, *Can. J. Bot.* 50:199-204), sacarosa 10 g/l, mio-inositol 1 g/l y GELRITE[®] 2,5 g/l con el pH ajustado 5,8 antes de tratamiento en autoclave. Una vez que los brotes alcanzan la parte superior del tubo, las plántulas se transfieren a macetas de 10 cm que contienen aproximadamente 0,25 kg de mezcla de suelo comercial (Conrad Fafard, Inc., Springfield; mezcla de suelo n° 3), se humedecen bien, y se cubren con copas de plástico transparente durante 2-4 días. En la etapa de 3-5 hojas, las plantas se trasplantan a macetas de 19 litros (5 galones) y se cultivan hasta madurez.
- 10
- Se puede encontrar información adicional relacionada con la escisión dirigida, recombinación dirigida e integración dirigida en las publicaciones de solicitudes de patentes de Estados Unidos US-2003-0232410; US-2005-0026157; US-2005-0064474; US-2005-0208489 y US-2007-0134796.
- 15

REIVINDICACIONES

1. Un método para introducir una secuencia exógena en el genoma de una célula de planta, comprendiendo el método las etapas de:

5 (a) poner en contacto la célula con un vector de ADN donador que comprende una primera, segunda y quinta secuencias de ADN, y la quinta secuencia de ADN comprende la secuencia exógena del vector de ADN que se va a introducir y está interpuesta entre la primera y segunda secuencias del vector de ADN donador;

10 en donde la primera secuencia tiene al menos de 90% a 95% de homología con una tercera secuencia en el genoma de la célula de planta, la segunda secuencia tiene al menos de 90% a 95% de homología con una cuarta secuencia en el genoma de la célula de planta, en donde la tercera y cuarta secuencias en el genoma de la célula de planta son secuencias de ADN cromosómico y los bordes cercanos de la tercera y cuarta secuencias están separados por al menos un par de nucleótidos; y

15 (b) expresar una o más nucleasas en la célula, en donde la una o más nucleasas son cada una un par de proteínas de fusión, en donde cada proteína de fusión es una fusión entre el dominio de escisión de una endonucleasa de restricción FokI y un dominio de unión con dedos de zinc genéticamente modificado, y en donde la una o más nucleasas escinden el ADN cromosómico en una o más secuencias diana correspondientes entre la tercera y cuarta secuencias, en donde la una o más nucleasas escinden el ADN cromosómico en una localización de 0,4 a 3 kilopares de bases de la tercera o cuarta secuencias en el genoma de la célula de planta;

20 de modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación de la quinta secuencia de ADN en el genoma por recombinación homóloga.

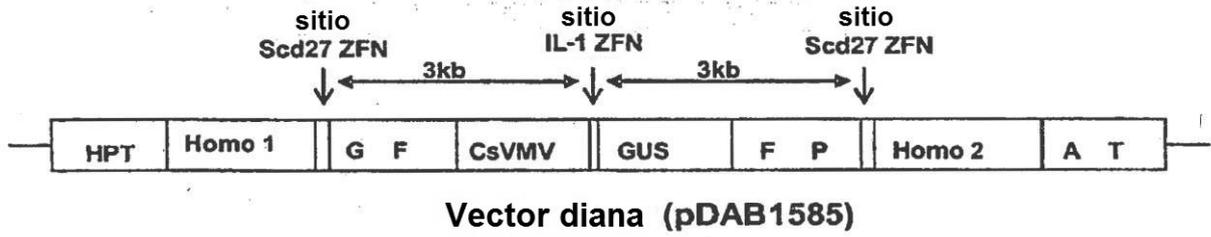
2. El método de la reivindicación 1, en donde la secuencia exógena es un marcador seleccionable.

25 3. El método de la reivindicación 2, en donde el marcador seleccionable se selecciona del grupo que consiste en la proteína verde fluorescente (GFP), β -glucuronidasa (GUS), fosfinotricina N-acetil-transferasa (PAT, BAR), neomicina fosfotransferasa, β -lactamasa, catecol dioxigenasa, α -amilasa, tirosinasa, β -galactosidasa, luciferasa, aequorina, EPSP sintasa, nitrilasa, acetolactato sintasa (ALS), dihidrofolato reductasa (DHFR), dalapon deshalogenasa y antranilato sintasa.

4. El método de la reivindicación 1, en donde la quinta secuencia comprende una o más secuencias reguladoras de la transcripción.

30 5. El método de la reivindicación 1, en donde la quinta secuencia comprende secuencias que codifican una proteína o una parte de una proteína.

6. El método de la reivindicación 1, en donde la quinta secuencia comprende un equivalente natural de una secuencia cromosómica mutante, o un equivalente mutante de una secuencia cromosómica natural.



Vector diana (pDAB1585)

