

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 654 923**

51 Int. Cl.:

**C12N 15/05** (2006.01)

**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.12.2009 PCT/US2009/006606**

87 Fecha y número de publicación internacional: **08.07.2010 WO10077319**

96 Fecha de presentación y número de la solicitud europea: **17.12.2009 E 09836506 (7)**

97 Fecha y número de publicación de la concesión europea: **01.11.2017 EP 2370575**

54 Título: **Integración dirigida en la localización Zp15**

30 Prioridad:

**17.12.2008 US 201946 P**

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

**15.02.2018**

73 Titular/es:

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

72 Inventor/es:

**BUTLER, HOLLY J.;  
CORBIN, DAVID R.;  
DOYON, YANNICK;  
GAO, ZHIFANG;  
SHUKLA, VIPULA K.;  
URNOV, FYODOR y  
WORDEN, SARAH E.**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 654 923 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

Integración dirigida en la localización *Zp15*

Declaración de derechos a invenciones realizadas bajo investigación promocionada federalmente

No aplicable

## 5 Campo de la invención

La presente descripción se encuentra en el campo de la ingeniería genómica de plantas, en particular la integración dirigida de un transgén en un gen *Zp15* de planta.

Antecedentes

10 La biotecnología se postula como una herramienta esencial para afrontar el reto de una creciente demanda global de producción de alimentos. Las estrategias convencionales para mejorar la productividad agrícola, p.ej., ingeniería genética para aumentar el rendimiento o la resistencia a plagas, se basan en el cultivo de mutaciones o en la introducción de nuevos genes en los genomas de especies de cultivo mediante transformación. Ambos procesos son inherentemente no específicos y relativamente ineficaces. Por ejemplo, los métodos convencionales de transformación de plantas suministran ADN exógeno que se integra en el genoma en posiciones aleatorias. De esta manera, para identificar y aislar líneas transgénicas con atributos deseables, es necesario generar miles de eventos únicos de integración aleatoria y posteriormente realizar un escrutinio de los individuos deseados. Como resultado, la modificación genética convencional de las características de una planta supone una tarea laboriosa, que consume mucho tiempo, e impredecible. Además, la naturaleza aleatoria de dichas integraciones hace difícil predecir si se han producido efectos pleiotrópicos debidos a una alteración no intencionada del genoma. Como resultado, la generación, aislamiento y caracterización de líneas de plantas con genes o rasgos modificados ha sido un proceso extremadamente laborioso e intensivo en costes, con una baja probabilidad de éxito.

25 La modificación genética dirigida solventa los retos logísticos de las prácticas convencionales en sistemas de plantas, y como tal ha sido una meta desde hace tiempo, aunque elusiva, tanto en la investigación de biología de plantas como en la biotecnología agrícola. Sin embargo, con la excepción del "tratamiento dirigido a genes" a través de una selección de fármacos positivos-negativos en arroz, o el uso de sitios de restricción pre-modificados, la modificación genómica dirigida en todas las especies vegetales, tanto en modelos como en cultivos, ha demostrado ser muy difícil hasta hace poco. Terada *et al.* (2002) *Nat Biotechnol* 20(10): 1030; Terada *et al.* (2007) *Plant Physiol* 144(2): 846; D'Halluin *et al.* (2008) *Plant Biotechnology J.* 6(1): 93.

30 Recientemente se han descrito métodos y composiciones para la ruptura dirigida de ADN genómico. Dichos eventos de ruptura dirigida se pueden usar, por ejemplo, para inducir una mutagénesis dirigida, para inducir eliminaciones dirigidas de secuencias de ADN celular, y para facilitar la recombinación dirigida en una localización cromosómica predeterminada. Véase, por ejemplo, las Publicaciones de Patente de Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474 y 20060188987, y la Publicación Internacional WO 2007/014275.

35 La Publicación de Patente de EE.UU. n° 20080182332 describe el uso de nucleasas con dedos de zinc (ZFNs, del inglés "zinc finger nucleases") no canónicas para la modificación dirigida de genomas de plantas, y la Solicitud de Patente de EE.UU. n° 12/284.888 describe la integración dirigida mediada por ZFN en una localización EPSPS de planta.

40 Sin embargo, sigue existiendo una necesidad de composiciones y métodos para la integración dirigida estable en localizaciones adicionales dentro del genoma de una planta para establecer modificaciones genéticas hereditarias estables en la planta y en su progenie.

Sumario

45 La presente memoria describe métodos y composiciones para expresar uno o más productos de una secuencia de ácido nucleico exógena (es decir, una proteína o una molécula de ARN) que ha sido integrada en un gen *Zp15* en un célula vegetal. La materia objeto de la invención se define en las reivindicaciones anexas. Tal como se muestra en la presente memoria, la integración de una o más secuencias exógenas en la localización *Zp15*, o en sus proximidades, no parece perjudicar la capacidad de la planta hospedante para regenerarse, florecer o producir semillas y, opcionalmente, permite la transmisión hereditaria de la(s) secuencia(s) exógena(s) a lo largo de las generaciones. Las secuencias de ácido nucleico pueden comprender, por ejemplo, uno o más genes o moléculas de ADNc, o cualquier tipo de secuencia codificadora o no codificadora, así como uno o más elementos de control (p.ej., promotores). Por ejemplo, se pueden integrar genes de tolerancia a herbicidas en dicha localización para producir plantas de cultivo con la resistencia a herbicida deseada. Las células que contienen ácidos nucleicos exógenos en la localización del *Zp15*, o en sus proximidades, también pueden contribuir al gametofito (línea germinal) y por tanto ser transmitidas a la progenie en generaciones sucesivas.

La integración de la secuencia de ácido nucleico en un gen *Zp15* se ve facilitada por la ruptura dirigida de la doble

cadena del genoma en la localización *Zp15* seleccionada. La ruptura es dirigida a un gen *Zp15* a través del uso de proteínas de fusión que comprenden un dominio de unión a ADN, tal como un dominio de unión a ADN de meganucleasa, un dominio de unión a ADN de cremallera de leucina, una proteína de dedo de zinc (ZFP), o combinaciones quiméricas de los anteriores, que es modificado para unirse a una secuencia dentro de la localización *Zp15* seleccionada, y un dominio de ruptura o un semi-dominio de ruptura. Dicha ruptura estimula la integración de secuencias de polinucleótido exógenas en el sitio de ruptura o en sus proximidades. La integración de secuencias exógenas puede proceder según mecanismos tanto dependientes de homología como independientes de homología.

En un aspecto, en la presente memoria se describen dominios de unión de ADN modificados (p.ej., ZFPs, meganucleasas, o cremalleras de leucina) que se unen a un sitio diana de un gen *Zp15*. El dominio de unión de ADN puede comprender, por ejemplo, cualquiera de los dominios de unión de ADN de dedo de zinc que comprenden las hélices de reconocimiento mostradas en la Tabla 1. Cualquiera de los dominios de ADN descritos en la presente memoria puede comprender además un dominio funcional, por ejemplo un dominio de ruptura o un semi-dominio de ruptura. En algunas realizaciones, el semi-dominio de ruptura puede proceder de una endonucleasa de restricción de Tipo IIS tal como *FokI* o *StuI*. En otras realizaciones, el dominio de ruptura puede comprender una endonucleasa albergadora, por ejemplo una endonucleasa albergadora con un dominio de unión de ADN modificado.

En otro aspecto, en la presente memoria se describen plantas o semillas que comprende una secuencia exógena integrada en la localización *Zp15*. En determinadas realizaciones, la secuencia exógena se integra en el gametofito de la planta.

En otro aspecto, en la presente memoria se describe un método para expresar el producto de una secuencia de ácido nucleico exógena en una célula vegetal, método que comprende: introducir la una o más secuencias de ácido nucleico exógenas en la célula vegetal, en donde las secuencias de ácido nucleico exógenas incluyen una o más regiones de homología con un gen *Zp15*; y romper una secuencia de cadena doble en un gen *Zp15* endógeno en el genoma de la célula vegetal usando al menos una proteína de fusión que comprende un dominio de ruptura y una proteína de dedo de zinc que comprende las regiones de hélice de reconocimiento tal como se muestra en una única fila de la Tabla 1, dando como resultado de este modo la integración de un polinucleótido que comprende la una o más secuencias exógenas en el genoma de la célula, en donde la ruptura de cadena doble se realiza (a) expresando una primera proteína de fusión en la célula, comprendiendo la primera proteína de fusión un primer dominio de unión de dedo de zinc y un primer semi-dominio de ruptura, en donde el primer dominio de unión de dedo de zinc ha sido diseñado para unirse a un primer sitio diana en la localización *Zp15*, o en sus proximidades, del genoma de la célula vegetal; y (b) expresando una segunda proteína de fusión en la célula, comprendiendo dicha segunda proteína de fusión un segundo dominio de unión de dedo de zinc y un segundo semi-dominio de ruptura, en donde el segundo dominio de unión de dedo de zinc se une a un segundo sitio diana en la localización *Zp15*, o en sus proximidades, del genoma de la célula vegetal, en donde el segundo sitio diana es diferente al primer sitio diana; y en donde la unión de la primera proteína de fusión con el primer sitio diana, y la unión de la segunda proteína de fusión con el segundo sitio diana posiciona los semi-dominios de ruptura de tal forma que el genoma de la célula vegetal en la localización *Zp15*, en sus proximidades, es sometido a ruptura expresando un producto de la una o más secuencias exógenas.

La secuencia de ácido nucleico exógena puede comprender una secuencia que codifica uno o más polipéptidos funcionales (p.ej., un ADNc), con o sin uno o más promotores y/o puede producir una o más secuencias de ARN (p.ej., a través de una o más casetes de expresión de ARNsh), que confieren los rasgos deseados a la planta. Dichos rasgos incluyen, aunque sin limitación, resistencia o tolerancia a herbicidas; resistencia o tolerancia a insectos; resistencia o tolerancia a enfermedades (víricas, bacterianas, fúngicas, de nematodos); tolerancia y/o resistencia al estrés, por ejemplo resistencia o tolerancia a la sequía, calor, frío, congelación, humedad excesiva, estrés salino; estrés oxidativo; incremento de rendimientos; contenido y composición de alimentos; apariencia física; esterilidad masculina; secado; capacidad de permanecer erguida; capacidad de ser prolífica; cantidad y calidad de almidón; cantidad y calidad de aceite; calidad y cantidad de proteínas; composición de aminoácidos; y similares. Por supuesto, se pueden emplear como se deseen cualesquier dos o más ácidos nucleicos exógenos de cualquier descripción, tal como aquellos que confieren resistencia a herbicidas, insectos, enfermedades (víricas, bacterianas, fúngicas, de nematodos) o sequía, esterilidad masculina, secado, capacidad de permanecer erguida; capacidad de ser prolífica, propiedades de almidón, cantidad y calidad de aceite, o los que aumentan el rendimiento o la calidad nutricional. En determinadas realizaciones, la secuencia de ácido nucleico comprende una secuencia que codifica una proteína de resistencia a herbicida (p.ej., el gen de *AAD-1* (ariloxialcanoato dioxigenasa), el gen de *AAD-12*, o el gen de fosfotricina acetil transferasa (*PAT*)) y/o los fragmentos funcionales de la misma. La expresión de la secuencia integrada puede ser llevada a cabo por un promotor ligado operativamente a la secuencia integrada. Alternativamente, la secuencia integrada no tiene promotor y la transcripción es llevada a cabo por el promotor endógeno *Zp15*.

En determinadas realizaciones, el polinucleótido comprende además una segunda secuencia de nucleótidos que es homóloga a una segunda secuencia del gen *Zp15*. La segunda secuencia de nucleótidos puede ser idéntica a la segunda secuencia del gen *Zp15*. Adicionalmente, en las realizaciones que comprenden una primera y una segunda secuencias de nucleótido, la primera secuencia de nucleótido puede ser idéntica a la primera secuencia del gen

*Zp15* y la segunda secuencia de nucleótidos puede ser homóloga, aunque no idéntica, a una segunda secuencia del gen *Zp15*. En cualquiera de los métodos descritos en la presente memoria, la primera y la segunda secuencias de nucleótido flanquean la secuencia exógena. En determinadas realizaciones, el polinucleótido es un plásmido. En otras realizaciones, el polinucleótido es una molécula de ADN lineal.

- 5 En otro aspecto, en la presente memoria se proporciona un método para integrar una o más secuencias de ácido nucleico exógenas en el genoma de una célula vegetal, método que comprende: introducir la una o más secuencias de ácido nucleico exógenas en la célula vegetal, en donde las secuencias de ácido nucleico exógenas incluyen una o más regiones de homología con un gen *Zp15*; y someter a ruptura una secuencia de cadena doble en un gen *Zp15* endógeno del genoma de la célula vegetal usando al menos una proteína de fusión que comprende un dominio de ruptura y una proteína de dedo de zinc que comprende las regiones de hélice de reconocimiento mostradas en una única fila de la Tabla 1, dando como resultado de este modo la integración de un polinucleótido que comprende la una o más secuencias exógenas en el genoma de la célula, en donde la ruptura de cadena doble se realiza (a) expresando una primera proteína de fusión en la célula, comprendiendo dicha primera proteína de fusión un primer dominio de unión de dedo de zinc y un primer semi-dominio de ruptura, en donde el primer dominio de unión de dedo de zinc ha sido diseñado para unirse a un primer sitio diana en la localización *Zp15*, o en sus proximidades, del genoma de la célula vegetal; y (b) expresando una segunda proteína de fusión en la célula, comprendiendo dicha segunda proteína de fusión un segundo dominio de unión de dedo de zinc y un segundo semi-dominio de ruptura, en donde el segundo dominio de unión de dedo de zinc se une a un segundo sitio diana en la localización *Zp15*, o en sus proximidades, en el genoma de la célula vegetal, en donde el segundo sitio diana es diferente del primer sitio diana; y en donde la unión de la primera proteína de fusión al primer sitio diana, y la unión de la segunda proteína de fusión al segundo sitio diana, posicionan los semi-dominios de tal modo que se produce la ruptura del genoma de la célula vegetal en la posición *Zp15*, o en sus proximidades.

En determinadas realizaciones, una secuencia exógena que codifica un polipéptido funcional es insertada en el gen *Zp15*.

- 25 En cualquiera de los métodos descritos en la presente memoria, el primer y el segundo semi-dominios de ruptura pueden proceder de una endonucleasa de restricción de Tipo IIS, por ejemplo, *FokI* o *StuI*. Además, en cualquiera de los métodos descritos en la presente memoria, al menos una de las proteínas de fusión pueden comprender una alteración en la secuencia de aminoácidos de la interfaz de dimerización del semi-dominio de ruptura, por ejemplo de tal modo que obligue a formarse a los heterodímeros de los semi-dominios de ruptura.
- 30 En cualquiera de los métodos descritos en la presente memoria, la célula vegetal puede comprender una célula vegetal monocotiledónea o dicotiledónea. En determinadas realizaciones, la célula vegetal es una planta de cultivo, por ejemplo maíz.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

- 35 **Figura 1:** muestra un ejemplo de resultados de análisis de secuencia de los productos de amplificación de *Zp15* procedentes de gADN Hill de maíz derivado de células sometidas a expresión transitoria del par ZFN nº25 (sitios de unión subrayados) y revela una inserción NHEJ de 6 pb (en negrita) en el sitio de ruptura esperado.

**Figura 2:** muestra un ejemplo de resultados de análisis de secuencia de los productos de amplificación de *Zp15* procedentes de gADN Hill de maíz derivado de células sometidas a expresión transitoria del par ZFN nº24 (sitios de unión subrayados) y revela una eliminación de 3 pb en el sitio de ruptura esperado.

- 40 **Figura 3:** es un esquema que muestra la construcción designada como pDAB7489.

**Figura 4:** es un esquema que muestra un ejemplo de casete de expresión de gen de tolerancia a herbicida que codifica un gen AAD.

**Figura 5:** es un esquema que muestra la construcción designada como pDAB7490.

- 45 **Figura 6:** muestra el alineamiento de un evento de integración dirigida (TI, del inglés "targeted integration") en donde se alinean las secuencias de maíz natural (WT, del inglés "wild-type"), del fragmento donante *Zp15*, así como de las regiones frontera 5' y 3' adjuntas a las secuencias donantes.

**Figura 7:** es un esquema que muestra la construcción designada como pDAB104101.

**Figura 8:** es un esquema que muestra la casete de expresión PAT.

**Figura 9:** es un esquema que muestra la construcción designada como pDAB104107.

- 50 **Figura 10:** es un esquema que muestra la construcción designada como pDAB104104.

**Figura 11:** es un esquema que muestra la construcción designada como pDAB104105.

**Figura 12:** es un esquema que muestra la construcción designada como pDAB104106.

**Figura 13:** es un esquema que muestra la construcción designada como pDAB104100.

**Figura 14:** es un esquema que muestra la construcción designada como pDAB104103.

**Figura 15:** es un esquema que muestra la construcción designada como pDAB104102.

#### DESCRIPCIÓN DETALLADA

5 La presente descripción se refiere a métodos y composiciones para la integración dirigida (TI) en un gen *Zp15* de planta, que se encuentra en el cromosoma 6 del maíz. Usando proteínas de fusión que comprenden dominios de unión de dedo de zinc y dominios de nucleasa, se puede ligar operativamente una secuencia insertada (donante) a un promotor exógeno o puede ir sin promotor. Si no tiene promotor, la transcripción del marco de lectura abierto integrado puede producirse desde el promotor de gen *Zp15* endógeno en los tejidos especificados por promotor. El uso de un donante sin promotor reduce la probabilidad de una integración aleatoria del donante y/o de una activación espuria de un gen endógeno por el promotor portado por el donante.

10 Las composiciones útiles para la ruptura dirigida y la recombinación en un gen *Zp15* incluyen proteínas de fusión que comprende un dominio de ruptura (o un semi-dominio de ruptura) y un dominio de unión a dedo de zinc, polinucleótidos que codifican dichas proteínas y combinaciones de polipéptidos y polinucleótidos que codifican polipéptidos. Un dominio de unión de dedo 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 puede diseñarse para unirse a cualquier secuencia del gen *Zp15*. La presencia de dicha proteína (o proteínas) de fusión en una célula da como resultado la unión de la(s) proteína(s) de fusión con su(s) sitio(s) de unión y la ruptura dentro del gen *Zp15* endógeno.

#### General

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

#### Definiciones

35 Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan de forma intercambiable y se refieren a un polímero de desoxirribonucleótidos o de ribonucleótidos, en conformación lineal o circular, y en forma de cadena sencilla o de cadena doble. Para los propósitos de la presente descripción, estos términos no deben considerarse como limitativos en relación a la longitud del polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que son modificados en las funciones de bases, azúcares y/o fosfatos (p.ej., las 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 será el par de base de T.

40 Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable para referirse a un polímero de residuos de aminoácido. El término también aplica a polímeros de aminoácido en los que uno o más aminoácidos son análogos químicos o derivados modificados de los aminoácidos naturales correspondientes.

45 "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 secuencia (p.ej., los contactos con residuos de fosfato en una cadena principal de ADN), siempre que la interacción en su conjunto sea específica de secuencia. Dichas interacciones generalmente se caracterizan por una constante de disociación ( $K_d$ ) de  $10^{-6} \text{ M}^{-1}$  o inferior. "Afinidad" se refiere a la fuerza de unión: un aumento de la afinidad de unión se correlaciona con una  $K_d$  menor.

50 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 (proteína de unión a ADN), a una molécula de ARN (proteína de unión a ARN) y/o a una molécula proteínica (proteína de unión a proteína). En el caso de una proteína de unión a proteína, se puede unir consigo misma (para formar homodímeros, homotrímeros, etc.) y/o puede unirse con una o más moléculas de una proteína o proteína diferentes. Una proteína de unión puede presentar más de un tipo de actividad de unión. Por ejemplo, las proteínas de dedos de zinc presentan actividad de unión a ADN, de unión a ARN y de unión a proteínas.

55 Una "proteína de unión a ADN con dedos de zinc" (o un dominio de unión) es una proteína, o un dominio dentro de

una proteína de mayor tamaño, que se une a ADN de una manera 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 está estabilizada por la coordinación de un ion de zinc. El término proteína de unión de ADN con dedos de zinc a menudo se abrevia como proteína con dedos de zinc o ZFP (del inglés "zinc finger protein").

5 Los dominios de unión con dedos de zinc pueden ser "modificados" para unirse a una secuencia de nucleótidos predeterminada. Los ejemplos no limitativos de métodos para modificar proteínas con dedos de zinc son el diseño y la selección. Una proteína con dedos de zinc diseñada es una proteína que no existe en la naturaleza cuyo diseño/composición es el resultado principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y de algoritmos de ordenador para procesar la información de una base de datos que almacena información de diseños de ZFP y de datos de unión existentes. Ver, por ejemplo, las Patentes de EE.UU. 6.140.081; 6.453.242 y 6.534.261; ver también WO 98/53058; WO 898/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

15 Una proteína con dedos de zinc "seleccionada" es una proteína que no existe en la naturaleza cuya producción es el resultado principalmente de un proceso empírico, tal como presentación de fagos, trampa de interacción o selección híbrida. Ver, p.ej., los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; WO 95/18431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197 y WO 02/099084.

20 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 de cadena sencilla o de cadena doble. El término "secuencia donante" se refiere a una secuencia de nucleótidos que es insertada en un genoma. Una secuencia donante puede tener cualquier longitud, por ejemplo una longitud entre 2 y 10.000 nucleótidos (o cual valor entero entre ambos o por encima), preferiblemente una longitud entre aproximadamente 100 y 1.000 nucleótidos (o cualquier número entero entre ambos), más preferiblemente una longitud entre aproximadamente 200 y 500 nucleótidos.

25 Una "secuencia homóloga, no idéntica" 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 homóloga y no idéntica a la secuencia del gen mutante. En determinadas realizaciones, el grado de homología entre las dos secuencias es suficiente para permitir la recombinación homóloga entre ambas, utilizando mecanismos celulares normales. Dos secuencias homólogas no idénticas pueden tener cualquier longitud y su grado de no homología puede ser tan pequeño como un único nucleótido (p.ej., para corrección de una mutación genómica puntual mediante recombinación homóloga dirigida) o tan grande como 10 o más kilobases (p.ej., para inserción de un gen en un sitio ectópico predeterminado de un cromosoma). No es necesario que los dos polinucleótidos que comprenden las secuencias homólogas no idénticas tengan la misma longitud. Por ejemplo, se puede usar un polinucleótido exógeno (es decir, polinucleótido donante) de entre 20 y 10.000 nucleótidos o pares de nucleótidos.

35 Las técnicas para determinar la identidad de secuencia de ácido nucleico y aminoácidos son conocidas en la técnica. Típicamente, dichas técnicas incluyen la determinación de la secuencia de nucleótidos del ARNm correspondiente a un gen y/o la determinación de la secuencia de aminoácidos codificada por el mismo, y la comparación de dichas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta manera. En general, la identidad se refiere a una correspondencia exacta de nucleótido-a-nucleótido o de aminoácido-a-aminoácido de dos secuencias de polinucleótidos o de polipéptido, respectivamente. Se pueden comparar dos o más secuencias (de polinucleótido o de aminoácido) para determinar su porcentaje de identidad. El porcentaje de identidad de dos secuencias, tanto si son secuencias de ácido nucleico como de aminoácidos, es el número de coincidencias exactas entre las dos secuencias alineadas dividido por la longitud de la secuencia más corta y multiplicado por 100. El algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981) proporciona un alineamiento aproximado para secuencias de ácido nucleico. Este algoritmo se puede aplicar a secuencias de aminoácido usando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 supl. 3: 353-358, National Biomedical Research Foundation, Washington, D.C., EE.UU., y normalizado por Gribskov, *Nucl. Acids Res.* 14(6): 6745-6763 (1986). Un ejemplo de implementación de este algoritmo para determinar el porcentaje de identidad de una secuencia es el proporcionado por el "Genetics Computer Group" (Madison, WI) en la aplicación de utilidad "BestFit". En la técnica se conocen de forma general programas adecuados para calcular el porcentaje de identidad o de similitud entre secuencias, por ejemplo, otro programa de alineamiento es el BLAST, usado con los parámetros por defecto. Por ejemplo, BLASTN y BLASTP pueden usarse con los parámetros por defecto: código genético = estándar; filtro = ninguno; cadena = ambas; valor de corte = 60; excepción = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; clasificado por = HIGH SCORE; Bases de datos = no redundantes, GenBank + EMBL + DDBJ + PDB + traducciones CDS de GenBank + Swiss protein + Spupdate + PIR. Los detalles de estos programas se pueden encontrar en internet. Con respecto a las secuencias descritas en la presente memoria, el rango de grados deseados de identidad de secuencia es de aproximadamente 80% a 100%, y cualquier valor de número entero entre ambos. Típicamente los porcentajes de identidad entre secuencias son al menos 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y lo más preferiblemente 98% de identidad de secuencia.

Alternativamente, el grado de similitud de secuencia entre polinucleótidos se puede determinar mediante hibridación de polinucleótidos en las condiciones que permitan la formación de dúplex estables entre regiones homólogas, seguida de digestión con nucleasa(s) específica(s) de cadena sencilla, y determinación de tamaño de los fragmentos digeridos. Dos secuencias de ácido nucleico, o dos secuencias de polipéptido, son sustancialmente homólogas una con respecto a la otra cuando las secuencias exhiben al menos un 70%-75% de identidad de secuencia a lo largo de una longitud definida de las moléculas según se determina usando los métodos anteriores, preferiblemente 80%-82%, más preferiblemente 85%-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y lo más preferiblemente 98%. Tal como se usa en la presente memoria, sustancialmente homólogas también se refiere a secuencias que presentan una identidad completa con respecto a una secuencia de ADN o de polipéptido especificada. Las secuencias de ADN que son sustancialmente homólogas se pueden identificar en un experimento de hibridación Southern, por ejemplo, en condiciones severas, según se definan para dicho sistema en particular. Las condiciones de hibridación apropiadas se encuentran dentro del estado del arte de la técnica. Ver, p.ej., Sambrook et al., ver antes; *Nucleic Acid Hybridization: A Practical Approach*, editores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

La hibridación selectiva de dos fragmentos de ácido nucleico puede determinarse como se indica a continuación. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta a la eficacia y la fortaleza de los eventos de hibridación entre dichas moléculas. Una secuencia de ácido nucleico idéntica parcialmente 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 puede determinarse usando ensayos de hibridación que son bien conocidos en la técnica (p.ej., tinción Southern (ADN), tinción Northern (ARN), hibridación en disolución, o similares, ver Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, Segunda Edición, (1989) Cold Spring Harbor, N.Y.). Dichos ensayos pueden llevarse a cabo usando varios grados de selectividad, por ejemplo, usando condiciones que varían entre baja severidad y alta severidad. Si se emplean condiciones de baja severidad, la ausencia de unión no específica puede determinarse usando una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda que presenta menos de aproximadamente 30% de identidad de secuencia con la molécula diana), de tal modo que, en ausencia de eventos de unión no específicos, la sonda secundaria no se hibridará con la diana.

Cuando se utiliza un sistema de detección basado en hibridación, se elige una sonda de ácido nucleico que sea complementaria a una secuencia de ácido nucleico de referencia, y a continuación por selección de las condiciones apropiadas la sonda y la secuencia de referencia se hibridan, o se unen, selectivamente una con la otra para formar una molécula dúplex. Una molécula de ácido nucleico que es capaz de hibridarse selectivamente con una secuencia de referencia en condiciones de hibridación moderadamente severas se hibrida típicamente 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 un 70% de identidad de secuencia con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación severas permiten típicamente la detección de secuencias de ácido nucleico de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencia superior a aproximadamente el 90-95% con respecto a la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de sonda/secuencia de referencia, en donde la sonda y la secuencia de referencia presentan un grado específico de identidad de secuencia, pueden determinarse como se conoce en la técnica (ver, por ejemplo, *Nucleic Acid Hybridization: A Practical Approach*, editores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

Las condiciones para la hibridación son bien conocidas por los especialistas en la técnica. La severidad de hibridación se refiere al grado en el cual las condiciones de hibridación desfavorecen la formación de híbridos que contienen nucleótidos no coincidentes, correlacionándose una mayor severidad con una menor tolerancia hacia híbridos no coincidentes. Los factores que afectan a la severidad de la hibridación son bien conocidos por los especialistas en la técnica, e incluyen, aunque sin limitación, la temperatura, el pH, la fuerza iónica y la concentración de disolventes orgánicos tales como, por ejemplo, formamida y dimetilsulfóxido. Como es sabido por los especialistas en la técnica, la severidad de la hibridación aumenta a mayores temperaturas, menor fuerza iónica y menores concentraciones de disolvente.

Con respecto a las condiciones de severidad para la hibridación, es bien conocido en la técnica que se pueden emplear numerosas condiciones equivalentes para establecer una severidad particular variando, por ejemplo, los siguientes factores: la longitud y la naturaleza de las secuencias, la composición de bases de las diversas secuencias, las concentraciones de sales y otros componentes de la disolución de hibridación, la presencia o la ausencia de agentes bloqueantes en las disoluciones de hibridación (p.ej., sulfato de dextrano, y polietilen glicol), los parámetros de temperatura y tiempo de reacción de hibridación, así como condiciones de lavado variables. La selección de un conjunto concreto de condiciones de hibridación se selecciona siguiendo métodos estándares en la técnica (ver, por ejemplo, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Segunda Edición, (1989) Cold Spring Harbor, N.Y.).