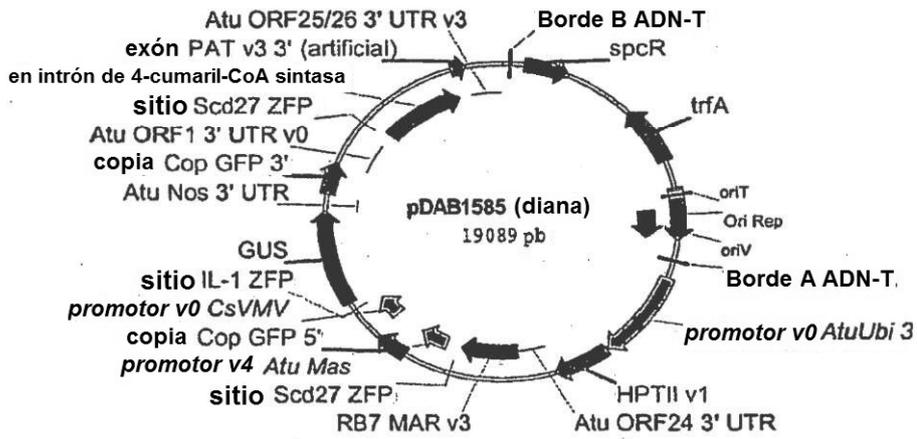


Figura 1

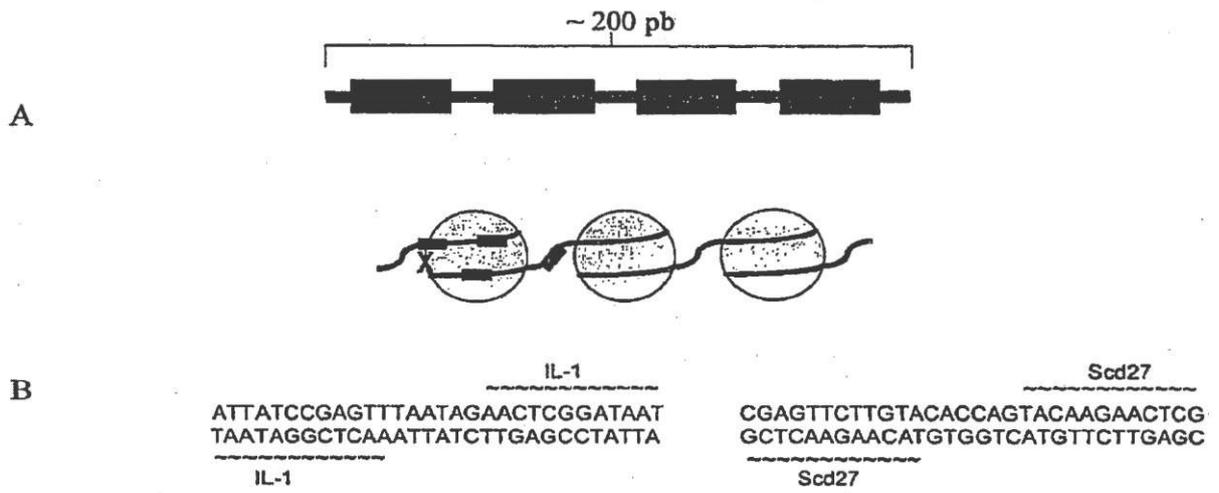


Figura 2

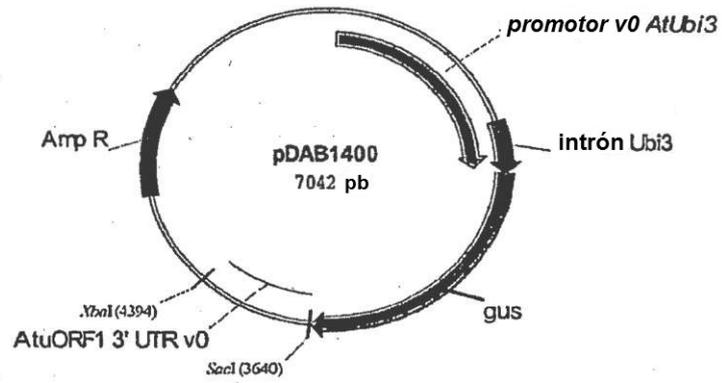


Figura 3

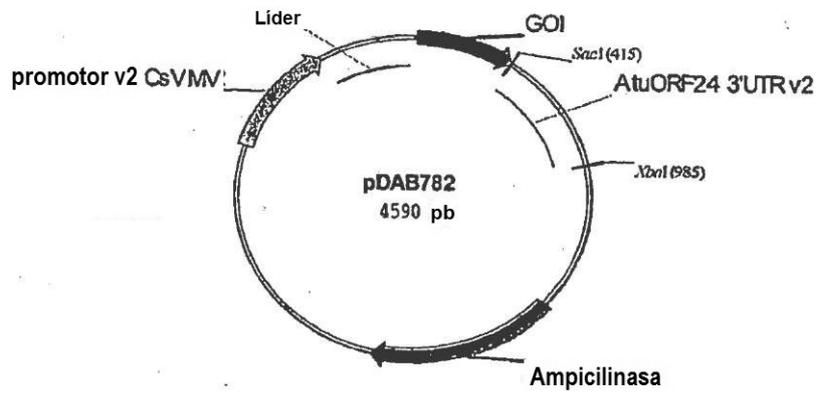


Figura 4

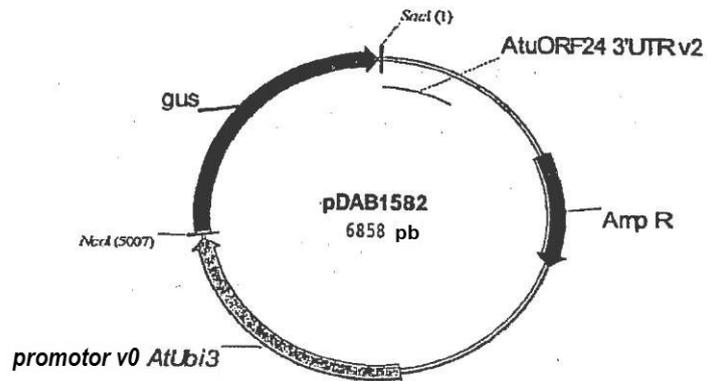


Figura 5

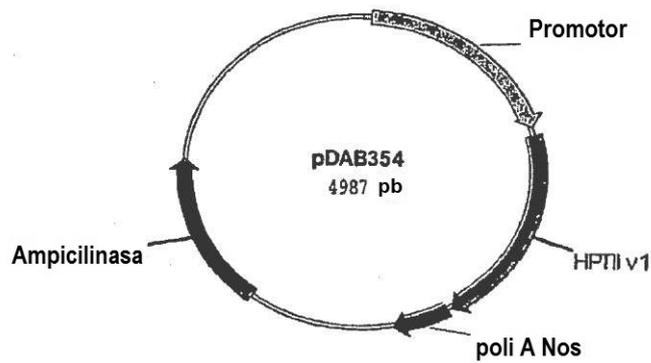


Figura 6

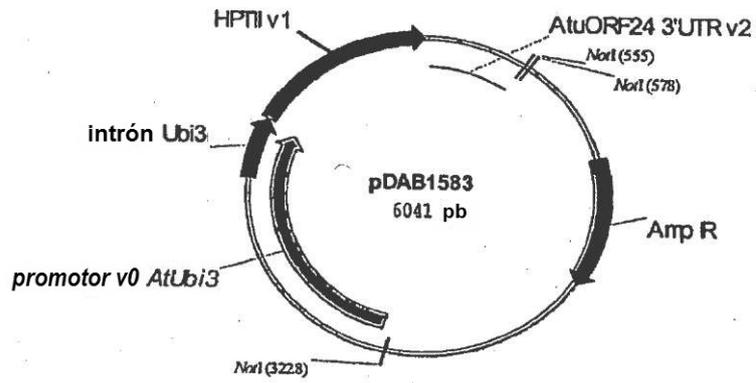


Figura 7

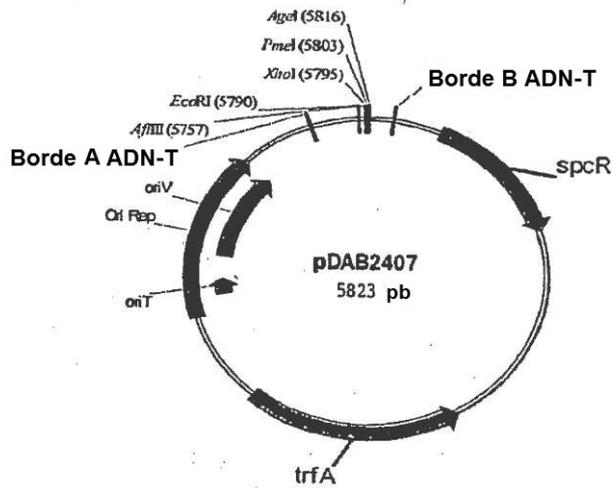


Figura 8

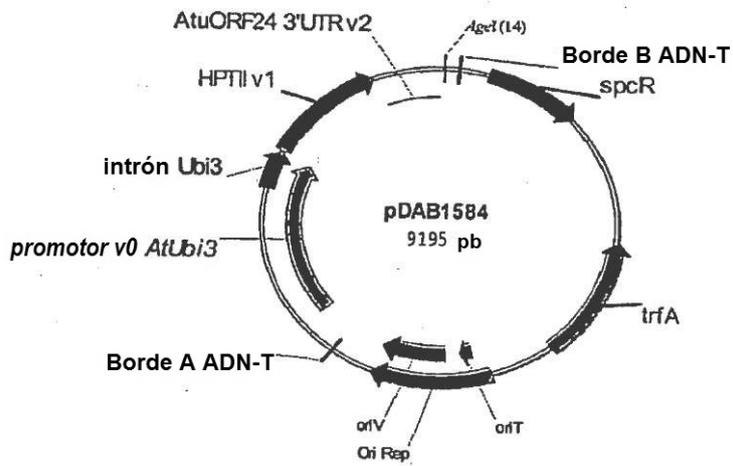


Figura 9

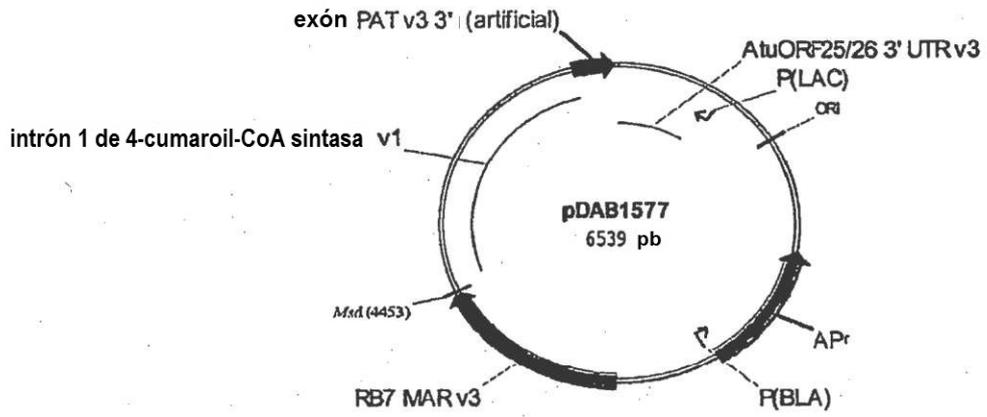


Figura 13

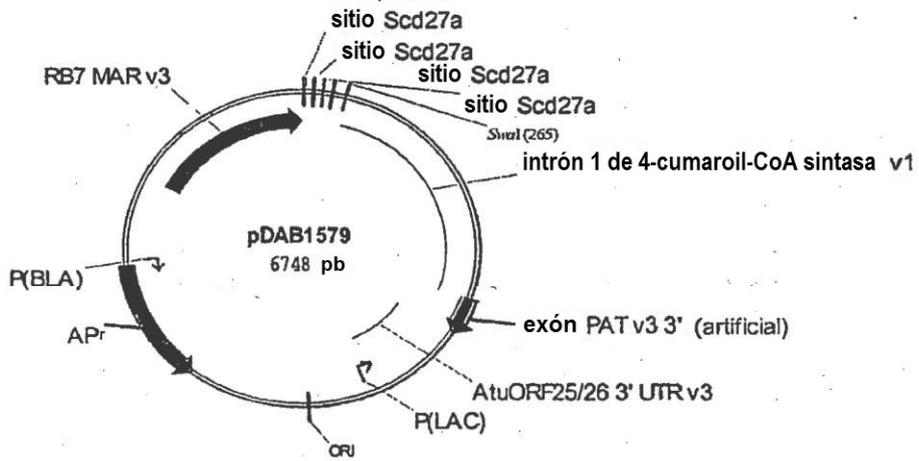


Figura 14

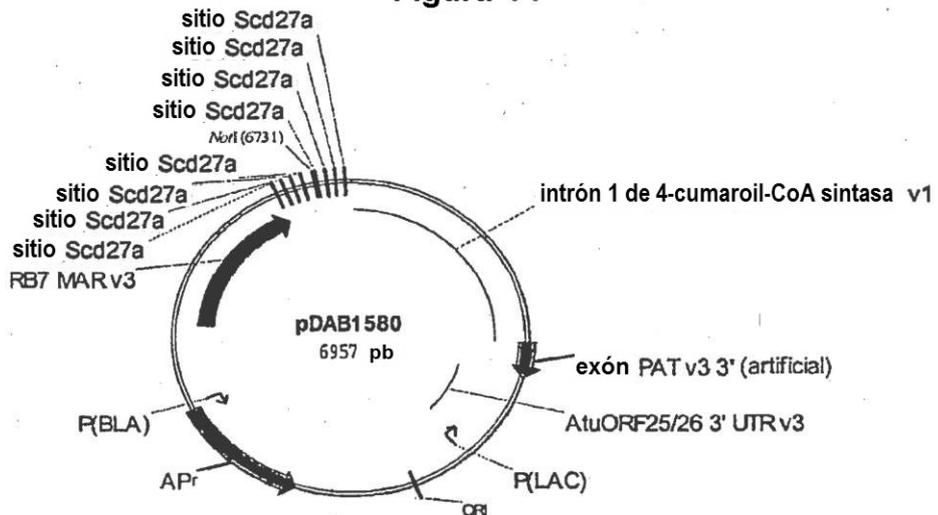


Figura 15

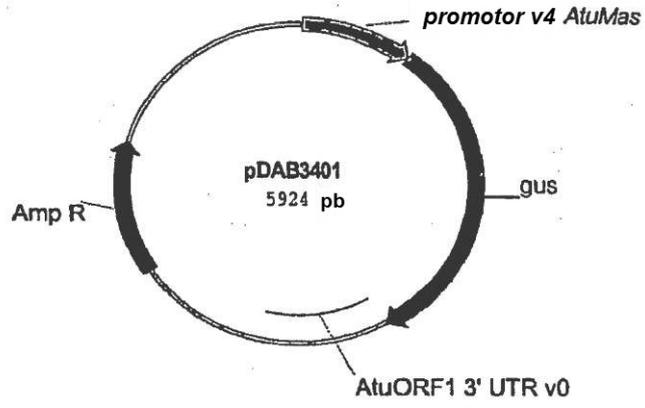


Figura 16

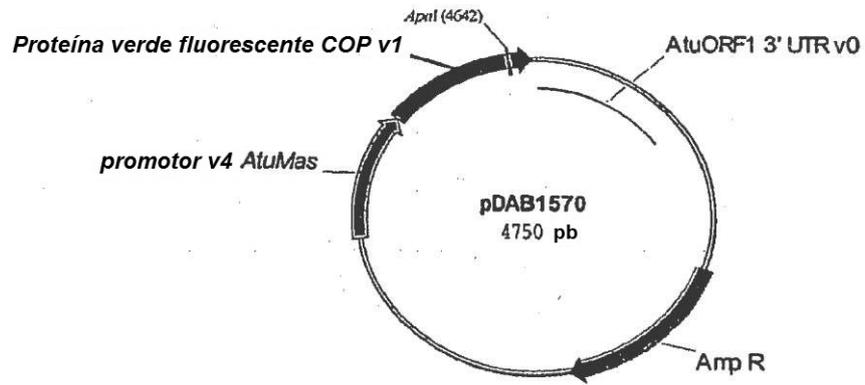


Figura 17

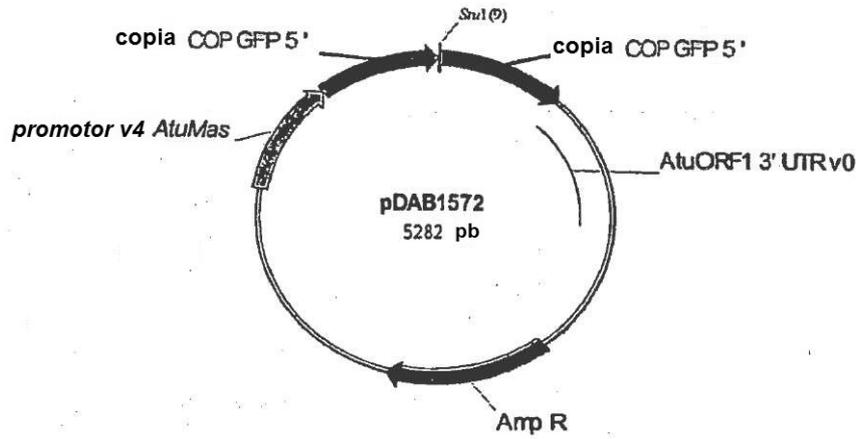


Figura 18

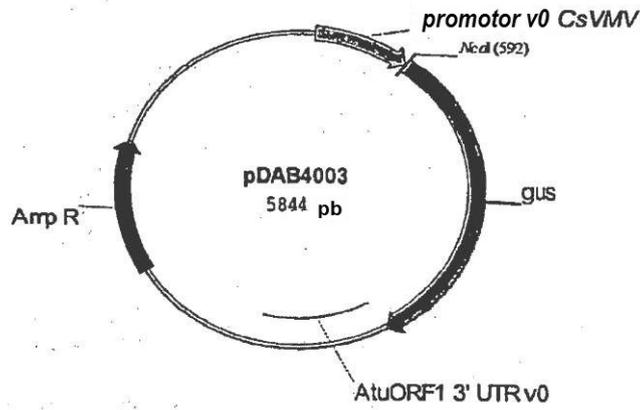