“Recombinación” se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, “recombinación homóloga (HR)” se refiere a la forma especializada en la que tiene lugar dicho intercambio, por ejemplo, durante la reparación de agujeros de doble cadena en las células. Este proceso requiere de homología de secuencia de nucleótidos, usa una molécula “donante” como plantilla para reparar

- una molécula “diana” (es decir, la que ha experimentado el agujero de cadena doble), y se conoce de forma diversa como “conversión génica no cruzada” o “conversión génica de tracto corto”, debido a que conduce a la transferencia de información genética desde el donante hasta la diana. Sin pretender establecer ninguna teoría particular, dicha transferencia puede implicar la corrección de discrepancias del ADN heterodúplex que se forman entre la diana rota y el donante, y/o la “maduración de cadenas dependiente de síntesis”, en la que el donante es usado para resintetizar información genética que se entrará a formar parte de la diana, y/o procesos relacionados. Dicha HR especializada a menudo da como resultado una alteración de la secuencia de la molécula diana de tal modo que una parte, o la totalidad, de la secuencia del polinucleótido donante es incorporada al polinucleótido diana.
- “Ruptura” se refiere a la ruptura de la cadena principal covalente de una molécula de ADN. La ruptura se puede iniciar mediante una variedad de métodos que incluyen, aunque sin limitación, la hidrólisis enzimática o química de un enlace de fosfodiéster. Es posible la ruptura tanto de cadena sencilla como de cadena doble, y la ruptura de cadena doble puede producirse como resultado de dos eventos de ruptura de cadena sencilla distintos. La ruptura de ADN puede dar como resultado la producción de extremos romos o de extremos escalonados. En determinadas realizaciones, se usan polipéptidos de fusión para la ruptura dirigida de ADN de cadena doble.
- Un “dominio de ruptura” comprende una o más secuencias de polipéptido que poseen actividad catalítica para la ruptura de ADN. Un dominio de ruptura puede estar contenido en una única cadena de polipéptido o la actividad de ruptura puede resultar de la asociación de dos (o más) polipéptidos.
- Un “semi-dominio de ruptura” es una secuencia de polipéptido que, en conjunción con un segundo polipéptido (idéntico o diferente) forma un complejo que presenta una actividad de ruptura (preferiblemente, actividad de ruptura de cadena doble).
- “Cromatina” es la estructura de nucleoproteína que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteína, que incluye proteínas cromosómicas histonas y no histonas. La mayoría de la cromatina celular eucariótica existe en forma de nucleosomas, en donde un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociados a un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y ADN ligado (de longitud variable dependiendo del organismo) se extiende entre los núcleos de nucleosoma. Generalmente, una molécula de histona H1 está asociada al ADN ligado. Para los propósitos de la presente descripción, el término “cromatina” pretende abarcar todos los tipos de nucleoproteína celular, tanto procariótica como eucariótica. La cromatina celular incluye cromatina tanto cromosómica como episómica.
- Un “cromosoma” es un complejo de cromatina que comprende la totalidad o una porción del genoma de una célula. El genoma de una célula a menudo se caracteriza por su cariotipo, que es la colección de todos los cromosomas que comprende el genoma de la célula. El genoma de una célula puede comprender uno o más cromosomas.
- Un “episoma” es un ácido nucleico replicante, un complejo de nucleoproteína u otra estructura que comprende un ácido nucleico que no forma parte del cariotipo cromosómico de una célula. Los ejemplos de episomas incluyen plásmidos y determinados genomas víricos.
- Una “región accesible” es un sitio de la cromatina celular en el que un sitio diana presente en el ácido nucleico puede unirse a una molécula exógena que reconoce dicho sitio diana. Sin pretender establecer ninguna teoría particular, se cree que una región accesible es aquella que no está empaquetada en la estructura nucleosomal. La estructura distintiva de una región accesible a menudo puede ser detectada por su sensibilidad a sondas químicas y enzimáticas, por ejemplo, nucleasas.
- Un “sitio diana” o una “secuencia diana” es una secuencia de ácido nucleico que define una porción de un ácido nucleico al cual se unirá la molécula de unión, siempre que se den las condiciones suficientes para que se produzca la unión. Por ejemplo, la secuencia 5'-GAATTC-3' es un sitio diana para la endonucleasa de restricción Eco RI.
- Una molécula “exógena” es una molécula que no está presente normalmente en una célula, pero que puede ser introducida en una célula mediante uno o más métodos genéticos, bioquímicos o de otro tipo. La “presencia normal en la célula” se determina con respecto a la etapa de desarrollo particular y a las condiciones ambientales de la célula. De esta manera, por ejemplo, una molécula que está presente solo durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula de músculo adulta. De forma similar, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula no inducida por choque térmico.
- Una molécula exógena puede comprender, por ejemplo, una secuencia codificadora para cualquier polipéptido o fragmento del mismo, una versión en funcionamiento de la molécula endógena malfunctionante o una versión malfunctionante de una molécula endógena que funciona con normalidad. Adicionalmente, una molécula exógena puede comprender una secuencia codificadora procedente de otra especie que es un ortólogo de un gen endógeno en la célula hospedante.
- Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como la generada en un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, carbohidrato, lípido, glicoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las anteriores moléculas, o cualquier complejo que comprende una o más de las anteriores moléculas. Los ácidos nucleicos incluyen ADN y ARN, pueden



ser de cadena sencilla o doble; pueden ser lineales, ramificados o circulares; y pueden tener cualquier longitud. Los ácidos nucleicos incluyen aquellos capaces de formar dúplex, así como los ácidos nucleicos capaces de formar tríplex. Ver, por ejemplo, las Patentes de EE.UU. nº 5.176.996 y 5.422.251. Las proteínas incluyen, aunque sin limitación, proteínas de unión a ADN, factores de transcripción, factores de remodelación de cromatina, proteínas de unión a ADN metiladas, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, quinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.

Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, p.ej., una proteína o ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma vírico infeccioso, un plásmido o episoma introducido en una célula, o un cromosoma que no está presente normalmente en la célula. Los métodos para la introducción de moléculas exógenas en células son conocidos por los especialistas en la técnica e incluyen, aunque sin limitación, la transferencia mediada por lípidos (es decir, liposomas, que incluyen lípidos neutros y catiónicos), la electroporación, la inyección directa, la fusión celular, el bombardeo de partículas, la coprecipitación de fosfato cálcico, la transformación de nanopartículas, la transferencia mediada por DEAE-dextrano y la transferencia mediada por vector vírico.

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

Tal como se usa en la presente memoria, el término "producto de un ácido nucleico exógeno" incluye tanto productos de polinucleótido como de polipéptido, por ejemplo, productos de transcripción (polinucleótidos tales como ARN) y productos de traducción (polipéptidos).

Una molécula de "fusión" es una molécula en la que dos o más moléculas de subunidad están unidas, por ejemplo, covalentemente. Las moléculas subunitarias pueden ser el mismo tipo químico de molécula, o pueden ser diferentes tipos químicos de moléculas. Los ejemplos del primer tipo de molécula de fusión incluyen, aunque sin limitación, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión a ADN de ZFP y un dominio de ruptura) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita más arriba). Los ejemplos del segundo tipo de molécula de fusión incluyen, aunque sin limitación, una fusión entre un ácido nucleico formador de tríplex y un polipéptido, y una fusión entre un aglomerante de ranura menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede dar ser el resultado de la administración de la proteína de fusión a la célula o de la administración 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. En la expresión de una proteína en una célula también pueden estar implicados la trans-división, la ruptura de polipéptido y la ligación de polipéptido. Los métodos para la administración de polinucleótidos y polipéptidos a células se presentan en otro apartado de la presente descripción.

Un "gen", para los propósitos de la presente descripción, incluye una región de ADN que codifica un producto génico (ver más adelante), así como todas las regiones de ADN que regulan la producción del producto génico, tanto si dichas secuencias reguladoras están adyacentes a secuencias codificadoras y/o transcritas o no. Por consiguiente, un gen incluye, aunque sin limitarse necesariamente a ello, secuencias promotoras, terminadores, secuencias reguladoras traduccionales tales como sitios de unión de ribosoma y sitios de entrada de ribosoma internos, potenciadores, silenciadores, aislantes, elementos de borde, orígenes de replicación, sitios de unión de matriz y regiones de control de localización.

"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, ARN interferente, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida mediante traducción de un ARNm. Los productos génicos también incluyen ARNs que son modificados por procesos tales como cortado, poliadenilación, metilación y edición, y proteínas modificadas mediante, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, 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, aunque sin limitación, la activación génica y la represión génica.

Las células "vegetales" incluyen, aunque sin limitación, células de plantas monocotiledóneas (monocots) o dicotiledóneas (dicots). Los ejemplos no limitativos de monocots incluyen plantas de cereal tales como maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, cebolla, plátano y coco. Los ejemplos no limitativos de dicots incluyen tabaco, tomate, girasol, algodón, remolacha, patata, lechuga, melón, soja, canola (colza) y alfalfa. Las células vegetales pueden proceder de cualquier parte y/o de cualquier etapa de desarrollo de la planta.

Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificadora dentro o adyacente a un gen, en el que es deseable unir una molécula exógena. La unión se puede

realizar para los propósitos de ruptura de ADN dirigida y/o de recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma organular (p.ej., mitocondrial, de cloroplasto) o un genoma vírico infeccioso, por ejemplo. Una región de interés puede estar dentro de la región codificadora de un gen, dentro de regiones no codificadoras transcritas tales como, por ejemplo, secuencias líder, secuencias tráiler o intrones, o dentro de regiones no transcritas, tanto por encima como por debajo de la región codificadora. Una región de interés puede ser tan pequeña como un único nucleótido o tener una longitud de hasta 2.000 pares de nucleótidos, o cualquier valor integral de pares de nucleótidos.

Los términos “enlace operación” y “ligado operativamente” (o “ligado operablemente”) se usan de modo intercambiable en referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en la que los componentes se disponen de tal modo que ambos componentes actúan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar en una función que es ejercida por al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora transcripcional, tal como un promotor, está ligado operativamente a una secuencia codificadora si la secuencia reguladora transcripcional controla el nivel de transcripción de la secuencia codificadora en respuesta a la presencia o la ausencia de uno o más factores reguladores transcripcionales. Una secuencia reguladora transcripcional generalmente está ligada operativamente en posición *cis* respecto a una secuencia codificadora, pero no es necesario que esté directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora transcripcional que está ligada operativamente a una secuencia codificadora incluso aunque no estén contiguas.

Con respecto a los polipéptidos de fusión, el término “ligado operativamente” puede referirse al hecho de que cada uno de los componentes realiza la misma función enlazado al otro componente como si no estuviera enlazado. Por ejemplo, con respecto a un polipéptido de fusión en el que el dominio de unión ADN de ZFP está fusionado a un dominio de ruptura, el dominio de unión a ADN de ZFP y el dominio de ruptura están enlazados operativamente si, en el polipéptido de fusión, la porción de dominio de unión a ADN de ZFP es capaz de unirse a su sitio diana y/o a su sitio de unión, y a su vez el dominio de ruptura es capaz de romper ADN en las proximidades del sitio diana.

Un “fragmento funcional” de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero que aún mantiene la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede poseer más, menos, o el mismo número de residuos que la correspondiente molécula nativa, y/o puede contener una o más sustituciones de aminoácido o nucleótido. Los métodos para determinar la función de un ácido nucleico (p.ej., función codificadora, capacidad para hibridarse a otro ácido nucleico) son bien conocidos en la técnica. De forma similar, los métodos para determinar la función de proteína son bien conocidos. Por ejemplo, la función de unión a ADN de un polipéptido se puede determinar, por ejemplo, mediante unión por filtración, cambio de movilidad electroforética, o ensayos de inmunoprecipitación. La ruptura de ADN puede determinarse mediante electroforesis de gel. Ver Ausubel *et al.*, más arriba. La capacidad de una proteína para interactuar con otra proteína puede determinarse, por ejemplo, mediante co-inmunoprecipitación, ensayos de dos-híbridos o complementación, tanto genética como bioquímica. Ver, por ejemplo, Fields *et al.* (1989) *Nature* **340**: 245-246; Patente de EE.UU. nº 5.585.245 y PCT WO 98/44350.

### Sitios diana

Los métodos y composiciones descritos incluyen proteínas de fusión que comprenden un semi-dominio de ruptura y un dominio de unión de dedos de zinc en el que el dominio de unión de dedos de zinc, mediante unión a una secuencia de una localización *Zp15* de planta, dirige la actividad del semi-dominio de ruptura a las proximidades de la secuencia y, por tanto, induce la ruptura (a saber, la ruptura de doble cadena) en *Zp15*. Como se ha establecido anteriormente en la presente descripción, se puede diseñar un dominio de dedos de zinc para que se una a virtualmente cualquier secuencia deseada. Por consiguiente, se puede diseñar uno o más dominios de unión de dedos de zinc para que se unan a una o más secuencias de un gen *Zp15* de planta. La expresión de dos proteínas de fusión, comprendiendo cada una un dominio de unión de dedos de zinc y un semi-dominio de ruptura, en una célula, efectúa la ruptura en el gen *Zp15*.

La selección de una secuencia en un *Zp15* para unión a un dominio de dedos de zinc (p.ej., un sitio diana) puede conseguirse, por ejemplo, según los métodos descritos en la Patente de EE.UU. compartida nº 6.453.242 (17 de septiembre de 2002), que también describe métodos para diseñar ZFPs para unirse a una secuencia seleccionada. Será evidente para los especialistas en la técnica que también se puede usar una simple inspección visual de una secuencia de nucleótidos para la selección de un sitio diana. Por consiguiente, se puede usar cualquier medio de selección de sitios diana en los métodos descritos en la presente memoria.

Para los dominios de unión a ADN de ZFP, los sitios diana generalmente están compuestos de una pluralidad de subsitios diana adyacentes. Un subsitio diana se refiere a la secuencia (normalmente un triplete de nucleótidos o un cuartete de nucleótidos que pueden solapar en un nucleótido con un cuartete adyacente) unida a un dedo de zinc individual. Ver, por ejemplo, el documento WO 02/077227. Si la cadena con la que la proteína de dedo de zinc hace la mayoría de los contactos se designa como la cadena diana “cadena de reconocimiento principal”, o “cadena de contacto principal”, algunas proteínas de dedos de zinc se unen a un triplete de tres bases de la cadena diana y a una cuarta base de la cadena no diana. Generalmente, un sitio diana tiene una longitud de al menos 9 nucleótidos y,

por consiguiente, está unido a un dominio de unión de dedos de zinc que comprende al menos tres dedos de zinc. Sin embargo, también es posible la unión de, por ejemplo, un dominio de unión de 4 dedos a un sitio diana de 12 nucleótidos, de un dominio de unión de 5 dedos a un sitio diana de 15 nucleótidos o de un dominio de unión de 6 dedos a un sitio diana de 18 nucleótidos. Como será evidente, también es posible la unión de dominios de unión más grandes (p.ej., de 7, 8, 9 o más dedos) a sitios diana de mayor tamaño.

No es necesario que un sitio diana sea múltiplo de tres nucleótidos. Por ejemplo, en los casos en los que se producen interacciones cruzadas de cadena (ver, p.ej., la Patente de EE.UU. 6.453.242 y el documento WO 02/077227), uno o más de los dedos de zinc individuales de un dominio de unión multi-dedo se pueden unir a subsitios de cuartete solapados. Como resultado, una proteína de tres dedos se puede unir a una secuencia de 10 nucleótidos, en donde el décimo nucleótido es parte de un cuartete unido a un dedo terminal, una proteína de cuatro dedos puede unirse a una secuencia de 13 nucleótidos, en donde el decimotercer nucleótido es parte de un cuartete unido a un dedo terminal, etc.

La longitud y la naturaleza de las secuencias ligando de aminoácidos entre los dedos de zinc individuales en un dominio de unión multi-dedo también afecta a la unión a una secuencia diana. Por ejemplo, la presencia de los denominados "ligandos no canónicos", "ligando largo" o "ligando estructurado" entre dedos de zinc adyacentes en un dominio de unión multi-dedo puede permitir a dichos dedos unirse a subsitios que no están inmediatamente adyacentes. Los ejemplos no limitativos de dichos ligados se describen, por ejemplo, en la Patente de EE.UU. nº 6.479.626 y en el documento WO 01/53480. Por consiguiente, se pueden separar uno o más subsitios, en un sitio diana para un dominio de unión de dedos de zinc, unos de los otros por 1, 2, 3, 4, 5 o más nucleótidos. Por mencionar un ejemplo, un dominio de unión de cuatro dedos puede unirse a un sitio diana de 13 nucleótidos que comprende, en secuencia, dos subsitios de 3 nucleótidos contiguos, un nucleótido intermedio y dos subsitios de triplete contiguos.

La distancia entre secuencias (p.ej., sitios diana) se refiere al número de nucleótidos o pares de nucleótido que intervienen entre dos secuencias, medida a partir de los bordes de las secuencias más próximas unas de otras.

En determinadas realizaciones en las que la ruptura depende de la unión de dos moléculas de fusión de dominio de dedo de zinc/semi-dominio de ruptura para separar sitios diana, los dos sitios diana pueden encontrarse en cadenas de ADN opuestas. En otras realizaciones, ambos sitios diana se encuentran sobre la misma cadena de ADN.

#### **Dominios de unión de ADN**

Se usa un dominio de unión de dedos de zinc en los métodos descritos en la presente memoria como dominio de unión de ADN.

Un dominio de unión de dedos de zinc comprende uno o más dedos de zinc. Miller *et al.* (1985) *EMBO J.* 4: 1609-1614; Rhodes (1993) *Scientific American* Feb.: 56-65; Patente de EE.UU. nº 6.453.242. Los dominios de unión de dedos de zinc descritos en la presente memoria generalmente incluyen 2, 3, 4, 5, 6 o incluso más dedos de zinc.

Típicamente, un dominio de dedos de zinc individual tiene una longitud de aproximadamente 30 aminoácidos. Los estudios estructurales han demostrado que cada dominio (estructura) de dedos de zinc contiene dos láminas beta (mantenidas en un giro beta que contiene los dos residuos de cisteína invariables) y una hélice alfa (que contiene los dos residuos de histidina invariables), que son mantenidos en una conformación particular a través de la coordinación de un átomo de zinc por las dos cisteínas y las dos histidinas.

Los dedos de zinc incluyen tanto dedos de zinc C<sub>2</sub>H<sub>2</sub> canónicos (es decir, aquellos en los que el ion de zinc está coordinado por dos residuos de cisteína y dos residuos de histidina) como dedos de zinc no canónicos tales como, por ejemplo, dedos de zinc C<sub>3</sub>H (aquellos en los que el ion de zinc está coordinado por tres residuos de cisteína y un residuo de histidina) y dedos de zinc C<sub>4</sub> (aquellos en los que el ion de zinc está coordinado por cuatro residuos de cisteína). Véase también el documento WO 02/057293 y también la Publicación de Patente de EE.UU. nº 20080182332 en relación a los ZFPs canónicos para uso en plantas.

Un dominio de unión de dedos de zinc diseñado puede presentar una nueva especificidad de unión, en comparación con una proteína de dedos de zinc natural. Los métodos de modificación incluyen, aunque sin limitación, el diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos de triplete (o cuartete) y secuencias de aminoácidos de dedos de zinc individuales, en donde cada secuencia de nucleótidos de triplete o cuartete está asociada a una o más secuencias de aminoácido de dedos de zinc que se unen a la secuencia de triplete o cuartete particular.

Los ejemplos de métodos de selección, que incluyen la presentación de fagos y los sistemas de dos híbridos, 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 patentes WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.

La potenciación de la especificidad de unión de los dominios de unión de dedos de zinc ha sido descrita, por ejemplo, en la patente compartida WO 02/077227.

Puesto que un dedo de zinc individual se une a una secuencia de tres nucleótidos (es decir, un triplete) (o a una secuencia de cuatro nucleótidos que puede solapar en un nucleótido con el sitio de unión de cuatro nucleótidos de un dedo de zinc adyacente), la longitud de una secuencia para la cual el dominio de unión de dedo de zinc se diseña para unirse (p.ej., una secuencia diana) determinará el número de dedos de zinc en un dominio de unión de dedos de zinc modificado. Por ejemplo, para ZFPs en los que las estructuras de dedos no se unen a subsitios solapantes, una secuencia diana de seis nucleótidos se une a un dominio de unión de dos dedos; una secuencia diana de nueve nucleótidos se une a un dominio de unión de tres dedos, etc. Como se indica en la presente memoria, los sitios de unión correspondientes a dedos de zinc individuales (es decir, subsitios) de un sitio diana no necesitan estar contiguos, sino que pueden estar separados por uno o varios nucleótidos, dependiendo de la longitud y la naturaleza de las secuencias de aminoácido entre los dedos de zinc (esto es, los ligandos inter-dedo) en un dominio de unión multi-dedo.

En un dominio de unión de dedos de zinc multi-dedo, los dedos de zinc adyacentes pueden estar separados por secuencias ligando de aminoácidos de aproximadamente 5 aminoácidos (denominadas ligandos inter-dedo "canónicos") o, alternativamente, por uno o más ligandos no canónicos. Ver, p.ej., las Patentes de EE.UU. compartidas nº 6.453.242 y 6.534.261. Para dominios de unión de dedos de zinc modificados que comprenden más de tres dedos, la inserción de ligandos inter-dedo más largos ("no canónicos") entre determinados de los dedos de zinc puede ser deseable en algunos casos, ya que puede aumentar la afinidad y/o la especificidad de la unión del dominio de unión. Ver, por ejemplo, la Patente de EE.UU. nº 6.479.626 y el documento WO 01/53480. Por consiguiente, los dominios de unión de dedos de zinc multi-dedo también pueden caracterizarse en relación a la presencia y localización de ligandos inter-dedo no canónicos. Por ejemplo, un dominio de unión de dedos de zinc de seis dedos que comprende tres dedos (enlazados por dos ligandos inter-dedo canónicos), un ligando largo y tres dedos adicionales (enlazados por dos ligandos inter-dedo canónicos) se clasifica como configuración 2x3. De forma similar, un dominio de unión que comprende dos dedos (con un ligando canónico entre ambos), un ligando largo y dos dedos adicionales (enlazados por un ligando canónico) se clasifica como configuración 2x2. Una proteína que comprende tres unidades de dos dedos (en cada una de las cuales los dos dedos están enlazados por un ligando canónico), y en la que cada unidad de dos dedos está enlazada a la unidad adyacente de dos dedos mediante un ligando largo, se clasifica como una configuración 3x2.

La presencia de un ligando inter-dedo no canónico o largo entre dos dedos de zinc adyacentes en un dominio de unión multi-dedo 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 en contacto con un dedo de zinc. Por ejemplo, un dominio de unión de dedos de zinc 2x2 puede unirse a dos secuencias de seis nucleótidos separadas por un nucleótido, es decir, se une a un sitio diana de 13 nucleótidos. Véase también Moore *et al.* (2001a) Proc. Natl. Acad. Sci. USA 98: 1432-1436; Moore *et al.* (2001b) Proc. Natl. Acad. Sci. USA 98: 1437-1441 y WO 01/53480.

Como se ha mencionado previamente, un subsitio diana es una secuencia de tres o cuatro nucleótidos que está unida a un único dedo de zinc. Para determinados propósitos, una unidad de dos dedos se denota como "módulo de unión". Se puede obtener un módulo de unión, por ejemplo, seleccionando dos dedos adyacentes en el contexto de una proteína multi-dedo (generalmente de tres dedos) que se une a una secuencia diana particular de seis nucleótidos. Alternativamente, se pueden construir módulos mediante ensamblaje de dedos de zinc individuales. Ver también los documentos WO 98/53057 y WO 01/53480.

La porción de dominio de ruptura de las proteínas de fusión descritas en la presente memoria puede obtenerse de cualquier endonucleasa o exonucleasa. Los ejemplos de endonucleasas a partir de las cuales se puede derivar el dominio de ruptura incluyen, aunque sin limitación, endonucleasas de restricción y endonucleasas albergadoras. Ver, por ejemplo, *2002-2003 Catalogue, New England Biolabs*, Beverly, MA; y Belfort *et al.* (1997) *Nucleic Acids Res.* **25**: 3379-3388. Se conocen otras enzimas adicionales que rompen ADN (p.ej., S1 Nucleasa; nucleasa de judía mung; ADNasa I pancreática; nucleasa micrococcal; endonucleasa HO de levadura; ver también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Se conocen ejemplos no limitantes de endonucleasas albergadoras y meganucleasas, que incluyen I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-Csml, I-PanI, I-Scell, I-Ppol, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII. Véase también la Patente de EE.UU. nº 5.420.032; la Patente de EE.UU. nº 6.833.252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25**: 3379-3388; Dujon *et al.* (1989) *Gene* **82**: 115-118; Perler *et al.* (1994) *Nucleic Acids Res.* **22**, 1125-1127; Jasin (1996) *Trends Genet.* **12**: 224-228; Gimble *et al.* (1996) *J. Mol. Biol.* **263**: 163-180; Argast *et al.* (1998) *J. Mol. Biol.* **280**: 345-353 y el catálogo de New England Biolabs. Se puede usar una o más de estas enzimas (o fragmentos funcionales de las mismas) como fuente de dominios de ruptura y semi-dominios de ruptura.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de producir unión específica de secuencia a ADN (en un sitio de reconocimiento), y ruptura de ADN en un sitio de unión, o en sus proximidades. Determinadas enzimas de restricción (p.ej., Tipo IIS) rompen ADN en los sitios eliminados del sitio de reconocimiento y presentan dominios separables de unión y ruptura. Por ejemplo, la enzima de Tipo IIS FokI cataliza la ruptura de cadena doble de ADN, a 9 nucleótidos de su sitio de reconocimiento en una cadena y a 13 nucleótidos de su sitio de reconocimiento en la otra. Ver, por ejemplo, las Patentes de EE.UU. 5.356.802; 5.436.150 y 5.487.994; así como Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89**: 4275-4279; Li *et al.*

(1993) *Proc. Natl. Acad. Sci. USA* **90**: 2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* **91**: 883-887; Kim *et al.* (1994b) *J. Biol. Chem.* **269**: 31.978-31.982. De esta manera, en una realización, las proteínas de fusión comprenden el semi-dominio de ruptura desde al menos una enzima de restricción de Tipo IIS y uno o más dominios de unión de dedos de zinc, que pueden haber sido modificados mediante ingeniería o no.

- 5 Un ejemplo de enzima de restricción de Tipo IIS, cuyo dominio de ruptura es separable del dominio de unión, es la *FokI*. Esta enzima particular es activa como dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95**: 10.570-10.575. Por consiguiente, para los propósitos de la presente descripción, la porción de la enzima *FokI* usada en las proteínas de fusión descritas se considera un semi-dominio de ruptura. Por tanto, para la ruptura dirigida de doble cadena y/o el reemplazamiento dirigido de secuencias celulares usando fusiones dedo de zinc-*FokI*, se pueden usar dos proteínas de fusión, que comprenden cada una un semi-dominio de ruptura de *FokI*, para reconstituir un dominio de ruptura catalíticamente activo. Alternativamente, también se puede usar una única molécula de polipéptido que contenga un dominio de unión de dedo de zinc y dos semi-dominios de ruptura de *FokI*. Los parámetros para la ruptura dirigida y la alteración dirigida de secuencia usando fusiones dedo de zinc-*FokI* se proporcionan en otro apartado de la presente descripción.
- 10
- 15 Un dominio de ruptura o un semi-dominio de ruptura puede ser cualquier porción de una proteína que retenga actividad de ruptura, o que retenga la capacidad de multimerizar (p.ej., dimerizar) para formar un dominio de ruptura funcional.

Los ejemplos de enzimas de restricción de Tipo IIS se describen en la Publicación Internacional compartida WO 2007/014275.

- 20 Para potenciar la especificidad de ruptura, los dominios de ruptura también pueden ser modificados. En determinadas realizaciones, se emplean variantes del semi-dominio de ruptura, dichas variantes minimizan o previenen la homodimerización de los semi-dominios de ruptura. Los ejemplos no limitativos de dichos semi-dominios de ruptura modificados se describen en detalle en el documento WO 2007/014275. Ver también los Ejemplos. En determinadas realizaciones, el dominio de ruptura comprende un semi-dominio de ruptura modificado (también denominados mutantes de dominio de dimerización) que minimiza o previene la homodimerización, y que son conocidos por los especialistas en la técnica y se describen, por ejemplo, en las Publicaciones de Patente de EE.UU. n° 20050064474 y 20060188987. Los residuos de aminoácido de las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de *FokI* son todos dianas para influir en la dimerización de los semi-dominios de ruptura de *FokI*. Ver, por ejemplo, las Publicaciones de Patente de EE.UU. n° 20050064474 y 20060188987; la Publicación de Patente Internacional WO 07/139898; Miller *et al.* (2007) *Nat. Biotechnol.* **25**(7): 778-785.
- 25
- 30

- También se pueden usar semi-dominios de ruptura modificados adicionales de la *FokI* que forman heterodímeros obligados en los ZFNs descritos en la presente memoria. En una realización, el primer semi-dominio de ruptura incluye mutaciones en los residuos de aminoácido de las posiciones 490 y 538 de *FokI* y el segundo semi-dominio de ruptura incluye mutaciones en los residuos de aminoácido 486 y 499.
- 35

En determinadas realizaciones, el dominio de ruptura comprende dos semi-dominios de ruptura, los cuales son ambos parte de un único polipéptido que comprende un dominio de unión, un primer semi-dominio de ruptura y un segundo semi-dominio de ruptura. Los semi-dominios de ruptura pueden presentar la misma secuencia de aminoácidos o diferentes secuencias de aminoácidos, siempre que funcionen para romper el ADN.