Figura 19

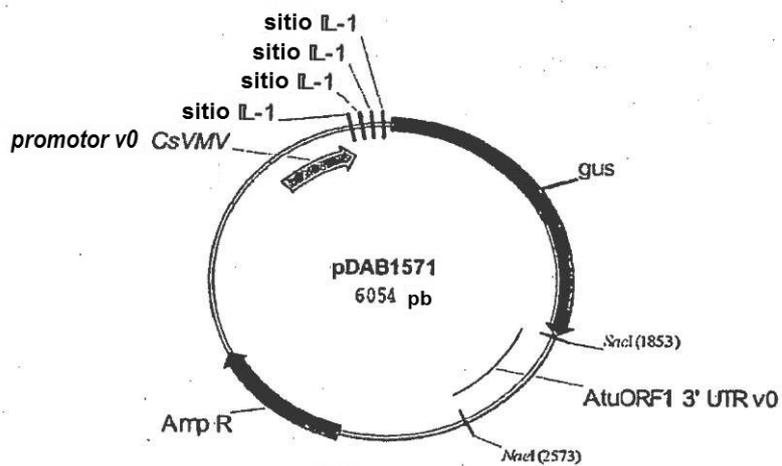


Figura 20

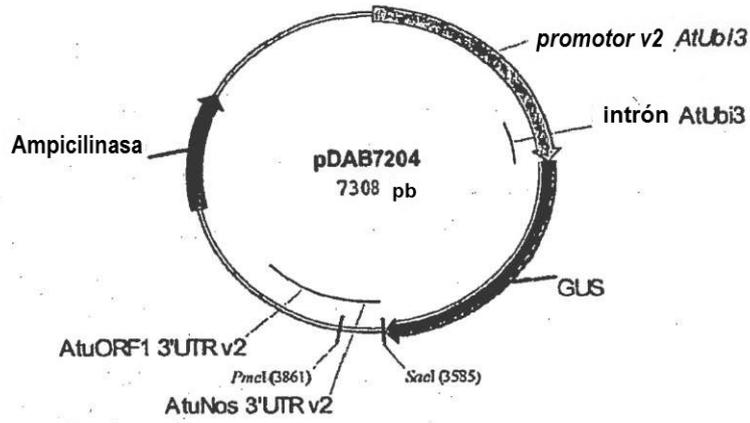


Figura 21

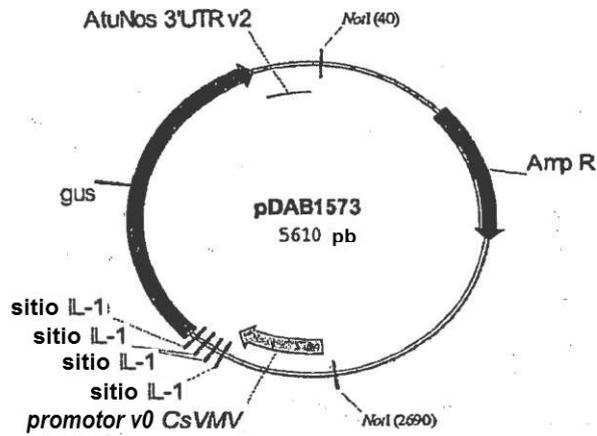


Figura 22

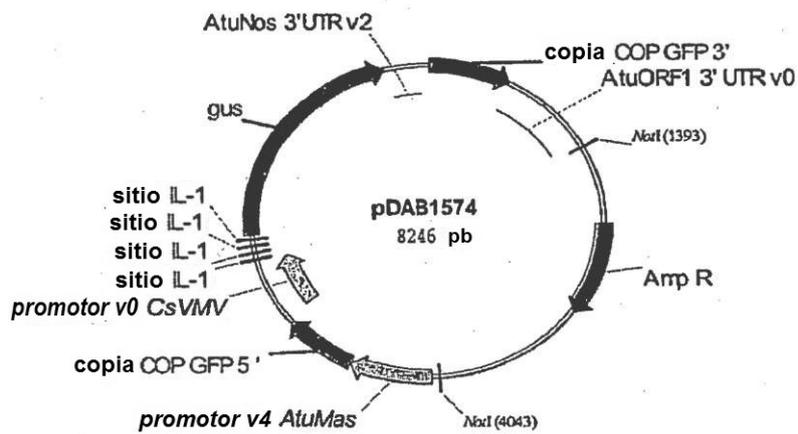


Figura 23

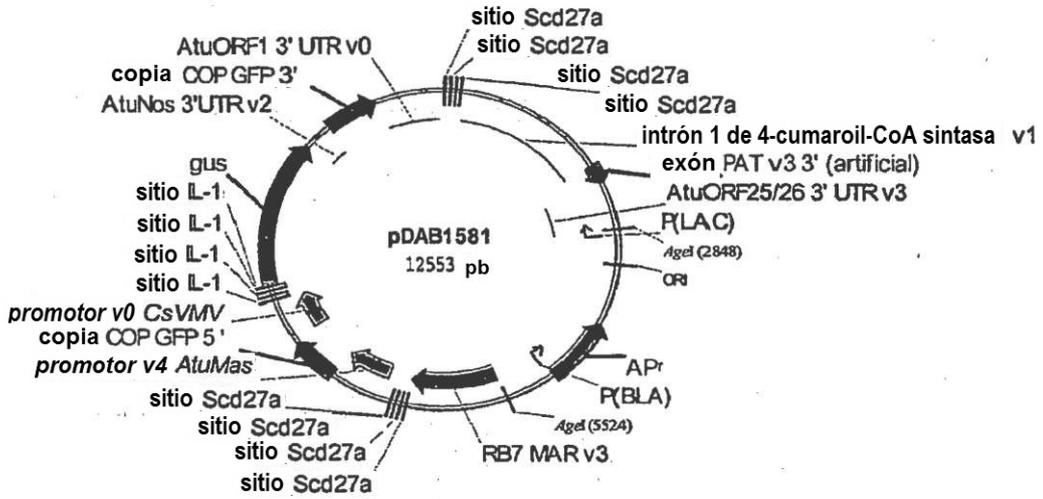


Figura 24

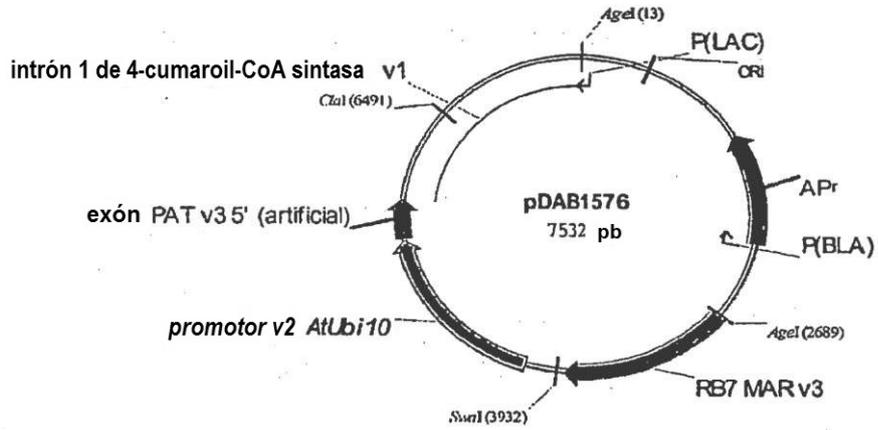


Figura 25

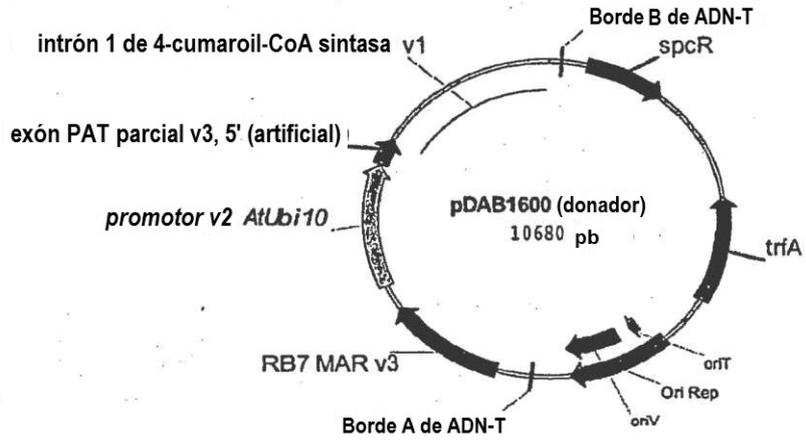
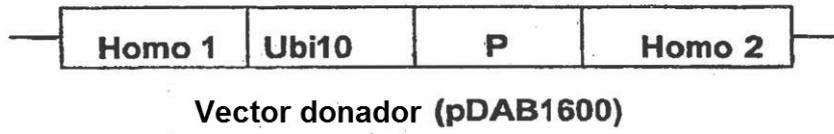


Figura 26