- 40 En general, se requieren dos proteínas de fusión para la ruptura si las proteínas de fusión comprenden semi-dominios de ruptura. Alternativamente, se puede usar una única proteína que comprende dos semi-dominios de ruptura. Los dos semi-dominios de ruptura pueden derivarse de la misma endonucleasa (o de fragmentos funcionales de la misma), o cada semi-dominio de ruptura puede derivarse de una endonucleasa diferente (o de fragmentos funcionales de la misma). Adicionalmente, los sitios diana para las dos proteínas de fusión preferiblemente se disponen, uno respecto del otro, de tal modo que la unión de las dos proteínas de fusión a sus respectivos sitios diana coloca los semi-dominios de ruptura en una orientación espacial relativa que permite que los semi-dominios de ruptura formen un dominio de ruptura funcional, p.ej., dimerizando. De esta manera, en determinadas realizaciones, los bordes próximos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, entre los dos sitios diana puede intervenir cualquier número entero de nucleótidos o pares de nucleótidos (p.ej., entre 2 y 50 nucleótidos o más). En general, el punto de ruptura se encuentra entre los sitios diana.
- 45
- 50

### Proteínas de fusión

Los métodos para el diseño y la construcción de proteínas de fusión (y de los polinucleótidos que las codifican) son conocidos por los especialistas en la técnica. Por ejemplo, los métodos para el diseño y la construcción de proteínas de fusión que comprenden dominios de unión a ADN (p.ej., dominios de dedos de zinc) y dominios reguladores o de ruptura (o semi-dominios de ruptura), y los polinucleótidos que codifican dichas proteínas de fusión, se describen en las Patentes de EE.UU. compartidas 6.453.242 y 6.534.261 y en las Publicaciones de Solicitud de Patente de EE.UU. 2007/0134796 y 2005/0064474.

55

En determinadas realizaciones, se construyen polinucleótidos que codifican las proteínas de fusión. Estos polinucleótidos pueden insertarse en un vector y el vector puede introducirse en una célula (ver más adelante una descripción adicional relativa a los vectores y métodos para introducir polinucleótidos en células).

5 En determinadas realizaciones de los métodos descritos en la presente memoria, una nucleasa de dedo de zinc comprende una proteína de fusión que comprende un dominio de unión de dedos de zinc y un semi-dominio de ruptura procedente de la enzima de restricción *FokI*, y dos de dichas proteínas de fusión son expresadas en una célula. La expresión de dos proteínas de fusión en una célula puede ser resultado de la administración de las dos proteínas a la célula; de la administración de una proteína y un ácido nucleico que codifica una de las proteínas a la célula; de la administración de dos ácidos nucleicos, que codifican cada uno una de las proteínas, a la célula; o de la  
10 administración de un único ácido nucleico, que codifica ambas proteínas, a la célula. En realizaciones adicionales, una proteína de fusión comprende una única cadena de polipéptido que comprende dos semi-dominios de ruptura y un dominio de unión de dedos de zinc. En este caso, una única proteína de fusión es expresada en una célula y, sin pretender establecer ninguna teoría, se cree que rompe el ADN como resultado de la formación de un dímero intramolecular de los semi-dominios de ruptura.

15 En determinadas realizaciones, los componentes de las proteínas de fusión (p.ej., las fusiones ZFP-*FokI*) están dispuestos de tal manera que el dominio de dedos de zinc está más cerca del extremo amino de la proteína de fusión, y el semi-dominio de ruptura está más cerca del extremo carboxi. Esto reproduce la orientación relativa del dominio de ruptura en dominios de ruptura dimerizantes naturales, tal como los derivados de la enzima *FokI*, en los que el dominio de unión a ADN está más próximo al extremo amino y el semi-dominio de ruptura está más próximo  
20 al extremo carboxi. En estas realizaciones, la dimerización de los semi-dominios de ruptura para formar una nucleasa funcional es producida por la unión de las proteínas de fusión a sitios de cadenas de ADN opuestas, estando los extremos 5' de los sitios de unión próximos entre ellos.

25 En realizaciones adicionales, los componentes de las proteínas de fusión (p.ej., fusiones ZFP-*FokI*) están dispuestos de tal manera que el semi-dominio de ruptura está más próximo al extremo amino de la proteína de fusión, y el dominio de dedos de zinc está más próximo al extremo carboxi. En estas realizaciones, la dimerización de los semi-dominios de ruptura para formar una nucleasa funcional es producida por la unión de las proteínas de fusión a los sitios de cadenas de ADN opuestas, estando los extremos 3' de los sitios de unión próximos entre ellos.

30 En otras realizaciones más adicionales, una primera proteína de fusión contiene el semi-dominio de ruptura más próximo al extremo amino de la proteína de fusión, y el dominio de dedos de zinc más próximo al extremo carboxi, y se dispone una segunda proteína de fusión de tal manera que el dominio de dedos de zinc está más próximo al extremo amino de la proteína de fusión, y el semi-dominio de ruptura está más próximo 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 de dedos de zinc más próximo al extremo carboxi localizado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio de dedos de zinc más próximo al extremo  
35 amino.

40 En determinadas realizaciones de las proteínas de fusión descritas, la secuencia de aminoácidos entre el dominio de dedos de zinc y el dominio de ruptura (o el semi-dominio de ruptura) se denota "ligando ZC". El ligando ZC debe distinguirse de los ligandos inter-dedo discutidos anteriormente. Véanse, p.ej., las Publicaciones de Patente de EE.UU. 20050064474A1 y 20030232410, y la Publicación de Patente Internacional WO 05/084190, para los detalles de la obtención de ligandos ZC que optimizan la ruptura.

El presente documento describe una ZFN que comprende una proteína de dedos de zinc que tiene una o más de las secuencias de aminoácido de hélice de reconocimiento mostradas en la Tabla 1. En la presente memoria también se describe un vector de expresión de ZFP que comprende una secuencia de nucleótidos que codifica una ZFP que tiene una o más de las hélices de reconocimiento mostradas en la Tabla 1.

#### 45 **Integración dirigida**

Los métodos y composiciones descritos pueden usarse para romper ADN en un gen *Zp15* de cromatina celular vegetal, lo que facilita la integración dirigida y estable de una secuencia exógena en la localización. Tal como se describe en la presente memoria, la pérdida de función de genes *Zp15* endógenos es bien tolerada por las células vegetales y las secuencias integradas en este gen son transcritas ampliamente y generan plantas con  
50 modificaciones en la línea germinal para la transmisión heredable de la secuencia integrada. Por consiguiente, el *Zp15* es un sitio deseable para la integración dirigida de secuencias exógenas.

Para la integración dirigida en *Zp15*, se diseña uno o más dominios de unión de ADN (a saber, ZFPs) para unirse a un sitio diana en el sitio de ruptura predeterminado, o en sus proximidades, y se expresa en una célula una proteína de fusión que comprende el dominio de unión a ADN diseñado y un dominio de ruptura. Tras la unión de la porción  
55 de unión a ADN (es decir, el dedo de zinc) de la proteína de fusión al sitio diana, el ADN se rompe mediante una ruptura de cadena doble, cerca del sitio diana, por acción del dominio de ruptura.

La presencia de una ruptura de cadena doble en la localización del *Zp15* facilita la integración de secuencias exógenas a través de recombinación homóloga. De esta manera, el polinucleótido que comprende la secuencia

exógena que se va a insertar en el gen *Zp15* incluirá una o más regiones de homología con un gen *Zp15* para facilitar la recombinación homóloga.

Se puede introducir cualquier secuencia de interés (secuencia exógena) en una localización *Zp15* tal como se describe en la presente memoria. Los ejemplos de secuencias exógenas incluyen, aunque sin limitación, cualquier secuencia codificadora de polipéptido (p.ej., ADNc), secuencias promotoras, potenciadoras y otras secuencias reguladoras (p.ej., secuencias de ARN interferentes, casetes de expresión de shARN, etiquetas de epítipo, genes marcadores, sitios de reconocimiento de enzimas de ruptura y diversos tipos de construcciones de expresión. Dichas secuencias se pueden obtener fácilmente usando técnicas de biología molecular estándares (clonación, síntesis, etc.) y/o están disponibles comercialmente.

Además de las moléculas de fusión descritas en la presente memoria, el reemplazamiento dirigido de una secuencia genómica seleccionada también implica la introducción de la secuencia de reemplazamiento (o donante). La secuencia donante puede introducirse en la célula antes, o concurrentemente, o posteriormente, a la expresión de la(s) proteína(s) de fusión. El polinucleótido donante contiene suficiente homología con respecto a *Zp15* como para soportar la recombinación homóloga (o reparación dirigida por homología) entre él y la secuencia genómica de *Zp15* con la que comparte homología. Aproximadamente 25, 50, 100, 200, 500, 750, 1.000, 1.500, 2.000 nucleótidos o más de homología de secuencia entre un donante y una secuencia genómica (o cualquier valor entero entre 10 y 2.000 nucleótidos, o más) apoyarán la recombinación homóloga entre ellas. En determinadas realizaciones, los brazos de homología tienen una longitud de menos de 1.000 pares de bases. En otras realizaciones, los brazos de homología tienen una longitud de menos de 750 pares de bases. Véase también la Solicitud de Patente Provisional de EE.UU. nº 61/124.047.

Las secuencias donantes pueden oscilar en longitud entre 10 y 5.000 nucleótidos (o cualquier valor entero de nucleótidos intermedio) o más. Será fácilmente evidente que la secuencia donante habitualmente no es idéntica a la secuencia genómica que reemplaza. Por ejemplo, la secuencia del polinucleótido donante puede contener uno o más cambios de bases individuales, inserciones, eliminaciones, inversiones o redistribuciones con respecto a la secuencia genómica, siempre que haya una homología suficiente con las secuencias cromosómicas. Alternativamente, una secuencia donante puede contener una secuencia no homóloga flanqueada por dos regiones de homología. Adicionalmente, las secuencias donantes pueden comprender una molécula vector que contiene secuencias que no son homólogas para la región de interés en la cromatina celular. Generalmente, la(s) región(es) homóloga(s) de una secuencia donante presentará(n) al menos un 50% de identidad de secuencia con respecto a una secuencia genómica con la que se desea la recombinación. En determinadas realizaciones, hay presente una identidad de secuencia del 60%, 70%, 80%, 90%, 95%, 98%, 99% o 99,9%. Puede haber presente cualquier valor de identidad de secuencia entre 1% y 100%, dependiendo de la longitud del polinucleótido donante.

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

Las moléculas donantes también pueden ser insertadas en la localización *Zp15* para servir como reservorio para un uso posterior. Por ejemplo, una molécula donante homóloga con respecto a un gen endógeno, pero que contiene una mutación de interés, puede ser insertada en la localización *Zp15*. A continuación, se pueden introducir ZFNs específicos del gen endógeno lo que romperá tanto la localización endógena como la molécula donante de la localización *Zp15* que contiene la mutación de interés. El DSB resultante en el genoma puede convertirse entonces en el sitio de integración para la molécula donante liberada desde la localización *Zp15*. De este modo, la eficacia de la integración dirigida de una secuencia donante en cualquier región de interés se puede aumentar enormemente, ya que el método no depende de la captación simultánea de ambos ácidos nucleicos que codifican los ZFNs y dichas secuencias donantes.

Las moléculas donantes también pueden insertarse en la localización *Zp15* para servir como sitio diana para posteriores inserciones. Por ejemplo, una molécula donante compuesta por secuencias de ADN que contienen sitios de reconocimiento para diseños adicionales de ZFN puede ser insertada en la localización *Zp15*. Posteriormente, se pueden generar diseños adicionales de ZFN y expresarse en células de tal modo que la molécula donante original se rompe y se modifica mediante reparación o recombinación homóloga. De esta manera, se pueden producir integraciones reiterativas de moléculas donantes en la localización *Zp15*.

Para simplificar los ensayos (p.ej., hibridación, PCR, digestión con enzima de restricción) para determinar la inserción con éxito de la secuencia donante, puede haber presentes determinadas diferencias de secuencia en la secuencia donante en comparación con la secuencia genómica de *Zp15*. Preferiblemente, si está localizada en una región codificadora, dichas diferencias de secuencia de nucleótidos no cambiarán la secuencia de aminoácidos, o harán cambios silenciosos de aminoácidos (es decir, cambios que no afectan a la estructura o la función de la proteína). El polinucleótido donante puede contener opcionalmente cambios en secuencias correspondientes a los sitios de unión del dominio de unión de ADN en la región de interés, para prevenir la ruptura de las secuencias donantes que han sido introducidas en la cromatina celular por recombinación homóloga.

El polinucleótido donante puede ser ADN o ARN, de cadena sencilla o de cadena doble, y puede ser introducido en una célula en forma lineal o circular. Si se introduce en forma lineal, los extremos de la secuencia donante pueden ser protegidos (p.ej., frente a la degradación exonucleolítica) mediante métodos conocidos por los especialistas en la técnica. Por ejemplo, se añade uno o más residuos de didesoxinucleótido en el extremo 3' de una molécula lineal y/o se ligan oligonucleótidos auto-complementarios en uno o en 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. Otros métodos adicionales para proteger polinucleótidos exógenos frente a la degradación incluyen, aunque sin limitación, la adición a grupo(s) aminoterminal(es) y el uso de enlaces internucleótidos modificados tales como, por ejemplo, fosforotioatos, fosforamidatos y residuos de O-metil ribosa o desoxirribosa.

Se puede introducir un polinucleótido 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, se pueden introducir polinucleótidos donantes como ácido nucleico puro, como ácido nucleico acomplejado con un agente tal como una nanopartícula, liposoma o poloxamer, o pueden administrarse con 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 cassava vein. Véase, p.ej., Chung *et al.* (2006) *Trends Plant Sci.* **11**(1): 1-4.

Parece ser que la presencia de una ruptura de cadena doble en una secuencia celular, acoplada a la presencia de una molécula de ADN exógena que presenta homología con respecto a una región adyacente a la ruptura, o de sus alrededores, activa mecanismos celulares que reparan la ruptura mediante transferencia de información de secuencia desde la molécula donante hasta la secuencia celular (p.ej., genómica o cromosómica); es decir, mediante procesos de reparación dirigida por homología, también conocidos como "conversión génica". Los métodos de los solicitantes combinan de forma ventajosa las potentes capacidades de ataque dirigido de las ZFPs modificadas con un dominio de ruptura (o semi-dominio de ruptura) para dirigirse específicamente a genes paralogos tales como los genes *Zp15* de tal manera que la ruptura de la secuencia diana produce una ruptura de cadena doble en la región del genoma en la que se desea la inserción de las secuencias exógenas.

Para la alteración de una secuencia cromosómica, no es necesario que la secuencia completa del donante sea copiada en el cromosoma, siempre que se copie una cantidad suficiente de la secuencia donante para efectuar la alteración de secuencia deseada.

La eficacia de la inserción de secuencias donantes mediante recombinación homóloga está relacionada inversamente con la distancia, en el ADN celular, entre la ruptura de cadena doble y el sitio en el que se desea la recombinación. En otras palabras, se observan mayores eficacias de recombinación homóloga cuando la ruptura de cadena doble 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 evento de recombinación deseado puede producirse a lo largo de un intervalo de la secuencia genómica), la longitud y la secuencia del ácido nucleico donante, junto con el(los) sitio(s) de ruptura, se seleccionan para obtener el evento de recombinación deseado. En los casos en los que el evento deseado se diseña para cambiar la secuencia de un único par de nucleótidos de una secuencia genómica, la cromatina celular se rompe en los primeros 10.000 nucleótidos de cada lado de dicho par de nucleótidos. En determinadas realizaciones, la ruptura se produce en los primeros 1.000, 500, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5 o 2 nucleótidos, o cualquier valor entero entre 2 y 1.000 nucleótidos, en cualquier lado del par de nucleótidos cuya secuencia se va a cambiar.

Como se ha detallado anteriormente, los sitios de unión para dos proteínas de fusión, que comprende cada una un dominio de unión de dedos de zinc y un semi-dominio de ruptura, pueden localizarse separados 5-8 o 15-18 nucleótidos, según se mide desde el borde de cada sitio de unión más próximo al otro sitio de unión, y la ruptura se produce entre los sitios de unión. Si la ruptura se produce en un sitio único o en múltiples sitios entre los sitios de unión es irrelevante, ya que las secuencias genómicas rotas son reemplazadas por las secuencias donantes. De esta manera, para una alteración eficaz de la secuencia de un único par de nucleótidos por recombinación dirigida, el punto medio de la región entre los sitios de unión se encuentra a menos 10.000 nucleótidos de dicho par de nucleótidos, preferiblemente a menos 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 un nucleótido, o en el par de nucleótidos de interés.

En determinadas realizaciones, un cromosoma homólogo puede servir como polinucleótido donante. De esta manera, por ejemplo, la corrección de una mutación en un heterocigoto se puede lograr mediante ingeniería de proteínas de fusión que se unen y rompen la secuencia mutante en un cromosoma, pero no rompen la secuencia natural del cromosoma homólogo. La ruptura de cadena doble sobre el cromosoma portador de la mutación estimula un proceso de "conversión génica" basado en homología en el que la secuencia natural del cromosoma homólogo es copiada en el cromosoma roto, restaurando así dos copias de la secuencia natural.

También se describen métodos y composiciones que pueden potenciar los niveles de recombinación dirigida que incluyen, aunque sin limitación, el uso de fusiones de dominio funcional de ZFP adicionales para activar la expresión de genes implicados en la recombinación homóloga, tales como, por ejemplo, genes de plantas del grupo de epistasis RAD54 (p.ej., *AtRad54*, *AtRad51*), y genes cuyos productos interactúan con los productos génicos



mencionados anteriormente. Véase, p.ej., Klutstein M., *et al. Genetics*, abril de 2008; **178**(4): 2389-97.

De forma similar, se pueden usar dominios de fusión funcionales de ZFP, 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). Ver, 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 los métodos para la activación y la represión de la expresión génica usando fusiones entre un dominio de unión de dedos de zinc y un dominio funcional, por ejemplo, en las Patentes compartidas de EE.UU. 6.534.261; 6.824.978 y 6.933.113. Otros métodos adicionales de represión incluyen el uso de oligonucleótidos antisentido y/o ARN interferente pequeño (siARN o ARNi) o shARNs dirigidos a la secuencia del gen que va a ser reprimido.

Como alternativa, o adicionalmente a la activación de la expresión de productos génicos implicados en la recombinación homóloga, las fusiones de dichas proteínas (o fragmentos funcionales de las mismas) con un dominio de unión de dedos de zinc dirigido a *Zp15*, pueden usarse para reclutar dichas proteínas (proteínas de recombinación) a la región de interés, aumentando de este modo su concentración local y estimulando adicionalmente los procesos de recombinación homóloga. Alternativamente, un polipéptido implicado en la recombinación homóloga como el descrito anteriormente (o un fragmento funcional del mismo) puede ser parte de una proteína de fusión triple que comprende un dominio de unión de dedos de zinc, un dominio de ruptura (o un semi-dominio de ruptura) y la proteína de recombinación (o un fragmento funcional de la misma). Las proteínas adicionales implicadas en la conversión génica y en la remodelación de cromatina relacionada con la recombinación, que pueden usarse en los métodos y composiciones mencionados anteriormente, 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 consiguen incrementos adicionales de la eficacia de recombinación dirigida, en células que comprende una molécula de fusión de dedo de zinc/nucleasa y una molécula de ADN donante, bloqueando las células de la fase G<sub>2</sub> del ciclo celular, cuando los procesos de reparación dirigidos por homología presentan la actividad máxima. Dicho arresto se puede lograr de una serie de formas. Por ejemplo, las células pueden ser tratadas con p.ej. fármacos, compuestos y/o moléculas pequeñas que influyen en la progresión del ciclo celular de tal modo que arresten a las células en la fase G<sub>2</sub>. Los ejemplos de moléculas de este tipo incluyen, aunque sin limitación, compuestos que afectan la polimerización de microtúbulos (p.ej., vinblastina, nocodazol, Taxo1), compuestos que interaccionan con ADN (p.ej., dicloruro de diamina de *cis*-plastino(II), Cisplatino, doxorubicina) y/o compuestos que afectan a la síntesis de ADN (p.ej., timidina, hidroxurea, L-mimosina, etoposide, 5-fluorouracilo). Se logran incrementos adicionales en la eficacia de recombinación mediante el uso de inhibidores de histona desacetilasa (HDAC) (p.ej., butirato sódico, tricostatina A) que alteran la estructura de cromatina para fabricar ADN genómico más accesible a la maquinaria de recombinación celular.

Métodos adicionales para el arresto del ciclo celular incluyen la sobre-expresión de proteínas que inhiben la actividad de las quinasas del ciclo celular de CDK, por ejemplo, introduciendo un ADNc que codifica la proteína en la célula o introduciendo en la célula un ZFP modificado que activa la expresión del gen que codifica la proteína. El arresto del ciclo celular también se consigue inhibiendo la actividad de ciclinas y CDKs, por ejemplo, usando métodos de ARNi (p.ej., Patente de EE.UU. n° 6.506.559) o introduciendo en la célula un ZFP modificado que reprime la expresión de uno o más genes implicados en la progresión del ciclo celular tal como, por ejemplo, ciclina y/o genes de CDK. Véase, p.ej., la Patente de EE.UU. compartida n° 6.534.261 para más información sobre los métodos de síntesis de proteínas de dedos de zinc modificadas para la regulación de la expresión génica.

Alternativamente, en determinados casos, la ruptura dirigida se lleva a cabo en ausencia de un polinucleótido donante (preferiblemente en fase S o G<sub>2</sub>), y la recombinación se produce entre cromosomas homólogos.

#### 45 **Vectores de expresión**

Un ácido nucleico que codifica una o más proteínas de fusión (es decir, ZFNs) como el descrito en la presente memoria puede clonarse en un vector para la transformación en células procarióticas o eucarióticas para replicación y/o expresión. Los vectores pueden ser vectores procarióticos, p.ej., plásmidos, o vectores lanzadera, vectores de insecto, o vectores eucarióticos. Un ácido nucleico que codifica una proteína de fusión también puede ser clonado en un vector de expresión, para la administración a una célula vegetal.

Para expresar las proteínas de fusión (es decir, ZFNs), las secuencias que codifican las proteínas de fusión típicamente son subclonadas en un vector de expresión que contiene un promotor para dirigir la transcripción. Los promotores bacterianos y eucarióticos adecuados son bien conocidos en la técnica y se describen, p.ej., en Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2ª edición, 1989; 3ª edición, 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y *Current Protocols in Molecular Biology* (Ausubel *et al.*, ver anterior). Los sistemas de expresión bacteriana para expresar la 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 se encuentran disponibles comercialmente. Los sistemas de expresión eucarióticos para células de mamífero, levaduras y células de insectos son bien conocidos por los especialistas en la técnica y también se encuentran disponibles

comercialmente.

El promotor usado para dirigir la expresión de un ácido nucleico que codifica una proteína de fusión depende de la aplicación particular. Por ejemplo, habitualmente se usa un promotor constitutivo fuerte adecuado a la célula hospedante para la expresión y la purificación de proteínas de fusión.

- 5 Por el contrario, cuando una proteína de fusión es administrada *in vivo* para la regulación de un gen vegetal (ver más adelante la sección "Administración de ácido nucleico a células vegetales"), se usa un promotor constitutivo o un promotor inducible, dependiendo del uso particular de la proteína de fusión. Los ejemplos no limitativos de promotores vegetales incluyen secuencias de promotor derivadas de ubiquitina-3 (ubi-3) de *A. thaliana* (Callis *et al.*, 1990, *J. Biol. Chem.* **265**: 12486-12493); de manopina sintasa ( $\Delta$ mas) de *A. tumifaciens* (Petolino *et al.*, Patente de EE.UU. nº 6.730.824); y/o de virus del mosaico de Cassava Vein (CsVMV) (Verdaguer *et al.*, 1996, *Plant Molecular Biology* **31**: 1129-1139). Ver también los Ejemplos.

- 15 Además del promotor, el vector de expresión habitualmente 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, tanto procarióticas como eucarióticas. De esta manera, una casete de expresión típica contiene un promotor ligado operativamente, p.ej., a una secuencia de ácido nucleico que codifica la proteína de fusión, y las señales requeridas, p.ej., para la poliadenilación eficiente del transcrito, la terminación transcripcional, los sitios de unión de ribosomas o la terminación de la traducción. Elementos adicionales de la casete pueden incluir, p.ej., potenciadores, señales de división heterólogas, y/o señal de localización nuclear (NLS).

- 20 El vector de expresión particular usado para transportar la información genética a la célula se selecciona en relación al uso pretendido para las proteínas de fusión, p.ej., expresión en plantas, animales, bacterias, hongos, protozoos, etc. (ver los vectores de expresión descritos más adelante). Los vectores de expresión bacterianos y de animales estándares son conocidos en la técnica y se describen detalladamente, por ejemplo, en la Publicación de Patente de EE.UU. 20050064474A1 y en las Publicaciones de Patente Internacionales WO05/084190, WO05/014791 y WO03/080809.

- 25 Se pueden usar métodos de transfección estándar para producir líneas celulares bacterianas, de mamífero, de levadura o de insecto que expresen grandes cantidades de proteína, que puede ser purificada a continuación empleando técnicas estándar (ver, p.ej., Colley *et al.*, *J. Biol. Chem.* **264**: 17619-17622 (1989); *Guide to Protein Purification*, en *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). La transformación de células eucarióticas y procarióticas se lleva a cabo según técnicas estándar (ver, p.ej., Morrison, *J. Bact.* **132**: 349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* **101**: 347-362 (Wu *et al.*, eds., 1983).

- 30 Se puede emplear cualquiera de los procedimientos bien conocidos para introducir secuencias de nucleótidos externas en dichas células hospedantes. Éstos incluyen el uso de transfección de fosfato cálcico, polibreno, fusión de protoplastos, electroporación, métodos ultrasónicos (p.ej., sonoporación), liposomas, microinyección, ADN desnudo, vectores de plásmido, vectores víricos, tanto episomales como integrativos, y cualquiera de los otros métodos bien conocidos para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético externo en una célula hospedante (ver, p.ej., Sambrook *et al.*, ver anterior). Solo es necesario que el procedimiento de ingeniería genética particular usado sea capaz de introducir con éxito al menos un gen en la célula hospedante capaz de expresar la proteína elegida.

#### Administración de ácido nucleico a células vegetales

- 40 Como se ha indicado anteriormente, se pueden introducir construcciones de ADN en una planta hospedante deseada (p.ej., en su genoma) mediante una variedad de técnicas convencionales. Para una revisión de dichas técnicas véase, por ejemplo, Weissbach & Weissbach, *Methods for Plant Molecular Biology* (1988, Academic Press, N.Y.) Sección VIII, páginas 421-463; y Grierson & Corey, *Plant Molecular Biology* (1988, 2ª Ed.), Blackie, Londres, Cap. 7-9.

- 45 Por ejemplo, la construcción de ADN puede introducirse directamente en el ADN genómico de la célula vegetal usando técnicas tales como la electroporación y la microinyección de protoplastos de células vegetales, o las construcciones de ADN se pueden introducir directamente en el tejido vegetal usando métodos biolíticos, tales como el bombardeo de partículas de ADN (véase, p.ej., Klein *et al.* (1987) *Nature* **327**: 70-73). Alternativamente, la construcción de ADN se puede introducir en la célula vegetal mediante transformación de nanopartículas (ver, p.ej., la Solicitud de Patente de EE.UU. nº 12/245.685).

- 50 Alternativamente, las construcciones de ADN se pueden combinar con regiones frontera/flanqueantes de ADN-T adecuadas y se pueden introducir en un vector hospedante convencional de *Agrobacterium tumefaciens*. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, que incluyen el desarme y el uso de vectores binarios, están bien descritas en la bibliografía científica. Ver, por ejemplo, Horsch *et al.* (1984) *Science* **233**: 496-498, y Fraley *et al.* (1983) *Proc. Nat'l. Acad. Sci. USA* **80**: 4803.

Adicionalmente, la transferencia génica se puede lograr usando bacterias que no son *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 cassva vein y/o virus del mosaico del tabaco. Ver, p.ej., Chung *et al.* (2006) *Trends Plant Sci.* **11**(1): 1-4.

Las funciones de virulencia del hospedante de *Agrobacterium tumefaciens* dirigirán la inserción de una cadena T que contiene la construcción y el marcador adyacente en el ADN de la célula vegetal cuando la célula sea infectada por las bacterias usando vector de ADN T binario (Bevan (1984) *Nuc. Acid Res.* **12**: 8711-8721) o el procedimiento de co-cultivo (Horsch *et al.* (1985) *Science* **227**: 1229-1231). Generalmente, el sistema de transformación de *Agrobacterium* se usa para modificar plantas dicotiledóneas (Bevan *et al.* (1982) *Ann. Rev. Genet* **16**: 357-384; Rogers *et al.* (1986) *Methods Enzymol.* **118**: 627-641). El sistema de transformación de *Agrobacterium* también puede usarse para transformar, así como para transferir, ADN a plantas y células de plantas monocotiledóneas. Ver 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**: 1677-1679; Boulton *et al.* (1989) *Plant Mol. Biol.* **12**: 31-40; y Gould *et al.* (1991) *Plant Physiol.* **95**: 426-434.