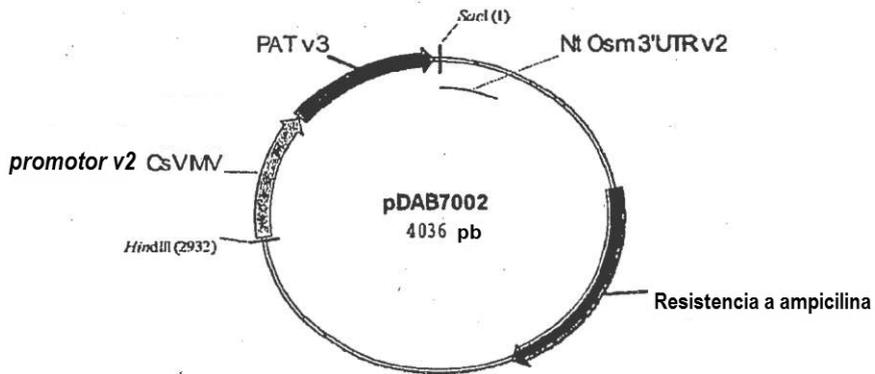


Figura 27

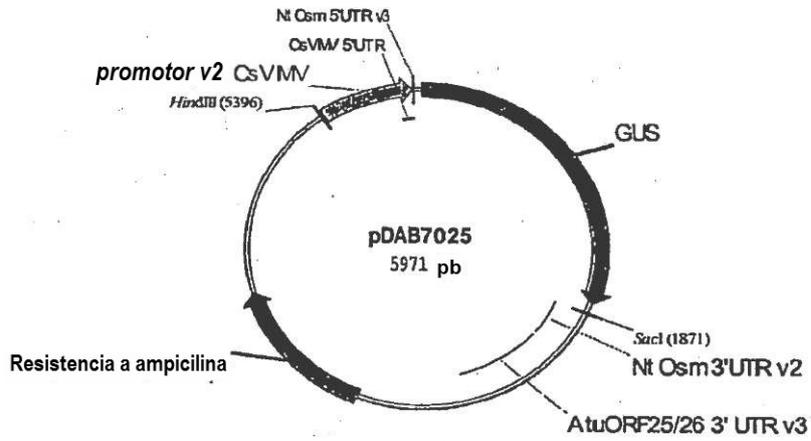


Figura 28

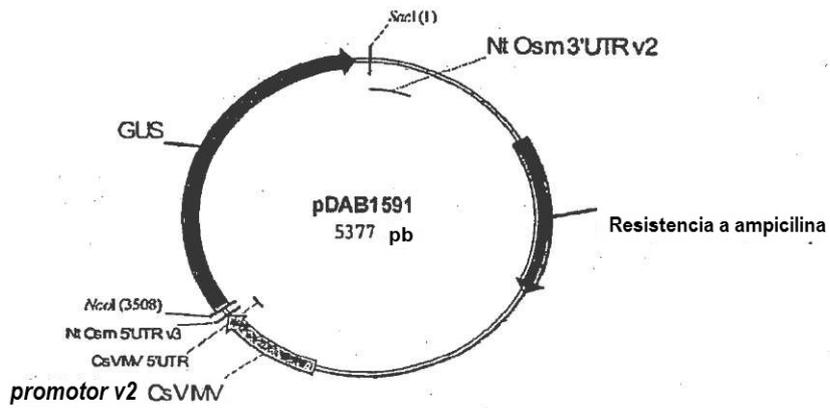


Figura 29

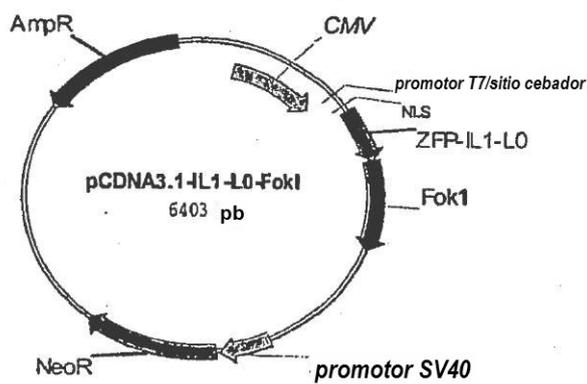


Figura 30

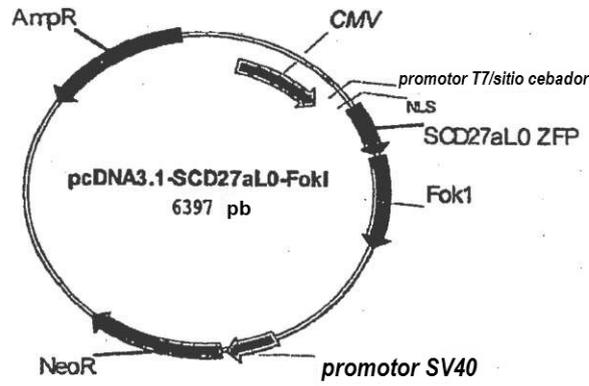


Figura 31

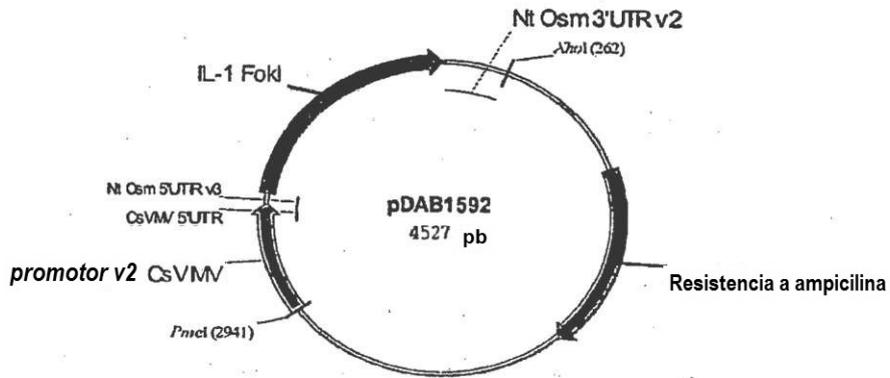


Figura 32

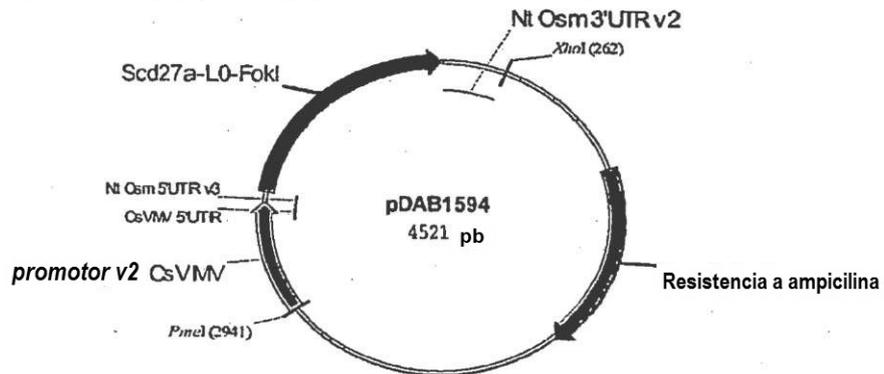


Figura 33

A



Vectores ZFN (pDAB1596, pDAB1598)

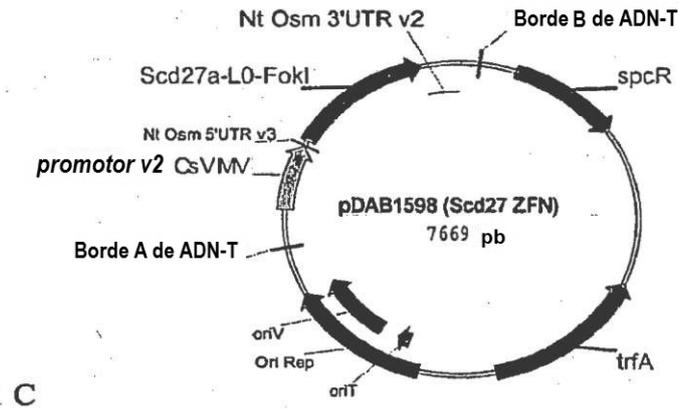
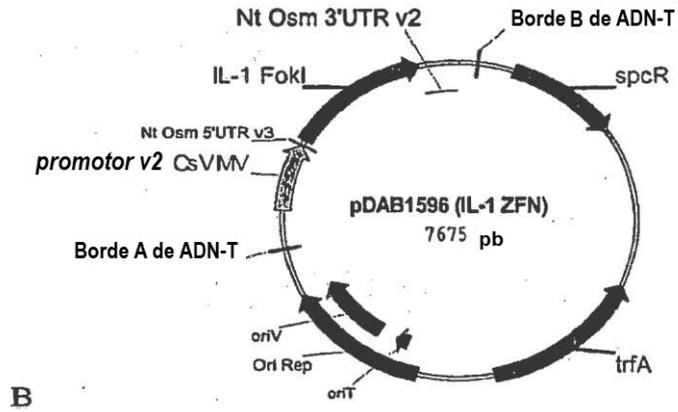


Figura 34

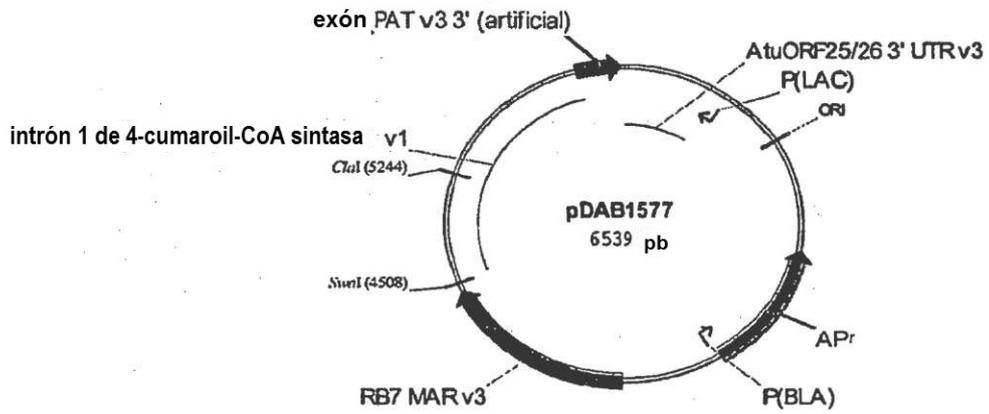


Figura 35

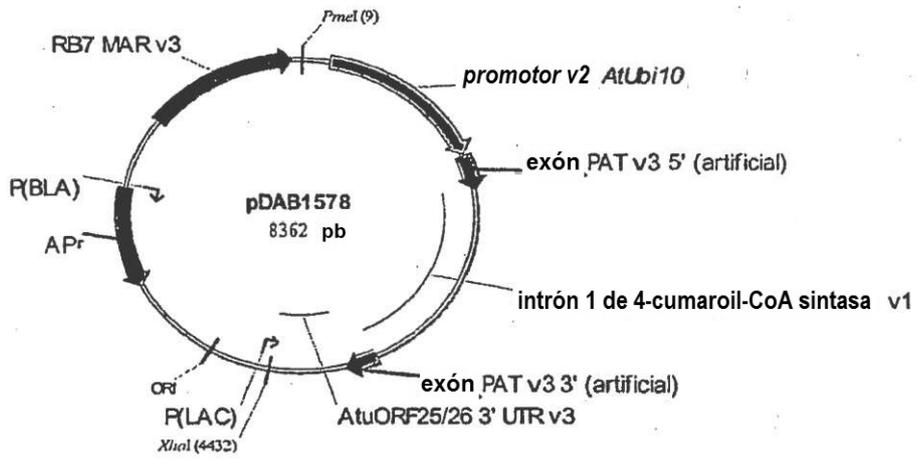
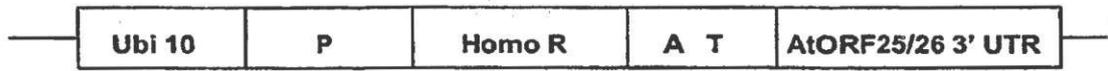


Figura 36



Vector de control positivo (pDAB1601)