Otros métodos de transferencia génica y de transformación alternativos incluyen, aunque sin limitación, la transformación de protoplastos a través de captación de ADN desnudo mediada por calcio, polietilén glicol (PEG) o electroporación (ver Paszkowski *et al.*, *EMBO J* **3**: 2717-2722, Potrykus *et al.* (1985) *Molec. Gen. Genet.* **199**: 169-177; Fromm *et al.* (1985) *Proc. Nat. Acad. Sci. USA* **82**: 5824-5828; y Shimamoto (1989) *Nature* **338**: 274-276) y electroporación de tejidos vegetales (D'Halluin *et al.* (1992) *Plant Cell* **4**: 1495-1505). Otros métodos adicionales para la transformación de células vegetales incluyen la microinyección, la captación de ADN mediada por carburo de silicio (Kaeppler *et al.* (1990) *Plant Cell Reporter* **9**: 415-418), y el bombardeo con microproyectiles (ver 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 (p.ej., en un gen *Zp15*) de un genoma de célula vegetal. Esto es útil en tanto en cuanto que la expresión de un transgén introducido en un genoma vegetal depende críticamente de su sitio de integración. Por consiguiente, los genes que codifican, p.ej., tolerancia a herbicidas, resistencia a insectos, nutrientes, antibióticos o moléculas terapéuticas pueden insertarse, mediante recombinación dirigida, en regiones de un genoma vegetal favorables para su expresión.

Las células vegetales transformadas que son producidas mediante cualquiera de las anteriores técnicas de transformación pueden cultivarse para regenerar una planta completa que posea el genotipo transformado y de esta manera el fenotipo deseado. Dichas técnicas de regeneración se basan en la manipulación de determinadas fitohormonas en un medio de crecimiento de cultivo de tejido, basado típicamente en un marcador biocida y/o herbicida que ha sido introducido junto con las secuencias de nucleótidos deseadas. La regeneración vegetal procedente de protoplastos cultivados se describe en Evans, *et al.*, "Protoplasts Isolation and Culture" en *Handbook of Plant Cell Culture*, pág. 124-176, Macmillian Publishing Company, Nueva 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, explantes, órganos, pólenes, embriones o partes de plantas. Dichas técnicas de regeneración se describen de forma general en Klee *et al.* (1987) *Ann. Rev. of Plant Phys.* **38**: 467-486.

Los ácidos nucleicos introducidos en una célula vegetal se pueden usar para conferir los rasgos deseados en prácticamente cualquier planta. Se pueden modificar una amplia variedad de plantas y de sistemas de células vegetales para obtener las características fisiológicas y agronómicas deseadas descritas en la presente memoria usando las construcciones de ácido nucleico de la presente descripción y los diversos métodos de transformación mencionados anteriormente. En las realizaciones preferidas, las plantas y las células vegetales diana para la modificación incluyen, aunque sin limitación, las de plantas monocotiledóneas y dicotiledóneas, tal como cultivos que incluyen cultivos de grano (p.ej., trigo, maíz, arroz, mijo, cebada), cultivos frutales (p.ej., tomate, manzana, pera, fresa, naranja), cultivos de forraje (p.ej., alfalfa), cultivos de tubérculos (p.ej., zanahoria, patata, remolacha, boniato), cultivos de verduras de hoja (p.ej., lechuga, espinaca); plantas de flor (p.ej., petunia, rosa, crisantemo), coníferas y pinos (p.ej., abeto, píceas); plantas usadas en fitorremediación (p.ej., plantas de acumulación de metales pesados); cultivos oleaginosos (p.ej., girasol, colza) y plantas usadas con fines experimentales (p.ej., *Arabidopsis*). De esta manera, los métodos y composiciones descritos presentan uso en un amplio rango de plantas que incluyen, aunque sin limitación, las especies de los géneros *Asparagus*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucurbita*, *Daucus*, *Erigeron*, *Glicina*, *Gossypium*, *Hordeum*, *Lactuca*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Nicotiana*, *Orychophragmus*, *Oryza*, *Persea*, *Phaseolus*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Solanum*, *Sorghum*, *Triticum*, *Vitis*, *Vigna* y *Zea*.

El especialista en la técnica reconocerá que después de que se ha incorporado de forma estable una secuencia exógena en plantas transgénicas y se ha confirmado que es operable, puede introducirse en otras plantas mediante cruce sexual. Se puede emplear cualquiera de una serie de técnicas de reproducción estándares, dependiendo de la especie que se vaya a cruzar.

Una célula vegetal, callo, tejido o planta transformados pueden identificarse y aislarse seleccionando o realizando un escrutinio del material vegetal modificado en función de los rasgos codificados por los genes marcadores presentes en el ADN transformante. Por ejemplo, la selección se puede llevar a cabo haciendo crecer el material vegetal modificado en un medio que contenga una cantidad inhibitoria del antibiótico o herbicida para el cual la construcción

génica transformante confiere resistencia. Además, las plantas y células vegetales transformadas también pueden ser identificadas mediante un escrutinio de las actividades de cualesquier genes marcadores visibles (p.ej., los genes de  $\beta$ -glucuronidasa, luciferasa, B o C1) que pueden estar presentes en las construcciones de ácido nucleico recombinantes. Dichas metodologías de selección y escrutinio son bien conocidas por los especialistas en la técnica.

- 5 También se pueden usar métodos físicos y bioquímicos para identificar transformantes de plantas o células vegetales que contienen las construcciones génicas insertadas. Dichos métodos incluyen, aunque sin limitación: 1) Análisis Southern o amplificación de PCR para detectar y determinar la estructura del injerto de ADN recombinante; 2) Análisis Northern blot, protección de ARNasa, extensión de cebador o amplificación PCR de transcriptasa inversa para detectar y examinar transcritos de ARN de las construcciones génicas; 3) ensayos enzimáticos para detectar actividad enzimática o de ribozimas, donde dichos productos génicos están codificados por la construcción génica; 4) electroforesis de gel de proteínas, técnicas Western blot, inmunoprecipitación o inmunoensayos ligados a enzima (ELISA), donde los productos de construcción génica son proteínas. También se pueden usar otras técnicas adicionales, tales como hibridación *in situ*, tinción enzimática e inmunotinción, para detectar la presencia o la expresión de la construcción recombinante en órganos y tejidos específicos de la planta. Los métodos para llevar a cabo todos estos ensayos son bien conocidos por los especialistas en la técnica.

Los efectos de manipulación génica usando los métodos descritos en la presente memoria pueden observarse, por ejemplo, mediante northern blots del ARN (p.ej., ARNm) aislado de los tejidos de interés. Típicamente, si hay presente ARNm o ha aumentado la cantidad de ARNm, se puede asumir que el correspondiente transgén está siendo expresado. Se puede usar otros métodos para medir la actividad génica y/o la actividad del polipéptido codificado. Se pueden usar diferentes tipos de ensayos enzimáticos, dependiendo del sustrato usado y del método de detección del aumento o descenso de un producto o sub-producto de reacción. Adicionalmente, se pueden medir los niveles de polipéptido expresado inmunológicamente, es decir, ELISA, RIA, EIA y otros ensayos basados en anticuerpos bien conocidos por los especialistas en la técnica, tal como mediante ensayos de detección electroforéticos (tanto con tinción como de western blot). Como ejemplo no limitativo, la detección de las proteínas AAD-1 y PAT usando un ensayo ELISA se describe en la Solicitud de Patente de EE.UU. n° 11/587.893. El transgén puede expresarse selectivamente en algunos tejidos de la planta o en algunas etapas del desarrollo, o el transgén puede expresarse sustancialmente en todos los tejidos de la planta, sustancialmente a lo largo del ciclo vital completo. Sin embargo, también es aplicable cualquier modo de expresión combinatoria.

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

Las proteínas de fusión (es decir, ZFNs) y los vectores de expresión que codifican proteínas de fusión se pueden administrar directamente a la planta para regulación génica, ruptura dirigida y/o recombinación. En determinadas realizaciones, la planta contiene múltiples genes diana paralogos. Es sabido que las plantas pueden contener múltiples genes paralogos. De esta manera, se puede administrar una o más proteínas de fusión diferentes o vectores de expresión que codifican proteínas de fusión diferentes a una planta a fin de actuar sobre uno o más genes *Zp15* de la planta.

La administración de cantidades efectivas se realiza por cualquiera de las rutas usadas normalmente para introducir proteínas de fusión en contacto definitivo con la célula vegetal que va a ser tratada. Los ZFPs se administran de cualquier manera adecuada, preferiblemente con vehículos aceptables. Los métodos adecuados de administración de dichos moduladores están disponibles y son bien conocidos por los especialistas en la técnica, y, aunque se puede usar más de una ruta para administrar una composición particular, a menudo una ruta particular proporciona una reacción más inmediata y más efectiva que otra ruta.

También se pueden usar vehículos y están determinados en parte por la composición particular que se esté administrando, así como por el método particular usado para administrar la composición. Por consiguiente, existe una amplia variedad de formulaciones adecuadas de vehículos que están disponibles.

#### EJEMPLOS

A continuación se incluyen ejemplos de realizaciones específicas para llevar a cabo la presente descripción. Los ejemplos se ofrecen con fines meramente ilustrativos, y no pretenden limitar el alcance de la presente descripción en modo alguno.

Se ha hecho hincapié en asegurar la precisión relativa a los números usados (p.ej., cantidades, temperaturas, etc.), pero se debería considerar la existencia de un determinado grado de error experimental y desviación.

#### **Ejemplo 1: Identificación y caracterización de la localización diana de *Zp15***

En base a los mapas genéticos disponibles públicamente (Lawrence, C., *et al.* (2004) *NAR* 32: 393-397; base de datos de internet del maíz) y al borrador de secuencia genómica del maíz, se eligió la localización *Zp15* en el brazo corto del cromosoma 6 como diana para la modificación usando ZFNs en base a la localización y la información

disponible sobre el gen *Zp15*. La estructura y la secuencia genómica de un gen que codifica la beta zeína de 15 kD (*Zp15*) del maíz han sido descritas y anotadas en el dominio público (Woo *et al.* (2001) *Plant Cell* **13**(10): 2297-2313). La secuencia correspondiente al gen *Zp15* se describe en el número de acceso de GenBank AF371264.

5 La secuencia genómica *Zp15* se usó para consultar las bases de datos de genoma de TIGR y Maize GDB usando algoritmos BLAST. Se identificaron varias secuencias con homología solapante a *Zp15* que incluyen, aunque sin limitación, dos elementos contiguos, (AZM5\_16782 y ZmGSSStuc11-12-04.8785.1) y varias ESTs (M72708, M13507, M12147, AY103640 y AF371264). En base a la secuencia de estos números de acceso, así como a la secuencia de *Zp15*, se diseñaron múltiples oligonucleótidos cortos para uso como cebadores de PCR usando el programa Primer3 (Rozen, S. y Skaletsky, H.J. (2000) *Primer3 on the WWW for general users and for biologist programmers*. En: Krawetz S, Misener S (eds.) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pág. 365-386; también disponible en internet).

Estos cebadores incluyen, aunque sin limitación, los siguientes oligonucleótidos de orientación directa:

P67F 5' – CGTATGAATTCATTGACAACC – 3' (SEQ ID NO: 1)  
 P68F 5' – ATGATCTATCTGTAAATCC – 3' (SEQ ID NO: 2)  
 15 P69F 5' – CGTCATGCAACGCAACATTCC – 3' (SEQ ID NO: 3)  
 P73F 5' – AAGAACATCACAAAGTTATGC – 3' (SEQ ID NO: 4)  
 P74F 5' – TCATGTGGATCCAAGGCATC – 3' (SEQ ID NO: 5)

Adicionalmente, los cebadores incluyen, aunque sin limitación, los siguientes oligonucleótidos en orientación inversa:

P70R 5' – ATGTGTGTCGTCTTACTGC – 3' (SEQ ID NO: 6)  
 20 P71R 5' – CAGTAGTAGGGCGGAATG – 3' (SEQ ID NO: 7)  
 P72R 5' – GGGCAGCTGGTACTG – 3' (SEQ ID NO: 8)  
 P75R 5' – CTATAATCGATGTAGAGC – 3' (SEQ ID NO: 9)  
 P76R 5' – CTATGCTTTGTCTATAGTCG – 3' (SEQ ID NO: 10)

25 Todos los cebadores de oligonucleótidos fueron sintetizados y adquiridos en Integrated DNA Technologies (IDT, Coralville, IA). Se llevaron a cabo amplificaciones de ADN<sub>g</sub> a partir de maíz de variedad Hi-II en un termociclador de PCR usando 30 ng de ADN<sub>g</sub>. Se aisló un fragmento de amplificación de 2.215 pb que correspondía al gen *Zp15* de Hi-II y se clonó en el plásmido pCR2.1 (Invitrogen, Carlsbad, CA). El análisis de secuencia de dicho fragmento reveló que la estructura genómica de *Zp15* de la variedad Hi-II de maíz contiene dos exones y un intrón pequeño de 31 pb (SEQ ID NO: 126). Los diseños de ZFN dirigidos a *Zp15* se centraron en las regiones codificadoras del gen *Zp15* de la variedad Hi-II de maíz.  
 30

## Ejemplo 2: Diseño de nucleasas de dedos de zinc dirigidas al gen *Zp15* de maíz

Para montar vectores de expresión para ZFNs, se planteó un esquema de clonación modular por pasos que es aplicable a cualquier par de genes que codifican ZFN seleccionado del archivo de librería o sintetizado desde cero.

35 Primero se realizó un escrutinio de ZFNs dirigidos a *Zp15* usando el escrutinio de ensayo de levadura como se describe en Doyon *et al.* (2008) *Nature Biotechnology* **26**(6): 702 y en la Solicitud de Patente de EE.UU. n° 12/284.887. Resumidamente, la localización *Zp15* completa se introdujo en la localización HO del genoma de levadura incipiente a fin de comparar directamente la actividad de ZFN en diferentes sitios de unión dentro del gen diana; los ZFNs fueron sometidos a escrutinio en términos de su capacidad para inducir un DSB en el gen indicador empleando un ensayo indicador (MEL1) tal como se describe en Doyon *et al.*  
 40

En base a los resultados de estos ensayos de sistema proxi, se confirmó que varios pares de ZFN evaluados eran capaces de inducir DSBs dentro de *Zp15*.

Tras el pre-escrutinio de levadura, los pares de ZFN fueron subclonados a continuación en vectores de expresión específicos de maíz. Tal como se describe en la Publicación de Patente de EE.UU. n° 20080182332, un vector que incluye segmentos rediseñados y sintetizados a partir de una señal de localización nuclear (NLS) derivada de maíz *op-2* y de un dominio de nucleasa *FokI* utilizando el modo de codón de maíz fue modificado con una única inserción de nucleótido (C) por debajo del sitio único *Xho I* para crear un sitio *Sac I* extra. Se modificó un vector similar para incluir la secuencia tartamuda ribosómica 2A procedente del virus *Thosea asigna*. Las casetas génicas que codifican ORFs de proteínas de dedos de zinc individuales fueron clonadas en estos vectores a través de sitios de restricción *Kpn I* y *BamH I*, y posteriormente se combinaron dos vectores a través de sitios de restricción *Bgl II/Xho I*, dando  
 50

lugar a una construcción intermedia que contenía una casete que incluye 2 dominios que codifican ZFN flanqueados por sitios de restricción *Nco* I y *Sac* I.

La casete *Nco* I/*Sac* I procedente de dicha construcción intermedia fue escindida mediante digestión con enzima de restricción y ligada en la cadena principal del plásmido pDAB3872, que contiene un promotor del gen de *ubiquitina-1* de maíz (Sharrock *et al.* (1992) *Plant Mol Biol.* **18**(4): 675) y secuencias terminadoras procedentes del gen de peroxidasa catiónica preferencial de raíz de maíz (Patente de EE.UU. nº 7.179.902).

Los plásmidos resultantes incluyen los genes de ZFN, más los marcadores seleccionables relevantes para el mantenimiento de plásmido y sitios attL flanqueantes para la manipulación conveniente usando el sistema GATEWAY™ de Invitrogen (Carlsbad, CA). Todas las construcciones de ZFN generadas usando este esquema de clonación fueron transformadas en células CH5α de *E. coli* y posteriormente fueron mantenidas en las condiciones de selección apropiadas.

La Tabla 1 muestra ejemplos de ZFNs dirigidos a *Zp15* usados para los experimentos de integración dirigida en la localización *Zp15*. La secuencia diana de ADN para el ZFN se muestra en la segunda columna (sitios diana de ADN indicados como letras en mayúsculas; nucleótidos no contactados indicados en minúsculas), y de la tercera a la sexta columna muestran la secuencia de aminoácidos de la región de reconocimiento (aminoácidos -1 a +6, con respecto al inicio de la hélice) de cada uno de los dedos de zinc (F1 a F4) de la proteína. También se proporciona en la primera columna de la Tabla 1 un número de identificación para cada proteína.

Tabla 1

Nombre ZFN	Sitio diana	F1	F2	F3	F4
11742	cgGGGCTGCAGGGCttg tacggcgctgg (SEQ ID NO:11)	DRSHLTR (SEQ ID NO:12)	RSDNLRE (SEQ ID NO:13)	RSDVLSE (SEQ ID NO:14)	RSAHLRSR (SEQ ID NO:15)
11743	gcAGGGGCAGGGCActc gcattgcagag (SEQ ID NO:16)	QSGSLTR (SEQ ID NO:17)	RSDHLTQ (SEQ ID NO:18)	DRSHLTR (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:18)
11750	ctGAGGCAGCCGCAgtg cagcccgcctgg (SEQ ID NO:19)	QSGDLTR (SEQ ID NO:20)	DRSDLSR (SEQ ID NO:21)	QSGDLTR (SEQ ID NO:20)	RSDNLTR (SEQ ID NO:22)
11753	acTCCGCGTAGGGGta cagcccgcgggc (SEQ ID NO:23)	RSDHLRSR (SEQ ID NO:24)	RSDNLTT (SEQ ID NO:25)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11754	acTCCGCGTAGGGGta cagcccgcgggc (SEQ ID NO:23)	RSDHLSE (SEQ ID NO:29)	RNDNRKN (SEQ ID NO:30)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11755	agCCGCAGtGCAGCCc gctggcgccggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	TSSTRKT (SEQ ID NO:33)
11756	agCCGCAGtGCAGCCc gctggcgccggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	RSSTRKE (SEQ ID NO:35)
11757	ccTCAGGTACTCCGcgt aggggtacagc (SEQ ID NO:36)	RSDTLSE (SEQ ID NO:37)	ARSTRTN (SEQ ID NO:38)	QSSHLTR (SEQ ID NO:39)	QSADRTK (SEQ ID NO:40)
11758	gcCCGCTGGCGGCGgc gccctactacgc (SEQ ID NO:41)	RSDDLTR (SEQ ID NO:26)	RSDDLTR (SEQ ID NO:26)	RSDTLA (SEQ ID NO:42)	RNQDRKT (SEQ ID NO:43)
11759	gcACTGCGGCTGCCctca ggtactccgcg (SEQ ID NO:44)	DRSDLSR (SEQ ID NO:21)	QSSDLRR (SEQ ID NO:45)	RSDDLTR (SEQ ID NO:26)	QSSDLTR (SEQ ID NO:32)
11760	cgCCGGGTGTGGGCag cagagcgccatg (SEQ ID NO:46)	DRSHLSR (SEQ ID NO:47)	RSDALAR (SEQ ID NO:48)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11761	cgCCGGGTGTGGGCag cagagcgccatg (SEQ ID NO:46)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11762	gcCGGGTGTGGGCagc cgagcgccatgt (SEQ ID NO:52)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	RSDLLR (SEQ ID NO:53)	RSDNLRE (SEQ ID NO:13)
11763	agTAGGGCGCCGCGc cagcgggctgca (SEQ ID NO:54)	RSDNLTT (SEQ ID NO:25)	DRSDLSR (SEQ ID NO:21)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)
11766	tgTGGGCAGCCGAGcgc ccatgtccagc (SEQ ID NO:55)	RSDNLAR (SEQ ID NO:56)	DRSDLSR (SEQ ID NO:21)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)

11767	tgTGGGCAGCCGAGcg ccatgtccagc (SEQ ID NO:55)	RSDNLSR (SEQ ID NO:58)	DNSTRKT (SEQ ID NO:59)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
11768	cgGCGTAGTAGGGCgc cgccgccagcgg (SEQ ID NO:60)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11769	cgGCGTAGTAGGGCgc cgccgccagcgg (SEQ ID NO:60)	RSDDLTR (SEQ ID NO:26)	RSDHLTQ (SEQ ID NO:18)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11770	caGCCGCTCCGGCAac agtgtgccagc (SEQ ID NO:64)	QSGSLTR (SEQ ID NO:17)	RSDDRKT (SEQ ID NO:49)	QSSDLSR (SEQ ID NO:65)	DRSDLSR (SEQ ID NO:21)
11771	acATGGCGcTCGGCTg cccacaccggc (SEQ ID NO:66)	QSSDLSR (SEQ ID NO:65)	RNDDRKK (SEQ ID NO:67)	RSDDLTR (SEQ ID NO:26)	RSDALTQ (SEQ ID NO:68)
11772	tgGCAGCCCAGGGTctc aaccccatggc (SEQ ID NO:69)	QSSHLTR (SEQ ID NO:39)	RSDNLRE (SEQ ID NO:13)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)
11773	caGCTGCTGCTGCTgct gcatcagagt (SEQ ID NO:70)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)
11774	ctGCCAGCTACCGcac caaccctgtg (SEQ ID NO:28)	RSDLSA (SEQ ID NO:34)	DNSNRK (SEQ ID NO:50)	RSDNLSE (SEQ ID NO:57)	ASKTRKN (SEQ ID NO:63)
11775	tgTACTGGTAGAGtcc accatggcgg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDVLSE (SEQ ID NO:14)	QSGSLTR (SEQ ID NO:17)
11776	tgTACTGGTAGAGtcc accatggcgg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDALSN (SEQ ID NO:72)	TSSARTT (SEQ ID NO:73)
11777	aaCCCCTGtGGCGTctc cgctgccattc (SEQ ID NO:74)	DRSALSR (SEQ ID NO:75)	DRSHLAR (SEQ ID NO:76)	RSDTLA (SEQ ID NO:42)	DRSTRTT (SEQ ID NO:77)
11778	gtGCGGTAGCTGGGca gctggactggt (SEQ ID NO:78)	RSDHLAR (SEQ ID NO:24)	QSSDLRR (SEQ ID NO:45)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11779	gtGCGGTAGCTGGGca gctggactggt (SEQ ID NO:78)	RSDHLST (SEQ ID NO:80)	HSDTRKK (SEQ ID NO:81)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11780	agGCGGGGcTTGACCa agttggaagcgg (SEQ ID NO:82)	RSDLSV (SEQ ID NO:83)	QNHQHRIN (SEQ ID NO:84)	RSDHLAR (SEQ ID NO:24)	RSDDLTR (SEQ ID NO:26)
11781	aaATGGAAAAACGcta aaattatgtg (SEQ ID NO:85)	RSDLSK (SEQ ID NO:86)	RNDHRKN (SEQ ID NO:87)	QRSNLVR (SEQ ID NO:88)	RSDALTQ (SEQ ID NO:68)
11782	ttGTGGTGCCAACGgga gccatgctcac (SEQ ID NO:89)	RSDTLAQ (SEQ ID NO:90)	QNATRIN (SEQ ID NO:91)	RSDALS (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11783	ttGTGGTGCCAACGgga gccatgctcac (SEQ ID NO:89)	RSDTLAQ (SEQ ID NO:90)	QKATRIT (SEQ ID NO:93)	RSDALS (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11784	ttGTGGTGcCAACGgga gccatgctcac (SEQ ID NO:89)	RSDHLSE (SEQ ID NO:29)	QANRKT (SEQ ID NO:94)	RSDALS (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11785	caATCACGCCGGTAgc ggggctagtat (SEQ ID NO:95)	QSGALAR (SEQ ID NO:79)	RSDDRKT (SEQ ID NO:49)	RSDTLAQ (SEQ ID NO:90)	DSSARKK (SEQ ID NO:96)

Será evidente que los ZFNs pueden ser insertados fácilmente en cadenas principales de C2H2 o C3H y que se puede usar una variedad de secuencias para unir la proteína de dedos de zinc y el dominio de ruptura. Ver la Publicación de Patente de EE.UU. 20080182332, particularmente la Tabla 6, en relación a dichas secuencias.

### 5 Ejemplo 3: Disrupción mediada por ZFN de *Zp15* en células de maíz

Se evaluó la inducción de DSB por los pares de ZFN. Se han identificado los pares de ZFN que son capaces de producir eficientemente DSB en su sitio diana pretendido en el gen *Zp15* endógeno. La naturaleza tendente al error de la reparación DSB mediante unión de extremos no homólogos (NHEJ), que se sabe que genera pequeñas eliminaciones/inserciones de ADN en el sitio de una ruptura inducida por ZFN, se utilizó para seleccionar pares de ZFN que se unan y rompan eficazmente el sitio diana del gen *Zp15* endógeno.

Se expresaron temporalmente ZFNs en células de maíz cultivadas y se llevó a cabo un análisis de secuencia de la localización diana en el sitio de ruptura predicho. Por ejemplo, se administró un plásmido pDAB7468 que codifica el par ZFN nº 25 (hélices de reconocimiento 11768 /11766 mostradas anteriormente) a través de una transformación mediada por WHISKERS™ en cultivos de células Hi-II de maíz tal como se describe en la Solicitud de Patente de EE.UU. nº 12/001.939. Después de 24 o 72 horas de expresión temporal, la secuencia diana de ZFN alterada

resultante fue amplificada a partir de ADN genómico aislado y clonada en el vector plásmido pCR2.1. El ADN<sub>g</sub> fue sometido a digestión de restricción usando la enzima *Bsu36I*, seguido de amplificación de la secuencia diana *Zp15* y clonación de los productos de PCR en el vector plásmido pCR2.1. Se analizaron las colonias individuales del producto de amplificación clonado mediante digestiones de restricción de ADN de plásmido seguido de electroforesis de gel de agarosa (se consideró que los productos de amplificación clonados que presentaron resistencia a la ruptura mediante la enzima de restricción *Bsu36I* contenían mutaciones que destruyen el sitio de restricción asociado al sitio de ruptura de ZFN).

El análisis de secuencia directo de 192 clones reveló una inserción de 6 pb (Figura 1). En otro ejemplo, el plásmido pDAB7467 que codifica el par ZFN nº24 (hélices de reconocimiento 11753/11750 mostradas anteriormente) se administró directamente a cultivos celulares de maíz y después de 24 o 72 horas de expresión temporal, se amplificó la diana de ZFN y se clonó en el vector plásmido pCR2.1. El análisis de secuencia directo de 192 clones reveló una eliminación de 3 pb (Figura 2) en el sitio preciso de ruptura. La inserción y la eliminación descritas en la presente memoria son el resultado de la reparación de NEHJ de un DSB inducido en el sitio diana e indican que las ZFNs 24 y 25 presentan actividad de ruptura en la localización *Zp15* en células de maíz.

Se llevó a cabo el mismo proceso usando la ZFN 25 o 28, pero en lugar de colonias de escrutinio por digestión con enzima de restricción, se secuenciaron directamente 192 clones independientes. Se detectó una inserción de 6 pb (Figura 2, parte superior).

Considerados en conjunto, estos datos demuestran que una exposición temporal a las ZFNs es suficiente para inducir un DSB dirigido en la localización *Zp15* en células de maíz cultivadas.

#### Ejemplo 4: Integración dirigida en la localización *Zp15*

Para evaluar si las ZFNs diseñadas con actividad de ruptura en *Zp15* podrían conducir la integración de secuencias exógenas, construimos moléculas de ADN donante que portan una casete génica autónoma que codifica un gen de resistencia a herbicida exógeno, *AAD-1* de *Sphingomonas herbicidovorans* (ATCC 700291). El *AAD-1* codifica la enzima ariloxialcanoato dioxigenasa y confiere resistencia a herbicidas de ariloxifenoxipropionato (patentes internacionales WO 2005/107437, WO2008141154 A2). Un especialista en la técnica apreciará que se podrían incorporar de forma similar otros ácidos nucleicos exógenos en moléculas de ADN donante, que incluyen, aunque sin limitación, otros genes de tolerancia a herbicida tal como el gen relacionado *AAD-12*. En este donante génico de tolerancia a herbicida, la secuencia promotora se deriva de actina de *O. sativa* (números de acceso de GenBank S44221 y X63830) y las secuencias terminadoras se derivan de L. lipasa de *Z. mays* (número de acceso de GenBank L35913).

##### A. Construcción de molécula de ADN donante

Las construcciones donantes que contienen regiones de homología con respecto a *Zp15* fueron generadas como se indica a continuación. Se diseñó una cadena principal de plásmido que contenía flancos de homología para el gen *Zp15* para permitir la integración de cualquier secuencia de ADN donante en el correspondiente sitio diana del gen *Zp15*. La cadena principal de plásmido ejemplificada aquí se originó con el vector de plásmido bases pBC SK(-) phagemid (3,4 kpb) (Stratagene, La Jolla, CA). Dicho proceso constó de cuatro etapas.

En primer lugar, el plásmido base se preparó linealizando 3 µg de pBC SK(-) usando las endonucleasas de restricción *Spe I* y *Sal I* (New England Biolabs, Beverly, MA). El vector de subclonación de 3,3 kpb digerido con *Spe I/Sal I*, pBC SK(-) fue escindido del gel y purificado según las instrucciones del fabricante usando el Kit de Extracción de Gel QIAQUICK (QIAGEN Inc., Valencia, CA).

En segundo lugar, los flancos de homología 5' y 3' fueron aislados de *Zp15* usando los siguientes cebadores de oligonucleótidos que fueron sintetizados por Integrated DNA Technologies, Inc. (Coralville, IA):

Zp15HRDonorNotI(205)F: 5' – GCGGCCGCATGCAAGAGCTGTTGATC – 3' (SEQ ID NO: 97)

Zp15HRDonorMfeI(1025)R: 5' – CAATTGCCGGCGTAGTAGGGCGCCGCCAGC – 3' (SEQ ID NO: 98)

Zp15HRDonorMfeI(1025)F: 5' – CAATTGGTGTGGCAGCCGAGCGCCATGTTCCAG – 3' (SEQ ID NO: 99)

Zp15HRDonorSalI(2270)R: 5' – GTCGACCGATACTGATGCGGACCGTCCACCTTGTC – 3' (SEQ ID NO: 100)

Las reacciones de amplificación de PCR se llevaron a cabo usando reactivos proporcionados con el kit de PCR LA TAQ (TaKaRa Biotechnology Inc. Otsu, Shiga, Japón). El cóctel de reacción de PCR consistió en: 5 µL de Tampón II (Mg<sup>2+</sup>) 10X LA PCR<sup>TM</sup>, 20 ng de plantilla de cadena doble (ADN genómico de maíz Hi-II), 10 pmol de cebador de oligonucleótido directo, 10 pmol de cebador de oligonucleótido inverso, 8 µL de mezcla dNTP (2,5 mM de cada uno de dNTP), 33,6 µL de H<sub>2</sub>O, 0,5 µL (2,5 unidades) de ADN polimerasa *TaKaRa LA Taq<sup>TM</sup>*, 1 gota de aceite mineral. Se usó los cebadores Zp15HRDonorNotI(205)F y Zp15HRDonorMfeI(1025)R para una reacción y los cebadores Zp15HRDonorMfeI(1025)F y Zp15HRDonorSalI(2270)R fueron usados para la segunda reacción. Las reacciones de PCR se llevaron a cabo usando un termociclador de ADN de 48 muestras Perkin-Elmer Cetus (Norwalk, CT) con las



- siguientes condiciones de ciclo: 94°C, 4 min/1 ciclo; 98°C 20 s, 65°C 1 min, 68°C 1 min/30 ciclos; 72°C, 5 min/1 ciclo; 4°C/mantener. Quince (15) µL de cada reacción de PCR fueron sometidos a electroforesis y los fragmentos amplificados fueron visualizados con luz UV y el tamaño de los fragmentos se estimó por comparación con una escala de ADN de 1 kpb. Los clones de plásmido esperados fueron diagnosticados por la presencia de fragmentos de ADN de 825 pb para el fragmento 5' o de 1.250 pb para el fragmento 3'. Dichos fragmentos fueron escindidos del gel y purificados según las instrucciones del fabricante usando un Kit de Extracción de Gel QIAquick (QIAGEN Inc., Valencia, CA). A continuación los fragmentos purificados fueron clonados en plásmido pCR2.1 usando el Kit TOPO TA CLONING® y transformados en células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante.
- Se identificaron colonias individuales que contenían el fragmento 5' de 825 pb o el fragmento 3' de 1.250 pb y se confirmaron mediante digestión de enzima de restricción y con los datos de secuenciamiento. Las colonias que contenían el fragmento 5' de 825 pb fueron confirmadas mediante una digestión con enzima de restricción de *Mfe* I y *Not* I (New England Biolabs, Beverly, MA). Las colonias que contenían el fragmento 3' de 1.250 pb fueron identificadas y confirmadas mediante digestión con enzima de restricción usando *Sal* I (New England Biolabs, Beverly, MA). Los clones de plásmido esperados fueron diagnosticados por la presencia de fragmentos de ADN insertados de 825 pb (fragmento 5') o 1.250 pb (fragmento 3') además del vector pCR@2.1 de 3,9 kpb. Se llevaron a cabo reacciones de secuenciamiento de cadena doble de clones de plásmido como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones fueron purificadas usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en los protocolos del fabricante. Las reacciones de secuencia fueron analizadas en un Sistema de Análisis de ADN Beckman-Coulter CEQ™ 2000 XL y la caracterización de nucleótidos se llevó a cabo usando SEQUENCHER™ versión 4.1.4 (Gene Codes Corporation, Ann Arbor, MI). Las secuencias de los fragmentos de homología 5' y 3' procedentes de *Zp15* se indican en la SEQ ID NO: 127 y la SEQ ID NO: 128.
- En tercer lugar, se ligó el flanco de homología 3' en el plásmido base como se describe a continuación. Los clones que contenían la secuencia de flanco de homología 3' correcta fueron digeridos con enzimas de restricción y un fragmento de ADN fue escindido del gel y purificado usando el Kit de Extracción de Gel QIAquick (QIAGEN Inc., Valencia, CA). Estos fragmentos fueron ligados en un plásmido base purificado, digerido previamente con *Spe* I / *Sal* I (ver antes), en una ratio vector:injerto de 1:5 usando 500 unidades de ADN ligasa T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 µL en unas condiciones de 16 h de incubación en un baño de agua a 16°C. Cinco (5) µL de la reacción de ligación fueron transformados a continuación en células químicamente competentes de *E. coli* One SHOT® Top 10 (Invitrogen Life Technologies, Carlsbad, CA) y se colocaron en medio que contenía selección antibiótica. Se aislaron colonias putativas y se digirieron con las enzimas de restricción *Spe* I y *Sal* I (New England Biolabs, Beverly, MA) para identificar los clones que contenían el fragmento 3' ligado.
- En cuarto lugar, se ligó el flanco de homología 5' en el plásmido que contenía el flanco de homología 3'. El plásmido que contenía el flanco de homología 3', descrito anteriormente en la etapa tres, fue digerido con las endonucleasas de restricción *Mfe* I y *Not* I (New England Biolabs, Beverly, MA) durante 1 h a 37°C. Las *Not* I/*Mfe* I digirieron el plásmido que contenía el flanco de homología 3', que fue escindido del gel y purificado según las instrucciones del fabricante usando el Kit de Extracción de Gel QIAquick (QIAGEN Inc., Valencia, CA).
- Se produjeron fragmentos aislados del donante de flanco de homología 5' generados en la digestión con enzimas de restricción usando las enzimas de restricción *Mfe* I y *Not* I y fueron ligados con el plásmido que contenía el flanco de homología 3' en una reacción de ligación de 20 µL usando una ratio vector:injerto de 1:5 y 500 unidades de ADN ligasa T4 (Invitrogen Life Technologies, Carlsbad, CA). Las reacciones de ligación fueron incubadas durante 16 h en un baño de agua de 16°C.
- Después de la ligación, se transformaron 5 µL de la reacción de ligación en células químicamente competentes MAX EFFICIENCY® DH5α™ (Invitrogen Life Technologies, Carlsbad, CA) siguiendo las recomendaciones del fabricante. Se seleccionaron colonias individuales y se aisló ADN de plásmido y se digirió con la enzima de restricción *Not* I (New England Biolabs, Beverly, MA) para identificar plásmidos que contuvieran un fragmento integrado del donante de flanco de homología 5'. El plásmido resultante recibió el nombre de pDAB7489 (Figura 3).
- Se construyó una casete de expresión génica de tolerancia a herbicida que comprende una unidad transcripcional de planta (PTU) que contiene secuencias de promotor, de gen de tolerancia a herbicida y de terminación de poliadenilación (poliA). La secuencia de promotor deriva de actina 1 de *O. sativa* (McElroy *et al.* (1990) *Plant Cell* 2, 163-171; Acceso GenBank S44221 y Acceso GenBank X63830). El gen de tolerancia a herbicida comprendía el gen *AAD-1* (ariloxialcanoato dioxigenasa), que confiere resistencia a herbicidas de ariloxifenoxipropionato (WO 2005/107437). La versión del gen utilizado fue la versión nº3, que incluye una secuencia optimizada de codón para expresión en plantas. Las secuencias terminadoras derivan de L. lipasa de *Z. mays* (clon de ADNc de lipasa de maíz con Número de Acceso GenBank L35913). Esta secuencia de maíz comprende la región no traducida 3'/región terminadora de la transcripción para el gen *AAD-1*). La casete de expresión de gen de tolerancia a herbicida se muestra en la Figura 4.
- Para generar esta casete, los siguientes cebadores de oligonucleótido fueron sintetizados por Integrated DNA Technologies, Inc. (Coralville, IA) en condiciones de salinidad estándar y diluidos en agua hasta una concentración

de 0,125 µg/µL:

OsActAad1v.3ZmLipMfeIF 5' – CAATTGGTCATTCATATGCTTGAGAAGAG – 3' (SEQ ID NO: 101).

OsActAad1v.3ZmLipMfeIR 5' – CAATTGAGCACTTAAAGATCTTTAGAAG – 3' (SEQ ID NO: 102)

5 Se llevaron a cabo reacciones de amplificación de PCR usando el Kit de PCR LA TAQ (*TaKaRa* Biotechnology Inc., Otsu, Shiga, Japón). El cóctel de reacción de PCR comprendía: 5 µL de Tampón II (Mg<sup>2+</sup>) 10X LA PCR™, 20 ng de plantilla de cadena doble (ADN plásmido pDAB3878), 10 pmol de cebador de oligonucleótido directo, 10 pmol de cebador de oligonucleótido inverso, 8 µL de mezcla dNTP (2,5 mM de cada), 33,5 µL de H<sub>2</sub>O, 0,5 µL (2,5 unidades) de ADN polimerasa *TaKaRa LA Taq*™, 1 gota de aceite mineral. Las reacciones PCR se llevaron a cabo usando un termociclador de ADN de 48 muestras Perkin-Elmer Cetus (Norwalk, CT) en las siguientes condiciones de ciclo: 10 94°C, 4 min/1 ciclo; 98°C 20 s, 55°C 1 min, 68°C 3 min/30 ciclos; 72°C, 5 min/1 ciclo; 4°C/mantener. Quince (15) µL de cada reacción de PCR fueron sometidos a electroforesis a 100 V durante 1 h en un gel de agarosa TAE suplementado con un 0,5% de bromuro de etidio. Los fragmentos amplificados fueron visualizados con luz UV y el tamaño de los fragmentos se estimó por comparación con una escala de ADN de 1 kpb.

15 Los productos de PCR esperados fueron diagnosticados por la presencia de un fragmento de ADN de 2,7 kpb (PTU de *AAD-1*). Dicho fragmento fue escindido del gel y purificado según las instrucciones del fabricante usando el Kit de Extracción de Gel QIAquick (QIAGEN Inc., Valencia, CA). A continuación los fragmentos purificados fueron clonados en plásmido pCR2.1 usando el Kit TOPO TA CLONING® (con el vector pCR@2.1) y células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante.

20 Se seleccionaron colonias individuales, se aisló el ADN de plásmido y se digirió con la enzima de restricción *Mfe* I (New England Biolabs, Beverly, MA). Los clones de plásmido esperados fueron diagnosticados a través de la presencia de un fragmento de ADN insertado de 2.674 pb (PTU *AAD-1*) además del vector de 3,9 kpb pCR@2.1. Se llevaron a cabo reacciones de secuenciamiento de cadena doble de los clones de plásmido tal como describe el fabricante usando el kit CEQ™ DTCS-Quick Start Kit (Beckman-Coulter, Palo Alto, CA). Las reacciones fueron purificadas usando los cartuchos de filtración de gel Performa DTR (Edge BioSystems, Gaithersburg, MD) tal como describen los protocolos del fabricante. Las reacciones de secuencia fueron analizadas en un Sistema de Análisis Beckman-Coulter CEQ™ 2000 XL DNA y se llevó a cabo la caracterización de nucleótidos usando el SEQUENCHER™ versión 4.1.4 (Gene Codes Corporation, Ann Arbor, MI).

30 El fragmento restringido procedente de un clon que contenía la secuencia de 2.674 pb correcta fue escindido del gel y purificado siguiendo las indicaciones del fabricante con un kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). A continuación este fragmento fue combinado en una reacción de ligación con pDAB7489 purificado (cadena principal de plásmido) que había sido digerido con la enzima de restricción *Mfe* I y posteriormente había sido desfosforilado. La ligación se llevó a cabo en las siguientes condiciones: ratio vector:injerto de 1:5 y 500 unidades de ADN ligasa T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 µL en condiciones de 16 h de incubación en un baño de agua a 16°C. A continuación se transformó cinco (5) µL de la reacción de ligación en 50 µL de células químicamente competentes MAX EFFICIENCY® DH5α™ (Invitrogen Life Technologies, Carlsbad, CA) y se llevaron a placa en las condiciones de selección descritas por el fabricante.

35 Se seleccionaron colonias individuales, se aisló el ADN de plásmido y se digirió con la enzima de restricción *Mfe* I (New England Biolabs, Beverly, MA). Los clones de plásmido esperados contenían los fragmentos de ADN de 2.674 pb (PTU *AAD-1*) y de 5.413 pb (vector pDAB7489). El plásmido resultante se denominó pDAB7490 (Figura 5).

40 Se generaron cultivos celulares embriogénicos de la variedad de maíz Hi-II (Armstrong et al. (1991) *Maize Genet Coop Newsletter* 65: 92-93), se mantuvieron y se sometieron a transformación simultánea de plásmidos que codifican ZFN24 y la molécula donante pDAB7490. La transformación y selección tejido calloso y la posterior regeneración de transformantes se describe en la Solicitud de Patente de EE.UU. nº 12/001.939.

45 Para más información en relación al protocolo de transformación y selección véase Petolino *et al.* (2000) *Plant Cell Rept.* 19: 781-786. Después de la antesis, las plantas fueron auto-polinizadas o cruzadas con la variedad de maíz DAS5XH751. Las semillas de la progenie resultante fueron recolectadas y secadas. La regeneración de callo en plantas de maíz fértiles, intactas, se describe en la Solicitud de Patente de EE.UU. nº 12/001.939, en particular en el Ejemplo 22.

#### B. Integración dirigida de la casete de gen *AAD-1* en la localización *Zp15*

50 De los eventos tolerantes de herbicida que contienen una molécula donante de ADN integrada que codifica una casete génica de tolerancia a herbicida, es de esperar que alguna proporción de dichos eventos sean producto de la integración dirigida de ADN donante en el sitio de la ruptura de cadena doble inducida por la ZFN. A fin de diferenciar estos eventos de integración dirigidos de los derivados de una integración aleatoria de la casete génica de tolerancia a herbicida, se utilizó una estrategia de genotipado basada en PCR usando una combinación de 55 cebadores de PCR específicos de genoma y después de cebadores de PCR específicos de genoma más cebadores de PCR específicos de donante.

El genotipado diferencial de la integración dirigida frente a la aleatoria del transgén *AAD-1* en todos los eventos transformados tolerantes a herbicida se llevó a cabo usando ensayos basados en PCR específicos de la localización *Zp15* y del gen *AAD-1*. En los ejemplos presentados en la presente memoria, todos los cebadores de oligonucleótido fueron sintetizados por Integrated DNA Technologies, Inc. (Coralville, IA) en condiciones estándar de desalación y diluidos con agua a una concentración de 100  $\mu$ M. Se diseñó el siguiente conjunto de cebadores de oligonucleótido directos e inversos para generar secuencias de ADN<sub>g</sub> específicas de la diana de gen *Zp15* que quedan fuera de los límites de las secuencias de ADN donante:

5 HB501f: 5' – AAGGTCCCAAATCTGAGGCATACTGTTGCT – 3' (SEQ ID NO: 103)

HB502r: 5' – GAGGTCCTATGCTTTGTCTATAGTCGGCAG – 3' (SEQ ID NO: 104)

10 También se diseñó un segundo conjunto de cebadores de oligonucleótido directo e inverso para incorporarse a una secuencia de ADN<sub>g</sub> específica para la diana génica *Zp15* fuera de los límites de las secuencias de ADN donantes, pero aun así anidada dentro del primer par:

HB503 f 5' – GGCATACTGTTGCTGCCCTGCTGGAA – 3' (SEQ ID NO: 105)

HB504 r 5' – GACACCTATAATCGATGTAGAGCCGAAGAG – 3' (SEQ ID NO: 106)

15 Se diseñaron adicionalmente cebadores de oligonucleótido directos e inversos para incorporarse específicamente al ADN donante correspondientes a la región codificadora del gen de tolerancia a herbicida *AAD-1*:

HB505f 5' – AGTCCACCCAGTGATCTCAGCACCA – 3' (SEQ ID NO: 107)

HB506f 5' – AGTGGCTGGACAGCTATTCTCTCAAAGCGT – 3' (SEQ ID NO: 108)

HB507r 5' – ACGCTTTGAGAGAATAGCTGTCCAGCCACT – 3' (SEQ ID NO: 109)

20 HB508r 5' – TGGTGCTGAGATCACTGGGGTGGACT – 3' (SEQ ID NO: 110)

Se llevaron a cabo dos reacciones de amplificación primaria distintas utilizando cebadores que se unen a la región genómica de *Zp15* y la molécula donante, dando lugar a un amplicón que extiende la frontera de la integración entre el genoma y el donante. La primera reacción se focalizó sobre la frontera 5' entre el genoma y el donante y usó el conjunto de cebadores HB501f y HB507r. La segunda reacción se focalizó en la frontera 3' entre el donante y el genoma y usó el conjunto de cebadores HB505f y HB502r. Se aisló ADN genómico de los sucesos de maíz Hi-II transformado. Las reacciones de amplificación de PCR primarias se llevaron a cabo usando reactivos suministrados en el kit de PCR LA TAQ (*TaKaRa* Biotechnology Inc., Otsu, Shiga, Japón). El cóctel de reacción PCR consistió en: 2,5  $\mu$ L de tampón 10X La Taq PCR<sup>TM</sup>, 40-200 ng de plantilla de ADN genómico de cadena doble, cebador de oligonucleótido directo 10  $\mu$ M, cebador de oligonucleótido inverso 10  $\mu$ M, 2  $\mu$ L mezcla dNTP (2,5 mM de cada), 16,25  $\mu$ L de H<sub>2</sub>O, 0,25  $\mu$ L (1,25 unidades) de ADN polimerasa LA Taq<sup>TM</sup>. Las reacciones de PCR se llevaron a cabo usando un Termociclador Peltier de ADN Engine Tetrad2, de 96 muestras (Hercules, CA), usando las siguientes condiciones de ciclo: 94°C, 2 min/1 ciclo; 94°C 30 s, 62°C 30 s, 68°C 5 min/30 ciclos; 4°C/mantener.

Los productos de la reacción de PCR primaria fueron diluidos a continuación 1:100 en H<sub>2</sub>O y se usaron como ADN plantilla para dos reacciones de PCR secundarias distintas. Las reacciones secundarias también utilizan cebadores que se unen en la región genómica de *Zp15* y la molécula donante, dando lugar a un amplicón que extiende la frontera de integración entre genoma y donante. La identidad de los cebadores específicos determina si la amplificación está focalizada en la frontera 5' o en la 3' entre el genoma y el donante. La primera reacción se focalizó sobre la frontera 5' entre el genoma y el donante y usó el conjunto de cebadores HB503 f y HB508r. La segunda reacción se focalizó en la frontera 3' entre el donante y el genoma y usó el conjunto de cebadores HB506f y HB504r. Ambas reacciones consistieron en lo siguiente: 2,5  $\mu$ L de tampón 10X La Taq PCR<sup>TM</sup>, 1  $\mu$ L de plantilla [dilución 1:50 de la 1ª reacción de PCR], cebador de oligonucleótido directo 10  $\mu$ M, cebador de oligonucleótido inverso 10  $\mu$ M, 2  $\mu$ L mezcla dNTP (2,5 mM de cada), 16,25  $\mu$ L de H<sub>2</sub>O, 0,25  $\mu$ L (1,25 unidades) de ADN polimerasa LA Taq<sup>TM</sup>. Las reacciones de PCR se llevaron a cabo usando un Termociclador Peltier de ADN Engine Tetrad2, de 96 muestras (Hercules, CA), usando las siguientes condiciones de ciclo: 94°C, 2 min/1 ciclo; 94°C 30 s, 62°C 30 s, 68°C 5 min/30 ciclos; 4°C/mantener. Se observaron los fragmentos de amplicón de PCR esperados de 2.180 pb para la frontera 5' o de 2.980 pb para la frontera 3'.

Dichos fragmentos fueron escindidos del gel y purificados según las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Los fragmentos purificados fueron clonados posteriormente en plásmido pCR2.1 usando el kit TOPO TA CLONING® (con el vector pCR®2.1) y células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante.

Se seleccionaron colonias individuales, se aisló ADN de plásmido y se digirió con la enzima de restricción *Eco* RI (New England Biolabs, Beverly, MA). Los clones de plásmido esperados fueron diagnosticados por la presencia de fragmentos de ADN insertados del tamaño apropiado además del vector pCR®2.1 de 3,9 kpb.

Se llevaron a cabo reacciones de secuenciamiento de cadena doble de clones de plásmido y se realizó la

caracterización de nucleótidos y alineamientos usando la versión 4.1.4 de SEQUENCHER™ (Gene Codes Corporation, Ann Arbor, MI).

Los datos de secuencia seleccionada derivados de un evento de integración dirigido (evento n° 147) de la casete génica de donante *AAD-1* insertada en el gen diana *Zp15* se muestran en el alineamiento de la Figura 6.

- 5 Los productos de amplificación de la PCR primaria focalizados en las fronteras 5' o 3' entre el genoma y el donante fueron sometidos a una amplificación secundaria también focalizada en las fronteras 5' o 3' entre el genoma y el donante. El alineamiento de los fragmentos clonados correspondientes a dichos productos de amplificación secundarios con la secuencia genómica de *Zp15* natural, así como la secuencia esperada de un evento de integración dirigido, claramente indica que se ha producido una integración precisa de ADN donante en el sitio diana.
- 10 La secuencia de nucleótidos de la localización genómica *Zp15*, la frontera genoma/donante, la secuencia de nucleótidos de las regiones donantes correspondientes a los flancos de homología de *Zp15* y la secuencia de nucleótidos de la casete de tolerancia a herbicida, fueron preservadas todas en los múltiples productos de PCR clonados derivados de este evento.

- 15 Por lo tanto, este evento representa un genoma en el que se ha producido la reparación dirigida por homología de una ruptura de cadena doble mediada por una ZFN y la integración dirigida de un ADN donante en un gen diana específico. En la Figura 6, se muestran los datos de alineamiento de secuencia derivados de un único callo de maíz transformado aislado representativo (evento n° 147).

Se han obtenido otros eventos transformados adicionales que representan integraciones dirigidas únicas, lo que demuestra que los métodos incluidos en la presente memoria son reproducibles en callos de maíz.

## 20 **Ejemplo 5: Integración dirigida de PAT en la localización *Zp15***

- Como ejemplo adicional de los métodos de integración dirigida de polinucleótidos exógenos seleccionados en localizaciones diana descritos en la presente memoria, se diseñaron y construyeron moléculas de ADN donante adicionales que portan una casete génica autónoma que codifica *PAT*, para evaluar la integración de una secuencia donante exógena dentro de la localización diana *Zp15* endógena. Se desplegaron ZFNs con actividad de ruptura específica de la diana génica *Zp15* para crear un DSB en dicha localización. A continuación se integraron moléculas de ADN donante que portan una casete génica autónoma que codifica *PAT*, procedente de *Streptomyces viridochromogenes*, y secuencias flanqueantes homólogas a *Zp15*, en el DSB inducido por ZFN de la diana *Zp15*. *PAT* codifica la enzima fosfinotricina acetil transferasa y confiere resistencia al compuesto herbicida fosfinotricina (PPT) por acetilación (Patente de EE.UU. n° 5.633.434). La fosfinotricina es el ingrediente activo de los herbicidas LIBERTY, BASTA e IGNITE. Se construyó la secuencia codificadora de *PAT* como una unidad de transcripción de planta (PTU) y contenía una secuencia promotora derivada de actina de *O. sativa* (números de acceso del GenBank S44221 y X63830) y una secuencia terminadora derivada de L. lipasa de *Z. mays* (número de acceso del GenBank L35913).
- 25
- 30

### A. Construcción de la molécula de ADN donante

- 35 Se generó sintéticamente una construcción donante de *Zp15* que contiene regiones de homología con *Zp15* como se indica a continuación. Se diseñó una región de homología de *Zp15* que comprende los nucleótidos 4595-5346 (secuencia de homología 5'; SEQ ID NO: 111) y los nucleótidos 21-796 (secuencia de homología 3'; SEQ ID NO: 112) a partir de pDAB7489. Esta región de homología incluyó un sitio de clonación *Mfe* I entre los elementos homólogos 5' y 3' y sitios de restricción *Not* I en los extremos 5' y 3'. Esta secuencia de ADN (SEQ ID NO: 113) se sintetizó y se insertó en el plásmido de tipo ColE1 resistente a kanamicina, pMK en los sitios de clonación *Sac* I y *Kpn* I (Gene Art Ag, Regensburg, Alemania). El plásmido donante resultante se designó pDAB104101 (Figura 7) y contiene flancos de homología correspondientes al gen *Zp15* para permitir la integración de cualquier secuencia de ADN de interés en el correspondiente sitio diana del gen *Zp15*.
- 40

- Se construyó una casete de expresión génica de tolerancia a herbicida que comprende una PTU completa que contiene un promotor, un gen de tolerancia a herbicida y secuencias de terminación de poli adenilación (poliA). La secuencia de promotor se deriva de actina 1 de *O. sativa* (McElroy *et al.* (1990) Plant Cell 2: 163-171). (Número de acceso del GenBank S44221 y número de acceso del GenBank X63830). El gen de tolerancia a herbicida fue *PAT*. Las secuencias terminadoras se derivaron de L. lipasa de *Z. mays* (clon de ADNc de lipasa de maíz con número de acceso del GenBank L35913). Esta secuencia de maíz comprende la región no traducida 3'/región de terminador de la transcripción correspondiente al gen *PAT*. La casete de expresión génica de tolerancia a herbicida *PAT* se muestra en la Figura 8.
- 45
- 50

La casete de gen *PAT* se amplificó a partir de un plásmido pDAB102256 mediante PCR usando cebadores sintetizados por Integrated DNA Technologies, Inc. (Coralville, IA):

DC001 5' – CCAGTGCAATTGGGTCATTCATATGCTTGAGAAG – 3' (SEQ ID NO: 114)

- 55 DC002 5' – CCAGTGCAATTGAATTCAGCACTTAAAGATCTTTAG – 3' (SEQ ID NO: 115)

Las reacciones de amplificación de PCR se llevaron a cabo usando ADN polimerasa PHUSION™ (New England Biolabs, Beverly, MA) con las siguientes condiciones de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 s, 60°C 20 s, 72°C 45 s/9 ciclos; 98°C 10 s, 72°C 60 s/24 ciclos; 72°C, 10 min/1 ciclo; 4°C/mantener. La reacción de PCR se analizó mediante electroforesis en un gel de agarosa de TAE al 1,0%.