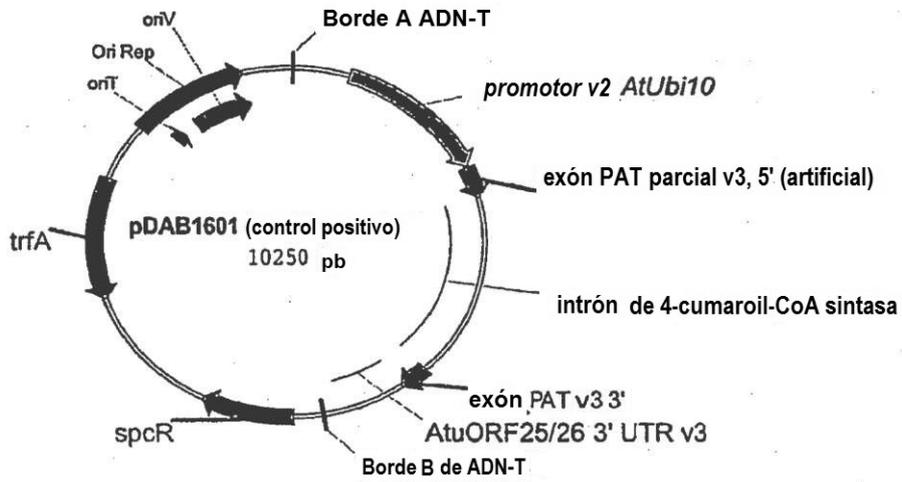


Figura 37

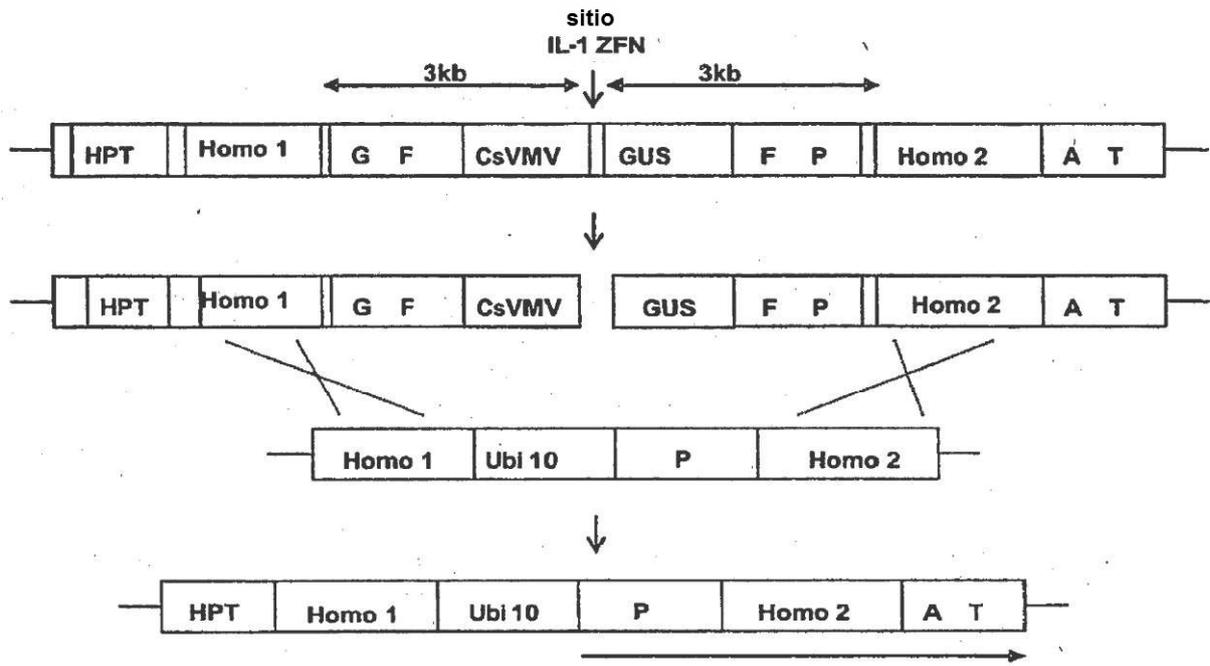


Figura 38

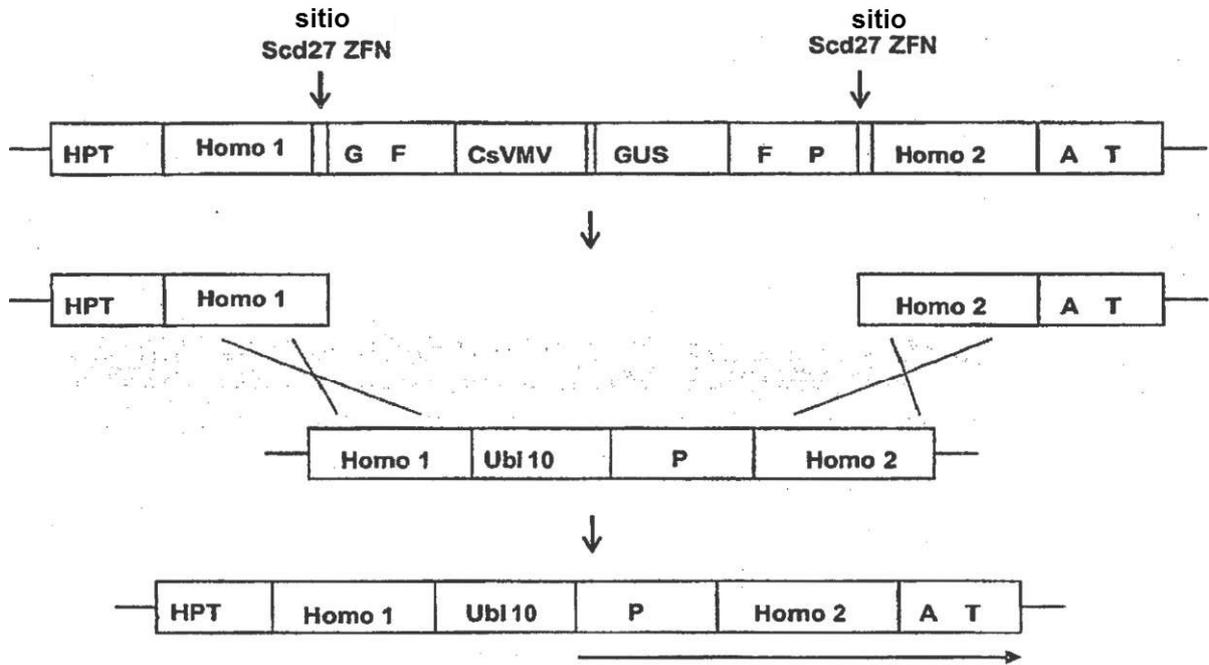


Figura 39

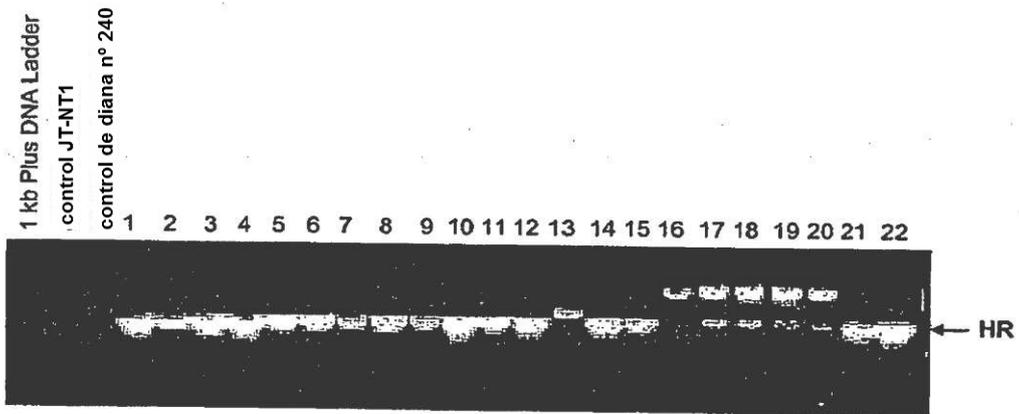


Figura 40

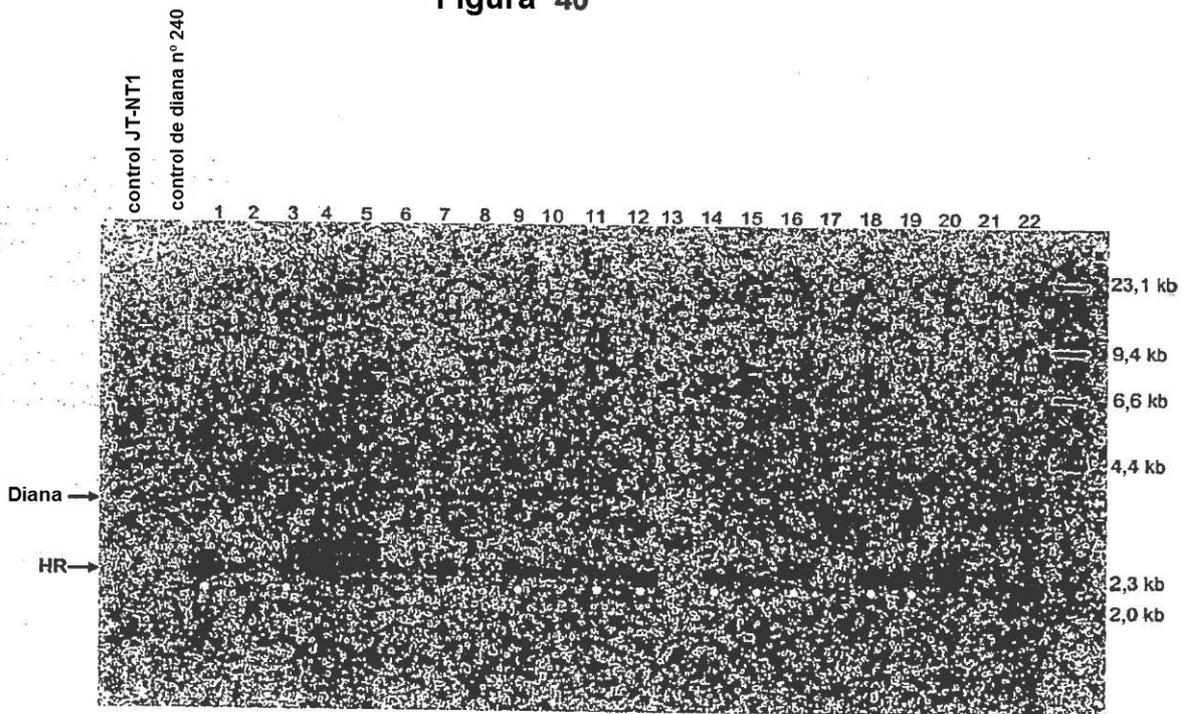


Figura 41

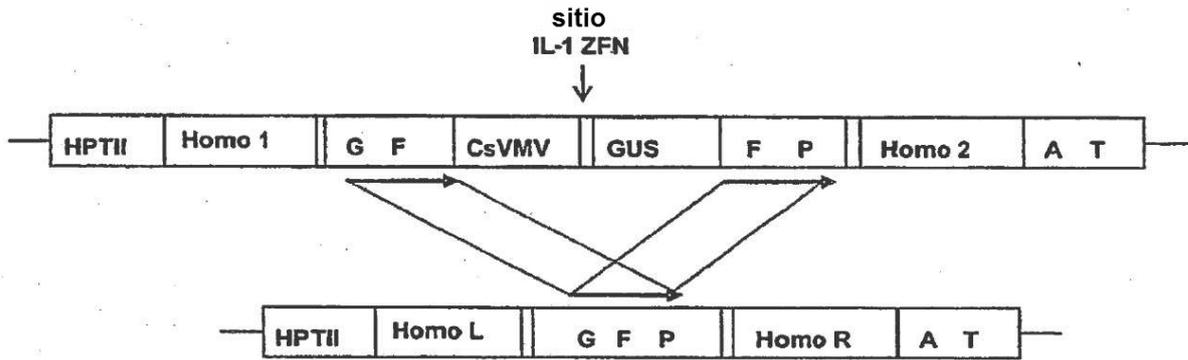