- 5 Los productos de PCR esperados fueron diagnosticados por la presencia de un fragmento de ADN de 2,3 kpb (PTU de *PAT*). Este fragmento se escindió y se purificó del gel siguiendo las indicaciones del fabricante usando el kit de extracción de gel QIAQUICK™ (QIAGEN Inc., Valencia, CA). A continuación los fragmentos purificados fueron clonados en el vector pCR-Blunt II-TOPO usando el kit de clonación de PCR Zero Blunt TOPO y células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA).
- 10 Se seleccionaron colonias individuales y se aisló ADN de plásmido y se sometió a digestión enzimática de restricción de ADN de plásmido y a análisis de secuencia. Los injertos de *PAT* clonados fueron secuenciados para demostrar la identidad y la fidelidad de secuencia de los productos de PCR clonados (SEQ ID NO: 116). Uno de dichos clones de plásmido fue designado pDAB104107 (Figura 9) y se usó a continuación como fuente de la casete del gen *PAT* para la inserción en los nuevos vectores donantes de homología de *Zp15*.
- 15 Se recuperó el fragmento génico de *PAT* de 2,3 kpb del pDAB104107 mediante digestión con *Mfe* I seguida de electroforesis de gel, escisión y purificación. También se digirió el plásmido pDAB104101 donante de homología de *Zp15* con *Mfe* I y se purificó en gel. La ligación del fragmento de gen de *PAT* en pDAB104101 dio lugar a clones en los que el gen *PAT* se insertó en el sitio *Mfe* I en cualquiera de las dos orientaciones con respecto a las secuencias génicas de *Zp15*, según se determina mediante digestión enzimática de restricción diferencial. pDAB104104 (Figura 10) comprendía el gen *PAT* insertado en la misma orientación transcripcional que el gen *Zp15*. pDAB104105 (Figura 11) comprendía el gen *PAT* insertado en la orientación opuesta respecto al gen *Zp15*.
- 20

#### B. Construcción de la molécula de ADN donante adicional

- Se generó otra construcción donante que contenía las regiones de homología a *Zp15* en la que la secuencia de homología *Zp15* 3' en pDAB7489 fue alterada mediante truncamiento, mientras que la secuencia de homología *Zp15* 5' en pDAB7489 se mantuvo igual. La región de homología 3' truncada fue generada a partir de pDAB7489 mediante PCR usando cebadores sintetizados por Integrated DNA Technologies, Inc. (Coralville, IA):
- 25

DC003 5' – CTAATCGTCTCGACTCGTCAAGCCCCGCCTTTAAAT – 3' (SEQ ID NO: 117)

DC004 5' – CTAATCCAATTGGTGTGGGCAGCCGAGCG – 3' (SEQ ID NO: 118)

- Se llevaron a cabo reacciones de amplificación de PCR usando ADN polimerasa PHUSION HOT START (New England Biolabs, Beverly, MA) con las siguientes condiciones de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 s, 72°C 15 s/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/mantener. La reacción de PCR se sometió a electroforesis en un gel de agarosa de TAE al 1,0%.
- 30

- Los productos de PCR esperados fueron diagnosticados por la presencia de un fragmento de ADN de 0,8 kpb (homología *Zp15* 3'). Este fragmento fue escindido de gel y purificado siguiendo las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). A continuación los fragmentos purificados fueron clonados en el vector pCR-Blunt II-TOPO usando el kit de clonación de PCR Zero Blunt TOPO y células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) según el protocolo del fabricante.
- 35

- Se seleccionaron colonias individuales y se sometieron a aislamiento de ADN de plásmido y a confirmación mediante digestión con enzima de restricción. Los injertos de homología 3' de *Zp15* fueron secuenciados para demostrar la identidad y la fidelidad de secuencia de los productos de PCR clonados (SEQ ID NO: 119). Uno de dichos clones de plásmido fue designado pDAB104106 (Figura 12) y se usó a continuación como fuente para la nueva secuencia de homología 3' de *Zp15* para la sustitución en pDAB7489 para crear un nuevo vector donante de homología *Zp15*.
- 40

- Se digirió el pDAB7489 secuencialmente con *Mfe* I y *Sal* I y se purificó en gel el fragmento de vector de 4,2 kpb. También se digirió el pDAB104106 con *Mfe* I y *Sal* I y se purificó en gel el fragmento de 0,8 kpb que comprendía la secuencia de homología 3' de *Zp15* truncada. La ligación y la transformación de estos fragmentos purificados en gel dio lugar a clones en los que la secuencia de homología 3' de *Zp15* truncada fue sustituida por la secuencia de homología 3' de *Zp15* original. Este plásmido se designó pDAB104100 (Figura 13) y se usó como receptor para el gen *PAT*.
- 45
- 50

- Se eliminó el gen *PAT* de pDAB104107 mediante digestión con *Mfe* I. Tras una electroforesis de gel, se purificó en gel el fragmento de gen *PAT* de 2,3 kpb. El plásmido donante de homología de *Zp15* pDAB104100 también fue digerido con *Mfe* I y purificado en gel. La ligación del gen *PAT* en pDAB104100 dio lugar a clones en los que el gen *PAT* fue insertado en el sitio *Mfe* I en cualquiera de las dos orientaciones con respecto al gen *Zp15*, según se determina mediante digestión con enzimas de restricción. El pDAB104103 (Figura 14) comprendía el gen *PAT* insertado en la misma orientación transcripcional que el gen *Zp15*. El pDAB104102 (Figura 15) comprendía el gen
- 55

*PAT* insertado en la orientación opuesta respecto al gen *Zp15*.

C. Transformación de maíz y recuperación de inserciones *PAT* dirigidas a *Zp15*.

Se generaron cultivos celulares embrionarios de la variedad de maíz Hi-II (Armstrong *et al.* (1991) *Maize Genet Coop Newsletter* **65**: 92-93), se mantuvieron y sometieron a transformación simultánea de plásmidos que codifican ZFN24 y la molécula donante. Las moléculas donantes incluyen las descritas en la presente memoria; pDAB104102, pDAB104103, pDAB104104 y pDAB104105. La transformación y la selección de tejido de callo y la posterior regeneración de transformantes se describen en la Solicitud de Patente de EE.UU. n° 12/001.939, en concreto en el Ejemplo 19. Para más información relativa al protocolo de transformación y selección, véase Petolino *et al.* (2000) *Plant Cell Rept.* **19**: 781-786. Después de la antesis, las plantas pueden auto-polinizarse o ser cruzadas con una variedad de maíz tal como la DAS5XH751. Las semillas de progenie resultantes pueden ser recolectadas y secadas, y las plantas procedentes de dichas semillas se pueden analizar para demostrar la heredabilidad de los eventos de integración dirigidos. La regeneración de callos en plantas de maíz intactas fértiles se describe en la Solicitud de Patente de los EE.UU. n° 12/001.939, en concreto en el ejemplo 22.

D. Identificación de inserciones de *PAT* dirigidas a *Zp15*

Las inserciones de *PAT* dirigidas a *Zp15* en tejido de callo transformado son detectadas mediante PCR. Se extrae ADN genómico de plantilla del tejido de callo mediante métodos bien conocidos y utilizados habitualmente, tal como el kit Plant DNEASY (QIAGEN Inc., Valencia, CA) o el método de Dellaporta (Dellaporta *et al.*, (1993) *Plant Mol. Biol. Rep.* **1**; 19-21). El uso de cebadores específicos de *PAT* en conjunción con los cebadores de secuencia flanqueantes de *Zp15* ya utilizados para detectar la integración dirigida de *AAD-1* en la localización *Zp15* da como resultado la amplificación de las uniones de inserción dirigidas a *PAT* en las regiones de homología 5' y 3' de *Zp15*. Los cebadores específicos de *PAT* pueden ser:

DC013 5' – CAATCGTAAGCGTTCCTAGCCTTCCAG – 3' (SEQ ID NO: 120)

DC014 5' – CTGGAAGGCTAGGAACGCTTACGATTG – 3' (SEQ ID NO: 121)

Específicamente, los cebadores HB501f o HB503f de la región genómica que flanquea la secuencia de homología 5' de *Zp15* de ADN donante se usan en conjunción con el cebador DC013 (SEQ ID NO: 120) en la región codificadora de la proteína *PAT* para detectar las intersecciones de injerto 5' *PAT* – *Zp15* cuando los ADNs donantes tienen el gen *PAT* en la orientación directa respecto al gen *Zp15*. Del mismo modo, los cebadores HB501f o HB503f se usan en conjunción con el cebador DC104 (SEQ ID NO: 121) en la región codificadora de la proteína *PAT* para detectar las intersecciones de injerto 5' *PAT* – *Zp15* cuando los ADNs donantes tienen el gen *PAT* en la orientación indirecta respecto al gen *Zp15*. Para la detección de las intersecciones de injerto 3' *PAT* – *Zp15*, se usan los cebadores HB502r o HB504r en conjunción con el cebador DC013 (SEQ ID NO: 120) para detectar la intersecciones de injerto cuando los ADNs donantes presentan el gen *PAT* en la orientación indirecta respecto al gen *Zp15*. Del mismo modo, se usan los cebadores HB502r o HB504r en conjunción con el cebador DC014 (SEQ ID NO: 121) para detectar la intersecciones de injerto cuando los ADNs donantes presentan el gen *PAT* en la orientación directa respecto al gen *Zp15*.

Las reacciones de amplificación de PCR se llevan a cabo usando ADN polimerasa PHUSION HOT START (New England Biolabs, Beverly, MA) con las siguientes condiciones de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 s, 72°C 15 s/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/mantener. Los productos de PCR se resuelven y se identifican usando electroforesis de gel de agarosa TAE. Los tamaños de los fragmentos de gel esperados para los productos de PCR procedentes de eventos de integración dirigida *PAT* – *Zp15* en callos transgénicos generados usando los diferentes ADNs donantes *PAT* – *Zp15* son como se indica a continuación:

HB501f/HB503f + DC013 (5') = 2,6 kpb (pDAB104103), 2,6 kpb (pDAB104104)

HB501f/HB503f + DC014 (5') = 1,6 kpb (pDAB104105), 1,7 kpb (pDAB104102)

HB502r/HB504r + DC013 (3') = 3,2 kpb (pDAB104105), 3,2 kpb (pDAB104102)

HB502r/HB504r + DC014 (3') = 1,2 kpb (pDAB104103), 1,2 kpb (pDAB104104)

Los productos de PCR que comprenden las intersecciones de integración dirigida *PAT* – *Zp15* 5' y 3' son clonados y secuenciados usando métodos estándar conocidos por el especialista en la técnica. Por ejemplo, los productos de PCR se purifican a partir del gel de agarosa y se clonan en el plásmido pCR-BluntII TOPO usando el kit TOPO BLUNT CLONING® y células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante.

Las intersecciones de integración clonadas son secuenciadas a continuación para demostrar que el gen *PAT* se inserta en el genoma de maíz en la localización *Zp15* mediante recombinación homóloga a través de las secuencias de homología 5' y 3' de *Zp15* que son incorporadas en los vectores de transformación donantes. Además de los cebadores específicos de vector TOPO, M13 directo y M13 inverso, se usan cebadores específicos de la casete

génica de PAT para obtener la secuencia completa de los clones de integración objetivo. Los cebadores de PCR SEQ ID NO: 120 y SEQ ID NO: 121 que son específicos de la secuencia codificadora de la proteína PAT también son usados como cebadores de secuenciación. Adicionalmente, también se pueden usar otros cebadores para el secuenciación. Éstos incluyen, aunque sin limitación, aquellos que son específicos del elemento promotor de actina de arroz de la casete de PAT:

5

DC-S1 5'-CCAACTGGACAATAGTCTCAAC-3' (SEQ ID NO: 122)

DC-S2 5'-CATCGCCACTATATACATACC-3' (SEQ ID NO: 123)

y aquellos que son específicos de la secuencia codificadora de la proteína PAT:

DC-S3 5'-CGTCTCAATGTAATGGTTAACG-3' (SEQ ID NO: 124)

10

DC-S4 5'-GCCCAGCGTAAGCAATACCAG-3' (SEQ ID NO: 125)

**Ejemplo 6: Heredabilidad de la integración dirigida de *AAD-1* en la localización *Zp15***

Se generó un evento de callo transgénico que portaba la casete génica de *AAD-1* dirigida a la localización *Zp15* (Evento 138). Las plantas  $T_0$  del evento 138 fueron regeneradas y cruzadas como hembras con machos DAS5XH751. Las semillas  $T_1$  resultantes fueron plantadas y se cultivaron plantas  $T_1$ .

15

Las plantas  $T_1$  fueron analizadas mediante PCR para demostrar la existencia de la integración dirigida *AAD-1* – *Zp15*. Las reacciones de amplificación de PCR fueron llevadas a cabo usando ADN polimerasa PHUSION HOT START (New England Biolabs, Beverly, MA) con las siguientes condiciones de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 s, 72°C 15 s/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/mantener. Las reacciones de PCR fueron analizadas mediante electroforesis en geles de agarosa TAE al 1,0%. Se extrajo ADN genómico de una planta  $T_1$  y se usó como plantilla de ADN en las reacciones de PCR con cebadores anidados diseñados para detectar la intersección de integración 5' *AAD-1* *Zp15*. Se llevaron a cabo reacciones de PCR primarias y secundarias usando los mismos cebadores usados antes en el análisis del evento de callo 147. Para la PCR primaria, se usaron los cebadores HB501f y HB507r. La PCR primaria dio lugar a una banda en el tamaño esperado de 2,2 kpb. Se diluyó una alícuota de la reacción de PCR primaria 1:100 y se usó como plantilla en la PCR secundaria usando cebadores anidados HB503f y HB508r. La PCR secundaria también dio lugar a una banda en el tamaño esperado de 2,2 kpb.

20

25

El producto de PCR secundaria de 2,2 kpb fue clonado en pCR-Blunt II-TOPO usando el kit de clonación ZERO BLUNT TOPO PCR y células de *E. coli* químicamente competentes ONE SHOT TOP10 (Invitrogen Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante. El ADN clonado fue secuenciado usando cebadores específicos de vector flanqueante (M13 directo y M13 inverso). Se observó que la secuencia era idéntica a la esperada para una integración dirigida de *AAD-1* en la localización *Zp15*.

30

## REIVINDICACIONES

1. Un método para integrar una o más secuencias de ácido nucleico exógenas en el genoma de una célula vegetal, método que comprende:

5 introducir la una o más secuencias de ácido nucleico exógenas en la célula vegetal, en donde las secuencias de ácido nucleico exógenas incluyen una o más regiones de homología con un gen *Zp15*;

y

romper una secuencia de cadena doble en un gen endógeno *Zp15* en el genoma de la célula vegetal usando al menos una proteína de fusión que comprende un dominio de ruptura y una proteína de dedos de zinc que comprende las regiones de hélice de reconocimiento mostradas en una única fila en la Tabla 1,

10 dando como resultado de este modo la integración de un polinucleótido que comprende la una o más secuencias exógenas en el genoma de la célula,

en donde la ruptura de cadena doble se realiza

15 (a) expresando una primera proteína de fusión en la célula, comprendiendo dicha primera proteína de fusión un primer dominio de unión de dedos de zinc y un primer semi-dominio de ruptura, en donde el primer dominio de unión de dedos de zinc ha sido diseñado para unirse a un primer sitio diana en la localización de *Zp15*, o en sus proximidades, del genoma de la célula vegetal; y

20 (b) expresar una segunda proteína de fusión en la célula, comprendiendo dicha segunda proteína de fusión un segundo dominio de unión de dedos de zinc y un segundo semi dominio de ruptura, en donde el segundo dominio de unión de dedos de zinc se une a un segundo sitio diana en la localización *Zp15*, o en sus proximidades, del genoma de la célula vegetal, en donde el segundo sitio diana es diferente del primer sitio diana; y en donde la unión de la primera proteína de fusión al primer sitio diana, y la unión de la segunda proteína de fusión al segundo sitio diana, posicionan los semi dominios de ruptura de tal modo que se rompe el genoma de la célula vegetal en la localización *Zp15*, o en sus proximidades.

25 2. El método de la reivindicación 1, que además comprende expresar un producto de la una o más secuencias exógenas.

3. El método de la reivindicación 1 ó 2, en donde los semi-dominios de ruptura son naturales o no naturales.

4. El método de cualquiera de las reivindicaciones 1 a 3, en donde la una o más secuencias de ácido nucleico exógenas comprenden una secuencia codificadora, una secuencia reguladora o un sitio diana para un dominio de unión de ADN.

30 5. El método de la reivindicación 4, en donde la secuencia codificadora codifica para un producto que confiere: resistencia a herbicida; tolerancia a herbicida; resistencia a insecto; tolerancia a insecto; resistencia a enfermedad; tolerancia a enfermedad; tolerancia a estrés; resistencia a estrés; un cambio en el estrés oxidativo; aumento del rendimiento de aceite; un cambio del contenido y constitución alimentaria; un cambio de la apariencia física; esterilidad masculina; secado; capacidad de permanecer erguida; capacidad de ser prolífica; un cambio en la cantidad y calidad del almidón; un cambio en la calidad del aceite; un cambio en la cantidad o calidad de las proteínas; un cambio en la composición de aminoácidos, o combinaciones de los mismos.

35 6. El método de cualquiera de las reivindicaciones 1 a 5, en donde el polinucleótido además comprende secuencias de nucleótido que son homólogas a las secuencias de la localización *Zp15*.

40 7. El método según la reivindicación 6, en donde las secuencias de nucleótidos homólogas flanquean la secuencia exógena.

8. El método de cualquiera de las reivindicaciones 1 a 7, en donde el polinucleótido además comprende un promotor.

9. El método de cualquiera de las reivindicaciones 1 a 8, en donde una o más de las secuencias exógenas integradas son transmitidas a la progenie en las generaciones subsiguientes.

45 10. El método de cualquiera de las reivindicaciones 1 a 9, en donde la célula vegetal es una célula vegetal monocotiledónea.

11. El método de la reivindicación 10, en donde la célula vegetal es una célula de maíz.

12. Una planta o parte de planta, que comprende una o más secuencias exógenas integradas en una localización *Zp15* mediante el método de la reivindicación 1.

50 13. Una semilla que comprende una o más secuencias exógenas integradas en una localización *Zp15* mediante el método de la reivindicación 1.



Zp15 genómico: GTGCAGCCCCGCTGGCGGGGGCCCTACTACGCCG:::GGTGTGGGCAGCCCGCCCATGTTCCAGCCGCTC  
Placa7 B10: GTGCAGCCCCGCTGGCGGGGGCCCTACTACGCCCGCAATTGGTGTGGGCAGCCCGAGCGCCCATGTTCCAGCCGCTC (+6)

11768

11766

### Figura 1

Zp15 genómico ATGGGCGCCGGGGCTGTACCCCTACGCCGAGTACCTGAGGCAGCCCGCATGCAGCCCCGCTGGCGGGCGCCCT  
Placa6 C12 ATGGGCGCCGGGGCTGTACCCCTACGGGAG::CTGAGGCAGCCCGCATGCAGCCCCGCTGGCGGGCGCCCT (Δ3)

11753

11750

### Figura 2

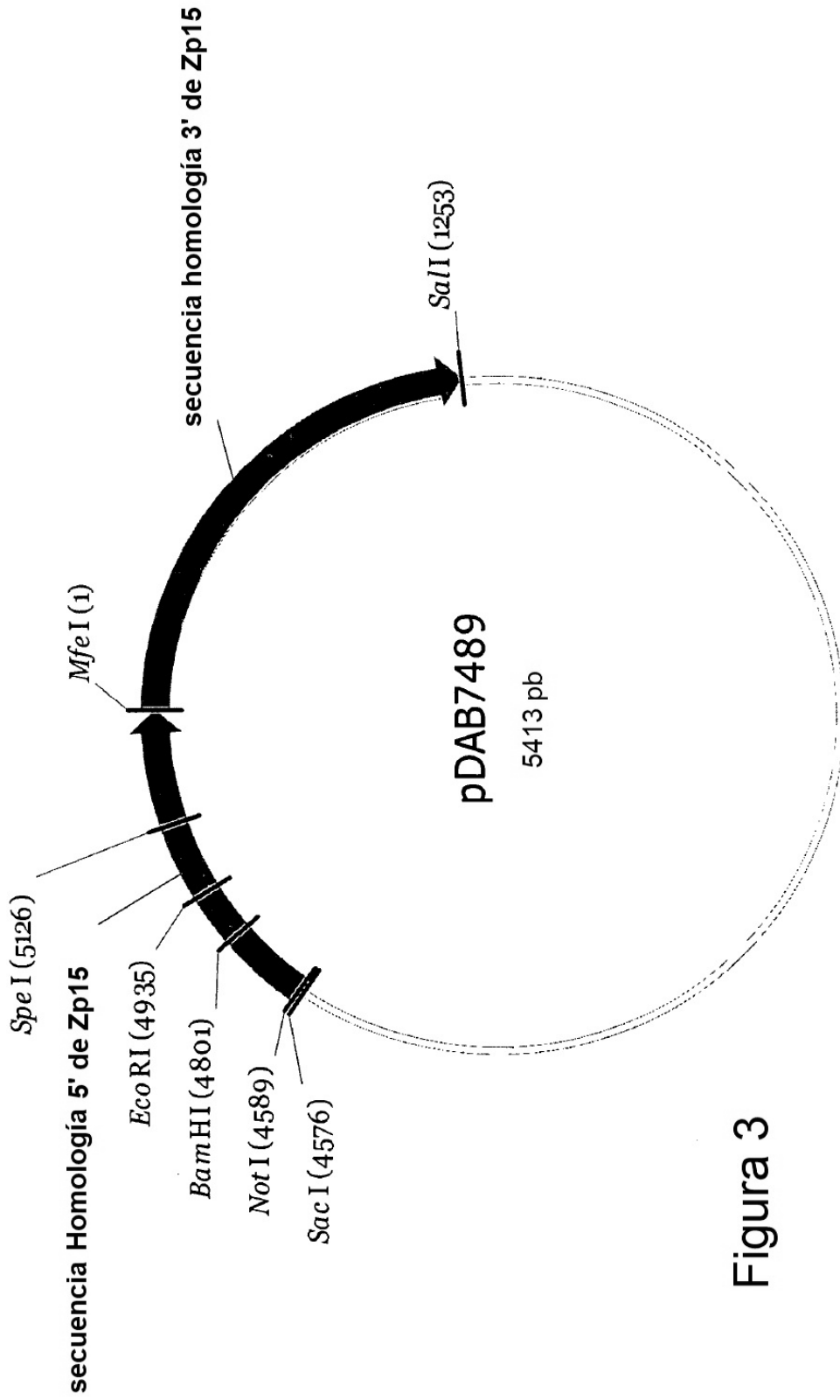


Figura 3

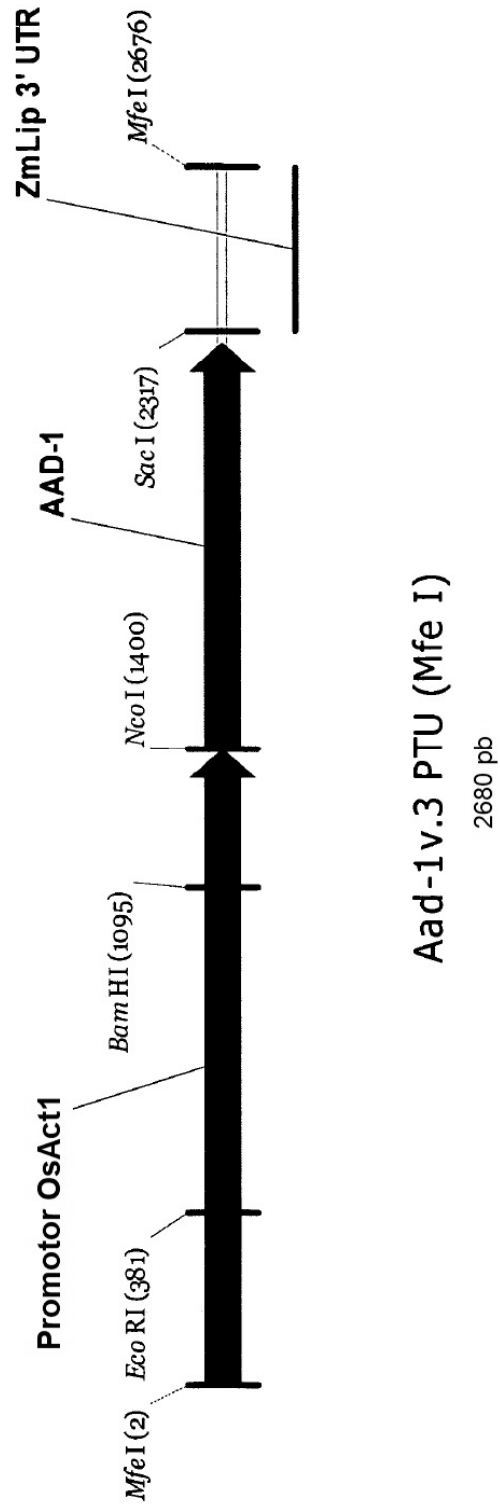


Figura 4

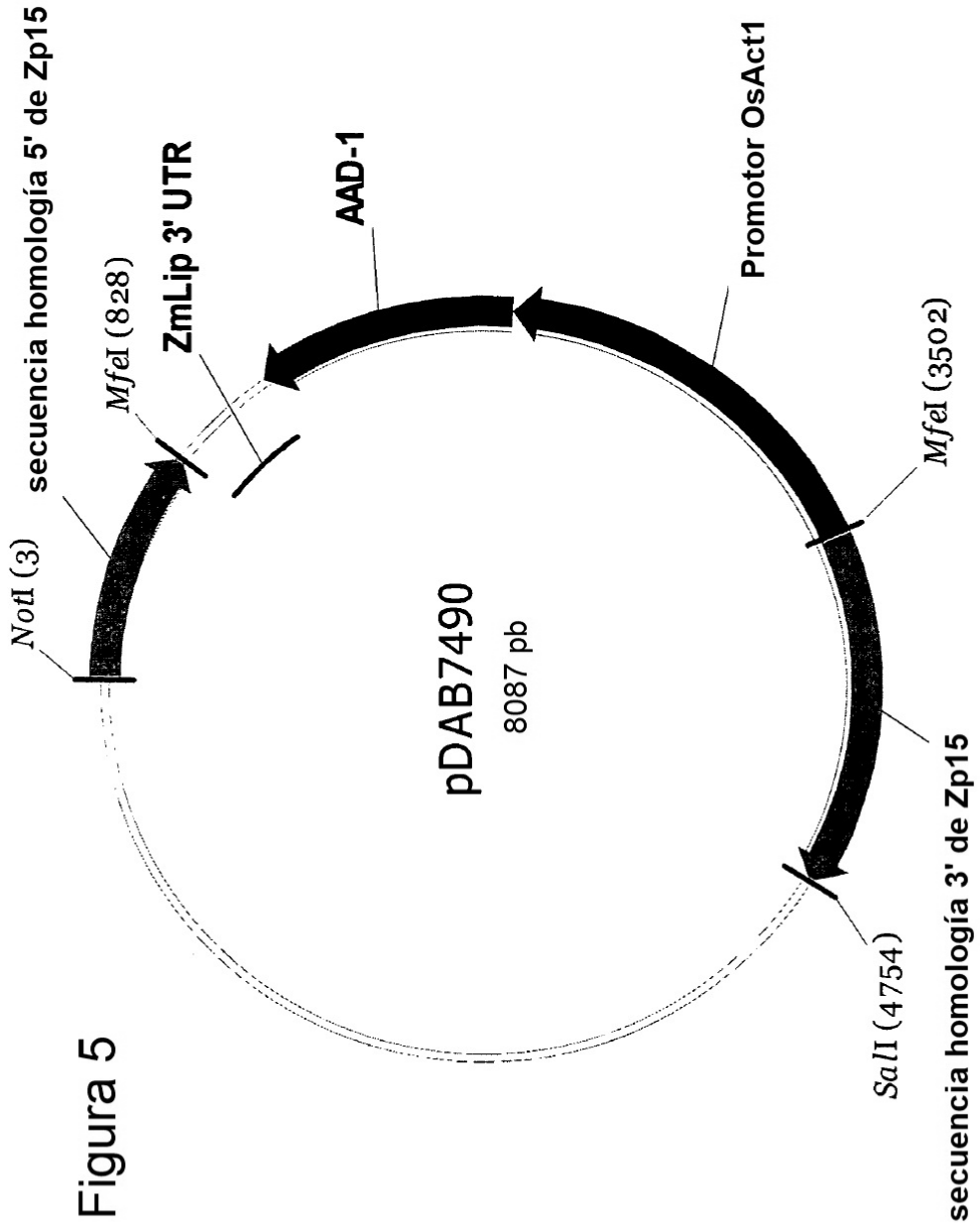


Figura 5

FIGURA 6

Localización Genómica Zp15 WT	(1)	<u>CATACTGTTGCTGCCCTGCTGGAATAAATGTGCTACTTCCCTGCCTG</u>
Fragmento donante Zp15	(1)	-----
Frontera_147_5'	(1)	<u>CATACTGTTGCTGCCCTGCTGGAATAAATGTGCTACTTCCCTGCCTG</u>
Frontera_147_3'	(1)	-----
		51 100
Localización Genómica Zp15 WT	(51)	<u>TTAAGGGAAAGGGTTTTGTTACAGATGTTACCTTGTAACTTGTACTTA</u>
Fragmento donante Zp15	(1)	-----
Frontera_147_5'	(51)	<u>TTAAGGGAAAGGGTTTTGTTACAGATGTTACCTTGTAACTTGTACTTA</u>
Frontera_147_3'	(1)	-----
		101 150
Localización Genómica Zp15 WT	(101)	<u>GTTTCATACTTGAATGAAGGTTCAATGGAACAAAAATATCTGCTGCATT</u>
Fragmento donante Zp15	(1)	-----
Frontera_147_5'	(101)	<u>GTTTCATACTTGAATGAAGGTTCAATGGAACAAAAATATCTGCTGCATT</u>
Frontera_147_3'	(1)	-----
		151 200
Localización Genómica Zp15 WT	(151)	<u>GCATGCAAGAGCTTTGATCATTAGAGCAATGTATTGGTTGTGAGTTT</u>
Fragmento donante Zp15	(1)	-----
Frontera_147_5'	(151)	<u>GCATGCAAGAGCTTTGATCATTAGAGCAATGTATTGGTTGTGAGTTT</u>
Frontera_147_3'	(1)	-----
		201 250
Localización Genómica Zp15 WT	(201)	<u>TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT</u>
Fragmento donante Zp15	(51)	<u>TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT</u>
Frontera_147_5'	(201)	<u>TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT</u>
Frontera_147_3'	(1)	-----
		251 300
Localización Genómica Zp15 WT	(251)	<u>ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATTTT</u>
Fragmento donante Zp15	(101)	<u>ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATTTT</u>
Frontera_147_5'	(251)	<u>ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATTTT</u>
Frontera_147_3'	(1)	-----
		301 350
Localización Genómica Zp15 WT	(301)	<u>TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCCTAA</u>
Fragmento donante Zp15	(151)	<u>TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCCTAA</u>
Frontera_147_5'	(301)	<u>TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCCTAA</u>
Frontera_147_3'	(1)	-----
		351 400
Localización Genómica Zp15 WT	(351)	<u>GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAA</u>
Fragmento donante Zp15	(201)	<u>GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAA</u>
Frontera_147_5'	(351)	<u>GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAA</u>
Frontera_147_3'	(1)	-----
		401 450
Localización Genómica Zp15 WT	(401)	<u>TATTGGTGTATATGTGCCTACAATGAAGTGAAGGTTGATGAGTCATGGTG</u>
Fragmento donante Zp15	(251)	<u>TATTGGTGTATATGTGCCTACAATGAAGTGAAGGTTGATGAGTCATGGTG</u>
Frontera_147_5'	(401)	<u>TATTGGTGTATATGTGCCTACAATGAAGTGAAGGTTGATGAGTCATGGTA</u>
Frontera_147_3'	(1)	-----
		451 500
Localización Genómica Zp15 WT	(451)	<u>ATGTGTAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT</u>
Fragmento donante Zp15	(301)	<u>ATGTGTAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT</u>
Frontera_147_5'	(451)	<u>ATGTGTAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT</u>
Frontera_147_3'	(1)	-----
		501 550
Localización Genómica Zp15 WT	(501)	<u>GACAACCCTTGACATGTAAGTTGATTCATATGTATAAGAAAGCTTAATG</u>
Fragmento donante Zp15	(351)	<u>GACAACCCTTGACATGTAAGTTGATTCATATGTATAAGAAAGCTTAATG</u>
Frontera_147_5'	(501)	<u>GACAACCCTTGACATGTAAGTTGATTCATATGTATAAGAAAGCTTAATG</u>
Frontera_147_3'	(1)	-----
		551 600
Localización Genómica Zp15 WT	(551)	<u>ATCTATCTGTAATCCAAATCCATGTAATGTTCCACGTGATGCAACG</u>
Fragmento donante Zp15	(401)	<u>ATCTATCTGTAATCCAAATCCATGTAATGTTCCACGTGATGCAACG</u>
Frontera_147_5'	(551)	<u>ATCTATCTGTAATCCAAATCCATGTAATGTTCCACGTGATGCAACG</u>
Frontera_147_3'	(1)	-----
		601 650
Localización Genómica Zp15 WT	(601)	<u>CAACATTCAAAACCATGGGTTGCAAGATGCTGCAGAATGCAAGCCATGG</u>
Fragmento donante Zp15	(451)	<u>CAACATTCAAAACCATGGGTTGCAAGATGCTGCAGAATGCAAGCCATGG</u>
Frontera_147_5'	(601)	<u>CAACATTCAAAACCATGGGTTGCAAGATGCTGCAGAATGCAAGCCATGG</u>
Frontera_147_3'	(1)	-----
		651 700
Localización Genómica Zp15 WT	(651)	<u>ATCATCTATAAATGGCTAGCTCCACATATGAAGTAGTCTCTATCATCAT</u>
Fragmento donante Zp15	(501)	<u>ATCATCTATAAATGGCTAGCTCCACATATGAAGTAGTCTCTATCATCAT</u>