Figura 42

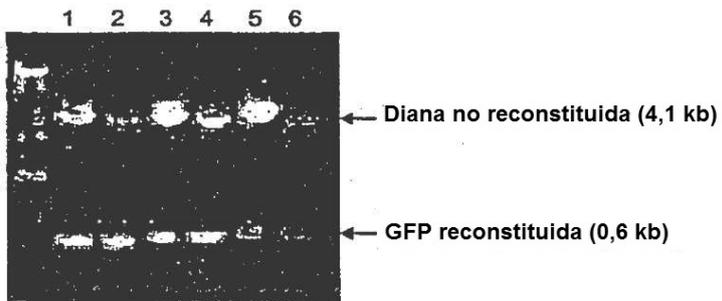


Figura 43

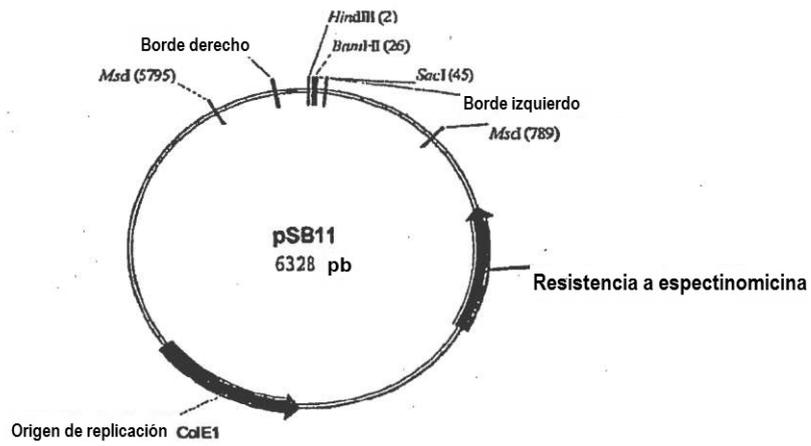


Figura 44

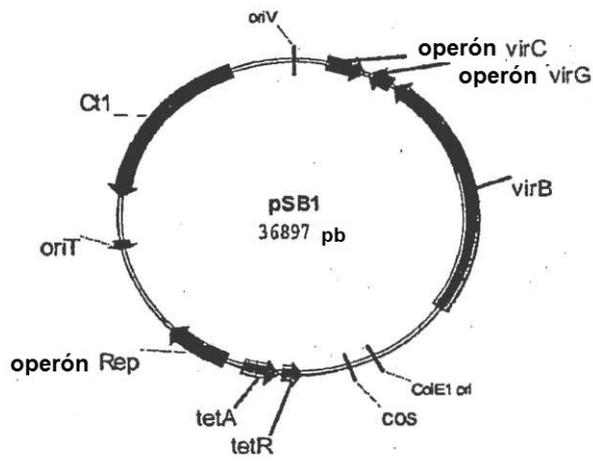


Figura 45

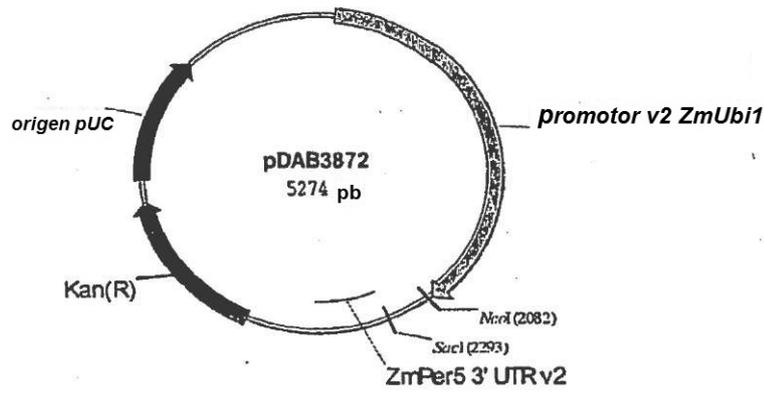


Figura 46

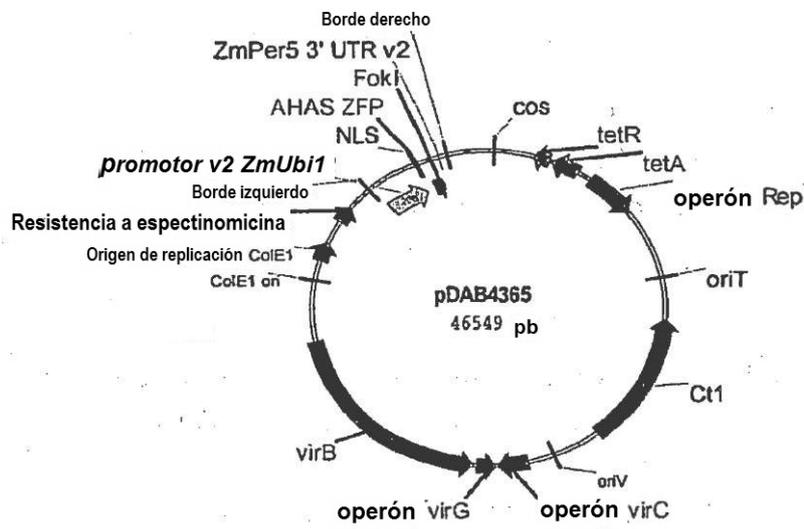


Figura 47