ES 2 654 923 T3

Frontera_147_5'	(651)	ATCATCTATAAATGGCTAGCTGCCACATATGAAGTAGTCTCTATCATCAT	
Frontera_147_3'	(1)	-----	
		701	750
Localización Genómica Zp15 WT	(701)	CCAATCCAGATCAGCAAAGGGGAGTCCGTAGAGAGGATCGTGGAAACAGA	
Fragmento donante Zp15	(551)	CCAATCCAGATCAGCAAAGGGGAGTCCGTAGAGAGGATCGTGGAAACAGA	
Frontera_147_5'	(701)	CCAATCCAGATCAGCAAAGGGGAGTCCGTAGAGAGGATCGTGGAAACAGA	
Frontera_147_3'	(1)	-----	
		751	800
Localización Genómica Zp15 WT	(751)	ACAGCATGAAGATCGTTCATCGTTCTCGTGGTGTGCCTGGCTCTGTCAAGCT	
Fragmento donante Zp15	(601)	ACAGCATGAAGATCGTTCATCGTTCTCGTGGTGTGCCTGGCTCTGTCAAGCT	
Frontera_147_5'	(751)	ACAGCATGAAGATCGTTCATCGTTCTCGTGGTGTGCCTGGCTCTGTCAAGCT	
Frontera_147_3'	(1)	-----	
		801	850
Localización Genómica Zp15 WT	(801)	GCCAGGGCTCTGCAATGCCAGATGCCCTGCCCTGGCGGGGCTGCAGGG	
Fragmento donante Zp15	(651)	GCCAGGGCTCTGCAATGCCAGATGCCCTGCCCTGGCGGGGCTGCAGGG	
Frontera_147_5'	(801)	GCCAGGGCTCTGCAATGCCAGATGCCCTGCCCTGGCGGGGCTGCAGGG	
Frontera_147_3'	(1)	-----	
		851	900
Localización Genómica Zp15 WT	(851)	CTTGTAGGGGGCTGGCGGGGGCTGAGGAGGATGATGGGGGGGGGGGGGG	
Fragmento donante Zp15	(701)	CTTGTAGGGGGCTGGCGGGGGCTGAGGAGGATGATGGGGGGGGGGGGGG	
Frontera_147_5'	(851)	CTTGTAGGGGGCTGGCGGGGGCTGAGGAGGATGATGGGGGGGGGGGGGG	
Frontera_147_3'	(1)	-----	
		901	950
Localización Genómica Zp15 WT	(901)	TGTACCGCTAGCGGGAGTACCTGAGGCAGCGGAGTGCAGCGCGCTGGGG	
Fragmento donante Zp15	(751)	TGTACCGCTAGCGGGAGTACCTGAGGCAGCGGAGTGCAGCGCGCTGGGG	
Frontera_147_5'	(901)	TGTACCGCTAGCGGGAGTACCTGAGGCAGCGGAGTGCAGCGCGCTGGGG	
Frontera_147_3'	(1)	-----	
		951	1000
Localización Genómica Zp15 WT	(951)	GGGGGGGCTACTAGGGG-----	
Fragmento donante Zp15	(801)	GGGGGGGCTACTAGGGGGCAATTGAGCAGTTAAAGATCTTTAGAAGAA	
Frontera_147_5'	(951)	GGGGGGGCTACTAGGGGGCAATTGAGCAGTTAAAGATCTTTAGAAGAA	
Frontera_147_3'	(1)	-----	
		1001	1050
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(851)	AGCAAAGCAATTTATTAAATACATAACAATGTCCAGGTAGCCGAGCTGAATT	
Frontera_147_5'	(1001)	AGCAAAGCAATTTATTAAATACATAACAATGTCCAGGTAGCCGAGCTGAATT	
Frontera_147_3'	(1)	-----	
		1051	1100
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(901)	ACAATAGGCAACTGCTCATATAATTCAAGAAAGCGAAGTAGTACACAAG	
Frontera_147_5'	(1051)	ACAATAGGCAACTGCTCATATAATTCAAGAAAGCGAAGTAGTACACAAG	
Frontera_147_3'	(1)	-----	
		1101	1150
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(951)	ATCCAGAAGCAAATAAAGCCCATAGGTACCAAAGCCCTACACAAGCAGCAA	
Frontera_147_5'	(1101)	ATCCAGAAGCAAATAAAGCCCATAGGTACCAAAGCCCTACACAAGCAGCAA	
Frontera_147_3'	(1)	-----	
		1151	1200
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1001)	CAGTCACTGCCAGTCCGGTGGGTCTTTAAAGCACAGGGGCTTGAGCCAG	
Frontera_147_5'	(1151)	CAGTCACTGCCAGTCCGGTGGGTCTTTAAAGCACAGGGGCTTGAGCCAG	
Frontera_147_3'	(1)	-----	
		1201	1250
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1051)	GGGATCCAGCTTGAACAAGCTTGGTAAAATTAAGCAAAGCAGAAGCAG	
Frontera_147_5'	(1201)	GGGATCCAGCTTGAACAAGCTTGGTAAAATTAAGCAAAGCAGAAGCAG	
Frontera_147_3'	(1)	-----	
		1251	1300
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1101)	ACACACGGCAAGGCAAGGCTTCTGATGGCGGGCCAAAGGCGGGGGGGGG	
Frontera_147_5'	(1251)	ACACACGGCAAGGCAAGGCTTCTGATGGCGGGCCAAAGGCGGGGGGGGG	
Frontera_147_3'	(1)	-----	
		1301	1350
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1151)	AGAAGGTAGGACGGACAGGCACAGGCTGGGAGCGAGCTCTAGGTGATTTAA	
Frontera_147_5'	(1301)	AGAAGGTAGGACGGACAGGCACAGGCTGGGAGCGAGCTCTAGGTGATTTAA	
Frontera_147_3'	(1)	-----	
		1351	1400
Localización Genómica Zp15 WT	(970)	-----	

ES 2 654 923 T3

Fragmento donante Zp15	(1201)	GGTAAGTACTCAGCGGGCAGGGCTAAGTCCAGCAACTGTGGTGGGAGTCA	
Frontera_147_5'	(1351)	GGTAAGTACTCAGCGGGCAGGGCTAAGTCCAGCAACTGTGGTGGGAGTCA	
Frontera_147_3'	(1)	-----	
		1401	1450
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1251)	AGTATCTGAACTTGGCAGCATAGTCAGGAACAGCACGGTGCATGGTGCAC	
Frontera_147_5'	(1401)	AGTATCTGAACTTGGCAGCATAGTCAGGAACAGCACGGTGCATGGTGCAC	
Frontera_147_3'	(1)	-----	
		1451	1500
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1301)	AAGTGTGCCAGACAAGGACTTGGTCTTTCCTTCCACCTCACAGGGCAAGT	
Frontera_147_5'	(1451)	AAGTGTGCCAGACAAGGACTTGGTCTTTCCTTCCACCTCACAGGGCAAGT	
Frontera_147_3'	(1)	-----	
		1501	1550
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1351)	GAAGTCAAATCTGGTGGCATGGTCAATAGAGGAAGTGAAGCAATGGCTTTG	
Frontera_147_5'	(1501)	GAAGTCAAATCTGGTGGCATGGTCAATAGAGGAAGTGAAGCAATGGCTTTG	
Frontera_147_3'	(1)	-----	
		1551	1600
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1401)	ATTCTGCATCTGTCTATGCCCTCAATTCCTGACAGTAGAGTTGATTCACA	
Frontera_147_5'	(1551)	ATTCTGCATCTGTCTATGCCCTCAATTCCTGACAGTAGAGTTGATTCACA	
Frontera_147_3'	(1)	-----	
		1601	1650
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1451)	TAAAGGCCCTTTCCTTCCAGAGCCAGGATGAGTCACAACCAAGGGATGGAC	
Frontera_147_5'	(1601)	TAAAGGCCCTTTCCTTCCAGAGCCAGGATGAGTCACAACCAAGGGATGGAC	
Frontera_147_3'	(1)	-----	
		1651	1700
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1501)	TGTCTCTCTGTCAACCAGCATCAACATCCATCACCTTGACTGAGGTGTTCG	
Frontera_147_5'	(1651)	TGTCTCTCTGTCAACCAGCATCAACATCCATCACCTTGACTGAGGTGTTCG	
Frontera_147_3'	(1)	-----	
		1701	1750
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1551)	TGAAGCCGACGGTTCCTGTCTTGGTAGAGGGAACCGAACACACGTTGTGGCA	
Frontera_147_5'	(1701)	TGAAGCCGACGGTTCCTGTCTTGGTAGAGGGAACCGAACACACGTTGTGGCA	
Frontera_147_3'	(1)	-----	
		1751	1800
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1601)	GAGTGCACAAAGGTTGAGCCCTTGGATGGTGGCTTGCATGGTGGAGACAA	
Frontera_147_5'	(1751)	GAGTGCACAAAGGTTGAGCCCTTGGATGGTGGCTTGCATGGTGGAGACAA	
Frontera_147_3'	(1)	-----	
		1801	1850
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1651)	GGTGTCCCAAGGTGTGTACATGAAAGGAAGCCAGTGTCTCCGCCATGGT	
Frontera_147_5'	(1801)	GGTGTCCCAAGGTGTGTACATGAAAGGAAGCCAGTGTCTCCGCCATGGT	
Frontera_147_3'	(1)	-----	
		1851	1900
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1701)	GAGGAAGATCTATGGCCCTCATCAGAACAGCAGCTGGAGGTGCATCAAGG	
Frontera_147_5'	(1851)	GAGGAAGATCTATGGCCCTCATCAGAACAGCAGCTGGAGGTGCATCAAGG	
Frontera_147_3'	(1)	-----	
		1901	1950
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1751)	AAAGTGGAGTCTGTGTGCCAGTCATCACC AATCACCCCTTCCAGACTCATTT	
Frontera_147_5'	(1901)	AAAGTGGAGTCTGTGTGCCAGTCATCACC AATCACCCCTTCCAGACTCATTT	
Frontera_147_3'	(1)	-----	
		1951	2000
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1801)	GGGTCTCTGCGGCATCATCTGAAGCTCTGGATAGCCCTTCAATGCTCTTGA	
Frontera_147_5'	(1951)	GGGTCTCTGCGGCATCATCTGAAGCTCTGGATAGCCCTTCAATGCTCTTGA	
Frontera_147_3'	(1)	-----	
		2001	2050
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1851)	GAAGAGGCACTGGATCAACTGGTCCAAACCTTCTGAGAATGGAATGTGG	
Frontera_147_5'	(2001)	GAAGAGGCACTGGATCAACTGGTCCAAACCTTCTGAGAATGGAATGTGG	
Frontera_147_3'	(1)	-----	
		2051	2100

ES 2 654 923 T3

Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1901)	TGCTCAITGGTIGATTGCTTGCCAGGAAAGTAGATGACTTGGTAAGTGTG	
Frontera_147_5'	(2051)	TGCTCAITGGTIGATTGCTTGCCAGGAAAGTAGATGACTTGGTAAGTGTG	
Frontera_147_3'	(1)	-----	
		2101	2150
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(1951)	GAAGGCATGCCAATATCTCATTCCAGGTGGTGTCAATCAAGTGGTTCGGTCA	
Frontera_147_5'	(2101)	GAAGGCATGCCAATATCTCATTCCAGGTGGTGTCAATCAAGTGGTTCGGTCA	
Frontera_147_3'	(1)	-----	
		2151	2200
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2001)	AGTCCACGCCAGTGTATCTCAGCAGCAAGGACACCAGTGAAGTGGCTGGACA	
Frontera_147_5'	(2151)	AGTCCACGCCAGTGTATCTCAGCAGCAAGGACACCAGTGAAGTGGCTGGACA	
Frontera_147_3'	(1)	-----	
		2201	2250
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2051)	GCTATTCTCTCAAAGCGTTGGGAGAGAGGGCTGAGGGCAGCATGAGCCAT	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2251	2300
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2101)	GGTCTACCTACAAAAAGCTCCGCACGAGGCTGCATTTGTACAAATCAT	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2301	2350
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2151)	GAAAAGAAAACTACCGATGAACAATGCTGAGGGATTCAAATCTACCCA	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2351	2400
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2201)	CAAAAAGAAGAAAGAAAGATCTAGCACATCTAAGCCTGACGAAGCAGCAG	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2401	2450
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2251)	AAATATATAAAAAATAAACCATAGTGCCCTTTTCCCTCTCTCTGATCT	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2451	2500
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2301)	TGTTTAGCATGGCGGAAATTTAAACCCCATCATCTCCCAACAACG	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2501	2550
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2351)	GCGGATCGCAGATCTACATCCGAGAGCCCCATTCCCCGCGAGATCCGGGC	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2551	2600
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2401)	CGGATCCACGCCGGCGAGAGCCCCAGCCGCGAGATCCCGCCCTCCCGCG	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2601	2650
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2451)	CACCGATCTGGGCGCGCACGAAGCCGCTCTCGCCACCCAAACTACCAA	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2651	2700
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2501)	GGCCAAAGATCGAGACGAGACGGAAAAAAGGAGAAAGGAAGAGG	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	
		2701	2750
Localización Genómica Zp15 WT	(970)	-----	
Fragmento donante Zp15	(2551)	AGAGGGCGGGGTGGTTACCGGCGCGCGGGGGAGGGGGGA	
Frontera_147_5'	(2177)	-----	
Frontera_147_3'	(1)	-----	



		2751	2800
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2601)	GGCGTCGTCCGGCAGCGAGGGGGGAGGAGGTGGAGGTGGTGGTGGTGGTGG	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(1)	-----	-----
		2801	2850
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2651)	GTGGTAGGGTTGGGGGATGGGAGGAGAGGGGGGGTATGTATATAGTGG	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(1)	-----	-----
		2851	2900
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2701)	CGATGGGGGGCGTTCTTTGGAAGCGGAGGGAGGGCCGGCCTCGTCGCTG	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(1)	-----	-----
		2901	2950
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2751)	GCTCGGATCCTCCTCGGTTTCCGGCCCCACGACCCGGACCCACCTGC	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(1)	-----	-----
		2951	3000
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2801)	TGTTTTTCTTTTTCTTTTTTTCTTTCTTTTTTTT	<u>TTTTTTGGGTGGGA</u>
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(1)	-----	<u>TTTTTTGGGTGGGA</u>
		3001	3050
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2851)	<u>GACGTGGGGTGGGTGGGACAACCTC</u> <u>CGGTCATAGTGGGGGG</u> <u>TGTGGA</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(15)	<u>GACGTGGGGTGGGTGGGACAACCTC</u> <u>CGGTCATAGTGGGGGG</u> <u>TGTGGA</u>	
		3051	3100
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2900)	<u>GACTATTGTCGAGTTGGCTGGACTGGGGTGGGTGGGTGGGTGGGTGG</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(65)	<u>GACTATTGTCGAGTTGGCTGGACTGGGGTGGGTGGGTGGGTGGGTGG</u>	
		3101	3150
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(2950)	<u>GGCTGGGGTTTGGTATGGATCGTGGATAGCACTTTGGGCTTTAGCAACTTT</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(115)	<u>GGCTGGGGTTTGGTATGGATCGTGGATAGCACTTTGGGCTTTAGCAACTTT</u>	
		3151	3200
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3000)	<u>AGGGGTGTTTTGTAATGTTTTGAGTCTAAGTTATCTTTATTTTTA</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(165)	<u>AGGGGTGTTTTGTAATGTTTTGAGTCTAAGTTATCTTTATTTTTA</u>	
		3201	3250
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3050)	<u>GTAGAAAAAATACCCATGCCGTGCAACGGGGAAAGCTATTTTAATCTTA</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(215)	<u>GTAGAAAAAATACCCATGCCGTGCAACGGGGAAAGCTATTTTAATCTTA</u>	
		3251	3300
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3100)	<u>TTATTGTTTCATTGTCAGAAATCGGCTGAATATATATTTTTCTCAAAATTT</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(265)	<u>TTATTGTTTCATTGTCAGAAATCGGCTGAATATATATTTTTCTCAAAATTT</u>	
		3301	3350
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3150)	<u>ATGTCAAATTAGCATATGGGTTTTTTTAAAGATATTTCTATACAAATCC</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(315)	<u>ATGTCAAATTAGCATATGGGTTTTTTTAAAGATATTTCTATACAAATCC</u>	
		3351	3400
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3200)	<u>CTCTGTATTTACAAAAGCAAAGCACTTAAAACGGGACTCAAAATACAGAT</u>	
Frontera_147_5'	(2177)	-----	-----
Frontera_147_3'	(365)	<u>CTCTGTATTTACAAAAGCAAAGCACTTAAAACGGGACTCAAAATACAGAT</u>	
		3401	3450
Localización Genómica Zp15 WT	(970)	-----	-----
Fragmento donante Zp15	(3250)	<u>ATGCATTTCCAAAAGCGAATAAAGTTAAAACCAATTCATACAAAATGA</u>	
Frontera_147_5'	(2177)	-----	-----

ES 2 654 923 T3

Frontera_147_3'	(415)	<u>ATGCATTTCCAAAAGCGAATAAACTTAAAAACCAATTCATACAAAATGA</u>	3451	3500
Localización Genómica Zp15 WT	(970)	-----		
Fragmento donante Zp15	(3300)	<u>CGTATCAAAGTACCGACAAAAACATCCTCAATTTTATAATAGTAGAAA</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(465)	<u>CGTATCAAAGTACCGACAAAAACATCCTCAATTTTATAATAGTAGAAA</u>	3501	3550
Localización Genómica Zp15 WT	(970)	-----		
Fragmento donante Zp15	(3350)	<u>GAGTAAATTTCACTTTGGGCCACCTTTTATTACCGATATTTTACTTTATA</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(515)	<u>GAGTAAATTTCACTTTGGGCCACCTTTTATTACCGATATTTTACTTTATA</u>	3551	3600
Localización Genómica Zp15 WT	(970)	-----		
Fragmento donante Zp15	(3400)	<u>CCACCTTTTAACTGATGTTTTCACCTTTTGACCAGGTAATCTTACCTTTGT</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(565)	<u>CCACCTTTTAACTGATGTTTTCACCTTTTGACCAGGTAATCTTACCTTTGT</u>	3601	3650
Localización Genómica Zp15 WT	(970)	-----		
Fragmento donante Zp15	(3450)	<u>TTTATTTTGGACTATCCCGACTCTCTCTCAAGCATATGAATGACCAAT</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(615)	<u>TTTATTTTGGACTATCCCGACTCTCTCTCAAGCATATGAATGACCAAT</u>	3651	3700
Localización Genómica Zp15 WT	(970)	<u>GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC</u>		
Fragmento donante Zp15	(3500)	<u>GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(665)	<u>GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC</u>	3701	3750
Localización Genómica Zp15 WT	(1020)	<u>CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCCGCAGCAGCTGCA</u>		
Fragmento donante Zp15	(3550)	<u>CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCCGCAGCAGCTGCA</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(715)	<u>CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCCGCAGCAGCTGCA</u>	3751	3800
Localización Genómica Zp15 WT	(1070)	<u>GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCTGTACG</u>		
Fragmento donante Zp15	(3600)	<u>GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCTGTACG</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(765)	<u>GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCTGTACG</u>	3801	3850
Localización Genómica Zp15 WT	(1120)	<u>AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC</u>		
Fragmento donante Zp15	(3650)	<u>AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(815)	<u>AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC</u>	3851	3900
Localización Genómica Zp15 WT	(1170)	<u>AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT</u>		
Fragmento donante Zp15	(3700)	<u>AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(865)	<u>AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT</u>	3901	3950
Localización Genómica Zp15 WT	(1220)	<u>CTACCAGTACCAGCTGCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG</u>		
Fragmento donante Zp15	(3750)	<u>CTACCAGTACCAGCTGCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(915)	<u>CTACCAGTACCAGCTGCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG</u>	3951	4000
Localización Genómica Zp15 WT	(1270)	<u>CTGCCATTCGCCCTACTACTGATTATGATATTTGGGAAATCTCCTCTA</u>		
Fragmento donante Zp15	(3800)	<u>CTGCCATTCGCCCTACTACTGATTATGATATTTGGGAAATCTCCTCTA</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(965)	<u>CTGCCATTCGCCCTACTACTGATTATGATATTTGGGAAATCTCCTCTA</u>	4001	4050
Localización Genómica Zp15 WT	(1320)	<u>TCCATCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT</u>		
Fragmento donante Zp15	(3850)	<u>TCCATCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(1015)	<u>TCCATCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT</u>	4051	4100
Localización Genómica Zp15 WT	(1370)	<u>TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAGTTT</u>		
Fragmento donante Zp15	(3900)	<u>TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAGTTT</u>		
Frontera_147_5'	(2177)	-----		
Frontera_147_3'	(1065)	<u>TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAGTTT</u>	4101	4150
Localización Genómica Zp15 WT	(1420)	<u>GGTTTTAATGAATCTATCGCCGCTTGATGTCATGATGGACAAATCAA</u>		
Fragmento donante Zp15	(3950)	<u>GGTTTTAATGAATCTATCGCCGCTTGATGTCATGATGGACAAATCAA</u>		

	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1115)	<u>GGTTTTAATGAATCTATCGGCCGCTTGATGCTATGATGGACAAATCAA</u>
			4151 4200
Localización Genómica Zp15 WT	(1470)		<u>GCTTCTCCTGTCAGGCATGTAATATTTCAAATCTCTATTGAGGCTCAA</u>
Fragmento donante Zp15	(4000)		<u>ACTTCTCCTGTCAGGCATGTAATATTTCAAATCTCTATTGAGGCTCAA</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1165)	<u>ACTTCTCCTGTCAGGCATGTAATATTTCAAATCTCTATTGAGGCTCAA</u>
			4201 4250
Localización Genómica Zp15 WT	(1520)		<u>ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAG</u>
Fragmento donante Zp15	(4050)		<u>ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAG</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1215)	<u>ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAG</u>
			4251 4300
Localización Genómica Zp15 WT	(1570)		<u>TTGAGTATAGAGTATAAAACATAAAGTCATGTGATTCTATTGGCTAGAT</u>
Fragmento donante Zp15	(4100)		<u>TTGAGTATAGAGTATAAAACATAAAGTCATGTGATTCTATTGGCTAGAT</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1265)	<u>TTGAGTATAGAGTATAAAACATAAAGTCATGTGATTCTATTGGCTAGAT</u>
			4301 4350
Localización Genómica Zp15 WT	(1620)		<u>AAGTGTAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATT</u>
Fragmento donante Zp15	(4150)		<u>AAGTGTAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATT</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1315)	<u>AAGTGTAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATT</u>
			4351 4400
Localización Genómica Zp15 WT	(1670)		<u>ACACATAATTTTAGCGTTTTTCCATTTAAAGCGGGGCTTGACGAAGT</u>
Fragmento donante Zp15	(4200)		<u>ACACATAATTTTAGCGTTTTTCCATTTAAAGCGGGGCTTGACGAAGT</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1365)	<u>ACACATAATTTTAGCGTTTTTCCATTTAAAGCGGGGCTTGACGAAGT</u>
			4401 4450
Localización Genómica Zp15 WT	(1720)		<u>GGAAGCCGTGGAACGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA</u>
Fragmento donante Zp15	(4250)		<u>GGAAGCCGTGGAACGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1415)	<u>GGAAGCCGTGGAACGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA</u>
			4451 4500
Localización Genómica Zp15 WT	(1770)		<u>GGGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACCTATTCAQ</u>
Fragmento donante Zp15	(4300)		<u>GAGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACCTATTCAQ</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1465)	<u>GAGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACCTATTCAQ</u>
			4501 4550
Localización Genómica Zp15 WT	(1820)		<u>GAGTAGTTTGATAACATTCGTACCCGAAAAGATTCTGTTAGATGGATG</u>
Fragmento donante Zp15	(4350)		<u>GAGTAGTTTGATAACATTCGTACCCGAAAAGATTCTGTTAGATGGATG</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1515)	<u>GAGTAGTTTGATAACATTCGTACCCGAAAAGATTCTGTTAGATGGATG</u>
			4551 4600
Localización Genómica Zp15 WT	(1870)		<u>CAGCAACTAGGATCTGGTGATAACTAGCCCGCTACCGCGTGATTGGTT</u>
Fragmento donante Zp15	(4400)		<u>CAGCAACTAGGATCTGGTGATAACTAGCCCGCTACCGCGTGATTGGTT</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1565)	<u>CAGCAACTAGGATCTGGTGATAACTAGCCCGCTACCGCGTGATTGGTT</u>
			4601 4650
Localización Genómica Zp15 WT	(1920)		<u>GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC</u>
Fragmento donante Zp15	(4450)		<u>GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1615)	<u>GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC</u>
			4651 4700
Localización Genómica Zp15 WT	(1970)		<u>CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG</u>
Fragmento donante Zp15	(4500)		<u>CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1665)	<u>CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG</u>
			4701 4750
Localización Genómica Zp15 WT	(2020)		<u>GCCTGCTCCGTCATGCAGGCTACACCCGGATAGTGACGGTAACCAATC</u>
Fragmento donante Zp15	(4550)		<u>GCCTGCTCCGTCATGCAGGCTACACCCGGATAGTGACGGTAACCAATC</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1715)	<u>GCCTGCTCCGTCATGCAGGCTACACCCGGATAGTGACGGTAACCAATC</u>
			4751 4800
Localización Genómica Zp15 WT	(2070)		<u>GTATGCCATTACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCC</u>
Fragmento donante Zp15	(4600)		<u>GTATGCCATTACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCC</u>
	Frontera_147_5'	(2177)	-----
	Frontera_147_3'	(1765)	<u>GTATGCCATTACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCC</u>
			4801 4850
Localización Genómica Zp15 WT	(2120)		<u>GAGCAACCAATCACGTGGTACCTGACCTAAGTAATGACCAGCAATAAAA</u>

ES 2 654 923 T3

Fragmento donante Zp15	(4650)	<u>GAGCAAGCAATCAGGTGGTACCTGACCTAAGTAATGACCGCAAAATAAAA</u>
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(1815)	<u>GAGCAAGCAATCAGGTGGTACCTGACCTAAGTAATGACCGCAAAATAAAA</u>
		4851 <span style="float:right">4900</span>
Localización Genómica Zp15 WT	(2170)	<u>GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCCGCATCAGTATCGCGT</u>
Fragmento donante Zp15	(4700)	<u>GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCCGCATCAGTATCG---</u>
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(1865)	<u>GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCCGCATCAGTATCGCGT</u>
		4901 <span style="float:right">4950</span>
Localización Genómica Zp15 WT	(2220)	<u>GCAGAGACAGTTAGGGTTCGGAGTTTGTGTGACGGTGTAGCTAAATT</u>
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(1915)	<u>GCAGAGACAGTTAGGGTTCGGAGTTTGTGTGACGGTGTAGCTAAATT</u>
		4951 <span style="float:right">5000</span>
Localización Genómica Zp15 WT	(2270)	<u>CGGGGAATTAACTCGGGAGATTGGTTTGTAAAGGGTCCAGACCGCTCGTC</u>
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(1965)	<u>CGGGGAATTAACTCGGGAGATTGGTTTGTAAAGGGTCCAGACCGCTCGTC</u>
		5001 <span style="float:right">5050</span>
Localización Genómica Zp15 WT	(2320)	<u>TATAAATATAAAGGAATACAGTTGATTGGGATAAACAATCGAACCTACAA</u>
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(2015)	<u>TATAAATATAAAGGAATACAGTTGATTGGGATAAACAATCGAACCTACAA</u>
		5051 <span style="float:right">5100</span>
Localización Genómica Zp15 WT	(2370)	<u>TCAATAAAATTTGCATTTTATCTTGTACATTTAGGAGTCGCTCTAGTTTA</u>
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(2065)	<u>TCAATAAAATTTGCATTTTATCTTGTACATTTAGGAGTCGCTCTAGTTTA</u>
		5101 <span style="float:right">5150</span>
Localización Genómica Zp15 WT	(2420)	<u>GTTCCTAGTTTAAAGCTCTCAATCGCCAAATTCCTGTTCCTTCGGCTCT</u>
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(2115)	<u>GTTCCTAGTTTAAAGCTCTCAATCGCCAAATTCCTGTTCCTTCGGCTCT</u>
		5151 <span style="float:right">5200</span>
Localización Genómica Zp15 WT	(2470)	<u>ACATCGATTATAGGTGTC</u> AGGTCGGCCTGCCGACTATAGACAAAGCATA
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(2165)	<u>ACATCGATTATAGGTGTC</u> -----
		5201
Localización Genómica Zp15 WT	(2520)	GGACCT
Fragmento donante Zp15	(4747)	-----
Frontera_147_5'	(2177)	-----
Frontera_147_3'	(2183)	-----

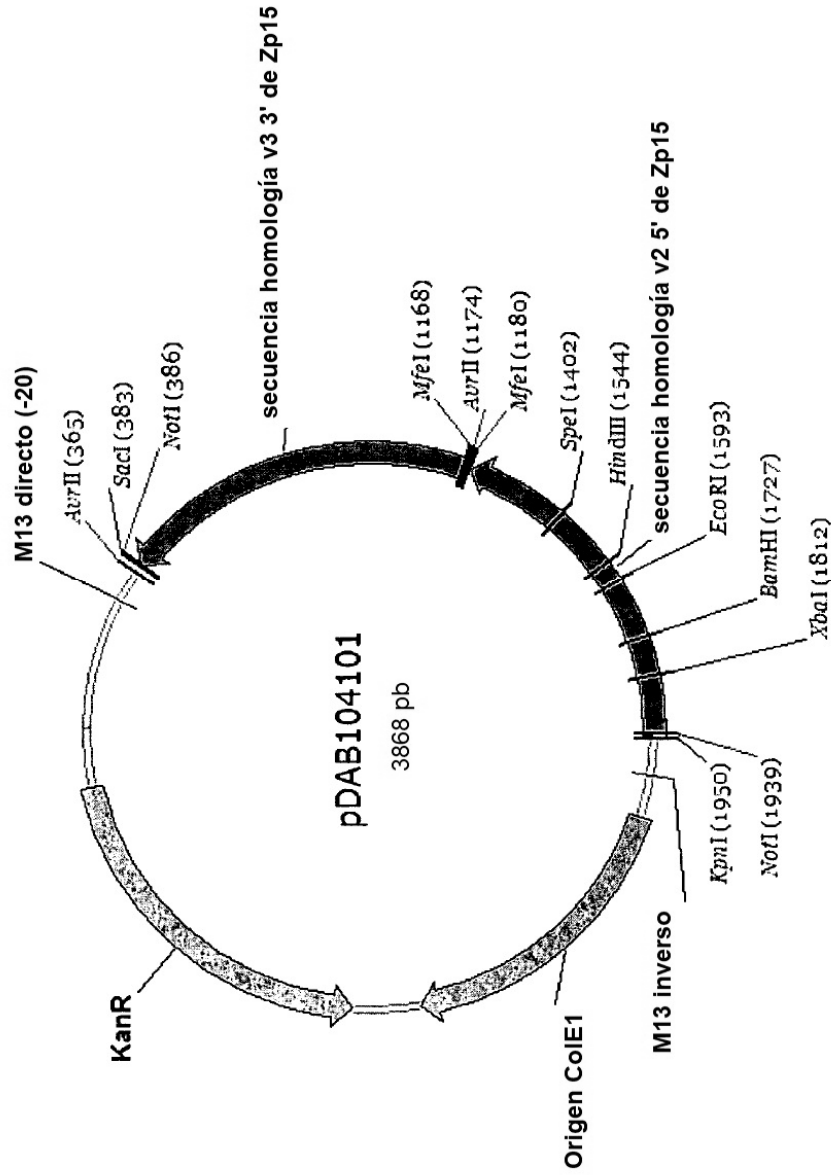
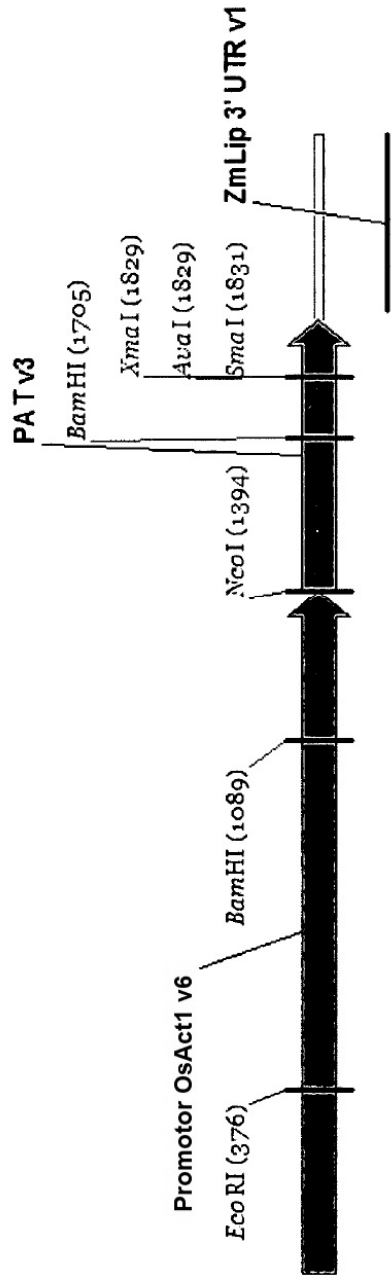


Figura 7



Casete de Gen PAT

2320 pb

Figura 8

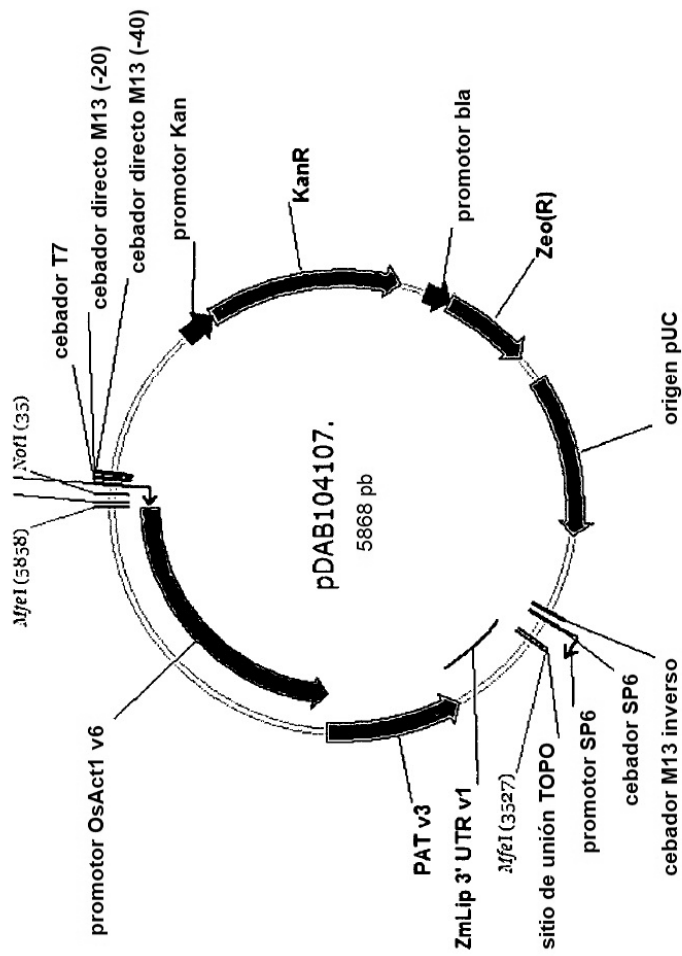


Figura 9

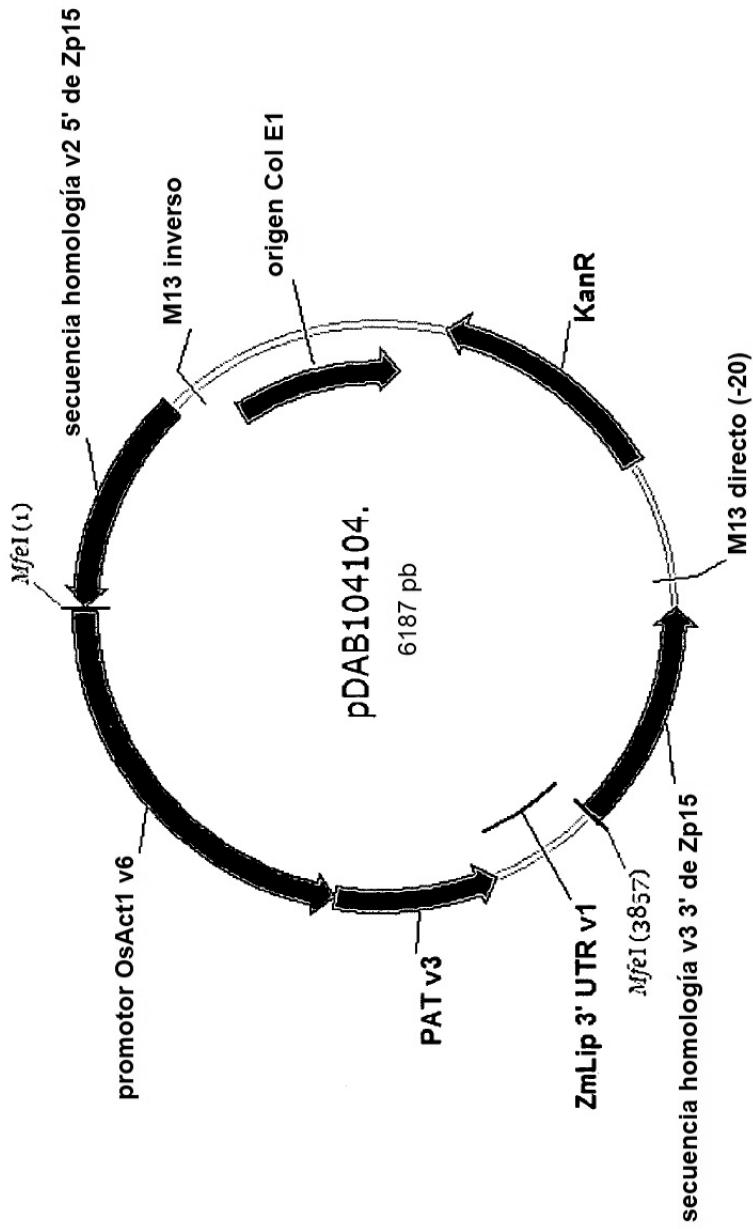


Figura 10



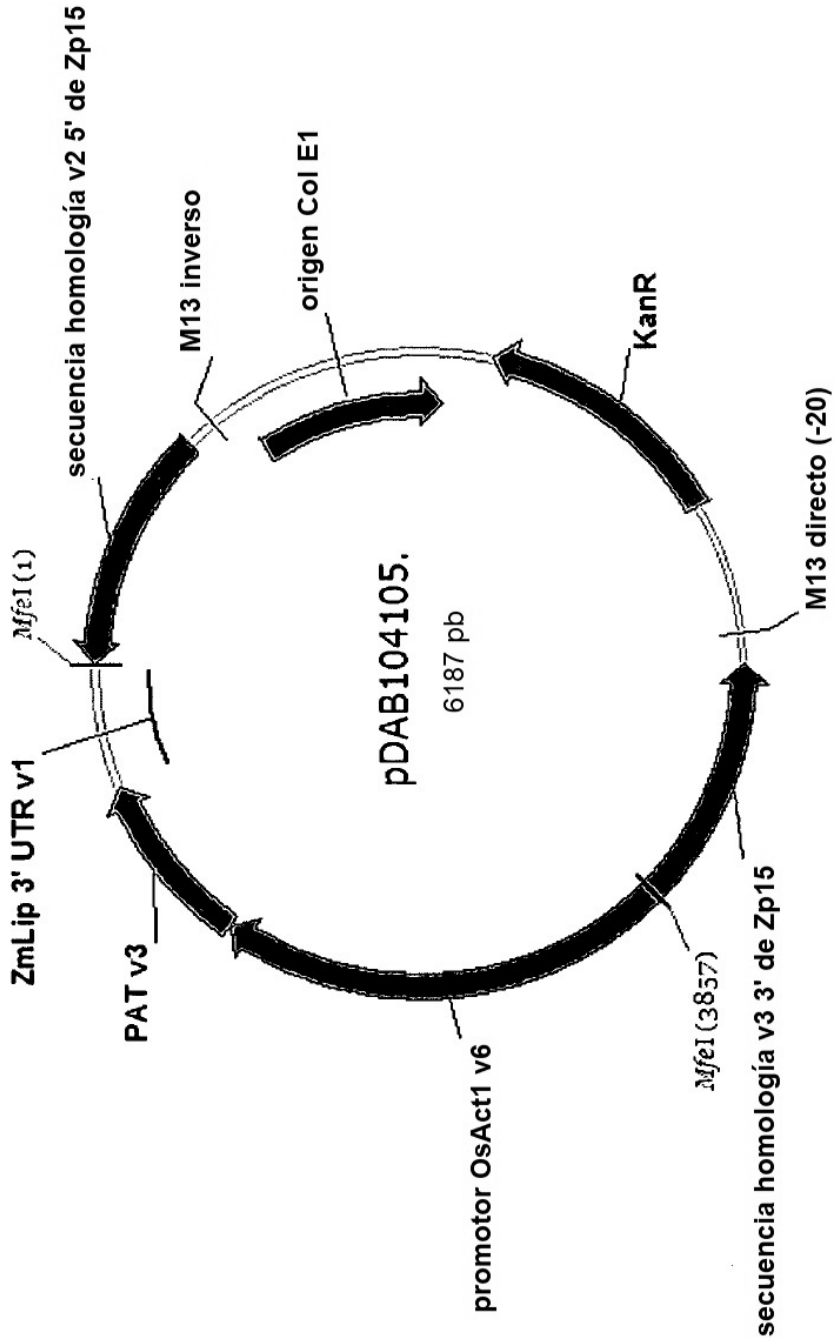


Figura 11

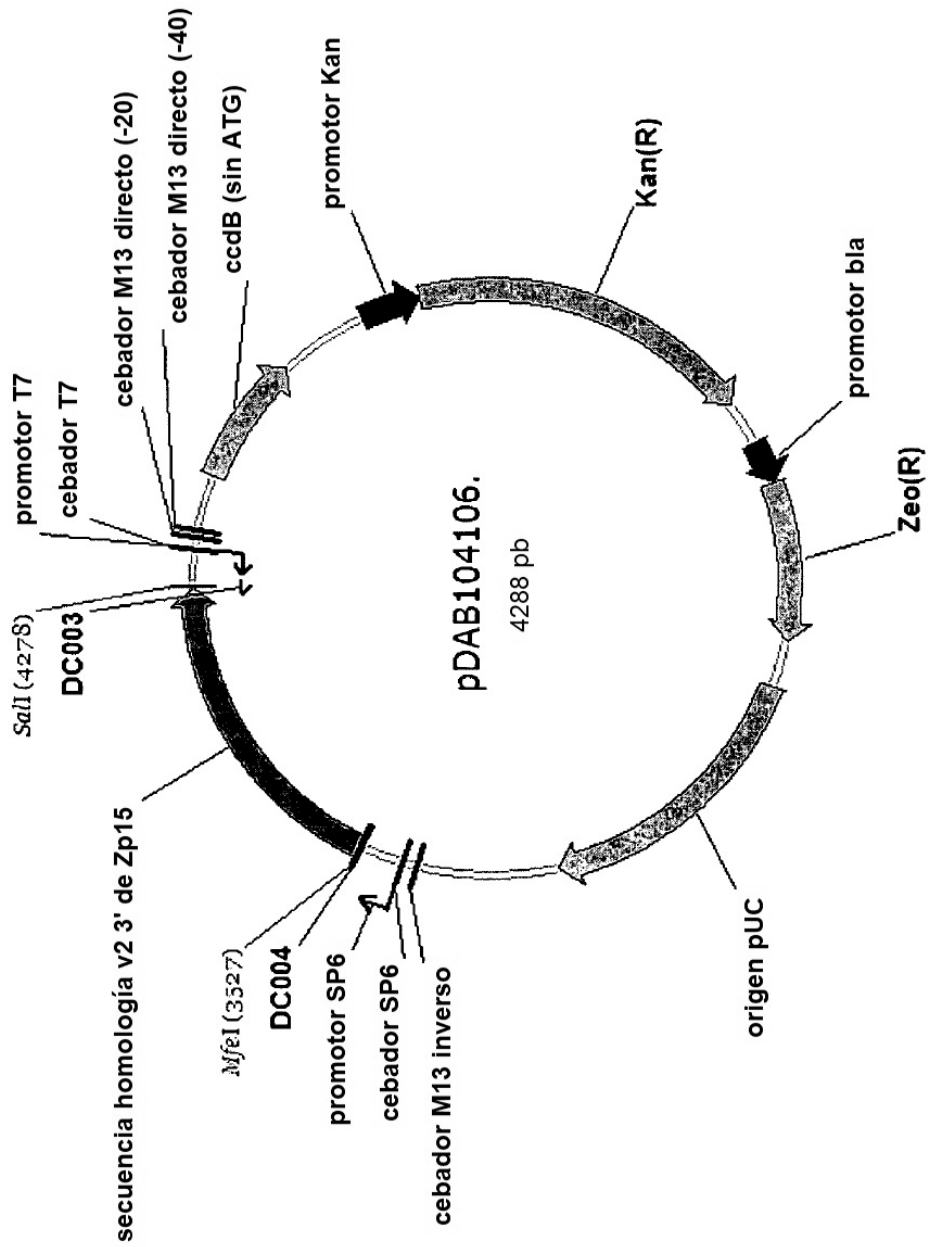


Figura 12

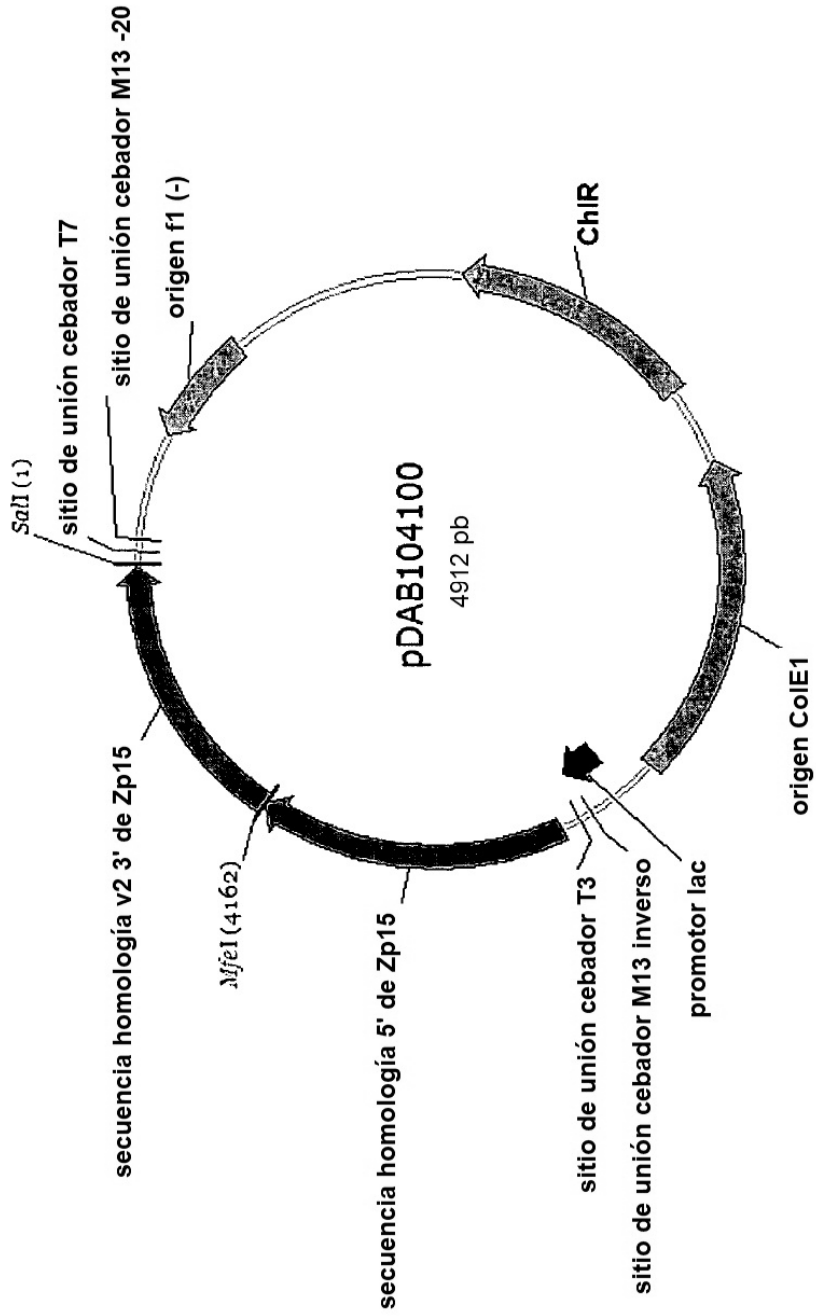


Figura 13

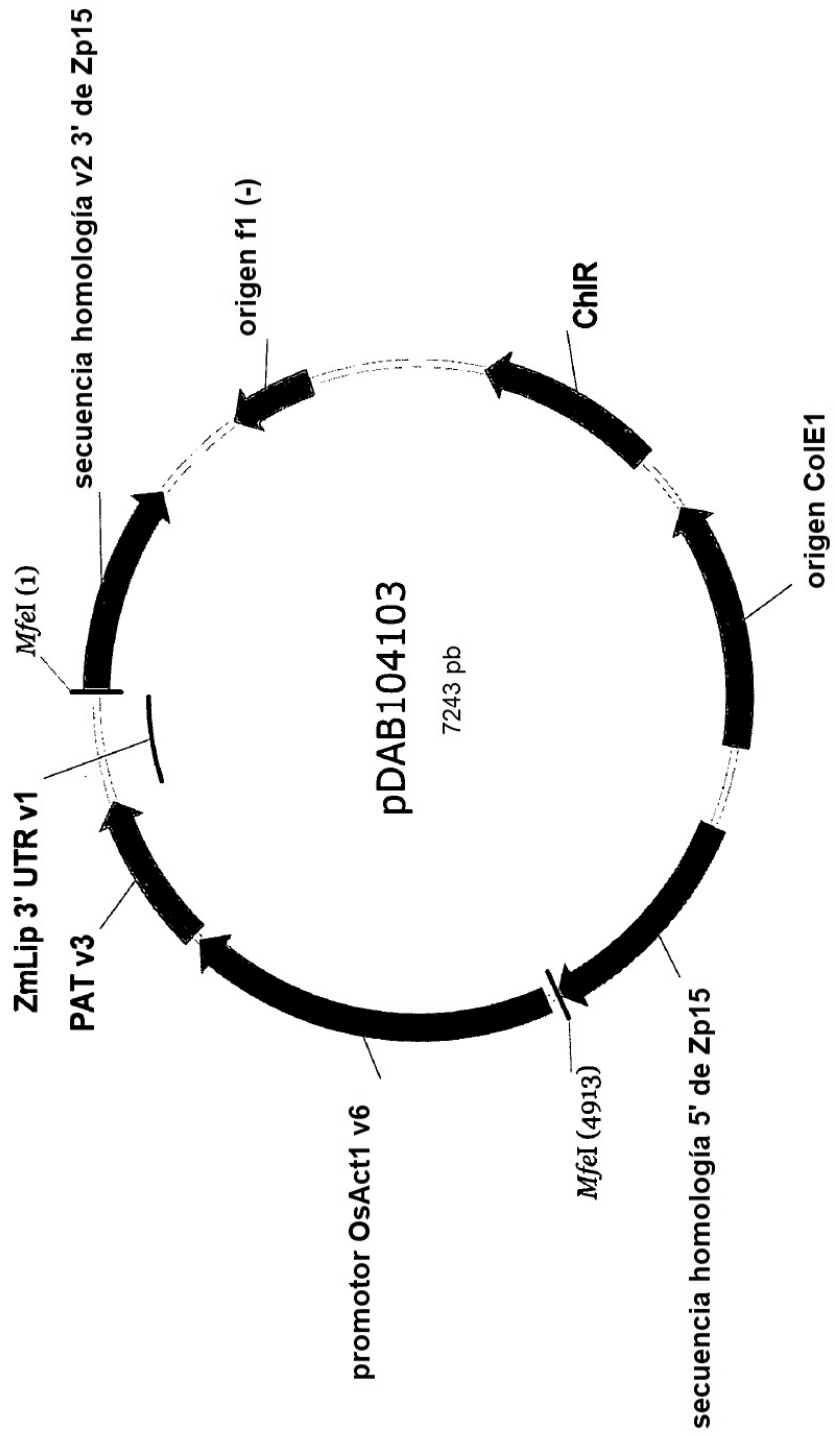


Figura 14

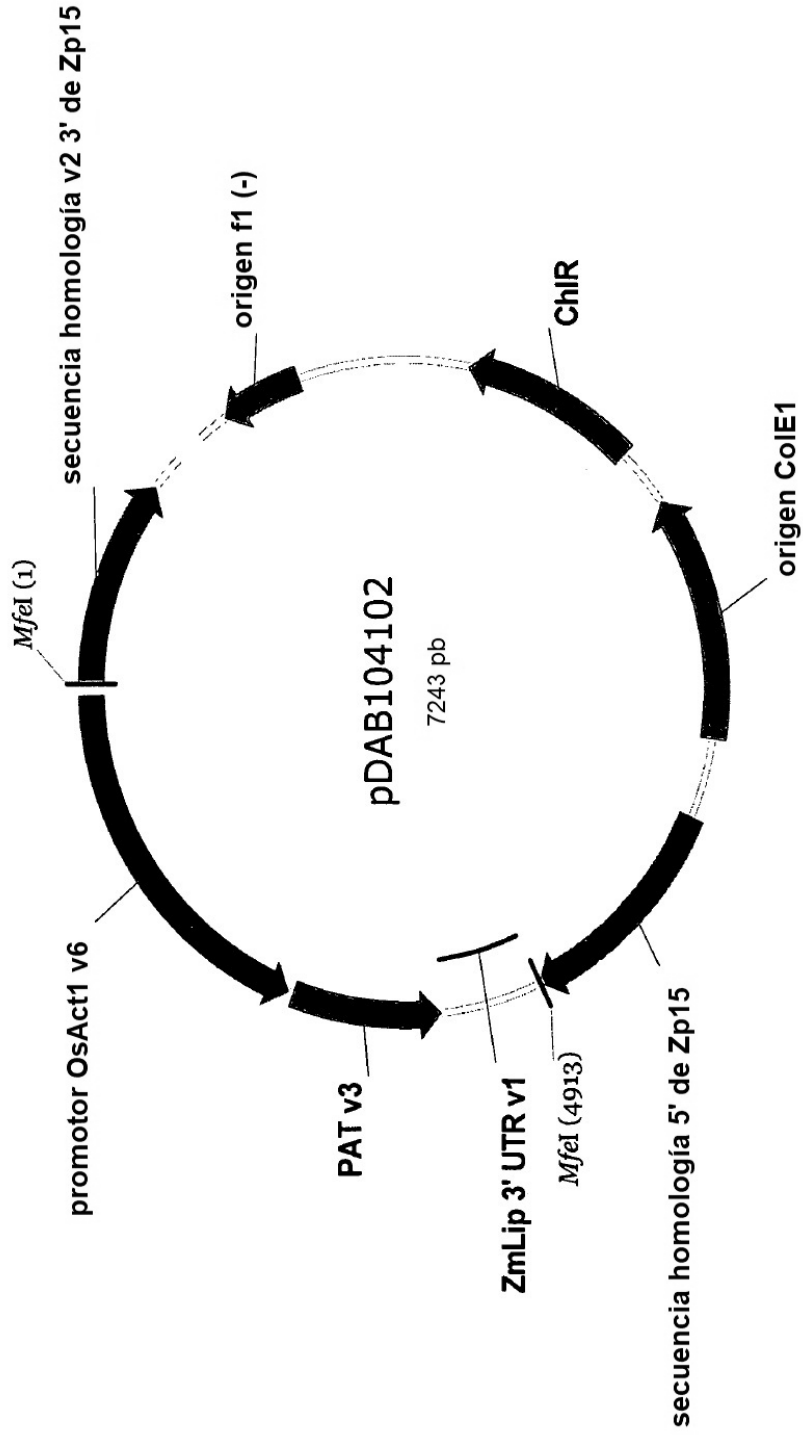


Figura 15